George Yin
Thaleia Zariphopoulou *Editors*

Stochastic Analysis, Filtering, and Stochastic Optimization

A Commemorative Volume to Honor Mark H. A. Davis's Contributions

Springer

Stochastic Analysis, Filtering, and Stochastic Optimization

George Yin • Thaleia Zariphopoulou
Editors

Stochastic Analysis, Filtering, and Stochastic Optimization

A Commemorative Volume to Honor Mark H. A. Davis's Contributions

Springer

Editors
George Yin
Department of Mathematics
University of Connecticut
Storrs, CT, USA

Thaleia Zariphopoulou
Department of Mathematics
University of Texas at Austin
Austin, TX, USA

ISBN 978-3-030-98518-9 ISBN 978-3-030-98519-6 (eBook)
https://doi.org/10.1007/978-3-030-98519-6

Mathematics Subject Classification (2020): 60Hxx, 60H30, 93E20, 49J55, 49K45

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Mark H.A. Davis was an eminent mathematician and so much more. He founded new theories and produced trailblazing results in stochastic analysis and stochastic optimization, and in mathematical finance. This volume is dedicated to his contributions in the first two areas.

Mark was born on April 25, 1945 and studied at Cambridge. He then left for Berkeley and obtained his Ph.D. under Pravin Varaiya in Electrical Engineering. His thesis was on dynamic programming in non-Markovian models, for which he developed a novel approach now known as optional Doob-Meyer decomposition theorem. This result opened up new ways to treat non-Markovian settings and created many new research lines. It initiated the martingale theory for stochastic optimization which became the main approach to study these problems. Much later, these results also influenced considerably the development of the field of mathematical finance.

Mark produced pioneering results in many areas in stochastic analysis and stochastic optimization, beyond the aforementioned martingale theory. It is difficult to list all of them, given the significant breadth and depth of his works. We note his ground-breaking contributions to the general theory of jump processes and the development of the novel pathwise non-linear filtering theory. Mark also developed the theory of piecewise deterministic processes which, besides their core contribution to stochastics, contributed considerably to the analysis of problems in actuarial science. Later on, Mark developed a deterministic approach to stochastic optimization using appropriate Lagrange multipliers. This method later became one of the main approaches to analyze optimization problems in financial mathematics.

Mark authored five books and co-authored six more on stochastic analysis, optimization and finance, and wrote close to two hundred other academic papers. He was Editor-in-Chief of Stochastics and Stochastics Reports (1978–1995), a founding Co-Editor of Mathematical Finance (1990–1993), an Associate Editor of the Annals of Applied Probability (1995–1998), an Associate Editor of Quantitative Finance (2000–2020), and an Associate Editor of the SIAM Journal of Financial Mathematics (2009–2020). He was also a highly influential editorial board member of the book series Springer Finance for 15 years, from 2001 to 2016.

Mark received the Naylor Prize in Applied Mathematics by the London Mathematical Society in 2002. He was also elected a Fellow of the Royal Statistical Society, a Fellow of the Institute of Mathematical Statistics, and an Honorary Fellow of the Institute of Actuaries.

Besides his towering academic stature, Mark was very much adored by our academic community. For those of us lucky enough to have known Mark, we will remember him most for his sharp and witty mind, provoking discussions, kindness and generosity, contagious laughter, and above all, old-fashioned academic nobility.

Mark passed away on March 18, 2020, at the age of 74. He had many interests beside academia. He enjoyed playing music and travelling, always accompanied by his beloved wife Jessica.

It is very difficult to capture the magnitude of Marks legacy. This volume honors him with eighteen papers by collaborators of his as well as by other academics whose research was very much influenced by his results. The papers cover a wide array of topics, offering new and survey results in stochastic analysis and stochastic control. We feel very honored to have been given the opportunity to edit this volume and we are so much grateful to all the authors who contributed their work. We are deeply indebted to Jan Obloj for providing us with the bibliography of Mark Davis, and to Jessica Smith-Davis for providing us with Mark's photo and the title page of Mark's dissertation. Finally, we thank Donna Chernyk and the Springer professionals for helping us to finalize the book.

George Yin Thaleia Zariphopoulou
October 2021

The original version of this book was revised: "The original version of the book was inadvertently published with incorrect abstracts in the chapters. This has now been amended". The correction to this chapter is available at https://doi.org/10.1007/978-3-030-98519-6_19.

Table of Contents

18

MD

University of California, Berkeley

Electronics Research Laboratory
Department of Electrical Engineering & Computer Sciences

Dynamic Programming Conditions for Partially Observable Stochastic Systems

BY

M. H. A. Davis

August, 1971

Bibliography of Mark H. A. Davis

Books

Davis, M. H. A. (1977b). *Linear estimation and stochastic control*. Chapman & Hall, London, Halsted Press [John Wiley & Sons], New York.

Davis, M. H. A. (1984a). *Lectures on stochastic control and nonlinear filtering* (Vol. 75). Published for the Tata Institute of Fundamental Research, Bombay; by Springer-Verlag, Berlin.

Davis, M. H. A. (1984b). *Lineinoe Otsenivanie i Stokhasticheskoe Upravlenie* [Translated from the English by M. V. Burnashev and A. A. Novikov, Translation edited and with a preface by A. N. Shiryaev]. "Nauka", Moscow.

Davis, M. H. A., & Vinter, R. B. (1985). *Stochastic modelling and control*. Chapman & Hall, London. https://doi.org/10.1007/978-94-009-4828-0

Davis, M. H. A., & Elliott, R. J. (1991). *Applied stochastic analysis*. Gordon & Breach Science Publishers.

Davis, M. H. A. (1993b). *Markov models and optimization* (Vol. 49). Chapman & Hall, London. https://doi.org/10.1007/978-1-4899-4483-2

Davis, M. H. A., Duffie, D., Fleming, W. H., & Shreve, S. (1995). *Mathematical finance* (Vol. 65). Springer-Verlag.

Davis, M., & Etheridge, A. (2006). *Louis Bachelier's theory of speculation: The origins of modern finance*. Princeton University Press.

Davis, M. H. A., & Lleo, S. (2014). *Risk-sensitive investment management* (Vol. 19). World Scientific.

Buchanan, D. L., & Davis, M. H. A. (2018). *Metals and energy finance: Application of quantitative finance techniques to the evaluation of minerals, coal and petroleum projects*. World Scientific Publishing.

Davis, M. H. A. (2019). *Mathematical finance: A very short introduction*. Oxford University Press, Oxford.

Published Papers

Davis, M. H. A., & Varaiya, P. P. (1972). Information states for linear stochastic systems. *J. Math. Anal. Appl.*, *37*, 384–402. https://doi.org/10.1016/0022-247X(72)90281-8

Davis, M. H. A. (1973a). On the existence of optimal policies in stochastic control. *SIAM J. Control*, *11*, 587–594.

Davis, M. H. A. (1973b). Optimal control of a degenerate Markovian system. In D. J. Bell (Ed.), *Recent mathematical developments in control*. Academic Press Inc.

Davis, M. H. A., & Varaiya, P. (1973). Dynamic programming conditions for partially observable stochastic systems. *SIAM J. Control*, *11*, 226–261. https://doi.org/10.1137/0311020

Davis, M. (1974). On the separation principle. *Proceedings of the 5th Symposium on Nonlinear Estimation and Applications*.

Davis, M. H. A., Segall, A., & Kailath, T. (1974). Estimation of point processes. *Proceedings of IFAC Symposium on Stochastic Control*.

Davis, M. H. A., & Varaiya, P. (1974). The multiplicity of an increasing family of σ-fields. *Ann. Probability*, *2*, 958–963. https://doi.org/10.1214/aop/1176996562

Davis, M. (1975a). The application of nonlinear filtering to fault detection in linear systems. *IEEE Transactions on Automatic Control*, *20*(2), 257–259. https://doi.org/10.1109/TAC.1975.1100908

Davis, M. H. A. (1975b). On stochastic differentiation. *Teor. Verojatnost. i Primenen.*, *20*(4), 887–892.

Segall, A., Davis, M. H. A., & Kailath, T. (1975). Nonlinear filtering with counting observations. *IEEE Trans. Inform. Theory*, *IT-21*, 143–149. https://doi.org/10.1109/tit.1975.1055360

Davis, M. H. A. (1976a). Martingales of Wiener and Poisson processes. *J. London Math. Soc. (2)*, *13*(2), 336–338. https://doi.org/10.1112/jlms/s2-13.2.336

Davis, M. H. A. (1976b). A note on the Poisson disorder problem. *Banach Center Publications*, *1*(1), 65–72.

Davis, M. H. A. (1976c). The representation of martingales of jump processes. *SIAM J. Control Optim.*, *14*(4), 623–638. https://doi.org/10.1137/0314041

Davis, M. H. A. (1976d). The separation principle in stochastic control via Girsanov solutions. *SIAM J. Control Optim.*, *14*(1), 176–188. https://doi.org/10.1137/0314015

Davis, M. H. A. (1976e). The structure of jump processes and related control problems. *Math. Programming Stud.*, (6), 2–14. https://doi.org/10.1007/bfb0120740

Davis, M., & Andreadakis, E. (1977). Exact and approximate filtering in signal detection: An example (corresp.) *IEEE Transactions on Information Theory*, *23*(6), 768–772. https://doi.org/10.1109/TIT.1977.1055783

Davis, M. H. A., & Elliott, R. (1977). Optimal control of a jump process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, *40*(3), 183–202. https://doi.org/10.1007/BF00736046

Wan, C. B., & Davis, M. H. A. (1977). The general point process disorder problem. *IEEE Trans. Inform. Theory*, *IT-23*(4), 538–540. https://doi.org/10.1109/tit.1977.1055748

Davis, M. H. A. (1978a). Detection, mutual information and feedback encoding: Applications of stochastic calculus. *Communication systems and random process theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977)*, 705–720. NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., No. 25.

Davis, M. H. A. (1978b). A direct proof of innovations/observations equivalence for Gaussian processes. *IEEE Trans. Inform. Theory*, *24*(2), 252–254. https://doi.org/10.1109/TIT.1978.1055845

Davis, M. H. A. (1978c). Martingale integrals and stochastic calculus. *Communication systems and random process theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977)*, 687–704. NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., No. 25.

Davis, M. H. A. (1978d). Nonlinear semigroups in the control of partially-observable stochastic systems. *Measure theory applications to stochastic analysis (Proc. Conf., Res. Inst. Math., Oberwolfach, 1977)*, *695*, 37–49. https://doi.org/10.1007/BFb0062653

Davis, M. H. A., & Wan, C. (1978). The principle of optimality for Markov jump processes. *IMA Conference on Analysis and Optimization of Stochastic Systems, Oxford*.

Davis, M. H. A. (1979a). Martingale methods in stochastic control. *Stochastic control theory and stochastic differential systems (Proc. Workshop, Deutsch. Forschungsgemeinsch., Univ. Bonn, Bad Honnef, 1979)*, *16*, 85–117. https://doi.org/10.1007/BFb0009377

Davis, M. H. A. (1979b). Pathwise solutions and multiplicative functionals in nonlinear filtering. *1979 18th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, *2*, 176–181. https://doi.org/10.1109/CDC.1979.270157

Davis, M. H. A., & Clark, J. M. C. (1979). On "predicted miss" stochastic control problems. *Stochastics*, *2*(3), 197–209. https://doi.org/10.1080/17442507908833126

Wan, C. B., & Davis, M. H. A. (1979). Existence of optimal controls for stochastic jump processes. *SIAM J. Control Optim.*, *17*(4), 511–524. https://doi.org/10.1137/0317037

Davis, M. (1980a). Review of "stochastic optimal control: The discrete case" by D. Bertsekas and S. E. Shreve. *IEEE Trans Automatic Control*, *25*, 1254–1255.

Davis, M. H. A. (1980b). Capacity and cutoff rate for Poisson-type channels. *IEEE Trans. Inform. Theory*, *26*(6), 710–715. https://doi.org/10.1109/TIT.1980.1056262

Davis, M. H. A. (1980c). Comments on: "Weaker conditions for innovations informational equivalence in the independent Gaussian case" [IEEE Trans. Automat. Control **24** (1979), no. 1, 63–69; MR 80h:93079] by E. Mosca [With a reply by Mosca]. *IEEE Trans. Automat. Control*, *25*(3), 607–609. https://doi.org/10.1109/TAC.1980.1102388

Davis, M. H. A. (1980d). Functionals of diffusion processes as stochastic integrals. *Math. Proc. Cambridge Philos. Soc.*, *87*(1), 157–166. https://doi.org/10.1017/S0305004100056590

Davis, M. H. A. (1980e). On a multiplicative functional transformation arising in nonlinear filtering theory. *Z. Wahrsch. Verw. Gebiete*, *54*(2), 125–139. https://doi.org/10.1007/BF00531444

Davis, M. H. A., & Kohlmann, M. (1980). Stochastic control by measure transformation: A general existence result. *Inform. Sci.*, *21*(3), 195–208. https://doi.org/10.1016/0020-0255(80)90029-8

Davis, M. H. A., & Wan, C. B. (1980). The principle of optimality for Markov jump processes. *Analysis and optimisation of stochastic systems (Proc. Internat. Conf., Univ. Oxford, Oxford, 1978)*, 21–33.

Davis, M. H. A., & Wellings, P. H. (1980). Computational problems in nonlinear filtering. *Analysis and optimization of systems (Proc. Fourth Internat. Conf., Versailles, 1980)*, *28*, 253–261. https://doi.org/10.1007/BFb0004045

Davis, M. (1981a). Review of "stochastic models, estimation and control" by P. S. Maybeck. *Proceedings of IEE (D)*, (128).

Davis, M. H. A. (1981b). Factorization of a multiplicative functional of nonlinear filtering theory. *Systems Control Lett.*, *1*(1), 49–53. https://doi.org/10.1016/S0167-6911(81)80012-6

Davis, M. H. A. (1981c). New approach to filtering for nonlinear systems. *Proc. IEE-D*, *128*(5), 166–172. https://doi.org/10.1049/ip-d.1981.0037

Davis, M. H. A. (1981d). Pathwise non-linear filtering. In M. Hazewinkel & J. Williams (Eds.), *Stochastic systems: The mathematics of filtering and identification and applications* (pp. 505–528). Springer. https://doi.org/10.1007/978-94-009-8546-9_25

Davis, M. H. A., & Marcus, S. I. (1981). An introduction to nonlinear filtering. In M. Hazewinkel & J. Williams (Eds.), *Stochastic systems: The mathematics of filtering and identification and applications* (pp. 53–75). Springer. https://doi.org/10.1007/978-94-009-8546-9_4

Elliott, R. J., & Davis, M. H. A. (1981). Optimal play in a stochastic differential game. *SIAM J. Control Optim.*, *19*(4), 543–554. https://doi.org/10.1137/0319033

Davis, M. H. A. (1982a). A pathwise solution of the equations of nonlinear filtering. *Teor. Veroyatnost. i Primenen.*, *27*(1), 160–167. https://doi.org/10.1137/1127017

Davis, M. H. A. (1982b). Stochastic control with noisy observations. *Advances in filtering and optimal stochastic control (Cocoyoc, 1982)* (pp. 79–90). Springer, Berlin. https://doi.org/10.1007/BFb0004527

Davis, M. H. A. (1982c). Stochastic control with tracking of exogenous parameters. *Stochastic differential systems (Bad Honnef, 1982)* (pp. 285–293). Springer, Berlin. https://doi.org/10.1007/BFb0044308

Gal'Čuk, L. I., & Davis, M. H. A. (1982). A note on a comparison theorem for equations with different diffusions. *Stochastics*, *6*(2), 147–149. https://doi.org/10.1080/17442508208833200

Davis, M. H. A. (1983a). The martingale theory of point processes and its application to the analysis of failure-time data. *Electronic systems effectiveness and life cycle costing (Norwich, 1982)* (pp. 135–155). Springer, Berlin. https://doi.org/10.1007/978-3-642-82014-4_9

Davis, M. H. A. (1983b). On a problem of D. R. Cox. *Fifth international conference on collective phenomena* (pp. 129–132). New York Acad. Sci., New York. https://doi.org/10.1111/j.1749-6632.1983.tb23310.x

Davis, M. H. A. (1984c). Optimization of piecewise-deterministic Markov models. *Advances in Applied Probability*, *16*(1), 14–14. https://doi.org/10.1017/S0001867800022011

Davis, M. H. A. (1984d). Piecewise-deterministic Markov processes: A general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, *46*(3), 353–388. https://doi.org/10.1111/j.2517-6161.1984.tb01308.x

Davis, M. H. A. (1985). Point process models: Some current techniques. *Biorhythms and Epilepsy: papers from a Workshop*, 59.

Davis, M. H. A., Spathopoulos, M. P., & Yoneyama, T. (1985). On the stochastic maximum principle for manifold-valued processes. *24th IEEE Conference on Decision and Control*, 1337–1342. https://doi.org/10.1109/CDC.1985.268725

Carvalhais, Z., & Davis, M. H. A. (1986). Optimal timing of capacity expansion. *J Economic Dynamics and Control*, 89–92. https://doi.org/10.1016/0165-1889(86)90023-0

Davis, M. (1986a). Review of "recursive estimation and control for stochastic systems" by Han-Fu Chen. *Proceedings of IEE (D)*.

Davis, M. H. A. (1986b). Control of piecewise-deterministic processes via discrete-time dynamic programming. *Stochastic differential systems (Bad Honnef, 1985)* (pp. 140–150). Springer, Berlin. https://doi.org/10.1007/BFb0041157

Davis, M. H. A., & Spathopoulos, M. P. (1986). Pathwise nonlinear filtering with observations on a manifold. *25th IEEE Conference on Decision and Control*, 1028–1029. https://doi.org/10.1109/CDC.1986.267531

Davis, M. (1987a). Optimal stochastic control: General aspects. In M. G. Singh (Ed.), *Encyclopaedia of systems and control*. Pergamon Press Oxford.

Davis, M. H. A. (1987b). Nonlinear filtering and stochastic flows. *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*, 1000–1010.

Davis, M. H. A. (1987c). Optimal control of piecewise-deterministic Markov processes. *Differential equations and applications, I, II (Russian) (Ruse, 1985)* (pp. 561–568). 'Angel Kanchev' Tech. Univ., Ruse.

Davis, M. H. A., Dempster, M. A. H., Sethi, S. P., & Vermes, D. (1987). Optimal capacity expansion under uncertainty. *Adv. in Appl. Probab.*, *19*(1), 156–176. https://doi.org/10.2307/1427378

Davis, M. H. A., & Gómez, G. L. (1987). The martingale maximum principle and the allocation of labour surplus [Eighth conference of the Society of Economic Dynamics and Control (Budapest, 1986)]. *J. Econom. Dynam. Control*, *11*(2), 241–247. https://doi.org/10.1016/0165-1889(87)90015-7

Davis, M. H. A., & Gómez M., G. L. (1987). The semimartingale approach to the optimal resource allocation in the controlled labour-surplus economy. *Stochastic processes—mathematics and physics, II (Bielefeld, 1985)* (pp. 36–74). Springer, Berlin. https://doi.org/10.1007/BFb0077347

Davis, M. H. A., & Spathopoulos, M. P. (1987). Pathwise nonlinear filtering for nondegenerate diffusions with noise correlation. *SIAM J. Control Optim.*, *25*(2), 260–278. https://doi.org/10.1137/0325016

Costa, O. L. V., & Davis, M. H. A. (1988). Approximations for optimal stopping of a piecewise-deterministic process. *Math. Control Signals Systems*, *1*(2), 123–146. https://doi.org/10.1007/BF02551405

Davis, M. H. A. (1988a). Local time on the stock exchange. *Stochastic calculus in application (Cambridge, 1987)* (pp. 4–28). Longman Sci. Tech., Harlow.

Davis, M. H. A. (1988b). The Wiener space derivative for functionals of diffusions on manifolds. *Nonlinearity*, *1*(1), 241–251. https://doi.org/10.1088/0951-7715/1/1/010

do Val, J. B. R., & Davis, M. H. A. (1988). Local optimality criteria for stochastic control of piecewise deterministic processes. *Rebrape*, *2*(2), 101–123.

Costa, O. L. V., & Davis, M. H. A. (1989). Impulse control of piecewise-deterministic processes. *Math. Control Signals Systems*, *2*(3), 187–206. https://doi.org/10.1007/BF02551384

Davis, M., & Zucker, M. S. (1989a). An image analysis method for closing the roll loop. *Proceedings of 28th IEEE Conference on Decision and Control*, 2371–2372.

Davis, M. H. A. (1989a). Anticipative LQG control. *IMA J. Math. Control Inform.*, *6*(3), 259–265. https://doi.org/10.1093/imamci/6.3.259

Davis, M. H. A. (1989b). Necessary conditions for optimality in the control of piecewise-deterministic Markov processes. *Proceedings of the 28th IEEE Conference on Decision and Control, Vol. 1–3 (Tampa, FL, 1989)*, 719–723. https://doi.org/10.1109/CDC.1989.70211

Davis, M. H. A. (1989c). A review of the statistical theory of signal detection (B. F. Schutz, Ed.). *Gravitational Wave Data Analysis*, 73–94. https://doi.org/10.1007/978-94-009-1185-7_6

Davis, M. H. A., & Hemerly, E. M. (1989). Order determination and adaptive control of ARX models using the PLS criterion. *Stochastic differential systems (Bad Honnef, 1988)* (pp. 91–101). Springer, Berlin. https://doi.org/10.1007/BFb0043776

Davis, M. H. A., & Spathopoulos, M. P. (1989). On the minimum principle for controlled diffusions on manifolds. *SIAM J. Control Optim.*, *27*(5), 1092–1107. https://doi.org/10.1137/0327058

Davis, M. H. A., & Zucker, M. (1989b). An image analysis method for vehicle stabilization. *Proceedings of the 28th IEEE Conference on Decision and Control,* 2371–2372 vol.3. https://doi.org/10.1109/CDC.1989.70599

Hemerly, E. M., & Davis, M. H. A. (1989a). Recursive order estimation of stochastic control systems. *Math. Systems Theory*, *22*(4), 323–346. https://doi.org/10.1007/BF02088305

Hemerly, E. M., & Davis, M. H. A. (1989b). Strong consistency of the PLS criterion for order determination of autoregressive processes. *Ann. Statist.*, *17*(2), 941–946. https://doi.org/10.1214/aos/1176347154

Lasdas, V., & Davis, M. H. A. (1989). A piecewise deterministic process approach to target motion analysis (sonar). *Proceedings of the 28th IEEE Conference on Decision and Control,* 1395–1396 vol.2. https://doi.org/10.1109/CDC.1989.70369

Davis, M. H. A., & Norman, A. R. (1990). Portfolio selection with transaction costs. *Math. Oper. Res.*, *15*(4), 676–713. https://doi.org/10.1287/moor.15.4.676

Davis, M. H. A. (1991c). Anticipative LQG control. II. *Applied stochastic analysis (London, 1989)* (pp. 205–214). Gordon & Breach, New York.

Davis, M. H. A., Dempster, M. A. H., & Elliott, R. J. (1991). On the value of information in controlled diffusion processes. *Stochastic analysis* (pp. 125–138). Academic Press, Boston, MA.

Davis, M. H. A., & Fotopoulos, P. (1991). On the approximate stochastic realisation problem. *Proceedings of the 30th IEEE Conference on Decision and Control*, 1698–1699. https://doi.org/10.1109/CDC.1991.261697

Davis, M. H. A., Lasdas, S., & Salmond, D. (1991). Clustering algorithms for bayesian fault detection in linear systems. *Proceedings of the 30th IEEE Conference on Decision and Control*, 674–675 vol.1. https://doi.org/10.1109/CDC.1991.261394

Davis, M., & Burstein, G. (1991). Anticipative stochastic control. *Proceedings of the 30th IEEE Conference on Decision and Control*, *2*, 1830–1835. https://doi.org/10.1109/CDC.1991.261728

Davis, M., & Panas, V. (1991). European option pricing with transactions costs. *Proceedings of the 30th IEEE Conference on Decision and Control*, *2*, 1299–1304. https://doi.org/10.1109/CDC.1991.261597

Hemerly, E. M., & Davis, M. H. A. (1991). Recursive order estimation of autoregressions without bounding the model set. *J. Roy. Statist. Soc. Ser. B*, *53*(1), 201–210. https://doi.org/10.1111/j.2517-6161.1991.tb01818.x

Davis, M. H. A. (1992). Dynamic optimization: A grand unification. *Proceedings of the 31st IEEE Conference on Decision and Control*, 2035–2036. https://doi.org/10.1109/CDC.1992.371439

Davis, M. H. A., & Burstein, G. (1992a). On the relation between deterministic and stochastic optimal control. *Stochastic partial differential equations and applications (Trento, 1990)* (pp. 124–157). Longman Sci. Tech., Harlow.

Davis, M. H. A., & Burstein, G. (1992b). On the relation between partially observed, stochastic optimal control and deterministic infinite dimensional optimal control. *Proceedings of the 31st IEEE Conference on Decision and Control*, 2400–2405 vol.2. https://doi.org/10.1109/CDC.1992.371360

Davis, M. H. A., & Burstein, G. (1992c). Random conservation laws and global solutions of nonlinear SPDE—application to the HJB SPDE of anticipative control. *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)* (pp. 43–53). Springer, Berlin. https://doi.org/10.1007/BFb0007319

Davis, M. H. A., & Burstein, G. (1992d). A deterministic approach to stochastic optimal control with application to anticipative control. *Stochastics Stochastics Rep.*, *40*(3-4), 203–256. https://doi.org/10.1080/17442509208833790

Davis, M. H. A., & Lasdas, V. (1992). Reducibility and unobservability of Markov processes: The linear system case. *IEEE Trans. Automat. Control*, *37*(4), 505–508. https://doi.org/10.1109/9.126587

Davis, M. H. A., & Zheng, W. (1992). A stochastic realisation approach to order determination in time series modeling. *Proceedings of the 31st IEEE Conference on Decision and Control*, 3690–3691 vol.4. https://doi.org/10.1109/CDC.1992.370960

Davis, M. (1993a). Translation of "Stratonovich integral". In M. Hazewinkel (Ed.), *Soviet encyclopaedia of mathematics*. Kluwer.

Davis, M. H. A., Panas, V. G., & Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM J. Control Optim.*, *31*(2), 470–493. https://doi.org/10.1137/0331022

Biron, M., & Davis, M. H. A. (1994). A stochastic realization approach to ar model order determination using partial autocorrelation coefficients. *Proceedings of 1994 IEEE International Symposium on Information Theory*, 297–. https://doi.org/10.1109/ISIT.1994.394721

Davis, M. (1994). Review of "controlled Markov processes and viscosity solutions" by W. H. Fleming and H. M. Soner. *Bull. AMS*, *31*, 75–85.

Davis, M. H. A., & Clark, J. M. C. (1994). A note on super-replicating strategies. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, *347*(1684), 485–494. https://doi.org/10.1098/rsta.1994.0058

Davis, M. H. A., & Howl, J. (1994). Exact computations for queueing network performance using piecewise-deterministic Markov models. *Proceedings of 1994 IEEE International Symposium on Information Theory*, 374–. https://doi.org/10.1109/ISIT.1994.395020

Davis, M. H. A., & Karatzas, I. (1994). A deterministic approach to optimal stopping. *Probability, statistics and optimisation* (pp. 455–466). Wiley, Chichester.

Davis, M. H. A., & Panas, V. G. (1994). The writing price of a European contingent claim under proportional transaction costs. *Mat. Apl. Comput.*, *13*(2), 115–157.

Davis, M. H. A., & Zervos, M. (1994). A problem of singular stochastic control with discretionary stopping. *Ann. Appl. Probab.*, *4*(1), 226–240. https://doi.org/10.1214/aoap/1177005209

Davis, M. H. A. (1995). European option pricing via a utility maximization method. In S. A. Zenios (Ed.), *Quantitative methods, super computers and AI in finance*. Stanley Thornes in association with UNICOM.

Davis, M. H. A., & Vellekoop, M. H. (1995). Permanent health insurance: A case study in piecewise-deterministic Markov modelling. *Schweiz. Verein. Versicherungsmath. Mitt.*, (2), 177–212.

Davis, M. H. A., & Zariphopoulou, T. (1995). American options and transaction fees. *Mathematical finance* (p. 47). Springer-Verlag.

Davis, M. H. A., & Zervos, M. (1995). A new proof of the discrete-time LQG optimal control theorems. *IEEE Trans. Automat. Control*, *40*(8), 1450–1453. https://doi.org/10.1109/9.402239

Davis, M. H. A., Beumee, J. G. B., & Lee, H. (1996). Interest rate models. *Introduction to Mitsubishi finance risk directory*. Risk Publications.

Davis, M. H. A., & Farid, M. (1996). A target recognition problem: Sequential analysis and optimal control. *SIAM J. Control Optim.*, *34*(6), 2116–2132. https://doi.org/10.1137/S0363012994273696

Davis, M. (1997). A note on the forward measure. *Finance and stochastics*, *2*(1), 19–28. https://doi.org/10.1007/s007800050030

Davis, M. H. A., & Howl, J. M. (1997). A Markovian analysis of the finite-buffer $M/D/1$ queue. *Proc. Roy. Soc. London Ser. A*, *453*(1964), 1947–1962. https://doi.org/10.1098/rspa.1997.0105

Davis, M. H. A., & Zheng, W. X. (1997). A new order estimation technique for time series modeling. *IEEE Trans. Automat. Control*, *42*(3), 400–403. https://doi.org/10.1109/9.557584

Davis, M. H. A., & Farid, M. (1998). Corrections to: "A target recognition problem: Sequential analysis and optimal control" [SIAM J. Control Optim. **34** (1996), no. 6, 2116–2132]. *SIAM J. Control Optim.*, *36*(4), 1481–1484. https://doi.org/10.1137/S0363012997320304

Davis, M. H. A., & Zervos, M. (1998). A pair of explicitly solvable singular stochastic control problems. *Appl. Math. Optim.*, *38*(3), 327–352. https://doi.org/10.1007/s002459900094

Davis, M. H. A., & Farid, M. (1999). Piecewise-deterministic processes and viscosity solutions. *Stochastic analysis, control, optimization and applications* (pp. 249–268). Birkhäuser Boston, Boston, MA. https://doi.org/10.1007/978-1-4612-1784-8_15

Farid, M., & Davis, M. H. A. (1999). Optimal consumption and exploration: A case study in piecewise-deterministic Markov modelling. *Ann. Oper. Res.*, *88*, 121–137. https://doi.org/10.1023/A:1018918027153

Davis, M. (2000a). Taking convertibles for a spin. *Derivatives Week*, *9*(19).

Davis, M. H. A. (2000b). Option valuation and hedging with basis risk. *System theory: Modeling, analysis and control (Cambridge, MA, 1999)* (pp. 245–

254). Kluwer Acad. Publ., Boston, MA. https://doi.org/10.1007/978-1-4615-5223-9_18

Davis, M., & Lo, V. (2001a). Infectious defaults. *Quantitative Finance*, *1*(4), 382. https://doi.org/10.1080/713665832

Davis, M., & Lo, V. (2001b). Modelling default correlation in bond portfolios. *Mastering risk*, *2*(1), 141–151.

Davis, M. H. A. (2001a). Mathematics of financial markets. In B. Engquist & W. Schmid (Eds.), *Mathematics unlimited — 2001 and beyond* (pp. 361–380). Springer, Berlin. https://doi.org/10.1007/978-3-642-56478-9_16

Davis, M. H. A. (2001b). Pricing weather derivatives by marginal value. *Quant. Finance*, *1*(3), 305–308. https://doi.org/10.1088/1469-7688/1/3/201

Davis, M. H. A., Schachermayer, W., & Tompkins, R. G. (2001a). Installment options and static hedging. In M. Kohlmann & S. Tang (Eds.), *Mathematical finance (Konstanz, 2000)* (pp. 131–139). Birkhäuser, Basel. https://doi.org/10.1007/978-3-0348-8291-0_12

Davis, M. H. A., Schachermayer, W., & Tompkins, R. G. (2001b). Pricing, no-arbitrage bounds and robust hedging of instalment options. *Quant. Finance*, *1*(6), 597–610. https://doi.org/10.1088/1469-7688/1/6/302

Fonseca, I., Grötschel, M., & Davis, M. H. A. (2001). Round table: The impact of mathematical research on industry and vice versa. In C. Cadacuberta, R. M. Miró-Roig, J. M. Ortega, & S. Xambó des Camps (Eds.), *Mathematical glimpses into the 21st century* (pp. 35–53).

Hodges, S., Roberts, G., Papaspiliopoulos, O., Sentana, E., Bingham, N., Cox, D., Nicolato, E., Venardos, E., Critchley, F., Davis, M. H. A., et al. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics — discussion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*, 208–241. https://doi.org/10.1111/1467-9868.00282

Davis, M. H. A., Schachermayer, W., & Tompkins, R. G. (2002). Installment options and static hedging. *Journal of Risk Finance*, *3*, 46–52.

Davis, M. H. A. (2004a). Complete-market models of stochastic volatility. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, *460*(2041), 11–26. https://doi.org/10.1098/rspa.2003.1233

Davis, M. H. A. (2004b). Valuation, hedging and investment in incomplete financial markets. *Applied mathematics entering the 21st century* (pp. 49–70). SIAM, Philadelphia, PA.

Davis, M. H. A., Schachermayer, W., & Tompkins, R. G. (2004). The evaluation of venture capital as an instalment option: Valuing real options using real options. In T. Dangl, M. Kopel, & W. Kürsten (Eds.), *Real options* (pp. 77–96). Gabler Verlag. https://doi.org/10.1007/978-3-663-12338-5_3

Pasik-Duncan, B., Elliott, R., & Davis, M. H. A. (2004). Guest editorial [Special issue on stochastic control methods in financial engineering]. *IEEE Trans. Automat. Control*, *49*(3), 321–323. https://doi.org/10.1109/TAC.2004.824472

Davis, M. H. A. (2005). Martingale representation and all that. *Advances in control, communication networks, and transportation systems* (pp. 57–68). Birkhäuser Boston, Boston, MA. https://doi.org/10.1007/0-8176-4409-1_4

Giampieri, G., Davis, M., & Crowder, M. (2005a). A hidden Markov model of default interaction. *Quantitative Finance*, *5*(1), 27–34.

Giampieri, G., Davis, M. H. A., & Crowder, M. (2005b). Analysis of default data using hidden Markov models. *Quant. Finance*, *5*(1), 27–34. https://doi.org/10.1080/14697680500039951

Davis, M. H. A. (2006). Optimal hedging with basis risk. *From stochastic calculus to mathematical finance* (pp. 169–187). Springer, Berlin. https://doi.org/10.1007/978-3-540-30788-4_8

Davis, M. H. A., & Johansson, M. P. (2006). Malliavin Monte Carlo Greeks for jump diffusions. *Stochastic Process. Appl.*, *116*(1), 101–129. https://doi.org/10.1016/j.spa.2005.08.002

Davis, M. H. A., & Esparragoza-Rodriguez, J. C. (2007). Large portfolio credit risk modeling. *Int. J. Theor. Appl. Finance*, *10*(4), 653–678. https://doi.org/10.1142/S0219024907004378

Davis, M. H. A., & Hobson, D. G. (2007). The range of traded option prices. *Math. Finance*, *17*(1), 1–14. https://doi.org/10.1111/j.1467-9965.2007.00291.x

Davis, M. H. A., & Mataix-Pastor, V. (2007). Negative Libor rates in the swap market model. *Finance Stoch.*, *11*(2), 181–193. https://doi.org/10.1007/s00780-006-0032-2

Davis, M. H. A., & Lleo, S. (2008). Risk-sensitive benchmarked asset management. *Quant. Finance*, *8*(4), 415–426. https://doi.org/10.1080/14697680701401042

Davis, M. H. A., & Obłój, J. (2008). Market completion using options. In Ł. Stettner (Ed.), *Advances in mathematics of finance* (pp. 49–60). Polish Acad. Sci. Inst. Math., Warsaw. https://doi.org/10.4064/bc83-0-4

Brody, D. C., Davis, M. H. A., Friedman, R. L., & Hughston, L. P. (2009). Informed traders. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, *465*(2104), 1103–1122. https://doi.org/10.1098/rspa.2008.0465

Davis, M. H. A., & Mataix-Pastor, V. (2009). Arbitrage-free interpolation of the swap curve. *Int. J. Theor. Appl. Finance*, *12*(7), 969–1005. https://doi.org/10.1142/S0219024909005543

Vellekoop, M., & Davis, M. (2009). An optimal investment problem with randomly terminating income. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, 3650–3655. https://doi.org/10.1109/CDC.2009.5400455

Becherer, D., & Davis, M. H. A. (2010). Arrow–Debreu prices. In R. Cont (Ed.), *Encyclopedia of quantitative finance*. Wiley. https://doi.org/10.1002/9780470061602.eqf04018

Björk, T., Davis, M. H. A., & Landén, C. (2010). Optimal investment under partial information. *Math. Methods Oper. Res.*, *71*(2), 371–399. https://doi.org/10.1007/s00186-010-0301-x

Davis, M., & Lleo, S. (2010). Risk sensitive investment management with affine processes: A viscosity approach. *Recent advances in financial engineering 2009* (pp. 1–41). https://doi.org/10.1142/9789814304078_0001

Davis, M. H. A. (2010a). Binomial tree. In R. Cont (Ed.), *Encyclopedia of quantitative finance*. Wiley. https://doi.org/10.1002/9780470061602.eqf05006

Davis, M. H. A. (2010b). Black–Scholes formula. In R. Cont (Ed.), *Encyclopedia of quantitative finance*. Wiley. https://doi.org/10.1002/9780470061602.eqf05004

Davis, M. H. A. (2010c). Options: Basic definitions. In R. Cont (Ed.), *Encyclopedia of quantitative finance*. Wiley. https://doi.org/10.1002/9780470061602.eqf05001

Davis, M. H. A., Guo, X., & Wu, G. (2010). Impulse control of multidimensional jump diffusions. *SIAM J. Control Optim.*, *48*(8), 5276–5293. https://doi.org/10.1137/090780419

Davis, M., & Lleo, S. (2011a). Fractional Kelly strategies for benchmarked asset management. In L. C. MacLean, E. O. Thorp, & W. T. Ziemba (Eds.), *The Kelly capital growth investment criterion: Theory and practice* (pp. 385–407). World Scientific. https://doi.org/10.1142/9789814293501_0027

Davis, M. H. A. (2011a). Contagion models in credit risk. *The Oxford handbook of credit derivatives*.

Davis, M. H. A. (2011b). Pathwise nonlinear filtering with correlated noise. *The Oxford handbook of nonlinear filtering* (pp. 403–424). Oxford Univ. Press, Oxford.

Davis, M. H. A., & Lleo, S. (2011b). Jump-diffusion risk-sensitive asset management I: Diffusion factor model. *SIAM J. Financial Math.*, *2*(1), 22–54. https://doi.org/10.1137/090760180

Andruszkiewicz, G., Davis, M. H. A., & Lleo, S. (2013). Taming animal spirits: Risk management with behavioural factors. *Ann. Finance*, *9*(2), 145–166. https://doi.org/10.1007/s10436-012-0217-y

Davis, M., & Lleo, S. (2013a). Fractional Kelly strategies in continuous time: Recent developments. In L. C. MacLean & W. T. Ziemba (Eds.), *Handbook of the fundamentals of financial decision making: Part II* (pp. 753–787). World Scientific.

Davis, M. H. A., & Lleo, S. (2013b). Black–Litterman in continuous time: The case for filtering. *Quantitative Finance Letters*, *1*(1), 30–35. https://doi.org/10.1080/21649502.2013.803794

Davis, M. H. A., & Lleo, S. (2013c). Jump-diffusion risk-sensitive asset management II: Jump-diffusion factor model. *SIAM J. Control Optim.*, *51*(2), 1441–1480. https://doi.org/10.1137/110825881

Davis, M. H. A., & Lleo, S. (2013d). Jump-diffusion risk-sensitive benchmarked asset management. *Stochastic programming* (pp. 97–127). World Sci. Publ., Hackensack, NJ. https://doi.org/10.1142/9789814407519_0005

Andruszkiewicz, G., Davis, M. H. A., & Lleo, S. (2014). Estimating animal spirits: Conservative risk calculation. *Quantitative Finance Letters*, *2*(1), 14–21. https://doi.org/10.1080/21649502.2014.946234

Davis, M. H. A., Obłój, J., & Raval, V. (2014). Arbitrage bounds for prices of weighted variance swaps. *Math. Finance*, *24*(4), 821–854. https://doi.org/10.1111/mafi.12021

Davis, M. H. A., & Lleo, S. (2015). Jump-diffusion asset-liability management via risk-sensitive control. *OR Spectrum*, *37*(3), 655–675. https://doi.org/10.1007/s00291-014-0371-x

Davis, M. H. A., & Pistorius, M. R. (2015). Explicit solution of an inverse first-passage time problem for Lévy processes and counterparty credit risk. *Ann. Appl. Probab.*, *25*(5), 2383–2415. https://doi.org/10.1214/14-AAP1051

Davis, M. H. A., & Yoshikawa, D. (2015a). A note on utility-based pricing. *Math. Financ. Econ.*, *9*(3), 215–230. https://doi.org/10.1007/s11579-015-0142-8

Davis, M. H. A., & Yoshikawa, D. (2015b). A note on utility-based pricing in models with transaction costs. *Math. Financ. Econ.*, *9*(3), 231–245. https://doi.org/10.1007/s11579-015-0143-7

Davis, M., & Lleo, S. (2016a). A simple procedure to incorporate predictive models in a continuous time asset allocation. *Quantitative Finance Letters*, *4*(1), 40–46.

Davis, M. H. A. (2016a). A Beaufort scale of predictability. In M. Podolskij, R. Stelzer, S. Thorbjrnsen, E. D. Almut, & A. E. D. Veraart (Eds.), *The fascination of probability, statistics and their applications. in honour of Ole E.Barndorff-Nielsen* (pp. 419–434). Springer, Cham. https://doi.org/10.1007/978-3-319-25826-3_19

Davis, M. H. A. (2016b). Model-free methods in valuation and hedging of derivative securities. *The handbook of post crisis financial modeling* (pp. 168–189). Springer. https://doi.org/10.1007/978-1-137-49449-8_7

Davis, M. H. A. (2016c). Verification of internal risk measure estimates. *Stat. Risk Model.*, *33*(3-4), 67–93. https://doi.org/10.1515/strm-2015-0007

Davis, M. H. A., & Lleo, S. (2016b). A simple procedure for combining expert opinion with statistical estimates to achieve superior portfolio performance. *The Journal of Portfolio Management*, *42*(4), 49–58. https://doi.org/10.3905/jpm.2016.42.4.049

Andruszkiewicz, G., Davis, M. H. A., & Lleo, S. (2017). Risk-sensitive investment in a finite-factor model. *Stochastics*, *89*(1), 89–114. https://doi.org/10.1080/17442508.2016.1139115

Davis, M. H. A. (2017). Discussion of “Elicitability and backtesting: Perspectives for banking regulation”. *Ann. Appl. Stat.*, *11*(4), 1886–1887. https://doi.org/10.1214/17-AOAS1041C

Davis, M. H. A., Obłój, J., & Siorpaes, P. (2018). Pathwise stochastic calculus with local times. *Ann. Inst. Henri Poincaré Probab. Stat.*, *54*(1), 1–21. https://doi.org/10.1214/16-AIHP792

Davis, M. H. A., & Lleo, S. (2020). Debiased expert forecasts in continuous-time asset allocation. *Journal of Banking & Finance*, *113*, 105759. https://doi.org/10.1016/j.jbankfin.2020.105759

Badikov, S., Davis, M. H. A., & Jacquier, A. (2021). Perturbation analysis of sub/super hedging problems [in this volume]. *Mathematical Finance*. https://doi.org/10.1111/mafi.12321

Davis, M. H. A., & Lleo, S. (2021). Risk-sensitive benchmarked asset management with expert forecasts [in this volume]. *Mathematical Finance*. https://doi.org/10.1111/mafi.12310

Unpublished Works

Davis, M. H. A. (1971). *Dynamic programming conditions for partially observable stochastic systems* [PhD thesis. UC Berkeley].

Davis, M. (1977a). *The martingale approach to stochastic control* (Report TRITA-MAT-1977-11). Royal Institute of Technology, Stockholm, Sweden.

Davis, M. (1991a). *Jump processes and their martingales* [Preprint available at https://www.duo.uio.no/bitstream/handle/10852/43612/1991-10.pdf].

Davis, M. (1991b). *Two quick derivations of the Black-Scholes option pricing formula* [Preprint available at https://www.duo.uio.no/bitstream/handle/10852/43637/1/Mark-Davis.pdf].

Davis, M. H. A., & Pistorius, M. (2010). *Quantification of counterparty risk via Bessel bridges* [Available at SSRN 1722604].

Crosby, J., & Davis, M. H. A. (2012). *Variance derivatives: Pricing and convergence* [Available at SSRN 2049278].

Davis, M. H. A. (2013). *Consistency of risk measure estimates* [Available at SSRN 2342279].

Davis, M. H. A., & Lleo, S. (2018). *Debiased expert opinions in continuous time asset allocation: Supplementary material* [Available at SSRN 3174744].

Obituaries and Legacy

Davis, M. H. A. (2011c, July 18). *RMS Research : Mark Davis, part 1*. https://www.youtube.com/watch?v=jlOAOizSd7k

Davis, M. H. A. (2011d, July 18). *RMS research : Mark Davis, part 2*. https://www.youtube.com/watch?v=HTpAVi0wyO8

Becherer, D., Nunno, G. D., Zervos, M., & Zheng, H. (2012). The Mark H. A. Davis festschrift: Stochastics, control and finance. https://doi.org/10.1080/17442508.2012.734073

Brigo, D., Jacquier, A., Obłój, J., & Zheng, H. (2020, March 28). *Obituary: Mark Davis*. https://www.bachelierfinance.org/forum/archives/7152

Capponi, A. (2020, April 16). *Obituary: Mark H. A. davis*. https://sinews.siam.org/Details-Page/obituary-mark-ha-davis

Protter, P. (2020, May 17). *Obituary: Mark H. A. Davis 1945–2020*. https://imstat.org/2020/05/17/obituary-mark-h-a-davis-1945-2020/

Vinter, R., & Zheng, H. (2020, September). Mark H.A. Davis: 1945 – 2020 [*LMS Newsletter, 490*]. https://www.lms.ac.uk/sites/lms.ac.uk/files/files/NLMS_490_for%20web2_1.pdf

Vinter, R., & Zheng, H. (2021, January). Obituary [*Stochastics, 93*, 1–2,]. https://doi.org/10.1080/17442508.2020.1869240

Vinter, R., & Zheng, H. (ongoing). Obituary: Professor Mark Davis [Imperial College. *N.B.: This page also displays personal tributes to Mark and allows to leave new ones*.]. https://www.imperial.ac.uk/news/196534/obituary-professor-mark-davis/

Control in Hilbert Space and First-Order Mean Field Type Problem

Alain Bensoussan, Hang Cheung, and Sheung Chi Phillip Yam

Abstract We extend the work [9] by two of the coauthors, which dealt with a deterministic control problem for which the Hilbert space could be generic and investigated a novel form of the 'lifting' technique proposed by P. L. Lions. In [9], we only showed the local existence and uniqueness of solutions to the FBODEs in the Hilbert space which were associated to the control problems with drift function consisting of the control only. In this article, we establish the global existence and uniqueness of the solutions to the FBODEs in Hilbert space corresponding to control problems with separable drift function which is nonlinear in state and linear in control. We shall also prove the sufficiency of the Pontryagin Maximum Principle and derive the corresponding Bellman equation. Finally, by using the 'lifting' idea as in [6, 7], we shall apply the result to solve the linear quadratic mean field type control problems, and to show the global existence of the corresponding Bellman equations.

Dedicated in memory of Mark Davis for his outstanding contribution in control theory.

Alain Bensoussan (Corresponding author)
International Center for Decision and Risk Analysis, Jindal School of Management, University of Texas at Dallas, 800 West Campbell Road, Richardson, TX 75080-3021, United States
School of Data Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong,
e-mail: Alain.Bensoussan@utdallas.edu

Hang Cheung
School of Data Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong,
e-mail: 1155032209@link.cuhk.edu.hk

Sheung Chi Phillip Yam
Department of Statistics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong,
e-mail: scpyam@sta.cuhk.edu.hk

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_1

1 Introduction

In recent years, Mean Field Game (MFG) and Mean Field Type Control Theory (MFTCT) are burgeoning. Carmona and Delarue [14] proved the existence of the general forward-backward systems of equations of McKean-Vlasov type using the probabilistic approach, and therefore obtained the classical solution to the master equation arisen from MFG. Their assumptions restricted their application to LQ models only. Cardaliaguet et al. [12] proved the existence of the classical solution to the master equation arisen from MFG by PDE techniques and the method of characteristics. To do so, they required the state space to be compact, and the Hamiltonian to be smooth, globally Lipschitz continuous and to satisfy a certain coercivity condition. Buckdahn et al. [11] adopted a similar approach to study forward flows, proving that the semigroup of a standard McKean-Vlasov stochastic differential equation with smooth coefficients is the classical solution of a particular type of master equation. A crucial assumption was made therein on the smoothness of the coefficients, which restricted the scope of applications. Gangbo and Mészáros in [19] constructed global solutions to the master equation in potential Mean Field Games, where displacement convexity was used as a substitution for the monotonicity condition. Besides the notion of classical solutions, Mou and Zhang in [26] gave a notion of weak solution of the master equation arisen from mean field games, using their results of mollifiers on the infinite dimensional space. More results can be found in the papers of Cosso and Pham [16], Pham and Wei [29] and Djete et al. [18], which concern the Bellman and Master equations of Mean Field Games and Mean Field Type Control Theory.

By Pontryagin Maximum Principle, MFG and MFTCT are deeply connected to mean field forward backward stochastic differential equations. Pardoux and Tang [27], Antonelli [2] and Hu and Yong [21] showed the existence and uniqueness of FBSDEs under small time intervals by a fixed point argument. For Markovian FBSDEs, to get rid of the small time issue, Ma et al. [24] employed the Four Step Scheme. They constructed decoupling functions by the use of the classical solutions of quasi-linear PDEs, hence non-degeneracy of the diffusion coefficient and the strong regularity condition on the coefficients were required. Another way to remove time constraints in Markovian FBSDEs was by Delarue [17]. Local solutions were patched together by the use of decoupling functions. PDE methods were used to bound the coefficients of the terminal function relative to the initial data in order for the problem to be well-posed. It was later extended to the case of non-Markovian FBSDEs by Zhang in [32]. Moreover, to deal with non-Markovian FBSDEs with arbitrary time length, there was the pure probabilistic method – method of continuation. It required monotonicity conditions on the coefficients. For seminal works one may consult [20, 28, 30, 31]. With the help of decoupling functions as in [17], but using a BSDE to control the terminal coefficient instead of PDEs, Ma et al. [25] covered most of the above cases, but in the case of codomain being $\mathbb{R}$. For mean field type FBSDE. A rather general existence result but with a restrictive assumption (boundedness of the coefficients with respect to the state variable) was first done in [13] by Carmona and Delarue. Taking advantage of the convexity of the underlying Hamiltonian and applying the continuation method, Carmona and Delarue extended their results in [14]. Bensoussan et al. [10] exploited the condition in [14] and gave weaker conditions for which the results in [14] still hold. By the method of continuation, Ahuja et al. [1] extended the above result to the FBSDEs which allow coefficients to be functionals of the processes. More details can be found in the monographs [15, 3] and [4, 5, 8].

We establish the global existence and uniqueness of the solutions to the FBODEs in Hilbert space corresponding to control problems with separable drift function which is nonlinear in state and linear in control. The result can be applied to solve linear quadratic mean field type control problems. We exploit the 'lifting to Hilbert space' approach suggested by P. L. Lions in [22, 23], but lift to another Hilbert space instead of the space of random variables. After lifting, the problems are akin to standard control problems, but the drawback is that they are in the infinite dimensional space. By the Pontryagin Maximum Principle, the control problems are reduced to FBODEs in the Hilbert space. In order to accommodate nonlinear settings, we make use of the idea of decoupling. By a Banach fixed point argument, we are able to locally find a decoupling function for the FBODEs. We then derive *a priori* estimates of the decoupling function and extend the solution from local to global as in Delarue [17] by the *a priori* estimates. Finally we apply our result to solve linear quadratic mean field type control problems and obtain their corresponding Bellman equations.

The rest of this article is organized as follows. In Section 2, we introduce the model in the Hilbert space. In Section 3, we express the related FBODE and define the decoupling function. *A priori* estimates of the decoupling function are derived in Section 4. In Section 5, we prove the local existence and uniqueness of the FBODE by using a Banach fixed point argument on the function space containing the decoupling function. In Section 6.1, we construct the global solution by our *a priori* estimates. We show the sufficiency of the Maximum Principle in Section 6.2 and write the corresponding Bellman function in Section 6.3. In Section 7, we apply our result in the Hilbert space as in [7], to solve the optimal control problem, and show the global existence to the corresponding Bellman equation.

2 The Model

2.1 Assumptions

Let $\mathcal{H}$ be a Hilbert space, with scalar product denoted by $(\cdot,\cdot)$. We consider a non-linear operator A, $x \in \mathcal{H} \mapsto A(x) \in \mathcal{H}$, such that

$$A(0) = 0. \tag{1}$$

We assume that $x \mapsto A(x)$ is C^1 and that $DA(x)(= D_x A(x)) \in \mathcal{L}(\mathcal{H};\mathcal{H})$, that means

$$D_xA(x)(\cdot) : y \in \mathcal{H} \mapsto D_xA(x)(y) = \lim_{\epsilon \to 0} \frac{A(x+\epsilon y) - A(x)}{\epsilon},$$

and it satisfies:

$$\text{The operator norm } ||DA(x)|| \leq \gamma. \tag{2}$$

By definition, we have the result:

$$(DA(x)y, z) = (D_x(A(x), z), y); \tag{3}$$

indeed, we can see this by noting that

$$\begin{aligned}(DA(x)y, z) &= \lim_{\epsilon \to 0} \left(\frac{1}{\epsilon}(A(x+\epsilon y) - A(x)), z \right) \\ &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}[(A(x+\epsilon y), z) - (A(x), z)] \\ &= (D_x(A(x), z), y),\end{aligned}$$

where the last step follows by differentiating the functional $(A(\cdot), z) : \mathcal{H} \to \mathbb{R}$. We also assume that $DA(x)$ is differentiable with a second derivative $D^2A(x) \in \mathcal{L}(\mathcal{H}; \mathcal{L}(\mathcal{H}; \mathcal{H}))$, similar to (3), such that

$$\begin{cases} (D_x(D_x(A(x), z), y), w) = (D^2_{xx}A(x)(y)w, z), \\ D_x(D_x(A(x), z), y) = (D^2_{xx}A(x)(y), z). \end{cases} \tag{4}$$

We assume the Lipschitz property:

$$||DA(x_1) - DA(x_2)|| \leq \frac{b|x_1 - x_2|}{1 + \max(|x_1|, |x_2|)}, \tag{5}$$

which implies

$$||D^2A(x)|| \leq \frac{b}{1+|x|}. \tag{6}$$

In the sequel, we shall make restrictions on the size of b.

We next consider $x \mapsto F(x)$ and $x \mapsto F_T(x)$, functionals on $\mathcal{H}$, which are C^2, with the properties:

$$\begin{cases} F(0) = 0, D_xF(0) = 0, \\ \nu|\xi|^2 \leq (D^2_{xx}F(x)\xi, \xi) \leq M|\xi|^2; \end{cases} \tag{7}$$

$$\begin{cases} F_T(0) = 0, D_xF_T(0) = 0, \\ \nu_T|\xi|^2 \leq (D^2_{xx}F_T(x)\xi, \xi) \leq M_T|\xi|^2, \end{cases} \tag{8}$$

and $\nu, \nu_T > 0$. $\mathcal{H}$ is the state space. In addition, there is a control space $\mathcal{V}$, also a Hilbert space and a linear bounded operator $B \in \mathcal{L}(\mathcal{V}; \mathcal{H})$, an invertible-self adjoint operator on $\mathcal{V}$, denoted by N. We assume that

$$(BN^{-1}B^{*}\xi,\xi) \geq m|\xi|^2,\ m > 0. \tag{9}$$

Remark 1 The assumption (1) and the first line assumptions (7), (8) are of course not necessary. It is just to simplify the calculations.

2.2 The Problem

We consider the following control problem. The state evolution is governed by the differential equation in $\mathcal{H}$:

$$\begin{cases} \dfrac{dx}{ds} = A(x) + Bv(s), \\ x(t) = x, \end{cases} \tag{10}$$

in which $v(\cdot)$ is in $L^2(t,T;\mathcal{V})$. It is easy to check that the state $x(\cdot)$ is uniquely defined and belongs to $H^1(t,T;\mathcal{H})$. We define the payoff functional:

$$J_{xt}(v(\cdot)) := \int_t^T F(x(s))ds + F_T(x(T)) + \frac{1}{2}\int_t^T (v(s), Nv(s))ds. \tag{11}$$

This functional is continuous and coercive. If $\mathcal{H}$ were $\mathbb{R}^n$, it would be classical that it has a minimum and thus we could write the necessary conditions of optimality. But the proof does not carry over to general Hilbert spaces. Moreover, since A is not linear, we do not have the convexity property, which would guarantee the existence and uniqueness of a minimum, and thus a solution of the necessary conditions of optimality. We shall then write the necessary conditions of optimality and prove directly the existence and uniqueness of a solution.

3 Necessary Conditions of Optimality

3.1 The System

It is standard to check the following system of forward-backward equations in $\mathcal{H}$:

$$\begin{cases} \dfrac{dy}{ds} = A(y) - BN^{-1}B^{*}z(s),\ t < s < T, \\ -\dfrac{dz}{ds} = (DA(y(s)))^{*}z(s) + DF(y(s)), \\ y(t) = x,\ z(T) = DF_T(y(T)). \end{cases} \tag{12}$$

The optimal state is $y(\cdot)$, and $z(\cdot)$ is the adjoint state. The optimal control is then:

$$u(s) = -N^{-1}B^*z(s). \tag{13}$$

The system (12) expresses the Pontryagin Maximum Principle. The objective is to study the system of Equations (12).

3.2 Decoupling

We set

$$z(t) = \Gamma(x,t). \tag{14}$$

It is standard to check that $z(s) = \Gamma(y(s), s)$. So $y(s)$ is the solution of the differential equation in $\mathcal{H}$:

$$\begin{cases} \dfrac{dy}{ds} = A(y) - BN^{-1}B^*\Gamma(y(s), s), \\ y(t) = x, \end{cases} \tag{15}$$

and $\Gamma(x, s)$ is the solution of the nonlinear partial differential equation:

$$\begin{cases} -\dfrac{\partial\Gamma}{\partial s} = D_x\Gamma(x)A(x) + (D_xA(x))^*\Gamma(x) - D_x\Gamma(x)BN^{-1}B^*\Gamma(x,s) + D_xF(x), \\ \Gamma(x,T) = D_xF_T(x). \end{cases} \tag{16}$$

If $A(x) = Ax$, $F(x) = \dfrac{1}{2}(x, Mx)$ and $F_T(x) = \dfrac{1}{2}(x, M_Tx)$, then $\Gamma(x, s) = P(s)x$, and $P(s)$ is solution of the Riccati equation:

$$\begin{cases} -\dfrac{dP}{ds} = P(s)A + A^*P(s) - P(s)BN^{-1}B^*P(s) + M, \\ P(T) = M_T. \end{cases} \tag{17}$$

4 *A Priori* Estimates

4.1 First Estimate

We state the first result:

Proposition 1 *We assume (1), (2), (5), (7), (8), (9) and*

$$\frac{b^2}{16} < (m-k)(\nu-k),\ 0 < k < \min(m,\nu), \tag{18}$$

then we have the a priori estimate:

$$|\Gamma(x,t)| \leq |x| \left(\frac{M_T^2}{\nu_T} + \frac{\gamma^2 + M^2}{k}(T-t) \right). \tag{19}$$

Proof From the system (12), we obtain:

$$\frac{d}{ds}(y(s), z(s)) = (A(y(s)) - BN^{-1}B^* z(s), z(s)) - ((DA(y(s)))^* z(s) + DF(y(s)), y(s)).$$

Integration yields:

$$(D_x F_T(y(T)), y(T)) + \int_t^T (BN^{-1}B^* z(s), z(s))ds + \int_t^T (D_x F(y(s)), y(s))ds$$
$$= (x, z(t)) + \int_t^T (A(y(s)) - DA(y(s))y(s), z(s))\, ds. \tag{20}$$

We note that

$$|A(x) - DA(x)x| \leq \frac{b}{2}|x|; \tag{21}$$

indeed, $A(x) - DA(x)x = \int_0^1 (DA(\theta x) - DA(x))x\, d\theta$, and from the assumption (5), we get:

$$|A(x) - DA(x)x| \leq \int_0^1 \frac{b|x|^2(1-\theta)}{1+|x|} d\theta,$$

which implies (21). Therefore, from (20), we obtain, using assumptions:

$$(x, z(t)) \geq \nu_T |y(T)|^2 + m \int_t^T |z(s)|^2 ds + \nu \int_t^T |y(s)|^2 ds - \frac{b}{2} \int_t^T |y(s)||z(s)|ds.$$

Using (18), we can state:

$$(x, z(t)) \geq \nu_T |y(T)|^2 + k \int_t^T (|y(s)|^2 + z(s)|^2)ds. \tag{22}$$

On the other hand, from the second equation (12), we write $z(t) = z(T) + \int_t^T ((DA(y(s)))^* z(s) + DF(y(s)))\, ds$, hence

$$(x.z(t)) = (x, DF_T(y(T)) + \int_t^T (DA(y(s))x, z(s))ds + \int_t^T (x, DF(y(s))ds,$$
$$(x.z(t)) \leq |x||z(t)| \leq |x|(M_T|y(T)| + \int_t^T \gamma|z(s)|ds + \int_t^T |y(t)|dt)$$
$$\leq \frac{1}{2}\left(\nu_T|y(T)|^2 + k\int_t^T (|y(s)|^2 + z(s)|^2)ds\right)$$
$$+ \frac{|x|^2}{2}\left(\frac{M_T^2}{\nu_T} + \frac{\gamma^2 + M^2}{k}(T-t)\right).$$

From this relation and (22), we get:

$$\nu_T|y(T)|^2 + k\int_t^T (|y(s)|^2 + z(s)|^2)ds \le |x|^2\left(\frac{M_T^2}{\nu_T} + \frac{\gamma^2 + M^2}{k}(T-t)\right).$$

Therefore,

$$|x||z(t)| \le |x|^2\left(\frac{M_T^2}{\nu_T} + \frac{\gamma^2 + M^2}{k}(T-t)\right),$$

and the result follows. We write

$$\alpha_t = \frac{M_T^2}{\nu_T} + \frac{\gamma^2 + M^2}{k}(T-t). \tag{23}$$

Note that in the system (12), we can write

$$|z(s)| \le \alpha_s|y(s)|. \tag{24}$$

4.2 Second Estimate

The second estimate concerns the gradient $D_x\Gamma(x,t)$. We have the following result:

Proposition 2 *We make the assumptions of Proposition 1 and*

$$\nu - b\alpha_0 > 0, \tag{25}$$

then we have the a priori estimate:

$$||D_x\Gamma(x,t)|| \le \frac{M_T^2}{\nu_T} + \frac{\gamma^2}{m}(T-t) + \int_t^T \frac{(M + b\alpha_s)^2}{\nu - b\alpha_s}ds. \tag{26}$$

Proof We differentiate the system (12) with respect to x. We denote

$$\mathcal{Y}(s) = D_x y(s), \mathcal{Z}(s) = D_x z(s). \tag{27}$$

Differentiating (12), we can write, by recalling notation (4):

$$\frac{d}{ds}\mathcal{Y}(s)\xi = D_x A(y(s))\mathcal{Y}(s)\xi - BN^{-1}B^*\mathcal{Z}(s)\xi, \tag{28}$$

$$-\frac{d}{ds}\mathcal{Z}(s)\xi = (D_{xx}^2 A(y(s))\mathcal{Y}(s)(\xi), z(s)) + (DA(y(s)))^*\mathcal{Z}(s)\xi + D_{xx}^2 F(y(s))\mathcal{Y}(s)\xi,$$

$$\mathcal{Y}(t)\xi = \xi,\ \mathcal{Z}(T)\xi = D_{xx}^2 F_T(y(T))\mathcal{Y}(T)\xi. \tag{29}$$

We compute $\frac{d}{ds}(\mathcal{Y}(s)\xi, \mathcal{Z}(s)\xi)$ and then integrate. We obtain that

$$
\begin{aligned}
(\mathcal{Z}(t)\xi,\xi) = (D^2_{xx}F_T(y(T))\mathcal{Y}(T)\xi,\mathcal{Y}(T)\xi) + \int_t^T (BN^{-1}B^*\mathcal{Z}(s)\xi,\mathcal{Z}(s)\xi)ds \\
+ \int_t^T (D^2_{xx}F(y(s))\mathcal{Y}(s)\xi,\mathcal{Y}(s)\xi)ds \\
+ \int_t^T (D^2_{xx}A(y(s))\mathcal{Y}(s)(\xi)\mathcal{Y}(s)\xi, z(s)) \\
\geq \nu_T|\mathcal{Y}(T)\xi|^2 + m\int_t^T |\mathcal{Z}(s)\xi|^2 ds + \int_t^T (\nu - b\alpha_s)|\mathcal{Y}(s)\xi|^2 ds.
\end{aligned}
\tag{30}
$$

Also, from the second line of (28),

$$
|\mathcal{Z}(t)\xi| \leq M_T|\mathcal{Y}(T)\xi| + \int_t^T (M + b\alpha_s)|\mathcal{Y}(s)\xi|ds + \gamma\int_t^T |\mathcal{Z}(s)\xi|ds. \tag{31}
$$

Combining (30) and (31) as in Proposition 1, we conclude that

$$
|\mathcal{Z}(t)\xi| \leq |\xi|\left(\frac{M_T^2}{\nu_T} + \frac{\gamma^2}{m}(T-t) + \int_t^T \frac{(M+b\alpha_s)^2}{\nu - b\alpha_s}ds\right).
$$

Since $\mathcal{Z}(t)\xi = D_x\Gamma(x,t)$, the result (26) follows immediately. The proof is complete.□

We shall call

$$
\beta_t = \frac{M_T^2}{\nu_T} + \frac{\gamma^2}{m}(T-t) + \int_t^T \frac{(M+b\alpha_s)^2}{\nu - b\alpha_s}ds. \tag{32}
$$

Since

$$
\Gamma(x,t) = \int_0^1 D_x\Gamma(\theta x,t)x\,d\theta,
$$

we also have:

$$
|\Gamma(x,t)| \leq \beta_t|x|, \tag{33}
$$

so in fact,

$$
\begin{cases}
|\Gamma(x,t)| \leq \min(\alpha_t,\beta_t)|x|, \\
||D_x\Gamma(x,t)|| \leq \beta_t.
\end{cases}
\tag{34}
$$

5 Local Time Solution

5.1 Fixed Point Approach

We want to solve (12) by a fixed point approach. Suppose we have a function $\lambda(x,t)$ with values in $\mathcal{H}$ such that:

$$\begin{cases} |\lambda(x,t)| \le \mu_t |x|, \\ ||D_x \lambda(x,t)|| \le \rho_t, \end{cases} \tag{35}$$

where μ_t and ρ_t are bounded functions on $[T-h,T]$, for some convenient h. These functions will be chosen conveniently in the sequel, with $\mu_t < \rho_t$. We then solve

$$\begin{cases} \dfrac{d}{ds} y(s) = A(y(s)) - BN^{-1}B^* \lambda(y(s),s), \\ y(t) = x. \end{cases} \tag{36}$$

This differential equation defines uniquely $y(s)$, thanks to the assumptions (35). We then define

$$\Lambda(x,t) := D_x F_T(y(T)) + \int_t^T (DA(y(s)))^* \lambda(y(s),s)ds + \int_t^T D_x F(y(s))ds. \tag{37}$$

We want to show that μ_t and ρ_t can be chosen such that

$$|\Lambda(x,t)| \le \mu_t |x|, \ ||D_x \Lambda(x,t)|| \le \rho_t, \tag{38}$$

and that the map $\lambda \mapsto \Lambda$ has a fixed point. This will be only possible when t remains close to T, namely $T-h < t < T$, with h small.

5.2 Choice of Functions μ_t and ρ_t

From (36), we obtain:

$$\frac{d}{ds}|y(s)| \le \left|\frac{d}{ds} y(s)\right| \le (\gamma + ||BN^{-1}B^*|| \mu_s)|y(s)|,$$

which implies

$$|y(s)| \le |x| \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*|| \mu_\tau) d\tau\right), \tag{39}$$

and thus from (37) it follows that

$$|\Lambda(x,t)| \le M_T |y(T)| + \int_t^T (M + \gamma \mu_s)|y(s)| ds.$$

Using (39), we obtain:

$$|\Lambda(x,t)| \leq |x| \left(M_T \exp\left(\int_t^T (\gamma + ||BN^{-1}B^*||\mu_\tau) d\tau \right) \right.$$
$$\left. + \int_t^T (M + \gamma\mu_s) \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\mu_\tau) d\tau \right) ds \right).$$

To obtain the first inequality (38), we must choose the function μ_t such that

$$\mu_t = M_T \exp\left(\int_t^T (\gamma + ||BN^{-1}B^*||\mu_\tau) d\tau \right)$$
$$+ \int_t^T (M + \gamma\mu_s) \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\mu_\tau) d\tau \right) ds. \tag{40}$$

So μ_t must be solution of the differential equation of Riccati type:

$$\begin{cases} \dfrac{d}{dt}\mu_t = -||BN^{-1}B^*||\mu_t^2 - 2\gamma\mu_t - M, \\ \mu_T = M_T. \end{cases} \tag{41}$$

To proceed, we need to assume that

$$\gamma^2 < M||BN^{-1}B^*||, \tag{42}$$

and we define μ_t bt the formula:

$$\arctan \frac{\mu_t ||BN^{-1}B^*|| + \gamma}{\sqrt{M||BN^{-1}B^*|| - \gamma^2}}$$
$$= \arctan \frac{M_T ||BN^{-1}B^*|| + \gamma}{\sqrt{M||BN^{-1}B^*|| - \gamma^2}} + \left(\sqrt{M||BN^{-1}B^*|| - \gamma^2} \right) (T - t). \tag{43}$$

For $h > 0$, define θ_h with

$$\arctan \frac{\theta_h ||BN^{-1}B^*|| + \gamma}{\sqrt{M||BN^{-1}B^*|| - \gamma^2}}$$
$$= \arctan \frac{M_T ||BN^{-1}B^*|| + \gamma}{\sqrt{M||BN^{-1}B^*|| - \gamma^2}} + \left(\sqrt{M||BN^{-1}B^*|| - \gamma^2} \right) h. \tag{44}$$

The number h must be small enough to guarantee that

$$\arctan \frac{M_T ||BN^{-1}B^*|| + \gamma}{\sqrt{M||BN^{-1}B^*|| - \gamma^2}} + \left(\sqrt{M||BN^{-1}B^*|| - \gamma^2} \right) h < \frac{\pi}{2}. \tag{45}$$

Formula (43) defines uniquely μ_t for $T - h < t < T$. It is decreasing in t, with $M_T < \mu_t < \theta_h$.

Therefore, for $T - h < t < T$, we have defined by (37) a function $\Lambda(x,t)$ which satisfies the first condition (38), with μ_t defined by equation (43). We turn now to the definition of ρ_t. Define $\mathcal{Y}(s) = D_x y(s)$, see (36). We have:

$$\begin{cases} \dfrac{d}{ds}\mathcal{Y}(s) = \left(DA(y(s)) - BN^{-1}B^* D_x\lambda(y(s),s)\right)\mathcal{Y}(s), \\ \quad \mathcal{Y}(t) = I. \end{cases} \tag{46}$$

We obtain, by techniques already used:

$$||\mathcal{Y}(s)|| \leq \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\rho_\tau)d\tau\right). \tag{47}$$

We then differentiate $\Lambda(x,t)$ in x, see (37). We get:

$$\begin{aligned} D_x\Lambda(x,t) = D_{xx}^2 F_T(y(T))\mathcal{Y}(T) + \int_t^T (D_{xx}^2 A(y(s))\mathcal{Y}(s), \lambda(y(s),s))ds, \\ + \int_t^T (D_x A(y(s)))^* D_x\lambda(y(s),s)\mathcal{Y}(s)ds + \int_t^T D_{xx}^2 F(y(s))\mathcal{Y}(s)ds, \end{aligned}$$

and we obtain:

$$||D_x\Lambda(x,t)|| \leq M_T||\mathcal{Y}(T)|| + \int_t^T (M + b\mu_s + \gamma\rho_s)||\mathcal{Y}(s)||ds.$$

Since $T - h < t < T$, we can majorize, using also (47), to obtain:

$$\begin{aligned} ||D_x\Lambda(x,t)|| \leq M_T \exp\left(\int_t^T (\gamma + ||BN^{-1}B^*||\rho_s)ds\right) \\ + \int_t^T (M + b\theta_h + \gamma\rho_s)\exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\rho_\tau)d\tau\right)ds. \end{aligned} \tag{48}$$

We are thus led to looking for ρ_t solution of

$$\begin{aligned} \rho_t = M_T \exp\left(\int_t^T (\gamma + ||BN^{-1}B^*||\rho_s)ds\right) \\ + \int_t^T (M + b\theta_h + \gamma\rho_s)\exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\rho_\tau)d\tau\right)ds. \end{aligned} \tag{49}$$

This equation is similar to the one defining μ_t, see (40), with the change of M into $M + b\theta_h$. Hence, by analogy with (43), we can assert that:

$$\begin{aligned} \arctan\frac{\rho_t||BN^{-1}B^*|| + \gamma}{\sqrt{(M+b\theta_h)||BN^{-1}B^*|| - \gamma^2}} = \arctan\frac{M_T||BN^{-1}B^*|| + \gamma}{\sqrt{(M+b\theta_h)||BN^{-1}B^*|| - \gamma^2}} \\ + \left(\sqrt{(M+b\theta_h)||BN^{-1}B^*|| - \gamma^2}\right)(T-t). \end{aligned} \tag{50}$$

In order to get a bounded solution for ρ_t, we need that the right hand side of (50) be smaller than $\frac{\pi}{2}$. We need to restrict h more than with (45), namely:

$$\arctan \frac{M_T||BN^{-1}B^*||+\gamma}{\sqrt{M||BN^{-1}B^*||-\gamma^2}}+\left(\sqrt{(M+b\theta_h)||BN^{-1}B^*||-\gamma^2}\right)h<\frac{\pi}{2}. \tag{51}$$

Then the function ρ_t is well defined on $(T-h,T]$, by formula (50) and the function $\Lambda(x,t)$ defined by (37), for $t\in(T-h,T]$ satisfies (38) if $\lambda(x,t)$ satisfies (35). We also claim that

$$\rho_t > \mu_t. \tag{52}$$

Indeed, ρ_t satisfies the Riccati equation:

$$\begin{cases} \dfrac{d}{dt}\rho_t = -||BN^{-1}B^*||\rho_t^2 - 2\gamma\rho_t - (M+b\theta_h), \\ \rho_T = M_T, \end{cases} \tag{53}$$

and comparing (41) and (53), it is standard to show the property (52).

5.3 Contraction Mapping

We define the space of functions $(x,t)\in\mathcal{H}\times(T-h,T)\mapsto\lambda(x,t)\in\mathcal{H}\times(T-h,T)$, equipped with the norm:

$$||\lambda||_h = \sup_{x\in\mathcal{H},t\in(T-h,T)} \frac{|\lambda(x,t)|}{|x|}. \tag{54}$$

This space is a Banach space, denoted by $\mathcal{B}_h$.. We next consider the convex closed subset of $\mathcal{B}_h$. of functions such that:

$$|\lambda(x,t)|\le\mu_t|x|,\ ||D_x\lambda(x,t)||\le\rho_t, \forall t\in(T-h,T], \tag{55}$$

where μ_t and ρ_t are defined by (43) and (50), respectively. The subset (55) is denoted by C_h. The map $\lambda\mapsto\Lambda$, defined by (36) and (37), is defined from C_h to C_h. We want to show that it leads to a contraction.

Let $\lambda^1(x,t)$, $\lambda^2(x,t)$ in C_h and the corresponding functions $\Lambda^1(x,t)$, $\Lambda^2(x,t)$, which also belong to C_h. Let $y^1(s), y^2(s)$ be the solutions of (36) corresponding to λ^1,λ^2. We call $\widetilde{y}(s)=y^1(s)-y^2(s)$. We have:

$$\begin{cases} \dfrac{d}{ds}\widetilde{y}(s) = A(y^1(s)) - A(y^2(s)) - BN^{-1}B^*(\lambda^1(y^1(s))-\lambda^2(y^2(s))), \\ \widetilde{y}(t)=0, \end{cases}$$

hence

$$\frac{d}{ds}|\widetilde{y}(s)| \leq \gamma|\widetilde{y}(s)| + ||BN^{-1}B^*||\,|\lambda^1(y^1(s)) - \lambda^2(y^2(s))|.$$

Next,

$$\begin{aligned}
&|\lambda^1(y^1(s)) - \lambda^2(y^2(s))| \\
&\leq |\lambda^1(y^1(s)) - \lambda^1(y^2(s))| + |\lambda^1(y^2(s)) - \lambda^2(y^2(s))| \\
&\leq \rho_s|\widetilde{y}(s)|| + ||\lambda^1 - \lambda^2||_h|x| \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\mu_\tau)d\tau\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{d}{ds}|\widetilde{y}(s)| &\leq (\gamma + ||BN^{-1}B^*||\rho_s)|\widetilde{y}(s)| \\
&\quad + ||BN^{-1}B^*||\,|x|\,||\lambda^1 - \lambda^2||_h \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\mu_\tau)d\tau\right).
\end{aligned}$$

We obtain that

$$\begin{aligned}
&|\widetilde{y}(s)| \exp\left(-\int_t^s (\gamma + ||BN^{-1}B^*||\rho_\tau)d\tau\right) \\
&\leq ||BN^{-1}B^*||\,|x|\,||\lambda^1 - \lambda^2||_h \int_t^s \exp\left(-||BN^{-1}B^*|| \int_t^\tau (\rho_\theta - \mu_\theta)d\theta\right) d\tau,
\end{aligned}$$

which implies:

$$|\widetilde{y}(s)| \leq h||BN^{-1}B^*||\,|x|\,||\lambda^1 - \lambda^2||_h \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\rho_\tau)d\tau\right). \tag{56}$$

We next have from the definition of the map $\Lambda(x,t)$ that:

$$\begin{aligned}
\Lambda^1(x,t) - \Lambda^2(x,t) = DF_T(y^1(T)) - DF_T(y^2(T)) \\
+ \int_t^T \left(DA^*(y^1(s))\lambda^1(y^1(s)) - DA^*(y^2(s))\lambda^2(y^2(s))\right) ds \\
+ \int_t^T (DF(y^1(s)) - DF(y^2(s)))ds.
\end{aligned} \tag{57}$$

We have:

$$\begin{aligned}
&|DA^*(y^1(s))\lambda^1(y^1(s)) - DA^*(y^2(s))\lambda^2(y^2(s))| \\
&\leq (b\theta_h + \gamma\rho_s)|\widetilde{y}(s)| + \gamma|x|\,||\lambda^1 - \lambda^2||_h \exp\left(\int_t^s (\gamma + ||BN^{-1}B^*||\mu_\tau)d\tau\right).
\end{aligned}$$

So, from (57), we obtain:

$$|\Lambda^1(x,t)-\Lambda^2(x,t)| \leq M_T|\widetilde{y}(T)| + \int_t^T (M+b\theta_h+\gamma\rho_s)|\widetilde{y}(s)|ds$$
$$+\gamma|x|\,||\lambda^1-\lambda^2||_h \int_t^T \exp\left(\int_t^s (\gamma+||BN^{-1}B^*||\mu_\tau)d\tau\right) ds, \tag{58}$$

and from (56):

$$\begin{aligned}
&|\Lambda^1(x,t)-\Lambda^2(x,t)| \\
&\leq |x|\,|\lambda^1-\lambda^2||_h h \times \Bigg[||BN^{-1}B^*|| \left(M_T \exp\left(\int_t^T (\gamma+||BN^{-1}B^*||\rho_\tau)d\tau\right)\right. \\
&\quad \left. + \int_t^T (M+b\theta_h+\gamma\rho_s)\left(\int_t^s (\gamma+||BN^{-1}B^*||\rho_\tau)d\tau\right) ds\right)\Bigg] \\
&\quad +\gamma|x|\,|\lambda^1-\lambda^2||_h \int_t^T \exp\left(\int_t^s (\gamma+||BN^{-1}B^*||\mu_\tau)d\tau\right) ds,
\end{aligned}$$

then from the definition of ρ_t (see (49)), we obtain:

$$\begin{aligned}
&|\Lambda^1(x,t)-\Lambda^2(x,t)| \\
&\leq |x|\,|\lambda^1-\lambda^2||_h h \left(\rho_t||BN^{-1}B^*|| + \gamma \exp\left(\int_{T-h}^T (\gamma+||BN^{-1}B^*||\mu_\tau)d\tau\right)\right).
\end{aligned} \tag{59}$$

Similarly to the definition of θ_h (see (44)), we define the quantity σ_h by the formula:

$$\begin{aligned}
\arctan \frac{\sigma_h||BN^{-1}B^*||+\gamma}{\sqrt{(M+b\theta_h)||BN^{-1}B^*||-\gamma^2}} &= \arctan \frac{M_T||BN^{-1}B^*||+\gamma}{\sqrt{(M+b\theta_h)||BN^{-1}B^*||-\gamma^2}} \\
&\quad + \left(\sqrt{(M+b\theta_h)||BN^{-1}B^*||-\gamma^2}\right) h.
\end{aligned} \tag{60}$$

From (50), we see that $M_T < \rho_t < \sigma_h$. Therefore from (59),

$$||\Lambda^1-\Lambda^2||_h \leq |\lambda^1-\lambda^2||_h\, h\left(\sigma_h||BN^{-1}B^*|| + \gamma\exp\left(h(\gamma+||BN^{-1}B^*||\theta_h)\right)\right). \tag{61}$$

Using the fact that $\theta_h \to M_T$ as $h \to 0$, equation (60) shows that $\sigma_h \to M_T$ as $h \to 0$. We deduce that:

$$h\left(\sigma_h||BN^{-1}B^*|| + \gamma\exp\left(h(\gamma+||BN^{-1}B^*||\theta_h)\right)\right) \to 0, \text{ as } h \to 0. \tag{62}$$

We can restrict h such that

$$h\left(\sigma_h||BN^{-1}B^*|| + \gamma\exp\left(h(\gamma+||BN^{-1}B^*||\theta_h)\right)\right) < 1, \tag{63}$$

and thus for h sufficiently small, the map $\lambda \mapsto \Lambda$ is paradoxical and leads to a contradiction. We can summarize the results in the following theorem:

Theorem 1 *We assume* (42)*. We choose h small enough to satisfy conditions (45), (51), (63). For $T-h<t<T$, there exists one and only one solution of the system of forward-backward equations (12). We have also one and only one solution of equation (16) on the same interval.*

6 Global Solution

6.1 Statement of Results

We have proven in Theorem 1 the existence and uniqueness of a local solution of the system (12). We want to state that this solution is global, under the assumptions of Proposition 2.

Theorem 2 *We make the assumptions of Proposition 2 and (42). The local solution defined in Theorem 1 can be extended. Thus there exists one and only one solution of the system (12) on any finite interval* $[0,T]$*, and there exists one and only one solution of equation (16) on any finite interval* $[0,T]$.

Proof Defining by $\Gamma(x,t)$ the fixed point obtained in Theorem 1, it is the unique solution of the parabolic equation:

$$\begin{cases} -\dfrac{\partial \Gamma}{\partial t} = D_x\Gamma(x)A(x) + (D_xA(x))^*\Gamma(x) - D_x\Gamma(x)BN^{-1}B^*\Gamma(x,s) \\ \qquad\qquad + D_xF(x), \quad T-h<t<T, \\ \Gamma(x,T) = D_xF_T(x), \end{cases} \tag{64}$$

with h restricted as stated in Theorem 1. We also have the estimates:

$$\begin{cases} |\Gamma(x,t)| \le \min(\alpha_t, \beta_t)|x|, \\ ||D_x\Gamma(x,t)|| \le \beta_t, \end{cases} \tag{65}$$

with

$$\begin{cases} \alpha_t = \dfrac{M_T^2}{\nu_T} + \dfrac{\gamma^2+M^2}{k}(T-t), \\ \beta_t = \dfrac{M_T^2}{\nu_T} + \dfrac{\gamma^2}{m}(T-t) + \displaystyle\int_t^T \dfrac{(M+b\alpha_s)^2}{\nu - b\alpha_s}ds. \end{cases} \tag{66}$$

These estimates follow from the *a priori* estimates stated in Proposition 1 and 2. They do not depend on h. Now we want to extend (64) for $t<T-h$. To avoid confusion, we define

$$U_{T-h}(x) := \Gamma(x, T-h). \tag{67}$$

We set $M_{T-h} = \beta_0$. We can then state:

$$\begin{cases} |U_{T-h}(x)| \leq M_{T-h}|x|, \\ ||D_x U_{T-h}(x)|| \leq M_{T-h}, \end{cases} \tag{68}$$

and we consider the parabolic equation:

$$\begin{cases} -\dfrac{\partial \Gamma}{\partial t} = D_x\Gamma(x)A(x) + (D_xA(x))^*\Gamma(x) - D_x\Gamma(x)BN^{-1}B^*\Gamma(x,s) \\ \qquad + D_xF(x), \quad t < T-h, \\ \Gamma(x, T-h) = U_{T-h}(x). \end{cases} \tag{69}$$

We associate to this equation the system:

$$\begin{cases} \dfrac{dy}{ds} = A(y) - BN^{-1}B^*z(s), t < s < T-h, \\ -\dfrac{dz}{ds} = (DA(y(s)))^*z(s) + DF(y(s)), \\ y(t) = x, \; z(T-h) = U_{T-h}(y(T-h)). \end{cases} \tag{70}$$

Proceeding like in Theorem 1, we can solve this system on an interval $[T-h-l, T-h]$, for a sufficiently small $l \neq h$. The difference is due to the fact that $M_{T-h} \neq M_T$. So in (64), we can replace $T-h$ by $T-h-l$. This time the estimates on $\Gamma(x, T-h-l)$ and $D_x\Gamma(x, T-h-l)$ are identical to those of $\Gamma(x, T-h)$ and $D_x\Gamma(x, T-h)$, thanks to the *a priori* estimates. So the intervals we can extend further will have the same length. Clearly, this implies that we can extend (64) up to $t = 0$. So, we obtain the global existence and uniqueness of equation (16) on $[0, T]$. The proof is complete.□

6.2 Optimal Control

In Theorem 2, we have obtained the existence and uniqueness of the solution of the pair $(y(s), z(s))$ of the system (12), for any $t \in [0, T]$. We want now to check that the control $u(s)$ defined by (13) is solution of the control problem (10), (11), and that the optimal control is unique.

Theorem 3 *Under the assumptions of Theorem 2, the control $u(\cdot)$ defined by (13) is the unique optimal control for the problem (10), (11).*

Proof Let $v(\cdot)$ be another control. We shall prove that

$$J(u(\cdot) + v(\cdot)) \geq J(u(\cdot)), \tag{71}$$

which will prove the optimality of $u(\cdot)$. We define by $y_v(\cdot)$ the state corresponding to the control $u(\cdot) + v(\cdot)$. It is the solution of

$$\begin{cases} \dfrac{d}{ds}y_v(s) = A(y_v(s)) + B(u(s) + v(s)), \\ y_v(t) = x, \end{cases} \tag{72}$$

and we have:

$$J(u(\cdot)+v(\cdot)) = \int_t^T F(y_v(s))ds + F_T(y_v(T)) + \frac{1}{2}\int_t^T (u(s)+v(s), N(u(s)+v(s)))ds,$$

and

$$\begin{aligned} &J(u(\cdot)+v(\cdot)) - J(u(\cdot)) \\ =& \int_t^T (F(y_v(s)) - F(y(s)))ds + F_T(y_v(T)) - F_T(y(T)) \\ &+ \frac{1}{2}\int_t^T (v(s), Nv(s))ds + \int_t^T (Nu(s), v(s))ds. \end{aligned}$$

We denote $\tilde{y}_v(s) := y_v(s) - y(s)$. It satisfies:

$$\begin{cases} \dfrac{d}{ds}\tilde{y}_v(s) = A(y_v(s)) - A(y(s)) + Bv(s), \\ \quad \tilde{y}_v(t) = 0. \end{cases} \tag{73}$$

Then,

$$\begin{aligned} &J(u(\cdot)+v(\cdot)) - J(u(\cdot)) \\ =& \int_t^T (D_x F(y(s)), \tilde{y}_v(s))ds \\ &+ \int_t^T \int_0^1 \int_0^1 \theta \left(D_{xx}^2 F(y(s) + \lambda\theta\tilde{y}_v(s))\tilde{y}_v(s), \tilde{y}_v(s) \right) ds d\lambda d\theta \\ &+ (D_x F_T(y(T)), \tilde{y}_v(T)) + \int_0^1 \int_0^1 \theta \left(D_{xx}^2 F_T(y(T) + \lambda\theta\tilde{y}_v(T))\tilde{y}_v(T), \tilde{y}_v(T) \right) d\lambda d\theta \\ &+ \frac{1}{2}\int_t^T (v(s), Nv(s))ds - \int_t^T (z(s), Bv(s))ds. \end{aligned}$$

From the assumptions (7), we can write:

$$\begin{aligned} J(u(\cdot)+v(\cdot)) - J(u(\cdot)) \geq& \int_t^T \left(-\frac{d}{ds}z(s) - DA^*(y(s))z(s), \tilde{y}_v(s) \right) ds \\ &+ \frac{\nu}{2}\int_t^T |\tilde{y}_v(s)|^2 ds + (z(T), \tilde{y}_v(T)) \\ &+ \frac{\nu_T}{2}|\tilde{y}_v(T)|^2 + \frac{1}{2}\int_t^T (v(s), Nv(s))ds \\ &- \int_t^T \left(z(s), \frac{d}{ds}\tilde{y}_v(s) - (A(y_v(s)) - A(y(s))) \right) ds, \end{aligned}$$

which reduces to:

$$J(u(\cdot)+v(\cdot))-J(u(\cdot)) \geq \frac{\nu}{2}\int_t^T |\tilde{y}_v(s)|^2 ds + \frac{\nu_T}{2}|\tilde{y}_v(T)|^2 + \frac{1}{2}\int_t^T (v(s), Nv(s))ds$$
$$+\int_t^T (z(s), A(y_v(s)) - A(y(s)) - DA(y(s))\tilde{y}_v(s))\, ds. \tag{74}$$

Note that

$$|(z(s), A(y_v(s)) - A(y(s)) - DA(y(s))\tilde{y}_v(s))| \leq \frac{b|z(s)||\tilde{y}_v(s)|^2}{2(1+|y(s)|)} \leq \frac{b\alpha_s}{2}|\tilde{y}_v(s)|^2.$$

Finally, we can state that

$$J(u(\cdot)+v(\cdot))-J(u(\cdot)) \geq \frac{1}{2}\int_t^T (\nu - b\alpha_s)|\tilde{y}_v(s)|^2 ds + \frac{\nu_T}{2}|\tilde{y}_v(T)|^2 \frac{1}{2}\int_t^T (v(s), Nv(s))ds. \tag{75}$$

Thanks to the assumption (25), the right hand side of (75) is positive, which proves (71) and completes the proof of the result. □

6.3 Bellman Equation

We have proven, under the assumptions of Theorem 2, that the control problem (10), (11) has a unique solution $u(\cdot)$. Defining the value function

$$V(x,t) := \inf_{v(\cdot)} J_{xt}(v(\cdot)) = J_{xt}(u(\cdot)), \tag{76}$$

we can state that:

$$V(x,t) = \int_t^T F(y(s))ds + F_T(y(T)) + \frac{1}{2}\int_t^T (BN^{-1}B^*\Gamma(y(s),s), \Gamma(y(s),s))ds, \tag{77}$$

with

$$\begin{cases} \dfrac{d}{ds}y(s) = A(y(s)) - BN^{-1}B^*\Gamma(y(s),s), \\ \quad y(t) = x. \end{cases} \tag{78}$$

We first have:

Proposition 3 *We have the following property:*

$$\Gamma(x,t) = D_x V(x,t). \tag{79}$$

Proof Since the minimum of $J_{xt}(v(\cdot))$ is attained in the unique value $u(\cdot)$, we can rely on the envelope theorem to claim that:

$$(D_x V(x,t), \xi) = \int_t^T (D_x F(y(s)), \mathcal{X}(s)\xi)ds + (D_x F_T(y(T)), \mathcal{X}(T)\xi), \tag{80}$$

in which $\mathcal{X}(s)$ is the solution of

$$\begin{cases} \dfrac{d}{ds}\mathcal{X}(s) = D_x A(y(s))\mathcal{X}(s), \\ \quad \mathcal{X}(t) = I. \end{cases}$$

Recalling the equation (12) for $z(s)$ and performing integration by parts in (80), the result $(D_x V(x,t), \xi) = (\Gamma(x,t), \xi)$ is easily obtained. This proves the result (79). □

We can then obtain the Bellman equation for the value function $V(x,t)$.

Theorem 4 *We make the assumptions of Theorem 2. The function $V(x,t)$ is the unique solution of*

$$\begin{cases} -\dfrac{\partial V}{\partial t} - (D_x V, A(x)) + \dfrac{1}{2}(D_x V, BN^{-1}B^* D_x V) = F(x), \\ V(x,T) = F_T(x). \end{cases} \tag{81}$$

Proof We know that $V(x,t)$ is Gâteaux differentiable in x, with the derivative $\Gamma(x,t)$. From (12), $\Gamma(x,t)$ is continuous in t. From equation (77), we can write:

$$\begin{aligned} V(x,t) - V(x,t+\epsilon) = &\int_t^{t+\epsilon} F(y(s))ds + \frac{1}{2}\int_t^{t+\epsilon} (BN^{-1}B^*\Gamma(y(s),s), \Gamma(y(s),s))ds \\ &+ V(y(\epsilon), t+\epsilon) - V(x,t+\epsilon). \end{aligned} \tag{82}$$

We then have:

$$\begin{aligned} &V(y(\epsilon), t+\epsilon) - V(x,t+\epsilon) \\ &= V\left(x + \int_t^{t+\epsilon} A(y(s))ds - \int_t^{t+\epsilon} BN^{-1}B^*\Gamma(y(s),s)ds, t+\epsilon\right) - V(x,t+\epsilon) \\ &= \left(\Gamma(x,t+\epsilon), \int_t^{t+\epsilon} A(y(s))ds - \int_t^{t+\epsilon} BN^{-1}B^*\Gamma(y(s),s)ds\right) \\ &\quad + \int_0^1 \Bigg(\Gamma\left(x + \theta\int_t^{t+\epsilon} (A(y(s)) - BN^{-1}B^*\Gamma(y(s)))ds, t+\epsilon\right) - \Gamma(x,t+\epsilon), \\ &\qquad \int_t^{t+\epsilon} (A(y(s)) - BN^{-1}B^*\Gamma(y(s)))ds\Bigg) d\theta. \end{aligned} \tag{83}$$

Using the fact that $\Gamma(x,t)$ is uniformly Lipschitz in x and continuous in t, we obtain easily from (83) that:

$$\frac{V(y(\epsilon), t+\epsilon) - V(x,t+\epsilon)}{\epsilon} \to (\Gamma(x,t), A(x) - BN^{-1}B^*\Gamma(x,t)).$$

Then, dividing (82) by ϵ and letting ϵ tend to 0, we obtain the PDE (81), recalling (79). The initial condition in (81) is trivial. If we take the gradient in x of (81), we recognize equation (16). Since this equation has a unique solution, the solution of (81) is also unique (easy checking). This completes the proof. □

7 Application to Mean Field Type Control Theory

7.1 Wasserstein Space

Denote by $\mathcal{P}_2(\mathbb{R}^n)$ the Wasserstein space of Borel probability measures m on $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} |x|^2 dm(x) < \infty$, with the metric

$$W_2(\mu, \nu) = \sqrt{\inf\left\{ \int |x-y|^2 d\pi(x,y) : \pi \in \Pi(\mu,\nu) \right\}}, \tag{84}$$

where $\Pi(\mu,\nu)$ is the space of all Borel probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ whose first and second marginals are μ and ν respectively.

7.2 Functional Derivatives

Let F be a functional on $\mathcal{P}_2(\mathbb{R}^n)$. We recall the idea of the functional derivative here.

Definition 1 F is said to have a functional derivative if there exists a continuous function $\dfrac{dF}{dm} : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \to \mathbb{R}$, such that for some $c : \mathcal{P}_2(\mathbb{R}^n) \to [0, \infty)$ which is bounded on bounded subsets, we have

$$\left| \frac{dF}{dm}(m,x) \right| \leq c(m)(1+|x|^2) \tag{85}$$

and

$$F(m') - F(m) = \int_0^1 \int_{\mathbb{R}^n} \frac{dF}{dm}(m+\theta(m'-m))(x) d(m'-m)(x) d\theta. \tag{86}$$

We require also $\int_{\mathbb{R}^n} \frac{dF}{dm}(m,x) dm(x) = 0$ as it is unique up to a constant by definition.

Definition 2 F is said to have a second order functional derivative if there exists a continuous function $\dfrac{d^2F}{dm^2} : \mathcal{P}_2 \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that, for some $c : \mathcal{P}_2(\mathbb{R}^n) \to [0, \infty)$ which is bounded on bounded subsets, we have

$$\left| \frac{d^2F}{dm^2}(m,x,\tilde{x}) \right| \leq c(m)(1+|x|^2+|\tilde{x}|^2) \tag{87}$$

and

$$\begin{aligned}
&F(m') - F(m) \\
&= \int_{\mathbb{R}^n} \frac{dF}{dm}(m)(x) d(m'-m)(x) \\
&\quad + \int_0^1 \int_0^1 \theta \frac{d^2F}{dm^2}(m+\lambda\theta(m'-m))(x,\tilde{x}) d(m'-m)(x) d(m'-m)(\tilde{x}) d\lambda d\theta.
\end{aligned} \tag{88}$$

Again, we require that $\int_{\mathbb{R}^n} \frac{d^2F}{dm}(m,x,\tilde{x})dm(\tilde{x}) = 0$, for all $x \in \mathbb{R}^n$, and $\int_{\mathbb{R}^n} \frac{d^2F}{dm}(m,x,\tilde{x})$ $dm(x) = 0$, for all $\tilde{x} \in \mathbb{R}^n$, as it is unique up to a constant. Note also that

$$\frac{d^2F}{dm^2}(m)(x,\tilde{x}) = \frac{d^2F}{dm^2}(m)(\tilde{x},x). \tag{89}$$

We write $D\frac{dF}{dm}(m)(x)$ to mean differentiating with respect to x, and $D_1\frac{d^2F}{dm^2}(m)(x_1,x_2)$ and $D_2\frac{d^2F}{dm^2}(m)(x_1,x_2)$ to denote partial differentiation with respect to x_1 and x_2, respectively.

7.3 Mean Field Type Control Problems

We introduce the setting of a mean-field type control problem. Consider real-valued functions $f(x,m)$ and $h(x,m)$ defined on $\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$. We define

$$\begin{aligned}
F(m) &:= \int_{\mathbb{R}^n} f(x,m)dm(x), \\
F_T(m) &:= \int_{\mathbb{R}^n} h(x,m)dm(x).
\end{aligned}$$

Fix a $m \in \mathcal{P}_2(\mathbb{R}^n)$. Let $A, B : \mathbb{R}^n \to \mathbb{R}^n$ be matrices, and $N : \mathbb{R}^n \to \mathbb{R}^n$ be a self-adjoint invertible matrix. We make the following assumptions on f, h, B, N, A. We assume that

(A1) $\forall x \in \mathbb{R}^n$,

$$BN^{-1}B^*x \cdot x \geq m|x|^2, m > 0. \tag{90}$$

(A2) f is regular enough such that the following is justifiable. $\forall y \in \mathbb{R}^n$,

$$\nu|y|^2 \leq \frac{\partial^2 f}{\partial x^2}(x,m)y \cdot y \leq M|y|^2, \tag{91}$$

$$\nu|y|^2 \leq D_\xi^2 \frac{\partial f}{\partial m}(x,m)(\xi)y \cdot y \leq M|y|^2, \tag{92}$$

$$D_\xi \frac{\partial^2 f}{\partial x \partial m}(x,m)(\xi) = 0, \tag{93}$$

$$D_{\xi_1} D_{\xi_2} \frac{\partial^2 f}{\partial m^2}(x,m)(\xi_1,\xi_2) = 0. \tag{94}$$

(A3) h is regular enough such that the following is justifiable. $\forall y \in \mathbb{R}^n$,

$$\nu_T|y|^2 \leq \frac{\partial^2 h}{\partial x^2}(x,m)y \cdot y \leq M_T|y|^2, \tag{95}$$

$$\nu_T|y|^2 \leq D_\xi^2 \frac{\partial h}{\partial m}(x,m)(\xi)y \cdot y \leq M_T|y|^2, \tag{96}$$

$$D_\xi \frac{\partial^2 h}{\partial x \partial m}(x,m)(\xi) = 0, \tag{97}$$

$$D_{\xi_1} D_{\xi_2} \frac{\partial^2 h}{\partial m^2}(x,m)(\xi_1,\xi_2) = 0. \tag{98}$$

(A4) For the matrices, we have

$$|A| < M|BN^{-1}B^*|, \text{ with } |\cdot| \text{ the matrix 2-norm.} \tag{99}$$

The set of our feasible control is $L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))$, i.e.,

$$v_{\cdot,m,t}(\cdot) \in L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))$$

if and only if

$$\int_t^T \int_{\mathbb{R}^n} |v_{x,m,t}(s)|^2 dm(x)ds < \infty.$$

To each $v_{\cdot,m,t}(\cdot) \in L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))$ and $x \in \mathbb{R}^n$ we associate the state

$$x_{x,m,t}(s;v) := x + \int_t^s \Big[Ax_{x,m,t}(\tau;v) + Bv_{x,m,t}(\tau)\Big]d\tau. \tag{100}$$

Note that $x_{\cdot,m,t}(\cdot) \in L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))$. We define the objective functional on $L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))$ by

$$J_{m,t}(v) := \int_t^T F(x_{\cdot,m,t}(s;v)_{\#}m)ds + F_T(x_{\cdot,m,T}(s;v)_{\#}m) \tag{101}$$
$$+ \frac{1}{2}\int_t^T \int_{\mathbb{R}^n} v^*_{x,m,t}(\tau)Nv_{x,m,t}(\tau)dm(x)d\tau.$$

Thus the value function is

$$V(m,t) := \inf_{v \in L^2(t,T;L^2_m(\mathbb{R}^n;\mathbb{R}^n))} J_{m,t}(v). \tag{102}$$

7.4 The Hilbert Space $\mathcal{H}_m$ and the Push-Forward Map

We proceed as our previous works [6, 7].

7.4.1 Settings

Fix $m \in \mathcal{P}_2(\mathbb{R}^n)$, we define $\mathcal{H}_m := L^2_m(\mathbb{R}^n;\mathbb{R}^n)$, the set of all measurable vector field Φ such that $\int_{\mathbb{R}^n} |\Phi(x)|^2 dm(x) < \infty$. We equip $\mathcal{H}_m$ with the inner product

$$\langle X, Y \rangle_{\mathcal{H}_m} := \int_{\mathbb{R}^n} X(x) \cdot Y(x) dm(x). \tag{103}$$

Write the corresponding norm as $\|X\|_{\mathcal{H}_m} = \sqrt{\langle X, X \rangle_{\mathcal{H}_m}}$.

Definition 3 For $m \in \mathcal{P}_2$, $X \in \mathcal{H}_m$, define $X \otimes m \in \mathcal{P}_2$ as follow: for all $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $x \mapsto \dfrac{|\phi(x)|}{1+|x|^2}$ is bounded, define

$$\int \phi(x) d(X \otimes m)(x) := \int \phi(X(x)) dm(x). \tag{104}$$

Remark 2 This actually is the push-forward map as we are working on the deterministic case. We write as $X \otimes m$ to align with our treatment of the stochastic case in [7].

We recall several useful properties from [7].

Proposition 4 *We have the following properties:*

1. *Let X, $Y \in \mathcal{H}_m$, and suppose $X \circ Y \in \mathcal{H}_m$. Then $(X \circ Y) \otimes m = X \otimes (Y \otimes m)$.*
2. *If $X(x) = x$ is the identity map, then $X \otimes m = m$.*
3. *Let $X \in \mathcal{H}_m$, denote the space $L^2_X(t,T;\mathcal{H}_m)$ to be the set of all processes in $L^2(t,T;\mathcal{H}_m)$ that is adapted to $\sigma(X)$. There exists a natural linear isometry between $L^2_X(t,T;\mathcal{H}_m)$ and $L^2(t,T;\mathcal{H}_{X\otimes m})$.*

Proof Please refer to [7] Section 2 and Section 3. □

7.4.2 Extending the Domain of Functions to $\mathcal{H}_m$

The proofs in this section is standard, we therefore omit unless specified. Readers may refer to [7] Section 2. Let $F : \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$, we extend F to be a function on $\mathcal{H}_m$

by $X \mapsto F(X \otimes m)$, $\forall X \in \mathcal{H}_m$. When the domain is $\mathcal{H}_m$, we can talk about Gâteaux derivative. We actually have the following relation between the Gâteaux derivative on $\mathcal{H}_m$ and its functional derivative:

Proposition 5 *Let $F : \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}$ have a functional derivative $\frac{dF}{dm}$, and $x \mapsto \frac{dF}{dm}(m, x)$ is continuously differentiable in $\mathbb{R}^n$. Assume that $D\frac{dF}{dm}(m, x)$ is continuous in both m and x, and*

$$\left| D\frac{dF}{dm}(m)(x) \right| \leq c(m)(1 + |x|) \tag{105}$$

for some constant $c(m)$ depending only on m. Denote the Gâteaux derivative as $D_X F(X \otimes m)$, we have

$$D_X F(X \otimes m) = D\frac{dF}{dm}(X \otimes m)(X(\cdot)). \tag{106}$$

We now look at the second order Gâteaux derivative, denoted as $D_X^2 F(X \otimes m)$, note that $D_X^2 F(X \otimes m)$ is a bounded linear operator from $\mathcal{H}_m$ to $\mathcal{H}_m$.

Proposition 6 *In addition to the assumptions in Proposition 5, let F has a second order functional derivative $\frac{d^2F}{dm^2}(m)(x_1, x_2)$, assume also $D^2\frac{dF}{dm}(m)(x)$, $D_1\frac{d^2F}{dm^2}(m)(x_1, x_2)$, $D_2\frac{d^2F}{dm^2}(m)(x_1, x_2)$ and*
$D_1 D_2\frac{d^2F}{dm^2}(m)(x_1, x_2)$ exist and are continuous, such that

$$\left| D^2\frac{dF}{dm}(m)(x) \right| \leq d(m), \tag{107}$$

$$\left| D_1 D_2\frac{d^2F}{dm^2}(m)(x_1, x_2) \right| \leq d'(m), \tag{108}$$

where d, d′ are constants depending on m only, and $|\cdot|$ is the matrix 2-norm. Then we have:

$$\begin{aligned} D_X^2 F(X \otimes m)Y(x) &= D^2\frac{dF}{dm}(X \otimes m)(X(x))Y(x) \\ &\quad + \int_{\mathbb{R}^n} D_1 D_2\frac{d^2F}{dm^2}(X \otimes m)(X(x), X(x'))Y(x')dm(x'). \end{aligned} \tag{109}$$

Besides, we can view $F(X \otimes m)$ as $m \mapsto F(X \otimes m)$, in this case, we can talk about differentiation with respect to m, denote it as $\dfrac{\partial F}{\partial m}$. The following relation between $\dfrac{\partial F}{\partial m}$ and $\dfrac{dF}{dm}$ holds.

Proposition 7 *Let $F : \mathcal{P}_2(\mathbb{R}^n) \mapsto \mathbb{R}^n$ have a functional derivative and fix $X \in \mathcal{H}_m$. We have*

$$\frac{\partial F}{\partial m}(X \otimes m)(x) = \frac{dF}{dm}(X \otimes m)(X(x)). \tag{110}$$

Now let $A:\mathbb{R}^n\to\mathbb{R}^n$, we extend it as $\forall X\in\mathcal{H}_m$, $X\mapsto A(X)\in\mathcal{H}_m$,

$$A(X)(x)=A(X(x)). \tag{111}$$

It is trivial to see that if A^{-1} exists in $\mathbb{R}^n$, then $A^{-1}(X)(x)=A^{-1}(X(x))$ is the inverse of A in $\mathcal{H}_m$. So is the transpose of A, if A is a matrix. Again, we can talk about its Gâteaux derivative.

Proposition 8 *Let A to be continuously differentiable. Denote its derivative to be dA. Assume that there exists k such that $|dA(x)|\le k$ for all $x\in\mathbb{R}^n$, where $|\cdot|$ is the matrix 2-norm. Then for all $X,Y\in\mathcal{H}_m$, we have*

$$D_XA(X)Y(x)=dA(X(x))Y(x). \tag{112}$$

Proof Let $X,Y,H\in\mathcal{H}_m$, then

$$\begin{aligned}&\frac{1}{\epsilon}\Big\langle A(X+\epsilon Y)-A(X),H\Big\rangle_{\mathcal{H}_m}\\&=\frac{1}{\epsilon}\int_{\mathbb{R}^n}\Big[A(X(\xi)+\epsilon Y(\xi))-A(X(\xi))\Big]\cdot H(\xi)dm(\xi)\\&=\int_{\mathbb{R}^n}\int_0^1 dA(X(\xi)+\theta\epsilon Y(\xi))Y(\xi)\cdot H(\xi)d\theta dm(\xi)\\&\to\int_{\mathbb{R}^n}dA(X(\xi))Y(\xi)\cdot H(\xi)dm(\xi)=\Big\langle dA(X(\cdot))Y(\cdot),H\Big\rangle_{\mathcal{H}_m}.\end{aligned}$$

Proposition 9 *Let A be twice continuously differentiable. Denote its second derivative to be d^2A. Note that $d^2A(x)(a,b)\in\mathbb{R}^n$, and $d^2A(x)(a,b)=d^2A(x)(b,a)$. Assume that there exists $k(x)$ such that $\forall a,b\in\mathbb{R}^n$, $|d^2A(x)(a,b)|\le k(x)$, then we have*

$$d^2A(X)(Y,W)(x)=d^2A(X(x))(Y(x),W(x)). \tag{113}$$

Proof Let $X,Y,W,H\in\mathcal{H}_m$, then

$$\begin{aligned}&\frac{1}{\epsilon}\Big\langle D_XA(X+W)Y-D_XA(X)Y,H\Big\rangle_{\mathcal{H}_m}\\&=\frac{1}{\epsilon}\int_{\mathbb{R}^n}\Big[dA(X(\xi)+\epsilon W(\xi))Y(\xi)-dA(X(\xi))Y(\xi)\Big]\cdot H(\xi)dm(\xi)\\&=\int_{\mathbb{R}^n}\int_0^1 d^2A(X(\xi)+\theta\epsilon W(\xi))(Y(\xi),W(\xi))\cdot H(\xi)d\theta dm(\xi)\\&\to\int_{\mathbb{R}^n}d^2A(X(\xi))(Y(\xi),W(\xi))\cdot H(\xi)dm(\xi).\end{aligned}$$

7.5 Control Problem in the Hilbert Space $\mathcal{H}_m$

Recall the definitions of A, B, N, F, F_T in Section 7.3. Extend the functions as in Section 7.4.2. We assume (A1), (A2), (A3) and (A4). It is not hard to derive (9), (7) and (42) from the assumptions. Note that in our case, $b = 0$.
Now fix $X \in \mathcal{H}_m$ as our initial data. For given $v_{Xt} \in L^2_X(t,T;\mathcal{H}_m)$ (subscript X and t to address the measurability and starting time), consider the dynamics:

$$X(s) = X + \int_t^s \Big[AX(\tau) + Bv_{Xt}(\tau)\Big] d\tau. \tag{114}$$

Denote the process as $X_{Xt}(s) = X_{Xt}(s; v_{Xt})$. Define the cost functional:

$$J_{Xt}(v_{Xt}) := \int_t^T F(X_{Xt}(s) \otimes m) ds + F_T(X_{Xt}(T) \otimes m) + \frac{1}{2}\int_t^T \langle v_{Xt}(\tau), Nv_{Xt}(\tau)\rangle_{\mathcal{H}_m} d\tau, \tag{115}$$

and the value function is

$$V(X,t) := \inf_{v_{Xt} \in L^2_X(t,T;\mathcal{H}_m)} J_{Xt}(v_{Xt}). \tag{116}$$

This is in the form of our concerned model in Section 2, with the Hilbert space being $\mathcal{H}_m$.
While (114) is infinite dimensional, there is a finite dimensional view point of it. For $v_{Xt} \in L^2_X(t,T;\mathcal{H}_m)$, by Proposition 4, let $\tilde{v} \in L^2(t,T;\mathcal{H}_{X\otimes m})$ be the representative of v_{Xt}. Consider

$$x(s) = x + \int_t^s \Big[Ax(\tau) + B\tilde{v}(\tau, x)\Big] d\tau. \tag{117}$$

Denote the solution to be $x(s; x, \tilde{v}(\cdot, x))$. Then we have

$$X_{Xt}(s; v_{Xt})(x) = x(s; X(x), \tilde{v}(\cdot, X(x))).$$

We introduce the notation $X_{xt}(\cdot)$ with a lowercase letter for x to mean $x(\cdot; x, \tilde{v}(\cdot, x))$, and $v_{\cdot t}(s)$ to mean $\tilde{v}(s, \cdot)$. From above we can conclude that the law of $X_{Xt}(s; v_{Xt}(\cdot))$ is $x(s; \cdot, \tilde{v}(\cdot, \cdot)) \otimes (X \otimes m)$. Hence the cost functional (115) can be written as

$$
\begin{aligned}
&J_{Xt}(v_{Xt})\\
&=\int_t^T F(X_{Xt}(s)\otimes m)ds + F_T(X_{Xt}(T)\otimes m) + \frac{1}{2}\int_t^T \langle v_{Xt}(\tau), N v_{Xt}(\tau)\rangle_{\mathcal{H}_m} d\tau\\
&=\int_t^T F(x(s;\cdot,\tilde{v}(\cdot,\cdot))\otimes(X\otimes m))ds + F_T(x(T;\cdot,\tilde{v}(\cdot,\cdot))\otimes(X\otimes m))\\
&\quad+\frac{1}{2}\int_t^T \langle v_{Xt}(\tau), N v_{Xt}(\tau)\rangle_{\mathcal{H}_m} d\tau\\
&=: J_{X\otimes m,t},
\end{aligned}
\tag{118}
$$

that means J depends on X only through $X\otimes m$. Respectively,

$$
V(X,t) = \inf_{v_{Xt}\in L^2_X(t,T;\mathcal{H}_m)} J_{Xt}(v_{Xt}) = \inf_{v_{Xt}\in L^2_X(t,T;\mathcal{H}_m)} J_{X\otimes m,t}(v_{Xt}) =: V(X\otimes m,t).
\tag{119}
$$

7.6 Necessary and Sufficient Condition for Optimality

Assume (A1), (A2), (A3) and (A4), we conclude from Theorem 2 that there exists unique optimal control $\hat{v}_{Xt}(s) = -N^{-1}B^* Z_{Xt}(s)$, where $Z_{Xt}(s)$ together with $Y_{Xt}(s)$ are the unique solution of the system

$$
Y_{Xt}(s) = X + \int_t^s \left[AY_{Xt}(\tau) - BN^{-1}B^* Z_{Xt}(\tau)\right]d\tau,
\tag{120}
$$

$$
Z_{Xt}(s) = \int_s^T \left[(AY_{Xt}(\tau))^* Z_{Xt}(\tau) + D_X F(Y_{Xt}(\tau)\otimes m)\right] + D_X F_T(Y_{Xt}(T)\otimes m).
\tag{121}
$$

Again, because $L^2_X(t,T;\mathcal{H}_m)$ is isometric to $L^2(t,T;\mathcal{H}_{X\otimes m})$, there exists $Y_{\xi t}(s)$, $Z_{\xi t}(s)$ such that $Y_{Xt} = Y_{\xi t}|_{\xi=X}$ and $Z_{Xt} = Z_{\xi t}|_{\xi=X}$, $(Y_{\xi t}, Z_{\xi t})$ solving

$$
Y_{\xi t}(s) = \xi + \int_t^s \left[AY_{\xi t}(\tau) - BN^{-1}B^* Z_{\xi t}(\tau)\right]d\tau,
\tag{122}
$$

$$
\begin{aligned}
Z_{\xi t}(s) = &\int_s^T \left[(AY_{\xi t}(\tau))^* Z_{\xi t}(\tau) + D\frac{dF}{dm}(Y_{\cdot t}(\tau)\otimes(X\otimes m))(Y_{\xi t}(\tau))\right]\\
&+ D\frac{dF_T}{dm}(Y_{\cdot t}(T)\otimes(X\otimes m))(Y_{\xi t}(T)).
\end{aligned}
\tag{123}
$$

As $(Y_{\xi t}, Z_{\xi t})$ depends on m through $X\otimes m$, we write $(Y_{\xi,X\otimes m,t}, Z_{\xi,X\otimes m,t})$. We can write the value function as

$$
\begin{aligned}
V(X,t) &= \int_t^T F(Y_{Xt}(s) \otimes m)ds + F_T(Y_{Xt}(T) \otimes m) \\
&\quad + \frac{1}{2}\int_t^T \langle N^{-1}B^* Z_{Xt}(\tau), B^* Z_{Xt}(\tau)\rangle_{\mathcal{H}_m} d\tau \\
&= \int_t^T F(Y_{\cdot,X\otimes m,t}(s) \otimes (X \otimes m))ds + F_T(Y_{\cdot,X\otimes m,t}(T) \otimes (X \otimes m)) \\
&\quad + \frac{1}{2}\int_t^T \int_{\mathbb{R}^n} N^{-1}B^* Z_{\xi,X\otimes m,t}(\tau) \cdot B^* Z_{\xi,X\otimes m,t}(\tau) d(X \otimes m)(\xi) d\tau \\
&= V(X \otimes m, t).
\end{aligned} \tag{124}
$$

In particular, if we choose X to be the identity function, i.e., $X(x) = x$, recall that $X \otimes m = m$, there exists $(Y_{x,m,t}, Z_{x,m,t})$ solving

$$
Y_{x,m,t}(s) = x + \int_t^s \left[AY_{x,m,t}(\tau) - BN^{-1}B^* Z_{x,m,t}(\tau)\right] d\tau, \tag{125}
$$

$$
\begin{aligned}
Z_{x,m,t}(s) = \int_s^T &\left[(AY_{x,m,t}(\tau))^* Z_{x,m,t}(\tau) + D\frac{dF}{dm}(Y_{\cdot,m,t}(\tau) \otimes m)(Y_{x,m,t}(\tau))\right] \\
&+ D\frac{dF_T}{dm}(Y_{\cdot,m,t}(T) \otimes (X \otimes m))(Y_{x,m,t}(T)),
\end{aligned} \tag{126}
$$

which is the system of optimality condition for our mean field type control problem in Section 7.3. For the value function, we have

$$
\begin{aligned}
V(m,t) = \int_t^T &F(Y_{\cdot,m,t}(s) \otimes m)ds + F_T(Y_{\cdot,m,t}(T) \otimes m) \\
&+ \frac{1}{2}\int_t^T \int_{\mathbb{R}^n} N^{-1}B^* Z_{x,m,t}(\tau) \cdot B^* Z_{x,m,t}(\tau) dm(x) d\tau.
\end{aligned} \tag{127}
$$

7.7 Properties of the Value Function

We give the functional derivative of the value function V, and the relation between the solution of the FBSDE and V. As the proofs are standard, we omit here and readers may refer to Section 4 of [7].

Proposition 10 *Assume (A1), (A2), (A3), (A4). We have the following properties for the value function:*

1. By Proposition 3, we have

$$
D_X V(X \otimes m, t) = Z_{Xt}(t). \tag{128}
$$

2. We have

$$\begin{aligned}\frac{dV}{dm}(m,t)(x) = &\int_t^T \frac{dF}{dm}(Y_{\cdot,m,t}(s)\otimes m)(Y_{x,m,s}(s))ds \\ &+ \frac{dF_T}{dm}(Y_{\cdot,m,t}(T)\otimes m)(Y_{x,m,s}(T)) \\ &+ \frac{1}{2}\int_t^T N^{-1}B^* Z_{x,m,t}(\tau)\cdot B^* Z_{x,m,t}(\tau)d\tau.\end{aligned} \tag{129}$$

3. We have

$$D\frac{d}{dm}V(m,t)(x) = Z_{x,m,t}(t), \tag{130}$$

$$D_X V(X\otimes m,t) = D\frac{d}{dm}V(X\otimes m,t)(X) \tag{131}$$

4. Also, the feedback nature of Z in Y, i.e., for any $x \in \mathbb{R}^n$, $\forall s \in [t,T]$, we have

$$Z_{x,m,t}(s) = D\frac{d}{dm}V(Y_{\cdot,m,t}\otimes m, s)(Y_{x,m,t}(s)), \tag{132}$$

and for any $X \in \mathcal{H}_m$, $\forall s \in [t,T]$,

$$Z_{Xt}(s) = D_X V(Y_{Xt}(s)\otimes m, s). \tag{133}$$

7.8 Bellman Equation

Assume (A1), (A2), (A3), (A4). By Theorem 4, we deduce that for any $T > 0$, $V(X\otimes m,t)$ is the unique solution to the following Bellman equation:

$$\begin{cases}-\dfrac{\partial V}{\partial t}(X\otimes m,t) - \left\langle D_X V(X\otimes m,t), AX\right\rangle_{\mathcal{H}_m} \\ \qquad +\dfrac{1}{2}\left\langle D_X V(X\otimes m,t), BN^{-1}B^* D_X V(X\otimes m,t)\right\rangle_{\mathcal{H}_m} = F(X\otimes m), \\ V(X\otimes m) = F_T(X\otimes m).\end{cases} \tag{134}$$

As before, let X be the identity function, together with Proposition 10, we conclude that for any $T > 0$, $V(m,t)$ solves the following PDE on the space of probability measures:

$$\begin{cases}-\dfrac{\partial V}{\partial t}(m,t) - \displaystyle\int_{\mathbb{R}^n} D\frac{dV}{dm}(m,t)(x)\cdot Ax dm(x) \\ \qquad +\dfrac{1}{2}\displaystyle\int_{\mathbb{R}^n} D\frac{dV}{dm}(m,t)(x)\cdot BN^{-1}B^* D\frac{dV}{dm}(m,t)(x)dm(x) = F(m), \\ V(m,T) = F_T(m).\end{cases} \tag{135}$$

8 Acknowledgements

The first author's research is supported by grant from the National Science Foundation 1905449 and grant from the Hong Kong SAR RGC GRF 11303316. This work constitutes part of the second author's PhD study at City University of Hong Kong. The third author acknowledges the financial supports from HKGRF-14300717 with the project title "New kinds of Forward-backward Stochastic Systems with Applications", HKGRF-14300319 with the project title "Shape-constrained Inference: Testing for Monotonicity", HKGRF-14301321 with the project title "General Theory for Infinite Dimensional Stochastic Control: Mean Field and Some Classical Problems", and Germany/Hong Kong Joint Research Scheme Project No. G-HKU701/20 with the project title "Asymmetry in Dynamically correlated threshold Stochastic volatility model". He also thanks Columbia University for the kind invitation to be a visiting faculty member in the Department of Statistics during his sabbatical leave. He also recalled the unforgettable moments and the happiness shared with his beloved father and used this work in memory of his father's brave battle against liver cancer.

References

1. S. Ahuja, W. Ren, and T. Yang. Forward-backward stochastic differential equations with monotone functionals and mean field games with common noise. *Stochastic Processes and their Applications*, 129(10):3859–3892, 2019.
2. F. Antonelli. Backward-forward stochastic differential equations. *Ann. Appl. Probab.*, 3:777–793, 1993.
3. A. Bensoussan, J. Frehse, and P. Yam. *Mean Field Games and Mean Field Type Control Theory.* Springer, 2013.
4. A. Bensoussan, J. Frehse, and S. C. P. Yam. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées*, 2014.
5. A. Bensoussan, J. Frehse, and S. C. P. Yam. On the interpretation of the master equation. *Stochastic Processes and their Applications*, 127(7):2093–2137, 2017.
6. A. Bensoussan, P. J. Graber, and S. C. P. Yam. Stochastic control on space of random variables. *arXiv:1903.12602*, 2019.
7. A. Bensoussan, P. J. Graber, and S. C. P. Yam. Control on Hilbert spaces and application to mean field type control theory. *arXiv:2005.10770*, 2020.
8. A. Bensoussan, K. C. J. Sung, S.C.P. Yam, and S.P. Yung. Linear quadratic mean field games. *Journal of Optimization Theory and Applications*, 169(2):469–529, 2016.
9. A. Bensoussan and S. C. P. Yam. Control problem on space of random variables and master equation. *ESAIM: Control, Optimisation and Calculus of Variations*, 25:10, 2019.
10. A. Bensoussana, S. C. P. Yam, and Z. Zhang. Well-posedness of mean-field type forward-backward stochastic differential equations. *Stochastic Processes and their Applications*, 125(9):3327–3354, 2015.
11. R. Buckdahn, J. Li, S. Peng, and C. Rainer. Mean-field stochastic differential equation and associated PDEs. *The Annals of Probability*, 45:824–878, 2017.
12. P. Cardaliaguet, F. Delarue, J.M. Lasry, and P.L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games: (AMS-201)*, volume 201. Princeton University Press, 2019.

13. R. Carmona and F. Delarue. Mean field forward-backward stochastic differential equations. *Electron. Commun. Probab.*, 18, 2013.
14. R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.*, 43(5):2647–2700, 2015.
15. R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018.
16. A. Cosso and H. Pham. Zero-sum stochastic differential games of generalized McKean-Vlasov type. *Journal de Mathematiques Pures et Appliquees*, 129:180–212, 2019.
17. F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Processes and their Applications*, 99:209–286, 2002.
18. M.F. Djete, D. Possamai, and X. Tan. McKean-Vlasov optimal control: The dynamic programming principle. *arXiv:1907.08860*, 2019.
19. W. Gangbo and A.R. Mészáros. Global well-posedness of master equations for deterministic displacement convex potential mean field games. *arXiv:2004.01660*, 2020.
20. Y. Hu and S. Peng. Solution of forward-backward stochastic differential equations. *Probab. Theory and Related Fields*, 103:273–283, 1995.
21. Y. Hu and J. Yong. Forward-backward stochastic differential equations with nonsmooth coefficients. *Stochastic Processes and their Applications*, 87(1):93–106, 2000.
22. P. L. Lions. Lectures at college de france. *http://www.college-de-france.fr*.
23. P. L. Lions. Seminar at college de france. 2014.
24. J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields*, 98:339–359, 1994.
25. J. Ma, Z. Wu, D. Zhang, and J. Zhang. On well-posedness of Forward-Backward SDEs — a unified approach. *Annals of Applied Probability*, 25(4):2168–2214, 2015.
26. C. Mou and J. Zhang. Weak solutions of mean field game master equations. *arXiv:1903.09907*, 2019.
27. E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory and Related Fields*, 114(2):123–150, 1999.
28. S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.*, 37(3):825–843, 1999.
29. H. Pham and X. Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM Journal on Control and Optimization*, 55(2):1069–1101, 2017.
30. J. Yong. Finding adapted solutions of forward-backward stochastic differential equations: Method of continuation. *Probab. Theory and Related Fields*, 107(4):537–572, 1997.
31. J. Yong. Forward-backward stochastic differential equations with mixed initial-terminal conditions. *Trans. Amer. Math. Soc.*, 362(2):1047–1096, 2010.
32. J. Zhang. The wellposedness of FBSDEs. *Discrete & continuous dynamical systems - B*, 6(4):927–940, 2006.

Risk-Sensitive Markov Decision Problems under Model Uncertainty: Finite Time Horizon Case

Tomasz R. Bielecki, Tao Chen, and Igor Cialenco

Abstract In this paper we study a class of risk-sensitive Markovian control problems in discrete time subject to model uncertainty. We consider a risk-sensitive discounted cost criterion with finite time horizon. The used methodology is the one of adaptive robust control combined with machine learning.

1 Introduction

The main goal of this work is to study finite time horizon *risk-sensitive Markovian control problems* subject to model uncertainty in a discrete time setup, and to develop a methodology to solve such problems efficiently. The proposed approach hinges on the following main building concepts: incorporating model uncertainty through the *adaptive robust* paradigm introduced in [BCC$^+$19] and developing efficient numerical solutions for the obtained Bellman equations by adopting the *machine learning techniques* proposed in [CL19].

There exists a significant body of work on incorporating model uncertainty (or model misspecification) in stochastic control problems, and among some of the well-known and prominent methods we would mention the robust control approach [GS89, HSTW06, HS08], adaptive control [KV15, CG91], and Bayesian adaptive control [KV15]. A comprehensive literature review on this subject is beyond the scope of this paper, and we refer the reader to [BCC$^+$19] and references therein. In [BCC$^+$19] the authors proposed a novel adaptive robust methodology that solves

Tomasz R. Bielecki,
Department of Applied Mathematics, Illinois Institute of Technology,
e-mail: tbielecki@iit.edu Tao Chen,
Department of Mathematics, University of Michigan,
e-mail: chenta@umich.edu and Igor Cialenco
Department of Applied Mathematics, Illinois Institute of Technology,
e-mail: cialenco@iit.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_2

time-consistent Markovian control problems in discrete time subject to model uncertainty - the approach that we take in this study too. The core of this methodology was to combine a recursive learning mechanism about the unknown model with the underlying Markovian dynamics, and to demonstrate that the so called adaptive robust Bellman equations produce an optimal adaptive robust control strategy.

In contrast to [BCC$^+$19], where the considered optimization criterion was of the terminal reward type, in the present work, we also allow intermediate rewards and we use the discounted risk sensitive criterion. Accordingly, we derive a new set of adaptive robust Bellman equations, similar to those used in [BCC$^+$19].

Risk sensitive criterion has been broadly used both in the control oriented literature, as well as in the game oriented literature. We refer to, e.g., [BP03, DL14, BR17], and the references therein for insight into risk sensitive control and risk sensitive games both in discrete time and in continuous time.

The paper is organized as follows. In Section 2 we formulate the finite time horizon risk-sensitive Markovian control problem subject to model uncertainty that is studied here. Section 3 is devoted to the formulation and to study of the robust adaptive control problem that is relevant for the problem formulated in Section 2. This section presents the main theoretical developments of the present work. In Section 4 we formulate an illustrative example of our theoretical results that is rooted in the classical linear-quadratic-exponential control problem (see e.g. [HS95]). Next, using machine learning methods, in Section 5 we provide numerical solutions of the example presented in Section 4.

Finally, we want to mention that the important case of an infinite time horizon risk-sensitive Markovian control problem in discrete time subject to model uncertainty will be studies in a follow-up work.

2 Risk-sensitive Markovian discounted control problems with model uncertainty

In this section we state the underlying discounted risk-sensitive stochastic control problems. Let $(\Omega,\mathscr{F})$ be a measurable space, $T\in\mathbb{N}$ be a finite time horizon, and let us denote by $\mathscr{T}:=\{0,1,2,\ldots,T\}$ and $\mathscr{T}':=\{0,1,2,\ldots,T-1\}$. We let $\Theta\subset\mathbb{R}^d$ be a non-empty compact set, which will play the role of the parameter space throughout. We consider a random process $Z=\{Z_t,\ t=1,2\ldots\}$ on $(\Omega,\mathscr{F})$ taking values in $\mathbb{R}^m$, and we denote by $\mathbb{F}=(\mathscr{F}_t,\ t=0,2\ldots)$ its natural filtration, with $\mathscr{F}_0=\{\emptyset,\Omega\}$. We postulate that this process is observed by the controller, but the true law of Z is unknown to the controller and assumed to be generated by a probability measure belonging to a (known) parameterized family of probability distributions on $(\Omega,\mathscr{F})$, say $\mathbf{P}(\Theta)=\{\mathbb{P}_\theta,\theta\in\Theta\}$. As usually, $\mathbb{E}_{\mathbb{P}}$ will denote the expectation under a probability measure $\mathbb{P}$ on $(\Omega,\mathscr{F})$, and, for simplicity, we will write $\mathbb{E}_\theta$ instead of $\mathbb{E}_{\mathbb{P}_\theta}$. We denote by $\mathbb{P}_{\theta^*}$ the measure generating the true law of Z, and thus $\theta^*\in\Theta$ is the unknown true parameter. The sets Θ and $\mathbf{P}(\Theta)$ are known to the observer. Clearly, the model uncertainty may occur if $\Theta \neq \{\theta^*\}$, which we will assume to hold throughout.

We let $A \subset \mathbb{R}^k$ be a finite set,[1] and $S : \mathbb{R}^n \times A \times \mathbb{R}^m \to \mathbb{R}^d$ be a measurable mapping. An admissible control process φ is an $\mathbb{F}$-adapted process, taking values in A, and we will denote by $\mathscr{A}$ the set of all admissible control processes.

We consider an underlying discrete time controlled dynamical system with the state process X taking values in $\mathbb{R}^n$ and control process φ taking values in A. Specifically, we let

$$X_{t+1} = S(X_t, \varphi_t, Z_{t+1}), \quad t \in \mathscr{T}', \quad X_0 = x_0 \in \mathbb{R}^n. \tag{1}$$

At each time $t = 0, \ldots, T-1$, the running reward $r_t(X_t, \varphi_t)$ is delivered, where, for every $a \in A$, the function $r_t(\cdot, a) : \mathbb{R}^n \to \mathbb{R}_+$ is bounded and continuous. Similarly, at the terminal time $t = T$ the terminal reward $r_T(X_T)$ is delivered, where $r_T : \mathbb{R}^n \to \mathbb{R}_+$ is a bounded and continuous function.

Let $\beta \in (0,1)$ be a discount factor, and let $\gamma \neq 0$ be the risk sensitivity factor. The underlying discounted, risk-sensitive control problem is:

$$\sup_{\varphi \in \mathscr{A}} \frac{1}{\gamma} \ln \left(\mathbb{E}_{\theta^*} e^{\gamma \left(\sum_{t=0}^{T-1} \beta^t r_t(X_t, \varphi_t) + \beta^T r_T(X_T) \right)} \right) \tag{2}$$

subject to (1). Clearly, since θ^* is not known to the controller, the above problem can not be solved as it is stated. The main goal of this paper is formulate and solve the adaptive robust control problem corresponding to (2).

Remark 1 (i) The risk-sensitive criterion in (2) is in fact an example of application of the entropic risk measure, say $\rho_{\theta^*,\gamma}$, which is defined as

$$\rho_{\theta^*,\gamma}(\xi) := \frac{1}{\gamma} \ln \mathbb{E}_{\theta^*} e^{\gamma \xi},$$

where ξ is a random variable on $(\Omega, \mathscr{F}, P_{\theta^*})$ that admits finite moments of all orders.
(ii) It can be verified that

$$\rho_{\theta^*,\gamma}(\xi) = \mathbb{E}_{\theta^*}(\xi) + \frac{\gamma}{2} \mathbb{VAR}_{\theta^*}(\xi) + O(\gamma^2).$$

Thus, in case when $\gamma < 0$ the term $\frac{\gamma}{2} \mathbb{VAR}_{\theta^*}(\xi)$ can be interpreted as the risk-penalizing term. On the contrary, when $\gamma > 0$, the term $\frac{\gamma}{2} \mathbb{VAR}_{\theta^*}(\xi)$ can be viewed as the risk-favoring term.
(iii) In the rest of the paper we focus on the case $\gamma > 0$. The case $\gamma < 0$ can be treated in an analogous way.

[1] A will represent the set of control values, and we assume it is finite for simplicity, in order to avoid technical issues regarding the existence of measurable selectors.

3 The adaptive robust risk sensitive discounted control problem

We follow here the developments presented in [BCC+19]. The key difference is that in this work we deal with running and terminal costs.

In what follows, we will be making use of a recursive construction of confidence regions for the unknown parameter θ^* in our model. We refer to [BCC17] for a general study of recursive constructions of (approximate) confidence regions for time homogeneous Markov chains. Section 4 provides details of a specific such recursive construction corresponding to the example presented in that section. Here, we just postulate that the recursive algorithm for building confidence regions uses a Θ-valued and observed process, say $C = (C_t,\ t \in \mathbb{N}_0)$, satisfying the following abstract dynamics

$$C_{t+1} = R(t, C_t, Z_{t+1}), \quad t \in \mathbb{N}_0,\, C_0 = c_0 \in \Theta, \tag{3}$$

where $R : \mathbb{N}_0 \times \mathbb{R}^d \times \mathbb{R}^m \to \Theta$ is a deterministic measurable function. Note that, given our assumptions about process Z, the process C is $\mathbb{F}$-adapted. This is one of the key features of our model. Usually C_t is taken to be a consistent estimator of θ^*.

Now, we fix a confidence level $\alpha \in (0,1)$, and for each time $t \in \mathbb{N}_0$, we assume that an $(1-\alpha)$-confidence region, say $\Theta_t \subset \mathbb{R}^d$, for θ^*, can be represented as

$$\Theta_t = \tau(t, C_t), \tag{4}$$

where, for each $t \in \mathbb{N}_0$, $\tau(t,\cdot) : \mathbb{R}^d \to 2^{\Theta}$ is a deterministic set valued function, where, as usual, 2^{Θ} denotes the set of all subsets of Θ. Note that in view of (3) the construction of confidence regions given in (4) is indeed recursive. In our construction of confidence regions, the mapping $\tau(t,\cdot)$ will be a measurable set valued function, with compact values. It needs to be noted that we will only need to compute Θ_t until time $T-1$. In addition, we assume that for any $t \in \mathscr{T}'$, the mapping $\tau(t,\cdot)$ is upper hemi-continuous (u.h.c.). That is, for any $c \in \Theta$, and any open set E such that $\tau(t,c) \subset E \subset \Theta$, there exists a neighbourhood D of c such that for all $c' \in D$, $\tau(t,c') \subset E$ (cf. [Bor85, Definition 11.3]).

Remark 2 The important property of the recursive confidence regions constructed as indicated above is that, in many models, $\lim_{t\to\infty} \Theta_t = \{\theta^*\}$, where the convergence is understood $\mathbb{P}_{\theta^*}$ almost surely, and the limit is in the Hausdorff metric. This is not always the case though in general. In [BCC17] is shown that the convergence holds in probability, for the model setup studied there.

The sequence Θ_t, $t \in \mathscr{T}'$ represents learning about θ^* based on the observation of the history $(Y_0, Y_1 \ldots, Y_t)$, $t \in \mathscr{T}'$, where $Y_t = (X_t, C_t)$, $t \in \mathscr{T}$, is the augmented state process taking values in the augmented state space

$$E_Y = \mathbb{R}^n \times \Theta.$$

We denote by $\mathscr{E}_Y$ the collection of Borel measurable sets in E_Y.

In view of the above, if the control process φ is employed then the process Y has the following dynamics

$$Y_{t+1} = \mathbf{G}(t, Y_t, \varphi_t, Z_{t+1}), \quad t \in \mathscr{T}',$$

where the mapping $\mathbf{G} : \mathbb{N}_0 \times E_Y \times A \times \mathbb{R}^m \to E_Y$ is defined as

$$\mathbf{G}(t, y, a, z) = \big(S(x, a, z), R(t, c, z)\big), \tag{5}$$

with $y = (x, c) \in E_Y$.

We define the corresponding histories

$$H_t = (Y_0, \ldots, Y_t), \quad t \in \mathscr{T}', \tag{6}$$

so that

$$H_t \in \mathbf{H}_t = \underbrace{E_Y \times E_Y \times \ldots \times E_Y}_{t+1 \text{ times}}. \tag{7}$$

Clearly, for any admissible control process φ, the random variable H_t is $\mathscr{F}_t$-measurable. We denote by

$$h_t = (y_0, y_1, \ldots, y_t) = (x_0, c_0, x_1, c_1, \ldots, x_t, c_t) \tag{8}$$

a realization of H_t. Note that $h_0 = y_0 = (x_0, c_0)$.

A control process $\varphi = (\varphi_t,\ t \in \mathscr{T}')$ is called history dependent control process if (with a slight abuse of notation)

$$\varphi_t = \varphi_t(H_t),$$

where (on the right hand side) $\varphi_t : \mathbf{H}_t \to A$, is a measurable mapping. Given our above setup, any history dependent control process is $\mathbb{F}$–adapted, and thus, it is admissible. For any admissible control process φ and for any $t \in \mathscr{T}'$, we denote by $\varphi^t = (\varphi_k,\ k = t, \ldots, T-1)$ the 't-tail' of φ. Accordingly, we denote by $\mathscr{A}^t$ the collection of 't-tails' of φ. In particular, $\varphi^0 = \varphi$ and $\mathscr{A}^0 = \mathscr{A}$. The superscript notation applied to processes should not be confused with power function applied such as β^t.

Let $\psi_t : \mathbf{H}_t \to \Theta$ be a Borel measurable mapping such that $\psi_t(h_t) \in \tau(t, c_t)$, and let us denote by $\psi = (\psi_t,\ t \in \mathscr{T}')$ the sequence of such mappings, and by ψ^t the t-tails of the sequence ψ, in analogy to φ^t. The set of all sequences ψ, and respectively ψ^t , will be denoted by Ψ and Ψ^t, respectively.

Strategies φ and ψ are called *Markovian strategies or policies* if (with some abuse of notation)

$$\varphi_t = \varphi_t(Y_t), \quad \psi_t = \psi_t(Y_t),$$

where (on the right hand side) $\varphi_t : E_Y \to A$, and is a (Borel) measurable mapping, and $\psi_t : E_Y \to \Theta$ is a (Borel) measurable mapping satisfying $\psi_t(x, c) \in \tau(t, c)$.

In order to simplify all the following argument we limit ourselves to Markovian policies. In case of Markovian dynamics settings, such as ours, this comes without loss of generality, as there typically exist optimal Markovian strategies, if optimal strategies exist at all. Accordingly, $\mathscr{A}$ and Ψ are now sets of Markov strategies.

Next, for each $(t,y,a,\theta) \in \mathscr{T}' \times E_Y \times A \times \Theta$, we define a probability measure on $\mathscr{E}_Y$, given by

$$Q(B \mid t,y,a,\theta) = \mathbb{P}_\theta(Z_{t+1} \in \{z : \mathbf{G}(t,y,a,z) \in B\}) = \mathbb{P}_\theta\left(\mathbf{G}(t,y,a,Z_{t+1}) \in B\right), \quad (9)$$

for any $B \in \mathscr{E}_Y$ We assume that for every $t \in \mathscr{T}$ and every $a \in A$, we have that $Q(dy' \mid t,y,a,\theta)$ is a Borel measurable stochastic kernel with respect to (y,θ). This assumption will be strengthened later on.

Using Ionescu-Tulcea theorem (cf. [BR11, Appendix B]), for every $t = 0,\dots,T-1$, every t-tail $\varphi^t \in \mathscr{A}^t$ and every state $y_t \in E_Y$, we define the family $\mathscr{Q}^{\varphi^t,\Psi^t}_{y_t,t} = \{\mathbb{Q}^{\varphi^t,\psi^t}_{y_t,t},\ \psi^t \in \Psi^t\}$ of probability measures on the concatenated canonical space $\mathsf{X}^T_{s=t+1} E_Y$, with

$$\mathbb{Q}^{\varphi^t,\psi^t}_{y_t,t}(B_{t+1} \times \cdots \times B_T) := \int_{B_{t+1}} \cdots \int_{B_T} \prod_{u=t+1}^{T} Q(\mathrm{d}y_u \mid u-1, y_{u-1}, \varphi_{u-1}(y_{u-1}), \psi_{u-1}(y_{u-1})). \quad (10)$$

The *discounted, risk-sensitive, adaptive robust control problem* corresponding[2] to (2) is:

$$\sup_{\varphi^0 \in \mathscr{A}^0} \inf_{\mathbb{Q} \in \mathscr{Q}^{\varphi^0,\Psi^0}_{y_0,0}} \mathbb{E}_{\mathbb{Q}} e^{\gamma \sum_{t=0}^{T} \beta^t r_t(X_t,\varphi_t(Y_t))}, \quad (11)$$

where, for simplicity of writing, here and everywhere below, with slight abuse of notations, we set $r_T(x,a) = r_T(x)$. In next section we will show that a solution to this problem can be given in terms of the discounted adaptive robust Bellman equations associated to it.

3.1 Adaptive robust Bellman equation

Towards this end we aim our attention at the following adaptive robust Bellman equations

$$\begin{aligned} W_T(y) &= e^{\gamma\beta^T r_T(x)}, \quad y \in E_Y, \\ W_t(y) &= \max_{a\in A} \inf_{\theta \in \tau(t,c)} \int_{E_Y} W_{t+1}(y') e^{\gamma\beta^t r_t(x,a)} Q(\mathrm{d}y' \mid t,y,a,\theta), \\ &\qquad y \in E_Y,\ t = T-1,\dots,0, \end{aligned} \quad (12)$$

where we recall that $y = (x,c)$.

[2] Since $\gamma > 0$, we omit the factor $1/\gamma$.

Remark 3 Clearly, in (12), the exponent $e^{\gamma\beta^t r_t(x,a)}$ can be factored out, and W_t can be written as

$$W_t(y) = \max_{a\in A}\left(e^{\gamma\beta^t r_t(x,a)}\cdot \inf_{\theta\in\tau(t,c)}\int_{E_Y} W_{t+1}(y')Q(\mathrm{d}y' \mid t,y,a,\theta)\right).$$

Nevertheless, in what follows, we will keep similar factors inside of the integrals, mostly for the convenience of writing as well as to match the visual appearance of classical Bellman equations.

We will study the solvability of this system. We start with Lemma 1 below, where, under some additional technical assumptions, we show that the optimal selectors in (12) exist; namely, for any $t\in\mathscr{T}'$, and any $y=(x,c)\in E_Y$, there exists a measurable mapping $\varphi_t^*: E_Y\to A$, such that

$$W_t(y) = \inf_{\theta\in\tau(t,c)}\int_{E_Y} W_{t+1}(y')e^{\gamma\beta^t r_t(x,\varphi_t^*(y))}Q(\mathrm{d}y' \mid t,y,\varphi_t^*(y),\theta).$$

In order to proceed, for the sake of simplicity, we will assume that under measure $\mathbb{P}_\theta$, for each $t\in\mathscr{T}$, the random variable Z_t has a density with respect to the Lebesgue measure, say $f_Z(z;\theta),\ z\in\mathbb{R}^m$. In this case we have

$$\int_{E_Y} W_{t+1}(y')Q(\mathrm{d}y' \mid t,y,a,\theta) = \int_{\mathbb{R}^m} W_{t+1}(\mathbf{G}(t,y,a,z))f_Z(z;\theta)\,\mathrm{d}z,$$

where $\mathbf{G}(t,y,a,z)$ is given in (5).

Additionally, we take the standing assumptions:

(i) for any a and z, the function $S(\cdot,a,z)$ is continuous;
(ii) for each z, the function $f_Z(z;\cdot)$ is continuous;
(iii) for each $t\in\mathscr{T}'$, the function $R(t,\cdot,\cdot)$ is continuous.

Then, the following result holds true.

Lemma 1 *The functions* $W_t,\ t=T,T-1,\dots,0$, *are lower semi-continuous (l.s.c.), and the optimal selectors* $\varphi_t^*,\ t=T-1,\dots,0$, *realizing maxima in* (12) *exist.*

Proof Since r_T is continuous and bounded, so is the function W_T. Since $\mathbf{G}(T-1,\cdot,a,z)$ is continuous, then, $W_T(\mathbf{G}(T-1,\cdot,a,z))$ is continuous. Consequently, recalling again that $y=(x,c)$, for each a, the function

$$\begin{aligned} w_{T-1}(y,a,\theta) &:= \int_{\mathbb{R}} W_T(\mathbf{G}(T-1,y,a,z))e^{\gamma\beta^{T-1}r_{T-1}(x,a)}f_Z(z;\theta)\,\mathrm{d}z \\ &= e^{\gamma\beta^{T-1}r_{T-1}(x,a)}\int_{\mathbb{R}} e^{\gamma\beta^T r_T(S(x,a,z))}f_Z(z;\theta)\,\mathrm{d}z \end{aligned}$$

is continuous in (y,θ).

Next, we will apply [BS78, Proposition 7.33] by taking (in the notations of [BS78])

$$
\begin{aligned}
\mathrm{X} &= E_Y \times A = \mathbb{R}^n \times \Theta \times A, \quad \mathrm{x} = (y,a),\\
\mathrm{Y} &= \Theta, \quad \mathrm{y} = \theta,\\
\mathrm{D} &= \bigcup_{(y,a)\in E_Y\times A} \{(y,a)\} \times \tau(T-1,c),\\
f(\mathrm{x},\mathrm{y}) &= w_{T-1}(y,a,\theta).
\end{aligned}
$$

Note that in view of the prior assumptions, Y is metrizable and compact. Clearly X is metrizable. From the above, f is continuous, and thus lower semi-continuous. Since $\tau(T-1,\cdot)$ is compact-valued and u.h.c. on $E_Y \times A$, then according to [Bor85, Proposition 11.9], the set-valued function $\tau(T-1,\cdot)$ is closed, which implies that its graph D is closed [Bor85, Definition 11.5]. Also note that the cross section $\mathrm{D_x} = \mathrm{D}_{(y,a)} = \{\theta \in \Theta : (y,a,\theta) \in \mathrm{D}\}$ is given by $\mathrm{D}_{(y,a)}(T-1) = \tau(T-1,c)$. Hence, by [BS78, Proposition 7.33], the function

$$
\widetilde{w}_{T-1}(y,a) = \inf_{\theta\in\tau(T-1,c)} (w_{T-1}(y,a,\theta)), \quad (y,a) \in E_Y \times A,
$$

is l.s.c. Consequently, the function $\widehat{w}_{T-1}(y,a) = -\widetilde{w}_{T-1}(y,a)$ is u.s.c. (upper semi-continuous). Thus, by [BS78, Proposition 7.34], the function

$$
-W_{T-1}(y) = -\max_{a\in A} \widetilde{w}_{T-1}(y,a) = \min_{a\in A} \widehat{w}_{T-1}(y,a)
$$

is u.s.c., so that $W_{T-1}(y)$ is l.s.c. Moreover, since A is finite, there exists an optimal selector φ^*_{T-1}, that is $W_{T-1}(y) = \widetilde{w}_{T-1}(y, \varphi^*_{T-1}(y))$.

Proceeding to the next step, note that $W_{T-1}(\mathbf{G}(T-2,y,a,z))e^{\gamma\beta^{T-2} r_{T_2}(x,a)}$ is l.s.c. and positive, hence bounded from below. Therefore, according to [BS78, Proposition 7.31], the function

$$
w_{T-2}(y,a,\theta) = \int_{\mathbb{R}} W_{T-1}(\mathbf{G}(T-2,y,a,z))e^{\gamma\beta^{T-2} r_{T-2}(x,a)} f_Z(z;\theta)\,\mathrm{d}z
$$

is l.s.c.. The rest of the proof follows in the analogous way. □

Next, we will prove an auxiliary result needed to justify the mathematical operations conducted in the proof of the main result – Theorem 1. Define the functions U_t and U_t^* as follows: for $\varphi^t \in \mathscr{A}^t$ and $y \in E_Y$,

$$
U_t(\varphi^t, y) = e^{\gamma\beta^t r_t(x,\varphi_t(y))} \inf_{\mathbb{Q}\in\mathscr{Q}_{y,t}^{\varphi^t,\Psi^t}} \mathbb{E}_{\mathbb{Q}} e^{\gamma\sum_{k=t+1}^{T} \beta^k r_k(X_k,\varphi_k(Y_k))}, \quad t \in \mathscr{T}', \tag{13}
$$

$$
U_t^*(y) = \sup_{\varphi^t\in\mathscr{A}^t} U_t(\varphi^t, y), \quad t \in \mathscr{T}', \tag{14}
$$

$$
U_T^*(y) = e^{\gamma\beta^T r_T(x)}. \tag{15}
$$

We now have the following result.

Lemma 2 *For any $t \in \mathscr{T}'$, and for any $\varphi^t \in \mathscr{A}^t$, the function $U_t(\varphi^t, \cdot)$ is lower semi-ananlytic (l.s.a.) on E_Y. Moreover, there exists a sequence of universally measurable functions ψ_k^*, $k = t, \ldots, T-1$ such that*

$$U_t(\varphi^t, y) = e^{\gamma \beta^t r_t(x, \varphi_t(y))} \mathbb{E}_{\mathbb{Q}_{y,t}^{\varphi^t, \psi^{t,*}}} e^{\gamma \Sigma_{k=t+1}^T \beta^k r_k(X_k, \varphi_k(Y_k))}. \tag{16}$$

Proof According to (9), and using the definition of $\mathscr{Q}_{y,t}^{\varphi^t, \Psi^t}$, we have that

$$\begin{aligned} U_t(\varphi^t, y) = \inf_{\psi^t \in \Psi^t} \int_{E_Y} \cdots \int_{E_Y} & e^{\gamma \Sigma_{k=t}^T \beta^k r_k(x_k, \varphi_k(y_k))} \\ & Q(\mathrm{d}y_T | T-1, y_{T-1}, \varphi_{T-1}(y_{T-1}), \psi_{T-1}(y_{T-1})) \\ & \cdots Q(\mathrm{d}y_{t+1} | t, y, \varphi_t(y), \psi_t(y)). \end{aligned} \tag{17}$$

For a given policy $\varphi \in \mathscr{A}$, define the following functions on E_Y

$$\begin{aligned} V_T(y) &= e^{\gamma \beta^T r_T(x)}, \\ V_t(y) &= \inf_{\theta \in \tau(t,c)} \int_{E_Y} e^{\gamma \beta^t r_t(x, \varphi_t(y))} V_{t+1}(y') Q(\mathrm{d}y' | t, y, \varphi_t(y), \theta), \quad t \in \mathscr{T}'. \end{aligned}$$

We will prove recursively that the functions V_t are l.s.a. in y, and that

$$V_t(y) = U_t(\varphi^t, y), \quad t = 0, \ldots, T-1. \tag{18}$$

Clearly, V_T is l.s.a. in y.

Next, we will prove that $V_{T-1}(y)$ is l.s.a.. By our assumptions, the stochastic kernel $Q(\cdot | T-1, \cdot, \cdot, \cdot)$ is Borel measurable on E_Y given $E_Y \times A \times \Theta$, in the sense of [BS78, Definition 7.2]. Then, the integral $\int_{E_Y} V_T(y') Q(\mathrm{d}y' | T-1, y, a, \theta)$ is l.s.a. on $E_Y \times A \times \Theta$ according to [BS78, Proposition 7.48]. Now, we set (in the notations of [BS78])

$$\begin{aligned} \mathrm{X} &= E_Y \times A, \quad \mathrm{x} = (y, a) \\ \mathrm{Y} &= \Theta, \quad \mathrm{y} = \theta, \\ \mathrm{D} &= \bigcup_{(y,a) \in E_Y \times A} \{y, a\} \times \tau(T-1, c), \\ f(\mathrm{x}, \mathrm{y}) &= \int_{E_Y} V_T(y') \, Q(\mathrm{d}y' | T-1, y, a, \theta). \end{aligned}$$

Note that in view of our assumptions, X and Y are Borel spaces. The set D is closed (see the proof of Lemma 1) and thus analytic. Moreover, $\mathrm{D}_\mathrm{x} = \tau(T-1, c)$. Hence, by [BS78, Proposition 7.47], for each $a \in A$ the function

$$\inf_{\theta \in \tau(T-1,c)} \int_{E_Y} V_T(y') \, Q(\mathrm{d}y' | T-1, y, a, \theta)$$

is l.s.a. in y. Thus, it is l.s.a. in (y,a). Moreover, in view of [BS78, Proposition 7.50], for any $\varepsilon > 0$, there exists an analytically measurable function $\psi^{\varepsilon}_{T-1}(y,a)$ such that

$$\inf_{\theta\in\tau(T-1,c)}\int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\theta) = \int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\psi^{\varepsilon}_{T-1}(y,a)) + \varepsilon.$$

Therefore, for any fixed (y,a), we obtain a sequence $\{\psi^{1/n}_{T-1}(y,a), n\in\mathbb{N}\}$ such that

$$\lim_{n\to\infty}\int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\psi^{1/n}_{T-1}(y,a)) = \inf_{\theta\in\tau(T-1,c)}\int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\theta).$$

Due to the assumption that $\tau(T-1,c)$ is compact, there exists a convergent subsequence $\{\psi^{1/n_k}_{T-1}(y,a), k\in\mathbb{N}\}$ such that its limit $\psi^{*}_{T-1}(y,a)$ is universally measurable and satisfies

$$\int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\psi^{*}_{T-1}(y,a)) = \inf_{\theta\in\tau(T-1,c)}\int_{E_Y} V_T(y')\,Q(\mathrm{d}y'|T-1,y,a,\theta).$$

Clearly, the function $e^{\gamma\beta^{T-1}r_{T-1}(x,a)}$ is l.s.a. in (y,a). Thus, since $\varphi_{T-1}(y)$ is a Borel measurable function, using part (3) in [BS78, Lemma 7.30] we conclude that both $e^{\gamma\beta^{T-1}r_{T-1}(x,\varphi_{T-1}(y))}$ and $\inf_{\theta\in\tau(T-1,c)}\int_{E_Y} V_T(y')Q(\mathrm{d}y'|T-1,y,\varphi_{T-1}(y),\theta)$ are l.s.a. in y. Since both these functions are non-negative then, by part (4) in [BS78, Lemma 7.30], we conclude that V_{T-1} is l.s.a. in y. The proof that V_t is l.s.a. in y and ψ^*_t exists for $t=0,\dots,T-2$, follows analogously. We also obtain that

$$\int_{E_Y} V_t(y')\,Q(\mathrm{d}y'|t-1,y,a,\psi^{*}_{t-1}(y,a)) = \inf_{\theta\in\tau(t-1,c)}\int_{E_Y} V_t(y')\,Q(\mathrm{d}y'|t-1,y,a,\theta), \tag{19}$$

for any $t=1,\dots,T-1$.

It remains to verify (18). For $t=T-1$, by (17), we have

$$\begin{aligned} U_{T-1}(\varphi^{T-1},y) &= \inf_{\theta\in\tau(T-1,c)}\int_{E_Y} e^{\gamma\beta^{T-1}r_{T-1}(x,\varphi_{T-1}(y))}V_T(y') \\ &\qquad Q(\mathrm{d}y'|T-1,y,\varphi_{T-1}(y),\theta) \\ &= V_{T-1}(y). \end{aligned}$$

Therefore, $U_{T-1}(\varphi^{T-1},\cdot)$ is l.s.a.. Assume that for $t=1,\dots,T-1$, $U_t(\varphi^t,y) = V_t(y)$, and it is l.s.a.. Then, for any $y_{t-1}\in E_Y$, with the notation $\psi^{t-1} = (\psi_{t-1},\psi^t)$, we get

$$
\begin{aligned}
&U_{t-1}(\varphi^{t-1}, y_{t-1}) \\
&\quad = \inf_{(\psi_{t-1},\psi^t)\in\Psi^{t-1}} \int_{E_Y} \cdots \int_{E_Y} e^{\gamma \sum_{k=t-1}^{T-1} \beta^k r_k(x_k,\varphi_k(y_k)) + \gamma\beta^T r_T(x_T)} \\
&\qquad\qquad \prod_{k=t}^{T} Q(\mathrm{d}y_k | k-1, y_{k-1}, \varphi_{k-1}(y_{k-1}), \psi_{k-1}(y_{k-1})) \\
&\quad \geq \inf_{(\psi_{t-1},\psi^t)\in\Psi^{t-1}} \int_{E_Y} e^{\gamma\beta^{t-1} r_{t-1}(x_{t-1},\varphi_{t-1}(y_{t-1}))} V_t(y_t) \\
&\qquad\qquad Q(\mathrm{d}y_t | t-1, y_{t-1}, \varphi_{t-1}(y_{t-1}), \psi_{t-1}(y_{t-1})) \\
&\quad - \inf_{\theta\in\tau(t-1,c)} \int_{E_Y} e^{\gamma\beta^{t-1} r_{t-1}(x_{t-1},\varphi_{t-1}(y_{t-1}))} V_t(y_t) \\
&\qquad\qquad Q(\mathrm{d}y_t | t-1, y_{t-1}, \varphi_{t-1}(y_{t-1}), \psi_{t-1}(y_{t-1})) \\
&\quad = V_{t-1}(y_{t-1}).
\end{aligned}
$$

Next, fix $\varepsilon > 0$, and let $\psi^{t,\varepsilon}$ denote an ε-optimal selectors sequence starting at time t, namely

$$
\begin{aligned}
\int_{E_Y} \cdots \int_{E_Y} e^{\gamma \sum_{k=t}^{T} \beta^k r_k(x_k,\varphi_k(y_k))} \prod_{k=t+1}^{T} Q(\mathrm{d}y_k | k-1, y_{k-1}, \varphi_{k-1}(y_{k-1}), \psi^{t,\varepsilon}_{k-1}(y_{k-1})) \\
\leq U_t(\varphi^t, y_t) + \varepsilon.
\end{aligned}
$$

Consequently, for any $y_{t-1} \in E_Y$,

$$
\begin{aligned}
U_{t-1}(\varphi^{t-1}, y_{t-1}) &= \inf_{(\psi_{t-1},\psi^t)\in\Psi^{t-1}} \int_{E_Y} \cdots \int_{E_Y} e^{\gamma \sum_{k=t-1}^{T} \beta^k r_k(x_k,\varphi_k(y_k))} \\
&\qquad \prod_{k=t}^{T} Q(\mathrm{d}y_k | k-1, y_{k-1}, \varphi_{k-1}(y_{k-1}), \psi_{k-1}(y_{k-1})) \\
&\leq \inf_{\psi_{t-1}\in\tau(t-1,c)} \int_{E_Y} \cdots \int_{E_Y} e^{\gamma \sum_{k=t-1}^{T} \beta^k r_k(x_k,\varphi_k(y_k))} \\
&\quad \prod_{k=t+1}^{T} Q(\mathrm{d}y_k | k-1, y_{k-1}, \varphi_{k-1}(y_{k-1}), \psi^{t,\varepsilon}_{k-1}(y_{k-1})) \\
&\qquad \cdots Q(\mathrm{d}y_t | t-1, y_{t-1}, \varphi_{t-1}(y_{t-1}), \psi_{t-1}(y_{t-1})) \\
&\leq \inf_{\varphi_{t-1}\in\tau(t-1,c)} \int_{E_Y} U_t(\varphi^t, y_t) Q(\mathrm{d}y_t | t-1, y_{t-1}, \varphi_{t-1}(y_{t-1}), \psi_{t-1}(y_{t-1})) + \varepsilon \\
&= \inf_{\varphi_{t-1}\in\tau(t-1,c)} \int_{E_Y} V_t(y_t) Q(\mathrm{d}y_t | t-1, y_{t-1}, \varphi_{t-1}(y_{t-1}), \psi_{t-1}(y_{t-1})) + \varepsilon \\
&= V_{t-1}(y_{t-1}) + \varepsilon.
\end{aligned}
$$

Since ε is arbitrary, (18) is justified. In particular, $U_t(\varphi^t, \cdot)$ is l.s.a. for any $t \in \mathscr{T}'$. Finally, in view of (19), the equality (16) follows immediately. This concludes the proof. $\square$

Now we are in the position to prove the main result in this paper.

Theorem 1 *For $t = 0,\dots,T$, we have that*

$$U_t^* \equiv W_t. \tag{20}$$

Moreover, the policy φ^* derived in Lemma 1 is adaptive robust-optimal, that is

$$U_t^*(y) = U_t(\varphi^{t,*}, y), \quad t = 0,\dots,T-1. \tag{21}$$

Proof We proceed similarly as in the proof of [Iye05, Theorem 2.1], and via backward induction in $t = T, T-1,\dots,1,0$.

For $t = T$, clearly, $U_T^*(y) = W_T(y) = e^{\gamma\beta^T r_T(x)}$ for all $y \in E_Y$. For $t = T-1$ we have, for $y \in E_Y$,

$$\begin{aligned}
U_{T-1}^*(y) &= \sup_{\varphi^{T-1}=\varphi_{T-1}\in\mathscr{A}^{T-1}} \inf_{\theta\in\tau(T-1,c)} \int_{E_Y} e^{\gamma\beta^{T-1} r_{T-1}(x,\varphi_{T-1}(y))} W_T(y') \\
&\qquad Q(\mathrm{d}y' \mid T-1, y_{T-1}, \varphi_{T-1}(y), \theta) \\
&= \max_{a\in A} \inf_{\theta\in\tau(T-1,c)} \int_{E_Y} e^{\gamma\beta^{T-1} r_{T-1}(x,a)} W_T(y')\, Q(\mathrm{d}y' \mid T-1, y, a, \theta) \\
&= W_{T-1}(y).
\end{aligned}$$

From the above, using Lemma 1, we obtain that U_{T-1}^* is l.s.c. and bounded.

For $t = T-2,\dots,1,0$, assume that U_{t+1}^* is l.s.c. and bounded. Recalling the notation $\varphi^t = (\varphi_t, \varphi^{t+1})$, we thus have, $y \in E_Y$,

$$\begin{aligned}
U_t^*(y) &= \sup_{(\varphi_t,\varphi^{t+1})\in\mathscr{A}^t} \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(x,\varphi_t(y))} U_{t+1}(\varphi^{t+1}, y')\, Q(\mathrm{d}y' \mid t, y, \varphi_t(y), \theta) \\
&\le \sup_{(\varphi_t,\varphi^{t+1})\in\mathscr{A}^t} \inf_{\theta\in\tau(c_t,t)} \int_{E_Y} e^{\gamma\beta^t r_t(x,\varphi_t(y))} U_{t+1}^*(y')\, Q(\mathrm{d}y' \mid t, y, \varphi_t(y), \theta) \\
&= \max_{a\in A} \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(x,a)} U_{t+1}^*(y')\, Q(\mathrm{d}y \mid t, y_t, a, \theta) \\
&= \max_{a\in A} \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(y,a)} W_{t+1}(y')\, Q(\mathrm{d}y' \mid t, y, a, \theta) \\
&= W_t(y).
\end{aligned}$$

Now, fix $\varepsilon > 0$, and let $\varphi^{t+1,\varepsilon}$ denote an ε-optimal control strategy starting at time $t+1$, that is

$$U_{t+1}(\varphi^{t+1,\varepsilon}, y) \ge U_{t+1}^*(y) - \varepsilon, \quad y \in E_y.$$

Then, for $y \in E_Y$, we have

$$
\begin{aligned}
U_t^*(y) &= \sup_{(\varphi_t,\varphi^{t+1})\in\mathscr{A}^t}\ \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(x,\varphi_t(y))} U_{t+1}(\varphi^{t+1},y')\,Q(\mathrm{d}y' \mid t,y,\varphi_t(y),\theta) \\
&\geq \sup_{(\varphi_t,\varphi^{t+1})\in\mathscr{A}^t}\ \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(x,\varphi_t(y))} U_{t+1}(\varphi^{t+1,\varepsilon},y')\,Q(\mathrm{d}y' \mid t,y,\varphi_t(y),\theta) \\
&\geq \max_{a\in A}\ \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^t r_t(x,a)} U_{t+1}^*(y')\,Q(\mathrm{d}y' \mid t,y,a,\theta) - \varepsilon \\
&= \max_{a\in A}\ \inf_{\theta\in\tau(t,c)} \int_{E_Y} W_{t+1}(y')\,Q(\mathrm{d}y' \mid t,y,a,\theta) - \varepsilon \\
&= W_t(y) - \varepsilon.
\end{aligned}
$$

Since ε was arbitrary, the proof of (20) is done. In particular, we have that for any $t \in \mathscr{T}$, the function $U_t^*(\cdot)$ is l.s.c. as well as bounded.

It remains to justify the validity of equality (21). We will proceed again by (backward) induction in t. For $t = T-1$, using (20), we have that

$$
\begin{aligned}
U_{T-1}^*(y) = W_{T-1}(y) &= e^{\gamma\beta^{T-1} r_{T-1}(x,\varphi_{T-1}^*(y))} \\
&\qquad \inf_{\theta\in\tau(t,c)} \int_{E_Y} e^{\gamma\beta^T r_T(x')}\,Q(\mathrm{d}y' \mid T-1,y,\varphi_{T-1}^*(y),\theta) \\
&= e^{\gamma\beta^{T-1} r_{T-1}(x,\varphi_{T-1}^*(y))} \inf_{\mathbb{Q}\in\mathscr{Q}_{y,T-1}^{\varphi^{T-1,*},\psi^{T-1}}} \left(\mathbb{E}_{\mathbb{Q}} e^{\gamma\beta^T r_T(X_T)}\right) \\
&= U_{T-1}(\varphi^{T-1,*},y).
\end{aligned}
$$

Moreover, by Lemma 2, we get that

$$
U_{T-1}^*(y) = U_{T-1}(\varphi^{T-1,*},y) = \mathbb{E}_{\mathbb{Q}_{y,T-1}^{\varphi^{T-1,*},\psi^{T-1,*}}} e^{\gamma\beta^{T-1} r_{T-1}(x,\varphi_{T-1}^*(y)) + \gamma\beta^T r_T(X_T)}.
$$

For $t = T-2$, using again (20), Lemma 1, and Lemma 2, we have

$$
\begin{aligned}
U_{T-2}^*(y) = W_{T-2}(y) &= e^{\gamma\beta^{T-2} r_{T-2}(x,\varphi_{T-2}^*(y))} \\
&\quad \times \int_{E_Y} W_{T-1}(y')\,Q(\mathrm{d}y' \mid T-2,y,\varphi_{T-2}^*(y),\psi_{T-2}^*(y,\varphi_{T-2}^*(y))) \\
&= e^{\gamma\beta^{T-2} r_{T-2}(x,\varphi_{T-2}^*(y))} \\
&\quad \times \int_{E_Y} U_{T-1}(\varphi^{T-1,*},y')\,Q(\mathrm{d}y' \mid T-2,y,\varphi_{T-2}^*(y),\psi_{T-2}^*(y,\varphi_{T-2}^*(y))) \\
&= e^{\gamma\beta^{T-2} r_{T-2}(x,\varphi_{T-2}^*(y))} \\
&\quad \times \int_{E_Y} \left(\mathbb{E}_{\mathbb{Q}_{y',T-1}^{\varphi^{T-1,*},\psi^{T-1,*}}} e^{\gamma\beta^{T-1} r_{T-1}(x'),\varphi_{T-1}^*(y')) + \gamma\beta^T r_T(X_T)}\right) \\
&\qquad\qquad Q(\mathrm{d}y' \mid T-2,y,\varphi_{T-2}^*(y),\psi_{T-2}^*(y,\varphi_{T-2}^*(y))) \\
&= \mathbb{E}_{\mathbb{Q}_{y,T-2}^{\varphi^{T-2,*},\psi^{T-2,*}}} e^{\gamma\beta^{T-2} r_{T-2}(x,\varphi_{T-2}^*(y)) + \gamma\beta^{T-1} r_{T-1}(x'),\varphi_{T-1}^*(y')) + \gamma\beta^T r_T(X_T)}.
\end{aligned}
$$

Hence, we have that $U^*_{T-2}(y)$ is attained at $\varphi^{T-2,*}$, and therefore $U^*_{T-2}(y) = U_{T-2}(\varphi^{T-2,*},y)$. The rest of the proof of (21) proceeds in an analogous way. The proof is complete. □

4 Exponential Discounted Tamed Quadratic Criterion Example

In this section, we consider a linear quadratic control problem under model uncertainty as a numerical demonstration of the adaptive robust method. To this end, we consider the 2-dimensional controlled process

$$X_{t+1} = B_1 X_t + B_2 \varphi_t + Z_{t+1},$$

where B_1 and B_2 are two 2×2 matrices and Z_{t+1} is a 2-dimensional normal random variable with mean 0 and convariance matrix

$$\Sigma^* = \begin{pmatrix} \sigma_1^{*,2} & \sigma_{12}^{*,2} \\ \sigma_{12}^{*,2} & \sigma_2^{*,2} \end{pmatrix},$$

where $\sigma_1^{*,2}$, $\sigma_{12}^{*,2}$, and $\sigma_2^{*,2}$ are unknown. Given observations $Z_1,\ldots,Z_t$, we consider an unbiased estimator, say $\widehat{\Sigma}_t = \begin{pmatrix} \widehat{\sigma}^2_{1,t} & \widehat{\sigma}^2_{12,t} \\ \widehat{\sigma}^2_{12,t} & \widehat{\sigma}^2_{2,t} \end{pmatrix}$, of the covariance matrix Σ^*, given as

$$\widehat{\Sigma}_t = \frac{1}{t+1}\sum_{i=1}^{t} Z_i Z_i^\top,$$

which can be updated recursively as

$$\widehat{\Sigma}_t = \frac{t(t+1)\widehat{\Sigma}_{t-1} + t Z_t Z_t^\top}{(t+1)^2}.$$

With slight abuse of notations, we denote by Σ, Σ^*, and $\widehat{\Sigma}_t$ the column vectors

$$\begin{aligned} \Sigma^\top &= (\sigma_1^2, \sigma_{12}^2, \sigma_2^2) \\ \Sigma^{*,\top} &= (\sigma_1^{*,2}, \sigma_{12}^{*,2}, \sigma_2^{*,2}) \\ \widehat{\Sigma}_t^\top &= (\widehat{\sigma}^2_{1,t}, \widehat{\sigma}^2_{12,t}, \widehat{\sigma}^2_{2,t}). \end{aligned}$$

The corresponding parameter set is defined as

$$\Theta := \left\{ \Sigma^\top = (\Sigma_1, \Sigma_{12}, \Sigma_2) \in \mathbb{R}^3 : \ 0 \le \Sigma_1,\ \Sigma_2 \le \overline{\Sigma},\ \Sigma_{12}^2 \le \Sigma_1 \Sigma_2 \right\},$$

where $\overline{\Sigma}$ is some fixed positive constant. Note that the set Θ is a compact subset of $\mathbb{R}^3$.

Putting the above together and considering the augmented state process $Y_t = (X_t, \widehat{\Sigma}_t)$, $t \in \mathscr{T}$, and some finite control set $A \subset \mathbb{R}^2$, we get that the function S defined in (1) is given by

$$S(x,a,z) = B_1 x + B_2 a + z, \quad x,z \in \mathbb{R}^2,\ a \in A,$$

and the function $R(t,c,z)$ showing in (3) satisfies that

$$R(t,c,z) = (\bar{c}_1, \bar{c}_2, \bar{c}_3)^\top, \quad \begin{pmatrix} \bar{c}_1 & \bar{c}_3 \\ \bar{c}_3 & \bar{c}_2 \end{pmatrix} = \frac{(t+1)(t+2)\begin{pmatrix} c_1 & c_3 \\ c_3 & c_2 \end{pmatrix} + (t+1)zz^\top}{(t+2)^2},$$

where $z \in \mathbb{R}^2$, $t \in \mathscr{T}'$, $c = (c_1, c_2, c_3)$. Then, function $\mathbf{G}$ defined in (5) is specified accordingly.

It is well-known that $\sqrt{t+1}(\widehat{\Sigma}_t - \Sigma^*)$ converges weakly to 0-mean normal dsitribution with covariance matrix

$$M_\Sigma = \begin{pmatrix} 2\sigma_1^{*,4} & 2\sigma_1^{*,2}\sigma_{12}^{*,2} & 2\sigma_{12}^{*,4} \\ 2\sigma_1^{*,2}\sigma_{12}^{*,2} & \sigma_1^{*,2}\sigma_2^{*,2} + \sigma_{12}^{*,4} & 2\sigma_{12}^{*,2}\sigma_2^{*,2} \\ 2\sigma_{12}^{*,4} & 2\sigma_{12}^{*,2}\sigma_2^{*,2} & 2\sigma_2^{*,4} \end{pmatrix}.$$

We replace every entry in M_Σ with the corresponding estimator at time $t \in \mathscr{T}'$ and denote by $\widehat{M}_t(\widehat{\Sigma}_t)$ the resulting matrix. With probability one, the matrix $\widehat{M}_t(\widehat{\Sigma}_t)$ is positive-definite. Therefore, we get the confidence region for $\sigma_1^{*,2}$, $\sigma_{12}^{*,2}$, and $\sigma_2^{*,2}$ as

$$\tau(t,c) = \left\{ \Sigma \in \Theta : (t+1)(\Sigma - c)^\top \widehat{M}_t^{-1}(c)(\Sigma - c) \le \kappa \right\},$$

where κ is the $1-\alpha$ quantile of χ^2 distribution with 3 degrees of freedom for some confidence level $0 < \alpha < 1$.

We further take functions $r_T(x) = \min\{b_1, \max\{b_2, x^\top K_1 x\}\}$ and

$$r_t(x,a) = \min\{b_1, \max\{b_2, x^\top K_1 x + a^\top K_2 a\}\},$$

$t \in \mathscr{T}'$, where $x, a \in \mathbb{R}^2$, $b_1 > 0$, $b_2 < 0$, and K_1 and K_2 are two fixed 2-by-2 matrices with negative trace.

For this example, all conditions of the adaptive robust framework of Section 2 are easy to verify, except for the u.h.c. property of set-valued function $\tau(t,\cdot)$, which we establish in the following lemma.

Lemma 3 *For any $t \in \mathscr{T}'$, the set valued function $\tau(t,\cdot)$ is upper hemi-continuous.*

Proof Fix any $t \in \mathscr{T}'$ and $c_0 \in \Theta$. According to our earlier discussion, the matrix $\widehat{M}_t(c_0)$ is positive-definite. Hence, its inverse admits the Cholesky decomposition $\widehat{M}_t^{-1}(c_0) = L_t(c_0)L_t^\top(c_0)$. Consider the change of coordinate system via the linear transformation $\mathscr{L}c = L_t^\top(c_0)c$, and we name it system-$\mathscr{L}$. Let $E \subset \Theta$ be open and such that $\tau(t,c_0) \subset E$. Note that $\mathscr{L}\tau(t,c_0)$ is a closed ball centered at $\mathscr{L}c_0$ in the system-$\mathscr{L}$. Also, the mapping $\mathscr{L}$ is continuous and one-to-one, hence $\mathscr{L}E$ is an open set and $\mathscr{L}\tau(t,c_0) \subset \mathscr{L}E$. Then, we have that there exists an open ball $B_r(\mathscr{L}c_0)$ in the system-$\mathscr{L}$ centered at $\mathscr{L}c_0$ with radius r such that $\mathscr{L}\tau(t,c_0) \subset B_r(\mathscr{L}c_0) \subset \mathscr{L}E$.

Any ellipsoid centered at c' in the original coordinate system has representation $(c-c')^\top F(c-c') = 1$ which can be written as $(L_t^\top c - L_t^\top c')L^{-1}F(L^\top)^{-1}(L^\top c - L^\top c') = 1$. Hence, it is still an ellipsoid in the $\mathscr{L}$-system after transformation. To this end, we define on Θ a function $h(c) := \|\mathscr{L}c - \mathscr{L}c_0\| + \max\{r_i(c), i = 1,2,3\}$, where $\|\cdot\|$ is the Euclidean norm in the system-$\mathscr{L}$, and $r_i(c)$, $i = 1,2,3$, are the lengths of the three semi axes of the ellipsoid $\mathscr{L}\tau(t,c)$. It is clear that $r_i(c)$, $i = 1,2,3$ are continuous functions.

Next, it is straightforward to check that f is a non-constant continuous function. Therefore, we consider the set $D := \{c \in \Theta : h(c) < r\}$ and see that it is an open set in Θ and non-empty as $c_0 \in D$. Moreover, for any $c \in D$, we get that the ellipsoid $\mathscr{L}\tau(t,c) \subset B_r(\mathscr{L}c_0)$. Hence, $\tau(t,c) \subset E$, and we conclude that $\tau(t,\cdot)$ is u.h.c.. □

Thus, according to Theorem 1, the dynamic risk sensitive optimization problem under model uncertainty can be reduced to the Bellman equations given in (12):

$$W_T(y) = e^{\gamma\beta^T r_T(x)}, \tag{22}$$

$$W_t(y) = \sup_{a\in A} \inf_{\theta\in\tau(t,c)} \int_{\mathbb{R}^2} W_{t+1}(\mathbf{G}(t,y,a,z))e^{\gamma\beta^t(r_t(x,a))} f_Z(z;\theta)dz, \tag{23}$$

$$y = (x,c_1,c_2,c_3) \in E_Y,\ t = T-1,\dots,0,$$

where $f_Z(\cdot;\theta)$ is the density function for two dimensional normal random variable with mean 0 and covariance parameter θ. In the next section, using (22)-(23), we will compute numerically W_t by a machine learning based method. Note that the dimension of the state space E_Y is five in the present case, for which the traditional grid-based numerical method becomes extremely inefficient. Hence, we employ the new approach introduced in [CL19] to overcome the challenges met in our high dimensional robust stochastic control problem.

5 Machine Learning Algorithm and Numerical Results

In this section, we describe our machine learning based method and present the numerical results for our example. Similarly to [CL19], we discretize the state space the relevant state space in the spirit of the regression Monte Carlo method and adaptive design by creating a random (non-gridded) mesh for the process $Y = (X,C)$. Note that the component X depends on the control process, hence at each time t we randomly select from the set A a value of φ_t, and we randomly generate a value of Z_{t+1}, so to simulate the value of X_{t+1}. Next, for each t, we construct the convex hull of simulated Y_t and uniformly generate in-sample points from the convex hull to obtain a random mesh of Y_t. Then, we solve the equations (22)–(23), and compute the optimal trading strategies at all mesh points.

The key idea of our machine learning based method is to utilize a non-parametric value function approximation strategy called Gaussian process surrogate. For the purpose of solving the Bellman equations (22)–(23), we build GP regression model for the value function $W_{t+1}(\cdot)$ so that we can evaluate

$$\int_{\mathbb{R}^2} W_{t+1}(\mathbf{G}(t,y,a,z))e^{\gamma\alpha^t(r_t(x,a))}f_Z(z;\theta)dz.$$

We also construct GP regression model for the optimal control φ^*. It permits us to apply the optimal strategy to out-of-sample paths without actual optimization, which allows for a significant reduction of the computational cost.

As the GP surrogate for the value function W_t we consider a regression model $\widetilde{W}_t(y)$ such that for any $y^1, \ldots, y^N \in E_Y$, with $y^i \neq y^j$ for $i \neq j$, the random variables $\widetilde{W}_t(y^1), \ldots, \widetilde{W}_t(y^N)$ are jointly normally distributed. Then, given training data $(y^i, W_t(Y^i))$, $i = 1, \ldots, N$, for any $y \in E_Y$, the predicted value $\widetilde{W}_t(y)$, providing an estimate (approximation) of $W_t(y)$ is given by

$$\widetilde{W}(y) = \left(k(y,y^1),\ldots,k(y,y^N)\right)[\mathbf{K}+\varepsilon^2\mathbf{I}]^{-1}\left(W_t(y^1),\ldots,W_t(y^N)\right)^T,$$

where ε is a tuning parameter, $\mathbf{I}$ is the $N\times N$ identity matrix and the matrix $\mathbf{K}$ is defined as $\mathbf{K}_{i,j} = k(y^i,y^j)$, $i,\ j = 1, \ldots, N$. The function k is the kernel function for the GP model, and in this work we choose the kernel as the Matern-5/2. Fitting the GP surrogate $\widetilde{W}_t$ means to estimate the hyperparameters inside k through the training data $(y^i, W_t(y^i))$, $i = 1, \ldots, N$ for which we take $\varepsilon = 10^{-5}$. The GP surrogates for φ^* is obtained in an analogous way.

Given the mesh points $\{y_t^i,\ i = 1, \ldots, N_t,\ t \in \mathscr{T}'\}$, the overall algorithm proceeds as follows:

Part A: Time backward recursion for $t = T-1,\ldots,0$.

1. Assume that $W_{t+1}(y_{t+1}^i)$, and $\varphi_{t+1}^*(y_{t+1}^i) = (\varphi_{t+1}^{1,*}(y_{t+1}^i), \varphi_{t+1}^{2,*}(y_{t+1}^i))$, $i = 1,\ldots,N_t$, are numerically approximated as $\overline{W}_{t+1}(y_{t+1}^i)$, $\overline{\varphi}_{t+1}^{1,*}(y_{t+1}^i)$ and $\overline{\varphi}_{t+1}^{2,*}(y_{t+1}^i)$, $i = 1,\ldots,N_t$, respectively. Also suppose that the corresponding GP surrogates $\widetilde{W}_{t+1}$, $\widetilde{\varphi}_{t+1}^{1,*}$, and $\widetilde{\varphi}_{t+1}^{2,*}$ are fitted through training data $(y_{t+1}^i, \overline{W}_{t+1}(y_{t+1}^i))$, $(y_{t+1}^i, \overline{\varphi}_{t+1}^{1,*}(y_{t+1}^i))$, and $(y_{t+1}^i, \overline{\varphi}_{t+1}^{2,*}(y_{t+1}^i))$, $i = 1,\ldots,N_t$, respectively.
2. For time t, any $a \in A$, $\theta \in \tau(t,c)$ and each y_t^i, $i = 1, \ldots, N_t$, use one-step Monte Carlo simulation to estimate the integral

$$w_t(y,a,\theta) = \int_{\mathbb{R}^2} W_{t+1}(\mathbf{G}(t,y,a,z))e^{\gamma\alpha^t(r_t(x,a))}f_Z(z;\theta)dz.$$

For that, if $Z_{t+1}^1, \ldots, Z_{t+1}^M$ is a sample of Z_{t+1} drawn from the normal distribution corresponding to parameter θ, where $M > 0$ is a positive integer, then estimate the above integral as

$$\widetilde{w}_t(y,a,\theta) = \frac{1}{M}\sum_{i=1}^{M} \widetilde{W}_{t+1}(\mathbf{G}(t,y,a,Z^i_{t+1}))e^{\gamma\alpha^t(r_t(x,a))}.$$

3. For each y^i_t, $i=1, \ldots, N_t$, and any $a \in A$, compute

$$\overline{w}_t(y^i_t,a) = \inf_{\theta\in\tau(t,c)} \widetilde{w}_t(y^i_t,a,\theta).$$

4. Compute

$$\overline{W}_t(y^i_t) = \max_{a\in A} \overline{w}_t(y^i_t,a),$$

and obtain a maximizer $\overline{\varphi}^*_t(y^i_t) = (\overline{\varphi}^{1,*}_t(y^i_t), \overline{\varphi}^{2,*}_t(y^i_t))$, $i=1,\ldots,N_t$.
5. Fit a GP regression model for $V_t(\cdot)$ using the results from Step 4 above. Fit GP models for $\varphi^{1,*}_t(\cdot)$ and $\varphi^{2,*}_t(\cdot)$ as well; these are needed for obtaining values of the optimal strategies for out-of-sample paths in Part B of the algorithm.
6. Goto 1: Start the next recursion for $t-1$.

Part B: Forward simulation to evaluate the performance of the GP surrogates $\varphi^{1,*}_t(\cdot)$ and $\varphi^{2,*}_t(\cdot)$, $t=0,\ldots,T-1$, over the out-of-sample paths.

1. Draw $K>0$ samples of i.i.d. $Z^{*,i}_1,\ldots,Z^{*,i}_T$, $i=1,\ldots,K$, from the normal distribution corresponding to the assumed true parameter θ^*.
2. All paths will start from the initial state y_0. The state along each path i is updated according to $\mathbf{G}(t,y^i_t,\widetilde{\varphi}^*_t(y^i_t),Z^{*,i}_{t+1})$, where $\widetilde{\varphi}^*_t = (\widetilde{\varphi}^{1,*}_t, \widetilde{\varphi}^{2,*}_t)$ is the GP surrogate fitted in Part A. Also, compute the running reward $r_t(x^i_t, \widetilde{\varphi}^*_t(y^i_t))$.
3. Obtain the terminal reward $r_T(x^i_T)$, generated by $\widetilde{\varphi}^*$ along the path corresponding to the sample of $Z^{*,i}_1, \ldots, Z^{*,i}_T$, $i=1, \ldots, K$, and compute

$$W^{\mathrm{ar}} := \frac{1}{\gamma}\ln\left(\frac{1}{K}\sum_{i=1}^{K} e^{\gamma(\sum_{t=0}^{T-1}\beta^t r_t(x^i_t,\widetilde{\varphi}^*_t(y^i_t))+\beta^T r_T(x^i_T))}\right) \tag{24}$$

as an estimate of the performance of the optimal adaptive robust risk sensitive strategy φ^*.

For comparison, we also analyze the optimial risk sensitive strategies of the adaptive and strong robust control methods. In (23), if we take $\tau(t,c)=\{c\}$ for any t, then we obtain the adaptive risk sensitive strategy. On the other hand, by taking $\tau(t,c)=\Theta$ for any t and c, we get the strong robust strategy. We will compute W^{ad} and W^{sr} the risk sensitive criteria of adaptive and strong robust, respectively, in analogy to (24).

Next, we apply the machine learning algorithm described above by solving (22)–(23) for a specific set of parameters. In particular, we take: $T=10$ with one period of time corresponding to one-tenth of a year; the discount factor being equal to 0.3 or equivalently $\beta=0.3$; the initial state $X_0^\top=(2,2)$; the confidence level $\alpha=0.1$; in Part A of our algorithm the number of one-step Monte Carlo simulations is $M=100$; the number of forward simluations in Part B is taken $K=2000$; the control set A is approximated by the compact set $[-1,1]^2$; the relevant matrices are

$$B_1 = B_2 = \begin{pmatrix} 0.5 & -0.1 \\ -0.1 & 0.5 \end{pmatrix}, \quad K_1 = \begin{pmatrix} 0.7 & -0.2 \\ -0.2 & 0.7 \end{pmatrix}, \quad K_2 = \begin{pmatrix} -200 & 100 \\ 100 & -200 \end{pmatrix}.$$

The assumed true covariance matrix for Z_t, $t \in \mathscr{T}$, as well as initital guess are

$$\Sigma^* = \begin{pmatrix} 0.009 & 0.006 \\ 0.006 & 0.016 \end{pmatrix}, \quad \widehat{\Sigma}_0 = \begin{pmatrix} 0.00625 & 0.004 \\ 0.004 & 0.02025 \end{pmatrix},$$

respectively. The parameter set is chosen as $\Theta = \tau(0, c_0)$, where $c_0^\top = (0.00625, 0.004, 0.02025)$. For all three control approaches, we compute W^{ar}, W^{ad}, and W^{sr}, respectively, for the risk sensitive parameters $\gamma = 0.2$ and $\gamma = 1.5$.

Finally, we report on the computed values of the optimality criterion corresponding to three different methods: adaptive robust (AR), adaptive (AD) and strong robust (SR).

	W^{ar}	W^{ad}	W^{sr}
$\gamma = 0.2$	-319.81	-323.19	-329.53
$\gamma = 1.5$	-427.76	-427.97	-442.97

Table 1 Risk sensitive criteria for AR, AD, and SR.

Acknowledgements

Tomasz R. Bielecki and Igor Cialenco acknowledge support from the National Science Foundation grant DMS-1907568.

References

[BCC17] T. R. Bielecki, I. Cialenco, and T. Chen. Recursive construction of confidence regions. *Electron. J. Statist.*, 11(2):4674–4700, 2017.

[BCC+19] T. R. Bielecki, I. Cialenco, T. Chen, A. Cousin, and M. Jeanblanc. Adaptive Robust Hedging Under Model Uncertainty. *SIAM J. Control Optim.*, 57(2):925–946, 2019.

[Bor85] K. Border. *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, 9 edition, 1985.

[BP03] T. R. Bielecki and S.R. Pliska. Economic properties of the risk sensitive criterion for portfolio management. *Review of Accounting and Finance*, 2:3–17, 2003.

[BR11] N. Bäuerle and U. Rieder. *Markov decision processes with applications to finance*. Universitext. Springer, Heidelberg, 2011.

[BR17] N. Bäuerle and U. Rieder. Zero-sum risk-sensitive stochastic games. *Stochastic Processes and their Applications*, 127(2):622 – 642, 2017.

[BS78] D. P. Bertsekas and S. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press, 1978.

[CG91] H. F. Chen and L. Guo. *Identification and stochastic adaptive control*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.

[CL19] T. Chen and M. Ludkovski. A machine learning approach to adaptive robust utility maximization and hedging. *Preprint, arXiv:1912.00244*, 2019.

[DL14] M. Davis and S. Lleo. *Risk-Sensitive Investment Management*, volume 19 of *Advanced Series on Statistical Science & Applied Probability*. World Sci., 2014.

[GS89] I. Gilboa and D. Schmeidler. Maxmin expected utility with nonunique prior. *J. Math. Econom.*, 18(2):141–153, 1989.

[HS95] L. P. Hansen and T. J. Sargent. Discounted linear exponential quadratic Gaussian control. *IEEE Transactions on Automatic Control*, 40(5):968–971, 1995.

[HS08] P. L. Hansen and T. J. Sargent. *Robustness*. Princeton University Press, 2008.

[HSTW06] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. *J. Econom. Theory*, 128(1):45–90, 2006.

[Iye05] G. N. Iyengar. Robust Dynamic Programming. *Mathematics of Operations Research*, 30(2):257–280, 2005.

[KV15] P. R. Kumar and P. Varaiya. *Stochastic systems: estimation, identification and adaptive control*, volume 75 of *Classics in applied mathematics*. SIAM, 2015.

Optimal Control of Piecewise Deterministic Markov Processes

O.L.V. Costa and F. Dufour

Abstract This chapter studies the infinite-horizon continuous-time optimal control problem of piecewise deterministic Markov processes (PDMPs) with the control acting continuously on the jump intensity λ and on the transition measure Q of the process. Two optimality criteria are considered, the discounted cost case and the long run average cost case. We provide conditions for the existence of a solution to an integro-differential optimality equality, the so called Hamilton-Jacobi-Bellman (HJB) equation, for the discounted cost case, and a solution to an HJB inequality for the long run average cost case, as well as conditions for the existence of a deterministic stationary optimal policy. From the results for the discounted cost case and under some continuity and compactness hypothesis on the parameters and non-explosive assumptions for the process, we derive the conditions for the long run average cost case by employing the so-called vanishing discount approach.

1 Introduction

Piecewise Deterministic Markov Processes (PDMPs) were introduced by M.H.A. Davis in the seminal paper [9] as a general family of nondiffusion stochastic models, suitable to formulate an enormous variety of applications in operations research,

O.L.V. Costa
Departamento de Engenharia de Telecomunicações e Controle
Escola Politécnica da Universidade de São Paulo
CEP: 05508 010-São Paulo, Brazil.
e-mail: oswaldo@lac.usp.br

F. Dufour
Institut Polytechnique de Bordeaux
INRIA Bordeaux Sud Ouest, Team: ASTRAL
IMB, Institut de Mathématiques de Bordeaux, Université de Bordeaux, France
e-mail: francois.dufour@math.u-bordeaux.fr

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_3

engineering systems and management science. The general theory of the PDMPs, including a full characterization of the extended generator as well as its applications in several stochastic control problems, were elegantly and comprehensively presented in the book [11]. PDMPs are characterized by three local parameters: the flow ϕ, the jump rate λ, and the transition measure Q. Roughly speaking, the motion of a PDMP starting at the initial state x_0 follows a deterministic flow $\phi(x_0,t)$ until the first jump time T_1, which occurs either spontaneously in a Poisson-like fashion with rate λ or when the flow $\phi(x_0,t)$ hits the boundary of the state space. In either case the post-jump location of the process is selected by the transition measure $Q(.|\phi(x,T_1))$ and the motion restarts from this new point afresh. As presented in [11], a suitable choice of the state space and the local characteristics ϕ, λ, and Q can cover a great deal of problems in operations research, engineering systems and management science. It is worth pointing out that the presence of the boundary is crucial for the modeling of some optimization problems as, for instance, in queueing and inventory systems or maintenance-replacement models (see, for instance, the capacity expansion problem in [9], item (21.13), in which the boundary represents that a project is completed, and the jump in this case represents that investment is channelled immediately into the next project).

Broadly speaking there are two types of control for PDMPs, as pointed out by Davis in [11, page 134]: *continuous control*, in which the control variable acts at all times on the process through the characteristics (ϕ,λ,Q), and *impulse control*, used to describe control actions that intervene in the process by moving it to a new point of the state space at some specific times. The focus of this chapter will be on the former case, but considering that the control acts only on (λ,Q). Two performance criteria will be considered along this chapter: the so-called infinite horizon discounted cost case and the long run average cost case. Other criteria that can be found in the literature for the PDMPs include, for instance, the risk-sensitive control problem, as analyzed in [20] and [22].

It is worth pointing out that the main difficulty in considering the control acting also on the flow ϕ relies on the fact that in this situation the time which the flow takes to hit the boundary as well as the first order differential operator associated to the flow ϕ would depend on the control. For the discounted cost criterion this problem was nicely studied in [10] by rewriting the integral cost as a sum of integrals between two consecutive jump times of the PDMP, which yields to the one step cost function for a discrete-time Markov decision model. However this decomposition for the long run average cost is not possible. When compared with the so-called continuous-time Markov decision processes (see, for instance, [18, 16, 17, 19, 26, 33, 34]), it should be highlighted that the PDMPs are characterized by a drift motion between jumps, and forced jumps whenever the process hits the boundary, so that the available results for the continuous-time Markov decision processes cannot be applied to the PDMPs case.

Two kinds of approach can be pointed out for dealing with the discounted and long run average control problems of PDMPs. The first one would be to characterize the value function as a solution to the so called Hamilton-Jacobi-Bellman (HJB) equation associated with an imbedded discrete-stage Markov decision model, with the stages defined by the jump times T_n of the process. As a sample of works along this direction we can refer to [2, 3, 5, 8, 10, 11, 15, 30, 31] and the references therein. The key idea behind this approach is to find, at each stage, a control function that solves an imbedded deterministic optimal control problem. Usually the control strategy is chosen among the set of piecewise open loop policies, that is, stochastic kernels or measurable functions that depend only on the last jump time and post jump location. The second approach for these problems, which we will call the infinitesimal approach, is to characterize the optimal value function as the viscosity solution of the corresponding integro-differential HJB equation. As a sample of works using this kind of approach we can mention [7, 11, 12, 13, 14, 32] and the references therein.

This chapter adopts the infinitesimal approach to study the discounted and long run average control problems of PDMPs. The results presented in this chapter were mainly drawn from [7] and [6]. The goal is to provide conditions for the existence of a solution to integro-differential HJB equality and inequality, and for the existence of a deterministic stationary optimal policy, associated to the discounted and long run average control problems. These conditions are essentially related to continuity and compactness assumptions on the parameters of the problem, as well as some non-explosive conditions for the controlled process. In order to derive the results for the long run average control problem we apply the so-called vanishing discounted approach by adapting and combining arguments used in the context of continuous-time Markov decision processes (see [33]), and the results obtained for the infinite-horizon discounted optimal control problem.

The chapter is organized as follows. In sections 2 and 3 we present the notation, some definitions, the parameters defining the model, the construction of the controlled process, the definition of the admissible strategies, and the problem formulation. In section 4 we give the main assumptions and some auxiliary results. In sections 5 and 6 we present the main results related to the discounted and long run average control problems (see Theorems 2, 3 and 4) that provide sufficient conditions for the existence of a solution to a HJB equality (for the discounted case) and inequality (for the long run average case) and for the existence of a deterministic stationary optimal policy. Some proofs of the auxiliary results are presented in the Appendix.

2 Notation and definition

In this section we present the notation and some definitions that will be used throughout the chapter as well as the definition of the generalized inferior limit and its properties. The generalized limit will be used for the results related to the vanishing discounted approach to be considered in section 6.

We will denote by $\mathbb{N}$ the set of natural numbers including 0, $\mathbb{N}^* = \mathbb{N} - \{0\}$, $\mathbb{R}$ the set of real numbers, $\mathbb{R}_+$ the set of non-negative real numbers, $\mathbb{R}_+^* = \mathbb{R}_+ - \{0\}$, $\widehat{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. By *measure* we will always refer to a countably additive, $\mathbb{R}_+$-valued set function. For X a Borel space (i.e. a Borel-measurable subset of a complete and separable metric space) we denote by $\mathcal{B}(X)$ its associated Borel σ-algebra, and by $\mathcal{M}(X)$ ($\mathcal{P}(X)$ respectively) the set of measures (probability measures) defined on $(X, \mathcal{B}(X))$, endowed with the weak topology. We represent by $\mathcal{P}(X|Y)$ the set of stochastic kernels on X given Y where Y denotes a Borel space. For any set A, I_A denotes the indicator function of the set A, and for any point $x \in X$, δ_x denotes the Dirac measure defined by $\delta_x(\Gamma) = I_\Gamma(x)$ for any $\Gamma \in \mathcal{B}(X)$.

The space of Borel-measurable (bounded, lower semicontinuous respectively) real-valued functions defined on the Borel space X will be denoted by $\mathbb{M}(X)$ ($\mathbb{B}(X)$, $\mathbb{L}(X)$ respectively) and we set $\mathbb{L}_b(X) = \mathbb{L}(X) \cap \mathbb{B}(X)$. Moreover, the space of Borel-measurable, lower semicontinuous, $\widehat{\mathbb{R}}$-valued functions defined on the Borel space X will be denoted by $\widehat{\mathbb{L}}(X)$. For all the previous space of functions the subscript $_+$ will indicate the case of non-negative functions. The infimum over an empty set is understood to be equal to $+\infty$, and $e^{-\infty} = 0$.

As in [29], the definition of the generalized inferior limit is as follows:

Definition 1 Let X be a Borel space and let $\{w_n\}$, be a family of functions in $\mathbb{M}(X)$. The generalized inferior limit of the sequence $\{w_n\}$, denoted by $\underline{\lim}^g_{n\to\infty} w_n$ is defined as

$$\underline{\lim}^g_{n\to\infty} w_n(x) = \sup_{k\geq 1} \sup_{\epsilon>0} \Big(\inf_{m\geq k} \inf_{\{y:d(y,x)<\epsilon\}} w_m(y) \Big) \tag{1}$$

where $d(.,.)$ is the metric in X. For notational convenience, $\underline{\lim}^g_{n\to\infty} w_n$ will be denoted by w_*.

The following properties from the generalized inferior limit will be used in section 6 for the vanishing discounted approach.

Proposition 1 *Let $\{w_n\}$ be a sequence of nonnegative functions in $\mathbb{M}(X)$ and consider an arbitrary $x \in X$. In this case, $w_*(x)$ as defined in (1) satisfies the following properties:*

(i) For any sequence $\{x_n\}$ such that $x_n \to x$, it follows that $\underline{\lim}_{n\to\infty} w_n(x_n) \geq w_*(x)$, and there exists a sequence $\{x_n\}$ such that $x_n \to x$ and $\underline{\lim}_{n\to\infty} w_n(x_n) = w_*(x)$.

(ii) $w_* \in \mathbb{L}_+(X)$.

(iii) [Generalized Fatou's Lemma] Suppose that $\{\mu_n\}$ is a sequence of probability measures in $\mathcal{P}(X)$ and that $\{\mu_n\}$ converges weakly to a $\mu \in \mathcal{P}(X)$. Then

$$\underline{\lim}_{n\to\infty} \int_S w_n(x)\mu_n(dx) \geq \int_S w_*(x)\mu(dx). \tag{2}$$

Proof: For the proof of (i) see Lemma 4.1 in [4]. For (ii) see Lemma 3.1 in [25] and for (iii) see Lemma 3.2 in [25]. □

3 Problem formulation for the controlled PDMP

The goal of this section is to introduce the parameters defining the model, the construction of the controlled process, the definition of the admissible strategies, and the problem formulation. Since it follows closely sections 2 and 3 in [7] some details will be skipped.

3.1 Parameters of the model

We will consider the control model depending on the following elements:

- The state space $\mathbf{X}$, which we assume to be an open subset of $\mathbb{R}^d$ ($d \in \mathbb{N}^*$) with boundary represented by $\partial\mathbf{X}$.
- The flow $\phi(x,t) : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, associated with a given Lipschitz continuous vector field in $\mathbb{R}^d$, that is, $\phi(x,0) = x$ and $\phi(x,t+s) = \phi(\phi(x,s),t)$ for all $x \in \mathbb{R}^d$ and $(t,s) \in \mathbb{R}^2$.
- The so called active boundary defined as $\Xi = \{x \in \partial\mathbf{X} : x = \phi(y,t)$ for some $y \in \mathbf{X}$ and $t \in \mathbb{R}_+^*\}$. With some abuse of notation, we set $\overline{\mathbf{X}}$ as $\mathbf{X} \cup \Xi$, and for $x \in \overline{\mathbf{X}}$, we define
$$t^*(x) = \inf\{t \in \mathbb{R}_+ :\ \phi(x,t) \in \Xi\}.$$
The flow ϕ outside the space $\overline{\mathbf{X}}$ can be defined arbitrarily since it plays no role for the problem.

- The action space $\mathbf{A}$, assumed to be a Borel space, and the set of feasible actions in state $x \in \overline{\mathbf{X}}$, given by $\mathbf{A}(x)$, which is a nonempty measurable subset of $\mathbf{A}$. Define the set $\mathbf{K} = \mathbf{K}^i \cup \mathbf{K}^g$ with
$$\mathbf{K}^g = \{(x,a) \in \mathbf{X} \times \mathbf{A} : a \in \mathbf{A}(x)\} \in \mathcal{B}(\mathbf{X} \times \mathbf{A}),$$
$$\mathbf{K}^i = \{(x,a) \in \Xi \times \mathbf{A} : a \in \mathbf{A}(x)\} \in \mathcal{B}(\Xi \times \mathbf{A}).$$
It is assumed that $\mathbf{K}^g$ (respectively, $\mathbf{K}^i$) contains the graph of a measurable function from $\mathbf{X}$ (respectively, Ξ) to $\mathbf{A}$.
- The controlled jumps intensity λ which is a $\mathbb{R}_+$-valued measurable function defined on $\mathbf{K}$.
- The stochastic kernel Q on $\mathbf{X}$ given $\mathbf{K}$ satisfying $Q(\mathbf{X} \setminus \{x\}|x,a) = 1$ for any $(x,a) \in \mathbf{K}$. It describes the state of the process after any jump. In other words, if a jump governed by the intensity λ occurs in the current state $x \in \mathbf{X}$ and with action $a \in \mathbf{A}(x)$, then $Q(\cdot|x,a)$ describes the distribution of the state immediately after the jump. If $z \in \Xi$, that is, the current state is at the boundary then an action $b \in \mathbf{A}(z)$ is applied and the state of the process changes instantly according to the stochastic kernel Q.

It should be noticed that in the framework of continuous-time MDPs, the signed kernel on $\mathbf{X}$ given $\mathbf{K}$, defined by

$$q(dy|x,a) = \lambda(x,a)[Q(dy|x,a) - \delta_x(dy)] \tag{3}$$

is the (controlled) infinitesimal generator of the jump process. For $V \in \mathbb{M}(\mathbf{X})$ we set,

$$QV(x,a) = \int_{\mathbf{X}} V(y)Q(dy|x,a),\ \ (x,a) \in \mathbf{K}, \tag{4}$$
$$\lambda QV(x,a) = \lambda(x,a)QV(x,a),\ (x,a) \in \mathbf{K}^i,$$

provided that the integral in (4) exists. From (3) we have that

$$qV(x,a) = \lambda(x,a)[QV(x,a) - V(x)],\ (x,a) \in \mathbf{K}^i. \tag{5}$$

We conclude this sub-section with the following definition that will be used in the sequel.

Definition 2 The set of functions $g \in \mathbb{M}(\mathbf{X})$ which are absolutely continuous with respect to the flow ϕ on $[0,t^*(x)[$ (that is, the function $g(\phi(x,\cdot))$ is absolutely continuous on $[0,t^*(x)] \cap \mathbb{R}_+$) and such that $\lim_{t\to t^*(x)} g(\phi(x,t))$ exists whenever $t^*(x) < \infty$ will be denoted by $\mathbb{A}(\overline{\mathbf{X}})$. In this case the domain of definition of the mapping g can be extended to $\overline{\mathbf{X}}$ by setting $g(z) = \lim_{t\to t^*(x)} g(\phi(x,t))$ where $z = \phi(x,t^*(x))) \in \mathbf{\Xi}$. Lemma 2.2 in [8] shows that, for $g \in \mathbb{A}(\overline{\mathbf{X}})$, there exists a real-valued measurable function $\mathcal{X}g$ defined on $\mathbf{X}$ satisfying

$$g(\phi(x,t)) = g(x) + \int_{[0,t]} \mathcal{X}g(\phi(x,s))ds, \tag{6}$$

for any $t \in [0,t^*(x)[$. Notice that for $g \in \mathbb{A}(\overline{\mathbf{X}})$ the function $\mathcal{X}g$ satisfying (6) is not necessarily unique. The case of bounded functions in $\mathbb{A}(\overline{\mathbf{X}})$ will be denoted, as before, by $\mathbb{A}_b(\overline{\mathbf{X}})$.

3.2 Construction of the controlled process ξ_t

The canonical space Ω is defined by $\Omega = \bigcup_{n=0}^{\infty} \Omega_n \bigcup (\mathbf{X} \times (\mathbb{R}_+^* \times \mathbf{X})^{\infty})$ where $\Omega_n = \mathbf{X} \times (\mathbb{R}_+^* \times \mathbf{X})^n \times (\{\infty\} \times \{x^{\infty}\})^{\infty}$ and x^{∞} is an isolated artificial point corresponding to the case when no jumps occur in the future, endowed with its Borel σ-algebra denoted by $\mathcal{F}$. In that case, the process stays forever in x^{∞}, and so $t^*(x^{\infty}) = +\infty$. Set $\mathbf{X}_{\infty} = \mathbf{X} \cup \{x^{\infty}\}$ and $\overline{\mathbf{X}}_{\infty} = \overline{\mathbf{X}} \cup \{x^{\infty}\}$. We also extend the definition of ϕ on $\mathbf{X}_{\infty} \times \widehat{\mathbb{R}}_+$ as $\phi(x^{\infty},t) = x^{\infty}$ for any $t \in \widehat{\mathbb{R}}_+$ and also $\phi(x,t^*(x)) = x^{\infty}$ whenever $t^*(x) = \infty$ for $x \in \mathbf{X}$.

We set $\omega \in \Omega$ as

$$\omega = (x_0, \theta_1, x_1, \theta_2, x_2, \dots),$$

where $x_0 \in \mathbf{X}$ represents the initial state of the controlled point process ξ, and for $n \in \mathbb{N}^*$, the components $\theta_n > 0$ and x_n correspond to the time interval between two consecutive jumps and the value of the process ξ immediately after the jump. For the case $\theta_n < \infty$ and $\theta_{n+1} = \infty$, the trajectory of the controlled point process has only n jumps, and we put $\theta_m = \infty$ and $x_m = x^\infty$ (artificial point) for all $m \geq n+1$. Between jumps, the state of the process ξ moves according to the flow ϕ. The path up to $n \in \mathbb{N}$ is denoted by $h_n = (x_0, \theta_1, x_1, \theta_2, x_2, \ldots \theta_n, x_n)$, and the collection of all such paths is denoted by $\mathbf{H}_n$. We denote by $H_n = (X_0, \Theta_1, X_1, \ldots, \Theta_n, X_n)$ the n-term random history process taking values in $\mathbf{H}_n$ for $n \in \mathbb{N}$.

For $n \in \mathbb{N}$, set the mappings $X_n : \ \Omega \to \mathbf{X}_\infty$ by $X_n(\omega) = x_n$ and, for $n \geq 1$, the mappings $\Theta_n : \ \Omega \to \overline{\mathbb{R}}_+^*$ by $\Theta_n(\omega) = \theta_n$; $\Theta_0(\omega) = 0$. The sequence $(T_n)_{n\in\mathbb{N}^*}$ of $\overline{\mathbb{R}}_+^*$-valued mappings is defined on Ω by $T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega) = \sum_{i=1}^n \theta_i$ and $T_\infty(\omega) = \lim_{n\to\infty} T_n(\omega)$. The random measure μ associated with $(\Theta_n, X_n)_{n\in\mathbb{N}}$ is a measure defined on $\mathbb{R}_+^* \times \mathbf{X}$ by

$$\mu(\omega; dt, dx) = \sum_{n\geq 1} I_{\{T_n(\omega)<\infty\}} \delta_{(T_n(\omega), X_n(\omega))}(dt, dx).$$

The dependence on ω will be suppressed for notational convenience and it will be written $\mu(dt, dx)$ instead of $\mu(\omega; dt, dx)$. For $t \in \mathbb{R}_+$, define $\mathcal{F}_t = \sigma\{H_0\} \vee \sigma\{\mu(]0, s] \times B) : \ s \leq t, B \in \mathcal{B}(\mathbf{X})\}$. The controlled process $\{\xi_t\}_{t\in\mathbb{R}_+}$ is defined as:

$$\xi_t(\omega) = \begin{cases} \phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ x^\infty, & \text{if } T_\infty \leq t, \end{cases}$$

and it is easy to see that $(\xi_t)_{t\in\mathbb{R}_+}$ could be equivalently described by the sequence $(\Theta_n, X_n)_{n\in\mathbb{N}}$. As in [11], we set

$$p^*(dt) = I_{\{\xi_{t-}\in\Xi\}} \mu(dt, \mathbf{X})$$

which counts the number of jumps from the boundary of the controlled process ξ_t (see [11], sub-section 26).

3.3 Admissible strategies

Associated to the state x^∞ we consider a special action a^∞ and we set $\mathbf{A}_\infty = \mathbf{A} \cup \{a^\infty\}$; $\mathbf{A}_\infty(x^\infty) = \{a^\infty\}$ and $\mathbf{A}_\infty(x) = \mathbf{A}(x)$ for $x \in \overline{\mathbf{X}}$. We also extend the definition of λ and Q at the point (x^∞, a^∞) by defining $\lambda(x^\infty, a^\infty) = 0$ and $Q(\{x^\infty\}|x^\infty, a^\infty) = 1$. An admissible control strategy is a sequence $u = (\pi_n, \gamma_n)_{n\in\mathbb{N}}$ such that, for any $n \in \mathbb{N}$,

- $\pi_n \in \mathcal{P}(\mathbf{A}_\infty|\mathbf{H}_n \times \mathbb{R}_+^*)$ and satisfies $\pi_n(\mathbf{A}(\phi(x_n, t))|h_n, t) = 1$ for $h_n = (x_0, \ldots, \theta_n, x_n) \in \mathbf{H}_n$ and $t \in]0, t^*(x_n)[$.
- $\gamma_n \in \mathcal{P}(\mathbf{A}_\infty|\mathbf{H}_n)$ and satisfies $\gamma_n(\mathbf{A}(\phi(x_n, t^*(x_n)))|h_n) = 1$ for $h_n = (x_0, \ldots, \theta_n, x_n) \in \mathbf{H}_n$ and $t^*(x_n) < \infty$.

We will denote by $\mathcal{U}$ the set of admissible control strategies, and for $u = (\pi_n, \gamma_n)_{n\in\mathbb{N}} \in \mathcal{U}$ we denote by π and γ the random processes with values in $\mathcal{P}(\mathbf{A}_\infty)$

correspondingly as

$$\pi(da|t) = \sum_{n\in\mathbb{N}} I_{\{T_n<t\leq T_{n+1}\}}\pi_n(da|H_n, t-T_n)$$

and

$$\gamma(da|t) = \sum_{n\in\mathbb{N}} I_{\{T_n<t\leq T_{n+1}\}}\gamma_n(da|H_n),$$

for $t \in \mathbb{R}_+^*$. The processes π and γ are $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$-predictable random processes with values in $\mathcal{P}(\mathbf{A}_\infty)$. The following class of admissible strategies will be considered along this chapter. A control strategy $u \in \mathcal{U}$ is called *deterministic stationary*, if $\pi_n(\cdot|h_n,t) = \delta_{\varphi^s(\phi(x_n,t))}(\cdot)$ and $\gamma_n(\cdot|h_n) = \delta_{\varphi^s(\phi(x_n,t^*(x_n)))}(\cdot)$, where $\varphi^s : \overline{\mathbf{X}}_\infty \to \mathbf{A}_\infty$ is a measurable mapping satisfying $\varphi^s(y) \in \mathbf{A}(y)$ for any $y \in \overline{\mathbf{X}}$. By a slight abuse of notation, such a strategy will be just denoted by $u = \varphi^s$.

From Theorem 3.6 in [23] (or Remark 3.43, page 87 in [24]) we have that, for any admissible strategy $u \in \mathcal{U}$ and an initial state $x_0 \in \mathbf{X}$, there exists a probability $\mathbb{P}^u_{x_0}$ on $(\Omega,\mathcal{F})$ such that the restriction of $\mathbb{P}^u_{x_0}$ to $(\Omega,\mathcal{F}_0)$ is given by (see [7] for further details) $\mathbb{P}^u_{x_0}(\{X_0 = x_0\}) = 1$, and (see Lemma 3.1 in [7]) the predictable projection of the random measure μ with respect to $\mathbb{P}^u_{x_0}$ is given by $\nu = \nu_0 + \nu_1$, where, for $\Gamma \in \mathcal{B}(\mathbb{R}_+^* \times \mathbf{X})$,

$$\nu_0(\Gamma) = \int_\Gamma \int_{\mathbf{A}(\xi_s)} Q(dx|\xi_s, a)\lambda(\xi_s, a)\pi(da|s)ds,$$

$$\nu_1(\Gamma) = \int_\Gamma \sum_{n\in\mathbb{N}^*} I_{\{\xi_{T_n-}\in\Xi\}} \int_{\mathbf{A}(\xi_{T_n-})} Q(dx|\xi_{T_n}-, a)\gamma(da|T_n-)\delta_{T_n}(ds).$$

3.4 Problems formulation

We introduce in this section the infinite-horizon expected discounted and long run average continuous-time optimal control problems we will consider in this chapter, with the control acting continuously on the jump intensity λ and on the transition measure Q of the process (but not on the deterministic flow ϕ).

In what follows the running cost rate C^g is a real-valued measurable mapping defined on $\mathbf{K}$ and the boundary cost C^i is a real-valued measurable mapping defined on $\mathbf{K}$. We set $C^g(x^\infty, a^\infty) = C^i(x^\infty, a^\infty) = 0$. The associated infinite-horizon discounted criterion corresponding to an admissible control strategy $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{U}$, $u_n = (\pi_n, \gamma_n)$, is defined by

$$\mathcal{V}_\alpha(u, x_0) = \mathbb{E}^u_{x_0}\bigg[\int_{]0,+\infty[} e^{-\alpha s} \int_{\mathbf{A}(\xi_s)} C^g(\xi_s, a)\pi(da|s)ds\bigg] + \mathbb{E}^u_{x_0}\bigg[\int_{]0,+\infty[} e^{-\alpha s} \int_{\mathbf{A}(\xi_{s-})} C^i(\xi_{s-}, a)\gamma(da|s)p^*(ds)\bigg], \tag{7}$$

where $\alpha > 0$ is the discount factor. Similarly, the associated long run average criterion corresponding to an admissible control strategy $u \in \mathcal{U}$ is defined by

$$\mathcal{A}(u, x_0) = \overline{\lim_{t\to\infty}} \frac{1}{t}\bigg\{\mathbb{E}^u_{x_0}\bigg[\int_{]0,t[} \int_{\mathbf{A}(\xi_s)} C^g(\xi_s, a)\pi(da|s)ds\bigg] + \mathbb{E}^u_{x_0}\bigg[\int_{]0,t[} \int_{\mathbf{A}(\xi_{s-})} C^i(\xi_{s-}, a)\gamma(da|s)p^*(ds)\bigg]\bigg\}. \tag{8}$$

Definition 3 The optimization problems consist in minimizing the performance criterion $\mathcal{V}_\alpha(u, x_0)$ and $\mathcal{A}(u, x_0)$ within the class of admissible strategies $u \in \mathcal{U}$, where x_0 is the initial state. The optimal value functions will be denoted respectively by $\mathcal{V}^*_\alpha(x_0)$ and $\mathcal{A}^*(x_0)$, that is,

$$\mathcal{V}^*_\alpha(x_0) = \inf_{u\in\mathcal{U}} \mathcal{V}_\alpha(u, x_0), \quad \mathcal{A}^*(x_0) = \inf_{u\in\mathcal{U}} \mathcal{A}(u, x_0)$$

and $u \in \mathcal{U}$ will be an optimal strategy for the discounted (respectively, long run average) problem if $\mathcal{V}_\alpha(u, x_0) = \mathcal{V}^*_\alpha(x_0)$ (respectively, $\mathcal{A}(u, x_0) = \mathcal{A}^*(x_0)$).

4 Main assumptions and auxiliary results

The objective of this section is to introduce the assumptions and present some technical results that will be used along this chapter.

4.1 Main assumptions

Our approach requires that the process must be non-explosive and that the expected value of the number of jumps at the boundary up to a time $t \in \mathbb{R}_+$ must be bounded from above by an *affine* function in the variable t. One of the main goals of Assumption A is to ensure these properties.

Assumption A. There are constants $K \geq 0$ and $\varepsilon_1 > 0$ such that

(A1) For any $(x, a) \in \mathbf{K}^g$, $\lambda(x, a) \leq K$.
(A2) For any $(z, b) \in \mathbf{K}^i$, $Q(A_{\varepsilon_1}|z, b) = 1$ *where*

$$A_{\varepsilon_1} = \{x \in \mathbf{X} : t^*(x) > \varepsilon_1\}.$$

(A3) For any $(x,a) \in \mathbf{K}^g$, $Q(A(x)|x,a) = 1$ *where*

$$A(x) = \{y \in \mathbf{X} : t^*(y) \geq \min\{t^*(x), \varepsilon_1\}\}.$$

Assumptions B and C are classical hypotheses. They mainly ensure the existence of an optimal selector.

Assumption B.

(B1) For every $y \in \overline{\mathbf{X}}$ *the set* $\mathbf{A}(y)$ *is compact.*
(B2) The kernel Q *is weakly continuous (also called weak-Feller Markov kernel) on* $\mathbf{K}^g$.
(B3) The function λ *is continuous on* $\mathbf{K}^g$.
(B4) The flow ϕ *is continuous on* $\mathbb{R}_+ \times \mathbb{R}^p$.
(B5) The function t^* *is continuous on* $\overline{\mathbf{X}}$.

Assumption C.

(C1) The multifunction Ψ^g *from* $\mathbf{X}$ *to* $\mathbf{A}$ *defined by* $\Psi^g(x) = \mathbf{A}(x)$ *is upper semi-continous. The multifunction* Ψ^i *from* Ξ *to* $\mathbf{A}$ *defined by* $\Psi^i(z) = \mathbf{A}(z)$ *is upper semicontinous.*
(C2) The cost function C^g *(respectively,* C^i*) is bounded and lower semicontinuous on* $\mathbf{K}^g$ *(respectively,* $\mathbf{K}^i$*).*

Without loss of generality, we assume, from Assumption (C2), that the inequalities $|C^g| \leq K$ and $|C^i| \leq K$ are valid, where K is the same constant as in Assumption (A1).

4.2 Auxiliary results

We present in this subsection some auxiliary results that will be useful to study both the infinite-horizon discounted control problem as well as the long-run average cost control problem. The first result of this subsection, Lemma 1, shows that the controlled process is non-explosive and provides an upper bound for the sum of the expected values of $e^{-\alpha T_n}$ as well as an affine upperbound on t for the expected value on the number of jumps from the frontier up to a time t. This result requires only Assumption A.

Lemma 1 *If Assumption A is satisfied then there exist positive numbers* $M < \infty$, $c_0 < \infty$ *such that, for any control strategy* $u \in \mathcal{U}$ *and initial state* $x_0 \in \mathbf{X}$,

$$\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*} e^{-\alpha T_n}\Big] \leq M, \; \mathbb{P}^u_{x_0}(T_\infty < +\infty) = 0. \tag{9}$$

Furthermore for any $t \in \mathbb{R}_+$,

$$\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*} I_{\{T_n \leq t, \xi_{T_n^-} \in \Xi\}}\Big] \leq Mt + c_0. \tag{10}$$

Proof: For the proof of (9), see Lemma 4.1 in [7] and, for the proof of (10), see Lemma 3.1 in [6]. □

Recalling the definitions of $\mathcal{V}_\alpha$ and $\mathcal{A}$ (see equations (7) and (8) respectively), it is easy to get that for any control strategy $u \in \mathcal{U}$

$$|\mathcal{V}_\alpha(u,x_0)| \leq K(\frac{1}{\alpha} + \mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*} e^{-\alpha T_n}\Big]) \leq K(\frac{1}{\alpha} + M)$$

and

$$|\mathcal{A}(u,x_0)| \leq K\Big(1 + \overline{\lim_{t\to\infty}}\, \frac{1}{t}\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*} I_{\{T_n\leq t, \xi_{T_n^-}\in\Xi\}}\Big]\Big) \leq K(1+M),$$

by using Lemma 1 and the fact that $|C^g| \leq K$ and $|C^i| \leq K$ (see Assumption (C2)). Therefore, the mappings $\mathcal{V}_\alpha(u,\cdot)$ and $\mathcal{A}(u,\cdot)$ are well defined.

The next lemma will be useful to obtain the characterization of the value functions in terms of integro differential equations.

Lemma 2 *Consider a bounded from below real-valued measurable function F defined on $\mathbf{X}$ such that, for a real number $\beta > 0$, it satisfies*

$$\int_{[0,t^*(x)[} e^{-\beta s} F(\phi(x,s))ds < +\infty$$

for any $x \in \overline{\mathbf{X}}$, and a bounded from below real-valued measurable function G defined on Ξ. Then the real-valued mapping V defined on $\overline{\mathbf{X}}$ by

$$V(x) = \int_{[0,t^*(x)[} e^{-\beta s} F(\phi(x,s))ds + e^{-\beta t^*(x)} G(\phi(x,t^*(x)))$$

belongs to $\mathbb{A}(\overline{\mathbf{X}})$. Moreover there exists a bounded from below measurable function $\mathcal{X}V$ satisfying

$$-\beta V(x) + \mathcal{X}V(x) = -F(x),$$

for any $x \in \mathbf{X}$ and, furthermore, $V(z) = G(z)$ for any $z \in \Xi$.

Proof: See the Appendix.

For any function V in $\mathbb{M}(\overline{\mathbf{X}})$ bounded from below let us introduce the $\widehat{\mathbb{R}}$-valued mappings $\mathfrak{R}V$ and $\mathfrak{T}V$ defined on $\mathbf{X}$ and Ξ respectively by

$$\mathfrak{R}V(x) = \inf_{a\in\mathbf{A}(x)} \Big\{C^g(x,a) + qV(x,a) + KV(x)\Big\}, \tag{11}$$

$$\mathfrak{T}V(z) = \inf_{b\in\mathbf{A}(z)} \Big\{C^i(z,b) + QV(z,b)\Big\}, \tag{12}$$

where the constant K has been defined in Assumption (A1) and the transition kernel q in equation (3). Observe that qV and QV are well defined since by hypothesis V is bounded from below. Note also qV and QV may take the value $+\infty$. Finally, for any $\alpha \in [0,1]$, let us introduce the $\widehat{\mathbb{R}}$-valued function $\mathfrak{B}_\alpha V$ defined on $\overline{\mathbf{X}}$ by

$$\mathfrak{B}_\alpha V(y) = \int_{[0,t^*(y)[} e^{-(K+\alpha)t} \mathfrak{R}V(\phi(y,t))dt + e^{-(K+\alpha)t^*(y)} \mathfrak{T}V(\phi(y,t^*(y))). \quad (13)$$

Again, remark the integral term in (13) is well defined but may take the value $+\infty$. Moreover, since $|C^g| \le K$ and $|C^i| \le K$, we have clearly that $\mathfrak{R}V(x) \ge -Kc_0$ and $\mathfrak{T}V(z) \ge -c_0$ for some constant $c_0 > 0$. By using the definition of $\mathfrak{B}_\alpha V$

$$\mathfrak{B}_\alpha V(y) \ge -c_0(1 - e^{-Kt^*(y)}) - c_0 e^{-Kt^*(y)} = -c_0$$

for any $\alpha \in [0,1]$.

The next lemma provides important properties of the operators $\mathfrak{R}$, $\mathfrak{T}$ and $\mathfrak{B}_\alpha$.

Lemma 3 *Suppose that Assumptions A, B and C are satisfied. If $V \in \mathbb{L}(\overline{\mathbf{X}})$ is bounded from below then for any $\alpha \in [0,1]$ we have that*

$$\mathfrak{R}V \in \widehat{\mathbb{L}}(\mathbf{X}),\ \mathfrak{T}V \in \widehat{\mathbb{L}}(\boldsymbol{\Xi}),\ \mathfrak{B}_\alpha V \in \widehat{\mathbb{L}}(\overline{\mathbf{X}})$$

and all these functions are bounded from below.

Proof: See the Appendix.

For any $0 < \alpha < 1$, let us introduce

$$K_\alpha = \frac{K(1+K)(1 - e^{-(K+\alpha)\varepsilon_1}) + (K+\alpha)Ke^{-(K+\alpha)\varepsilon_1}}{\alpha(1 - e^{-(K+\alpha)\varepsilon_1})},$$
$$K_C = \frac{2K(1+K)}{1 - e^{-K\varepsilon_1}},$$

where K and ϵ_1 have been defined in Assumption A. Clearly, for any $0 < \alpha < 1$

$$0 < \alpha K_\alpha \le K_C. \quad (14)$$

The next lemma provides upper bounds and absolutely continuity properties of the operator $\mathfrak{B}_\alpha$.

Lemma 4 *Suppose that Assumptions A, B and C hold. Consider $V \in \mathbb{L}_b(\overline{\mathbf{X}})$ satisfying, for any $y \in \overline{\mathbf{X}}$,*

$$|V(y)| \le K_\alpha I_{A_{\varepsilon_1}}(y) + (K_\alpha + K) I_{A^c_{\varepsilon_1}}(y).$$

Then $\mathfrak{B}_\alpha V \in \mathbb{A}_b(\overline{\mathbf{X}})$ and for any $y \in \overline{\mathbf{X}}$,

$$|\mathfrak{B}_\alpha V(y)| \le K_\alpha I_{A_{\varepsilon_1}}(y) + (K_\alpha + K) I_{A^c_{\varepsilon_1}}(y).$$

Proof: See Lemma 5.4 in [7]. □

We conclude this section with the following result, which is a consequence of the so-called Dynkin formula associated with the controlled process $(\xi_t)_{t\in\mathbb{R}_+}$.

Theorem 1 *Suppose that Assumption A is satisfied and that the cost functions C^g and C^i are bounded (below or above). Then we have, for any strategy $u=(\pi_n,\gamma_n)\in\mathcal{U}$ and $(W,\mathcal{X}W)\in\mathbb{A}_b(\overline{\mathbf{X}})\times\mathbb{B}(\mathbf{X})$, that*

$$\begin{aligned}
&\mathcal{V}_\alpha(u,x_0)=W(x_0)+\mathbb{E}^u_{x_0}\Big[\int_{]0,+\infty[}e^{-\alpha s}\left[\mathcal{X}W(\xi_s)-\alpha W(\xi_s)\right]ds\Big]\\
&+\mathbb{E}^u_{x_0}\Big[\int_{]0,+\infty[}e^{-\alpha s}\int_{\mathbf{A}^g}\{C^g(\xi_s,a)\\
&+\int_{\mathbf{X}}W(y)Q(dy|\xi_s,a)\lambda(\xi_s,a)-W(\xi_s)\lambda(\xi_s,a)\}\pi(da|s)]ds\Big]\\
&+\mathbb{E}^u_{x_0}\Big[\sum_{n\in\mathbb{N}^*}I_{\{\xi_{T_n-}\in\Xi\}}e^{-\alpha T_n}\Big[\int_{\mathbf{A}^i}\{C^i(\xi_{T_n-},a)\\
&+\int_{\mathbf{X}}W(y)Q(dy|\xi_{T_n-},a)\}\gamma(da|T_n-)-W(\xi_{T_n-})\Big]\Big].
\end{aligned}\tag{15}$$

Proof: See Corollary 4.3 in [7]. □

5 The discounted control problem

Theorem 2 below presents sufficient conditions based on the three local characteristics of the process ϕ, λ, Q, and the semi-continuity properties of the set valued action space, for the existence of a solution for an integro-differential HJB optimality equation associated with the discounted control problem as well as conditions for the existence of an optimal selector. Moreover it shows that the solution of the integro-differential HJB optimality equation is in fact unique and coincides with the optimal value for the α-discounted problem, and the optimal selector derived in Theorem 2 yields an optimal deterministic stationary strategy for the discounted control problem.

Theorem 2 *Suppose Assumptions A, B and C are satisfied. Then there exist $W\in\mathbb{A}_b(\overline{\mathbf{X}})$ and $\mathcal{X}W\in\mathbb{B}(\mathbf{X})$ satisfying, for any $x\in\mathbf{X}$,*

$$-\alpha W(x)+\mathcal{X}W(x)+\inf_{a\in A^g(x)}\Big\{C^g(x,a)+qW(x,a)\Big\}=0,\tag{16}$$

and, for any $z\in\Xi$,

$$W(z)=\inf_{b\in A^i(z)}\Big\{C^i(z,b)+QW(z,b)\Big\}.\tag{17}$$

Moreover there is a measurable mapping $\widehat{\varphi}_\alpha : \overline{\mathbf{X}} \to \mathbf{A}$ *such that* $\widehat{\varphi}_\alpha(y) \in \mathbf{A}(y)$ *for any* $y \in \overline{\mathbf{X}}$ *and satisfying, for any* $x \in \mathbf{X}$,

$$C^g(x, \widehat{\varphi}_\alpha(x)) + qW(x, \widehat{\varphi}_\alpha(x)) = \inf_{a \in \mathbf{A}(x)} \Big\{ C^g(x,a) + qW(x,a) \Big\}, \tag{18}$$

and, for any $z \in \Xi$,

$$C^i(z, \widehat{\varphi}_\alpha(z)) + QW(z, \widehat{\varphi}_\alpha(z)) = \inf_{b \in \mathbf{A}(z)} \Big\{ C^i(z,b) + QW(z,b) \Big\}. \tag{19}$$

Furthermore we have that

a) the deterministic stationary strategy $\widehat{\varphi}_\alpha$ *is optimal for the* α*-discounted problem,*
b) the function $W \in \mathbb{A}_b(\overline{\mathbf{X}})$*, solution of (16)-(17), is unique and coincides with* $\mathcal{V}_\alpha^*(x) = inf_{u \in \mathcal{U}} \mathcal{V}_\alpha(u,x)$*, and*
c) $\mathcal{V}_\alpha^*(x)$ *satisfies*

$$|\mathcal{V}_\alpha^*(x)| \leq K_\alpha + K I_{A_{\varepsilon_1}^c}(x). \tag{20}$$

Proof: By Lemma 3, one can define recursively the sequence of functions $\{W_i\}_{i \in \mathbb{N}}$ in $\mathbb{L}_b(\overline{\mathbf{X}})$ as follows: $W_{i+1}(y) = \mathfrak{B}_\alpha W_i(y)$, for $i \in \mathbb{N}$ and $W_0(y) = -K_\alpha I_{A_{\varepsilon_1}}(y) - (K_\alpha + K) I_{A_{\varepsilon_1}^c}(y)$ for any $y \in \overline{\mathbf{X}}$. By using Lemma 4 and the definition of W_0, we obtain that $W_1(y) \geq W_0(y)$ for any $y \in \overline{\mathbf{X}}$. Now, note that the operator $\mathfrak{B}_\alpha$ is monotone, that is, $V_1 \leq V_2$ implies $\mathfrak{B}_\alpha V_1 \leq \mathfrak{B}_\alpha V_2$. Consequently, it can be shown by induction on i that the sequence $\{W_i\}_{i \in \mathbb{N}}$ is increasing and, from Lemma 4 and the definition of W_0, that for every $i \in \mathbb{N}$,

$$|W_{i+1}(x)| = |\mathfrak{B}_\alpha W_i(x)| \leq K_\alpha I_{A_{\varepsilon_1}}(x) + (K_\alpha + K) I_{A_{\varepsilon_1}^c}(x). \tag{21}$$

Therefore from (21) the sequence of functions $\{W_i\}_{i \in \mathbb{N}}$ is uniformly bounded, that is, for any $i \in \mathbb{N}$, $\sup_{y \in \overline{\mathbf{X}}} |W_i(y)| \leq K_\alpha + K$. As a result, $\{W_i\}_{i \in \mathbb{N}}$ converges to a mapping $W \in \mathbb{B}(\overline{\mathbf{X}})$. Since $\{W_i\}_{i \in \mathbb{N}}$ is an increasing sequence of lower semicontinuous functions, $W \in \mathbb{L}_b(\overline{\mathbf{X}})$, $KW_i + qW_i \in \mathbb{L}_b(\mathbf{K}^g)$, and so, $C^g + KW_i + qW_i \in \mathbb{L}_b(\mathbf{K}^g)$ by Assumption (C2). Therefore, combining Assumptions (B1) and (C1) and Lemma 2.1 in [28], it follows that $\lim_{i \to \infty} \mathfrak{R} W_i(x) = \mathfrak{R} W(x)$ for any $x \in \mathbf{X}$ and $\lim_{i \to \infty} \mathfrak{T} W_i(z) = \mathfrak{T} W(z)$ for any $z \in \Xi$. By using the bounded convergence Theorem, it implies that the mapping W satisfies the following equations

$$\begin{aligned} W(y) &= \mathfrak{B}_\alpha W(y) \\ &= \int_{[0, t^*(y)[} e^{-(K+\alpha)t} \mathfrak{R} W(\phi(y,t)) dt + e^{-(K+\alpha)t^*(y)} \mathfrak{T} W(\phi(y, t^*(y))), \end{aligned} \tag{22}$$

where $y \in \overline{\mathbf{X}}$. Applying Lemma 2 to the mapping W where the function F (respectively G) is given by $\mathfrak{R} W$ (respectively, $\mathfrak{T} W$), it yields that the function $W \in \mathbb{A}_b(\overline{\mathbf{X}})$ and satisfies

$$-(\alpha + K) W(x) + \mathcal{X} W(x) = -\inf_{a \in A^g(x)} \Big\{ C^g(x,a) + qW(x,a) + KW(x) \Big\},$$

for any $x \in \mathbf{X}$ and

$$W(z) = \inf_{b \in A^i(z)} \Big\{ C^i(z,b) + QW(z,b) \Big\},$$

for any $z \in \mathbf{\Xi}$. This shows the existence of $W \in \mathbb{A}_b(\overline{\mathbf{X}})$ and $\mathcal{X}W \in \mathbb{B}(\mathbf{X})$ satisfying equations (16) and (17).

Now, under Assumptions B and C, for any $x \in \mathbf{X}$ the mapping defined on $\mathbf{A}(x)$ by

$$a \to C^g(x,a) + \lambda(x,a)[QW(x,a) - W(x)] + KW(x)$$

is lower semicontinuous and since Ψ^g is upper semicontinuous, it follows from Proposition D.5 in [21] that there exists a measurable mapping $\varphi^g_\alpha : \mathbf{X} \to \mathbf{A}^g$ such that $\forall x \in \mathbf{X}$ $\varphi^g_\alpha(x) \in \mathbf{A}(x)$ and equation (18) holds. Similar arguments can be used to show the existence of a measurable mapping $\varphi^i_\alpha : \mathbf{\Xi} \to \mathbf{A}^i$ satisfying $\varphi^i_\alpha(z) \in \mathbf{A}(z)$ for any $z \in \mathbf{\Xi}$ and equation (19) holds. Therefore, the measurable mapping $\widehat{\varphi}_\alpha$ defined by $\widehat{\varphi}_\alpha(x) = \varphi^i_\alpha(x)$ for any $x \in \mathbf{X}$ and $\widehat{\varphi}_\alpha(z) = \varphi^i_\alpha(z)$ for any $z \in \mathbf{\Xi}$ satisfies the claim.

To show a) and b), notice that for an arbitrary control strategy $u \in \mathcal{U}$ we have, by using Theorem 1, that $\mathcal{V}_\alpha(u,x) \geq W(x)$ for any $x \in \mathbf{X}$ and also that $\mathcal{V}_\alpha(\widehat{\varphi},x) = W(x)$ for any $x \in \mathbf{X}$. Indeed from (16) and (17) we have that

$$\mathcal{X}W(\xi_s) - \alpha W(\xi_s) + \int_{\mathbf{A}^g} \{C^g(\xi_s,a) \\ + \int_{\mathbf{X}} W(y)Q(dy|\xi_s,a)\lambda(\xi_s,a) - W(\xi_s)\lambda(\xi_s,a)\}\pi(da|s)] \geq 0$$

and, for any $z \in \mathbf{\Xi}$,

$$\int_{\mathbf{A}^i} \{C^i(\xi_{T_n-},a) + \int_{\mathbf{X}} W(y)Q(dy|\xi_{T_n-},a)\}\gamma(da|T_n-) - W(\xi_{T_n-}) \geq 0$$

with equality whenever the strategy $\widehat{\varphi}$ is used. From (15) the terms inside the expected value are positive, being zero whenever the strategy $\widehat{\varphi}$ is used, which shows that $\mathcal{V}_\alpha(u,x) \geq W(x)$ and $\mathcal{V}_\alpha(\widehat{\varphi},x) = W(x)$ as desired. Finally from (21) we have c) since $\mathcal{V}^*_\alpha(x) = W(x) = \sup_{i \in \mathbb{N}} W_i(x)$. □

6 The average control problem

The objective of this section is to provide sufficient conditions to show the existence of a solution to an integro-differential HJB inequality as well as the existence on optimal selector. This results is proved by using the so-called vanishing discount approach. The second main result of this section (see Theorem 4) gives the existence of a deterministic stationary optimal policy for the infinite-horizon long run average continuous-time control problem according to Definition 3.

Let us introduce

$$m_\alpha = \inf_{x\in\overline{\mathbf{X}}} \mathcal{V}_\alpha^*(x), \;\; \rho_\alpha = \alpha m_\alpha, \tag{23}$$

$$h_\alpha(x) = \mathcal{V}_\alpha^*(x) - m_\alpha \ge 0, \tag{24}$$

where $x \in \overline{\mathbf{X}}$. In what follows we refer to section 2 for the definition of the generalized inferior limit $\underline{\lim}^g$. The following final assumption will be required.

Assumption D. $\underline{\lim}^g_{\alpha\to 0} h_\alpha(x) < \infty$ for all $x \in \overline{\mathbf{X}}$.

It is easy to show that there exist a sequence $\{\alpha_n\}$ satisfying $\lim_{n\to\infty}\alpha_n = 0$ and such that $\lim_{n\to\infty}\rho_{\alpha_n} = \rho$ for some $|\rho| \le K_C + K$. To see this, observe that by combining equations (14), (20) and (23) we obtain that for any $0 < \alpha < 1$

$$|\rho_\alpha| = |\alpha m_\alpha| \le \alpha | \inf_{x\in\overline{\mathbf{X}}} \mathcal{V}_\alpha^*(x)| \le \alpha \sup_{x\in\overline{\mathbf{X}}} |\mathcal{V}_\alpha^*(x)| \le \alpha K_\alpha + K \le K_C + K. \tag{25}$$

Let us introduce the function h_* given by

$$h_*(x) = \underline{\lim}^g_{n\to\infty} h_{\alpha_n}(x). \tag{26}$$

It is easy to see that $h_*(x) \ge 0$ since $h_\alpha(x) \ge 0$. Clearly,$h_*(x) < \infty$ by Assumption D and $h_* \in \mathbb{L}_+(\overline{\mathbf{X}})$ by using Proposition 1.

Before showing the main results of this section, we need the following technical result.

Lemma 5 *The function h_* defined in* (26) *satisfies the following inequality:*

$$h_*(x) \ge \int_{[0,t^*(x)[} e^{-Ks}(\Re h_*(\phi(x,s)) - \rho)ds + e^{-Kt^*(x)}\mathfrak{T} h_*(\phi(x,t^*(x))). \tag{27}$$

Proof: See the Appendix. □

The following theorem provides sufficient conditions for the existence of a solution and optimal selector to an integro-differential HJB inequality, associated to the long run average control problem.

Theorem 3 *Suppose that Assumptions A, B, C and D are satisfied. Then the following holds:*

a) There exist $H \in \mathbb{A}(\overline{\mathbf{X}}) \cap \mathbb{L}(\overline{\mathbf{X}})$ bounded from below satisfying

$$\rho \ge \mathcal{X} H(x) + \inf_{a\in A^g(x)} \Big\{ C^g(x,a) + qH(x,a) \Big\}, \tag{28}$$

for any $x \in \mathbf{X}$, and

$$H(z) \ge \inf_{b\in A^i(z)} \Big\{ C^i(z,b) + QH(z,b) \Big\}, \tag{29}$$

for any $z \in \Xi$.

b) There is a measurable mapping $\widehat{\varphi} : \overline{\mathbf{X}} \to \mathbf{A}$ *such that* $\widehat{\varphi}(y) \in \mathbf{A}(y)$ *for any* $y \in \overline{\mathbf{X}}$ *and satisfying*

$$C^g(x, \widehat{\varphi}(x)) + qH(x, \widehat{\varphi}(x)) = \inf_{a \in \mathbf{A}(x)} \Big\{ C^g(x,a) + qH(x,a) \Big\}, \tag{30}$$

for any $x \in \mathbf{X}$, *and*

$$C^i(z, \widehat{\varphi}(z)) + QH(z, \widehat{\varphi}(z)) = \inf_{b \in \mathbf{A}(z)} \Big\{ C^i(z,b) + QH(z,b) \Big\}, \tag{31}$$

for any $z \in \Xi$.

Proof: Let us introduce $H(x)$ as

$$H(x) = \int_{[0,t^*(x)[} e^{-Ks} (\mathfrak{R} h_*(\phi(x,s)) - \rho) ds + e^{-Kt^*(x)} \mathfrak{T} h_*(\phi(x, t^*(x))), \tag{32}$$

for all $x \in \overline{\mathbf{X}}$.

We will prove first item a). Observe that $H(x) = \mathfrak{B}_0 h_*(x) - \rho \int_{[0,t^*(x)[} e^{-Ks} ds$. Now by Lemma 3 it follows that H is bounded below and that $H \in \widehat{\mathbb{L}}(\overline{\mathbf{X}})$ since $h_* \in \mathbb{L}(\overline{\mathbf{X}})$ and t^* is continuous by Assumption (B5). Observe that equation (27) implies that $H(x) \leq h_*(x)$ showing that $H \in \mathbb{L}(\overline{\mathbf{X}})$.

A straightforward application of Lemma 2 shows that $H(x) \in \mathbb{A}(\overline{\mathbf{X}})$ and it also follows that there exists a bounded from below measurable function $\mathcal{X}H$ satisfying

$$-KH(x) + \mathcal{X}H(x) + \inf_{a \in \mathbf{A}(x)} \Big\{ C^g(x,a) + qh_*(x,a) + Kh_*(x) \Big\} = \rho \tag{33}$$

for any $x \in \mathbf{X}$ and

$$H(z) = \inf_{b \in \mathbf{A}(z)} \Big\{ C^i(z,b) + Qh_*(z,b) \Big\}, \tag{34}$$

for any $z \in \Xi$. Recalling that $h_*(x) \geq H(x)$, we obtainfrom (33) and (34) that for any $x \in \mathbf{X}$,

$$\begin{aligned} &\mathcal{X}H(x) + \inf_{a \in \mathbf{A}(x)} \Big\{ C^g(x,a) + qH(x,a) \Big\} \\ &\qquad \leq -KH(x) + \mathcal{X}H(x) + \inf_{a \in \mathbf{A}(x)} \Big\{ C^g(x,a) + qh_*(x,a) + Kh_*(x) \Big\} = \rho \end{aligned} \tag{35}$$

and for any $z \in \Xi$,

$$\inf_{b \in \mathbf{A}(z)} \Big\{ C^i(z,b) + QH(z,b) \Big\} \leq \inf_{b \in \mathbf{A}(z)} \Big\{ C^i(z,b) + Qh_*(z,b) \Big\} = H(z). \tag{36}$$

Combining equations (35), (36), we finally get that $H \in \mathbb{A}(\overline{\mathbf{X}}) \cap \mathbb{L}(\overline{\mathbf{X}})$ and satisfies equations (28) and (29) giving item a).

Item b) is an easy consequence of the fact that H is lower semicontinuous on $\overline{\mathbf{X}}$, Assumptions A, B, C and Proposition D.5 in [21]. □

The goal now is to establish a deterministic stationary optimal policy for the long run average control problem as defined in Definition 3, based on a solution for the integro-differential HJB inequality (28), (29) and its associated optimal selector (30), (31). In order to do that we introduce the following notation for a measurable selector φ, a function $W \in \mathbb{M}(\overline{\mathbf{X}})$ bounded from below, and any $x \in \mathbf{X}$,

$$
\begin{aligned}
\lambda^{\varphi}(x) &= \lambda(x,\varphi(x)),\ \ \Lambda^{\varphi}(x,t) = \int_0^t \lambda^{\varphi}(\phi(x,s))ds,\\
Q^{\varphi}W(x) &= QW(x,\varphi(x)),\ \ q^{\varphi}W(x) = qW(x,\varphi(x)),\\
\lambda^{\varphi}Q^{\varphi}W(x) &= \lambda(x,\varphi(x))QW(x,\varphi(x)),\\
C^{g,\varphi}(x) &= C^{g}(x,\varphi(x)),\ \ C^{i,\varphi}(z) = C^{z}(x,\varphi(z)),\ \ z \in \mathbf{\Xi}
\end{aligned}
$$

and for ρ, $\widehat{\varphi}$ as in Theorem 3,

$$
\begin{aligned}
G^{\widehat{\varphi}}W(x) &= \int_{]0,t^*(x)[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}\lambda^{\widehat{\varphi}}Q^{\widehat{\varphi}}W(\phi(x,s)))ds + e^{\Lambda^{\widehat{\varphi}}(x,t^*(x))}Q^{\widehat{\varphi}}W(\phi(x,t^*(x))),\\
L^{\widehat{\varphi}}W(x) &= \int_{]0,t^*(x)[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}W(\phi(x,s))ds,\\
\mathcal{L}^{\widehat{\varphi}}(x) &= \int_{]0,t^*(x)[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}ds,\\
P^{\widehat{\varphi}}W(x) &= e^{-\Lambda^{\widehat{\varphi}}(x,t^*(x))}W(\phi(x,t^*(x)),\\
\mathcal{T}^{\widehat{\varphi}}(\rho,W)(x) &= -\rho\mathcal{L}^{\widehat{\varphi}}(x) + L^{\widehat{\varphi}}C^{g,\widehat{\varphi}}(x) + P^{\widehat{\varphi}}C^{i,\widehat{\varphi}}(x) + G^{\widehat{\varphi}}W(x).
\end{aligned}
$$

We have the following auxiliary result.

Lemma 6 *For H and ρ, $\widehat{\varphi}$ as in Theorem 3 we have that*

$$H(x) \geq \mathcal{T}^{\varphi}(\rho,H)(x) \tag{37}$$

$$J_m^{\widehat{\varphi}}(t,x) \leq H(x) \tag{38}$$

where

$$
\begin{aligned}
J_m^{\widehat{\varphi}}(t,x) =& \mathbb{E}_x^{\widehat{\varphi}}\Big[\int_{]0,t\wedge T_m[}\Big[C^{g}(\xi_s,\widehat{\varphi})) - \rho\Big]ds\Big]\\
&+ \mathbb{E}_x^{\widehat{\varphi}}\Big[\int_{]0,t\wedge T_m[} C^{i}((\xi_{s-},\widehat{\varphi}))dp^*(s) + \mathcal{T}^{\widehat{\varphi}}(\rho,H)(\xi_{t\wedge T_m})\Big].
\end{aligned}
$$

Proof: See the Appendix. □

Theorem 4 *Suppose that Assumptions A, B, C and D are satisfied and consider $\widehat{\varphi}$ as in (30), (31). Then the deterministic stationary strategy $\widehat{\varphi}$ is optimal for the average cost problem and for any $x \in \mathbf{X}$,*

$$\rho = \mathcal{A}(\widehat{\varphi}, x) = \mathcal{A}^*(x). \tag{39}$$

Proof: Applying Proposition 4.6 in [8] it follows that $\overline{\lim}_{\alpha\downarrow 0}\alpha\mathcal{V}^*_\alpha(x) \le \mathcal{A}^*(x)$. Therefore,

$$\rho = \lim_{n\to\infty} \alpha_n \inf_{x\in\overline{\mathbf{X}}} \mathcal{V}^*_{\alpha_n}(x) \le \overline{\lim_{n\to\infty}}\, \alpha_n \mathcal{V}^*_{\alpha_n}(x) \le \mathcal{A}^*(x).$$

To get the reverse inequality, first observe that, since $\mathcal{T}^{\widehat{\varphi}}(\rho, H)$ is bounded from below by, say, $-c_0$, we obtain from Lemma 6 that

$$-c_0 + \mathbb{E}^{\widehat{\varphi}}_x\Big[\int_{]0,t\wedge T_m[} \big[C^g(\xi_s, \widehat{\varphi}))\big]ds + \int_{]0,t\wedge T_m[} C^i((\xi_{s-}, \widehat{\varphi}))dp^*(s)\Big] \le H(x) + \rho\mathbb{E}^{\widehat{\varphi}}_x(t\wedge T_m).$$

Taking the limit as m goes to infinity, this yields

$$-c_0 + \mathbb{E}^{\widehat{\varphi}}_x\Big[\int_{]0,t[} \big[C^g(\xi_s, \widehat{\varphi}))\big]ds + \int_{]0,t[} C^i((\xi_{s-}, \widehat{\varphi}))dp^*(s)\Big] \le H(x) + \rho t,$$

and so,

$$\mathcal{A}(x, \widehat{\varphi})(x) \le \rho.$$

However, $\mathcal{A}^*(x) \le \mathcal{A}(x, \widehat{\varphi})(x)$ giving the results. □

Acknowledgements The first author received financial support from CNPq (Brazilian National Council for Scientific and Technological Development), grant 304149/2019-5, FAPESP (São Paulo Research Foundation)/ Shell through the Research Centre for Gas Innovation, FAPESP Grant Proc. 2020/15230-5, project INCT, grant CNPq 465755/2014-3, FAPESP 2014/50851-0 and FUSP (University of São Paulo Support Foundation).

Appendix

In this appendix we present the proof of some auxiliary results needed along this chapter.

Proof of Lemma 2: Write $V_n(x) = \int_{[0,t^*(x)[} e^{-\beta s}F_n(\phi(x,s))ds + e^{-\beta t^*(x)}G_n(\phi(x,t^*(x)))$ for $x \in \overline{\mathbf{X}}$ with $F_n(x) = \min\{F(x), n\}$ and $G_n(x) = \min\{G(x), n\}$ on $\mathbf{X}$ (respectively, Ξ). Now, observe that for any $x \in \mathbf{X}$, $t^*(\phi(x,t)) = t^*(x) - t$, $\phi(\phi(x,t), t^*(\phi(x,t))) = \phi(x, t^*(x))$ and $\phi(\phi(x,t), s) = \phi(x, t+s)$, for any $(t,s) \in \mathbb{R}^2_+$ with $t + s \le t^*(x)$.Then, it can be easily shown by a change of variable that for any $x \in \mathbf{X}$ and $t \in [0, t^*(x)[$,

$$V_n(\phi(x,t)) = e^{\beta t}\int_{[t,t^*(x)[} e^{-\beta s}F_n(\phi(x,s))ds + e^{\beta t}e^{-\beta t^*(x)}G_n(\phi(x,t^*(x))),$$

and so,

$$V(\phi(x,t)) = e^{\beta t} \int_{[t,t^*(x)[} e^{-\beta s} F(\phi(x,s))ds + e^{\beta t} e^{-\beta t^*(x)} G(\phi(x,t^*(x))) \quad (40)$$

by the monotone convergence theorem. Consequently, the function $V(\phi(x,\cdot))$ is absolutely continuous on $[0,t^*(x)]\cap\mathbb{R}_+$ and so, $V \in \mathbb{A}(\overline{\mathbf{X}})$. Equation (40) implies that for any $x \in \mathbf{X}$ $\mathcal{X}V(\phi(x,t)) = \beta V(\phi(x,t)) - F(\phi(x,t))$, almost everywhere w.r.t. the Lebesgue measure on $[0,t^*(x)[$. This implies that $-\beta V(x) + \mathcal{X}V(x) = -F(x)$ for any $x \in \mathbf{X}$. Moreover, we have $V(z) = G(z)$ for any $z \in \Xi$, showing the result. □

Proof of the Lemma 3: Define $V_n(x) = \min\{V(x),n\}$ so that $V_n \in \mathbb{L}_b(\overline{\mathbf{X}})$. By using hypotheses (B2)-(B3) and the fact that λ is bounded by K on $\mathbf{K}^g$, we obtain that $qV_n + KV_n \in \mathbb{L}(\mathbf{K}^g)$, and so, by Assumption (C2) $C_0^g + qV_n + KV_n \in \mathbb{L}(\mathbf{K}^g)$. Therefore, combining Lemma 17.30 in [1] with Assumptions (B1) and (C1), it yields that $\mathfrak{R}V_n \in \mathbb{L}(\mathbf{X})$. By using the same arguments, it can be shown that $\mathfrak{T}V_n \in \mathbb{L}(\Xi)$.

Now consider $y \in \overline{\mathbf{X}}$ and a sequence $\{y_n\}_{n\in\mathbb{N}}$ in $\overline{\mathbf{X}}$ converging to y. By a slight abuse of notation, for any $y \in \mathbf{X}$, $I_{[0,t^*(y)[}(t)\ e^{-(K+\alpha)t}\ \mathfrak{R}V_n(\phi(y,t))$ denotes the function defined on $\mathbb{R}_+$ which is equal to $e^{-(K+\alpha)t}\mathfrak{R}V_n(\phi(y,t))$ on $[0,t^*(y)[$ and zero elsewhere. It can be shown easily by using the lower semicontinuity of the function $\mathfrak{R}V_n$ and the continuity of the flow ϕ that $\underline{\lim}_{n\to\infty} I_{[0,t^*(y_n)[}(t)\ e^{-(K+\alpha)t}\mathfrak{R}V_n(\phi(y_n,t)) \geq I_{[0,t^*(y)[}(t)\ e^{-(K+\alpha)t}\mathfrak{R}V_n(\phi(y,t))$, for any $t \in [0,t^*(y)[$. An application of Fatou's Lemma gives that

$$\underline{\lim}_{n\to\infty} \int_{[0,t^*(y_n)[} e^{-(K+\alpha)t}\mathfrak{R}V_n(\phi(y_n,t))dt \geq \int_{[0,t^*(y)[} e^{-(K+\alpha)t}\mathfrak{R}V_n(\phi(y,t))dt.$$

The case $t^*(y) = \infty$ is trivial. Now, if $t^*(y) < \infty$ then combining the lower semicontinuity of the function $\mathfrak{T}V$ with the continuity of the flow ϕ and t^* (see Assumptions (B4)-(B5)), it gives easily that

$$\underline{\lim}_{n\to\infty} e^{-(K+\alpha)t^*(y_n)}\mathfrak{T}V_n(\phi(y_n,t^*(y_n))) \geq e^{-(K+\alpha)t^*(y)}\mathfrak{T}V_n(\phi(y,t^*(y))),$$

showing the results hold for V_n, that is, $\mathfrak{R}V_n \in \mathbb{L}_b(\mathbf{X})$, $\mathfrak{T}V_n \in \mathbb{L}_b(\Xi)$, and $\mathfrak{B}_\alpha V_n \in \mathbb{L}_b(\overline{\mathbf{X}})$. From Proposition 10.1 in [27], it follows that $\mathfrak{R}V = \lim_{n\to\infty}\mathfrak{R}V_n \in \widehat{\mathbb{L}}(\mathbf{X})$ and similarly, $\mathfrak{T}V = \lim_{n\to\infty}\mathfrak{T}V_n \in \widehat{\mathbb{L}}(\Xi)$. Now, from the monotone convergence theorem, we have $\mathfrak{B}_\alpha V = \lim_{n\to\infty}\mathfrak{B}_\alpha V_n$, and so $\mathfrak{B}_\alpha V \in \widehat{\mathbb{L}}(\overline{\mathbf{X}})$. Clearly, these functions are bounded from below, giving the result. □

Proof of the Lemma 5: From Theorem 2 we have that $W(x) = \mathcal{V}_\alpha^*(x)$ satisfies (16) and (17), and thus from (23), (24) and after some algebraic manipulations we obtain that

$$-(\alpha+K)h_\alpha(x) + \mathcal{X}h_\alpha(x) + \inf_{a\in A^g(x)}\left\{C^g(x,a) + qh_\alpha(x,a) + Kh_\alpha(x)\right\} - \rho_\alpha = 0, \quad (41)$$

for any $x \in \mathbf{X}$,

$$h_\alpha(z) = \inf_{b\in A^i(z)} \Big\{C^i(z,b) + Qh_\alpha(z,b)\Big\}, \tag{42}$$

for any $z \in \Xi$. Moreover, according to Theorem 2 there exists a measurable selector $\widehat{\varphi}_\alpha : \overline{\mathbf{X}} \to \mathbf{A}$ satisfying $\widehat{\varphi}_\alpha(y) \in \mathbf{A}(y)$ for any $y \in \overline{\mathbf{X}}$ reaching the infimum in (41) and (42). Thus,

$$-(\alpha+K)h_\alpha(x) + \mathcal{X}h_\alpha(x) + C^g(x,\widehat{\varphi}_\alpha(x)) + qh_\alpha(x,\widehat{\varphi}_\alpha(x)) + Kh_\alpha(x) - \rho_\alpha = 0, \tag{43}$$

for any $x \in \mathbf{X}$,

$$h_\alpha(z) = C^i(z,b) + Qh_\alpha(z,\widehat{\varphi}_\alpha(z)) \tag{44}$$

for any $z \in \Xi$. Taking the integral of (43) along the flow $\phi(x,t)$, we get from (43) and (44) (see [8]) that for any $x \in \mathbf{X}$,

$$h_\alpha(x) = \int_{[0,t^*(x)[} e^{-(K+\alpha)t} (\Re h_\alpha(\phi(x,t)) - \rho_\alpha)dt + e^{-(K+\alpha)t^*(x)} \mathfrak{T} h_\alpha(\phi(x,t^*(x))), \tag{45}$$

where we recall that

$$\Re h_\alpha(y) = C^g(y,\widehat{\varphi}_\alpha(y)) + qh_\alpha(y,\widehat{\varphi}_\alpha(y)) + Kh_\alpha(y), \quad y \in \mathbf{X}, \tag{46}$$

$$\mathfrak{T} h_\alpha(z) = C^i(z,\widehat{\varphi}_\alpha(z)) + Qh_\alpha(z,\widehat{\varphi}_\alpha(z)), \quad z \in \Xi. \tag{47}$$

According to Proposition 1 (i), we can find a sequence $\{x_n\} \in \mathbf{X}$ such that $x_n \to x$ and $\underline{\lim}_{n\to\infty} h_{\alpha_n}(x_n) = h_*(x)$. In what follows set, for notational simplicity, $x_n(t) = \phi(x_n,t)$, $x(t) = \phi(x,t)$, $a_n(t) = \widehat{\varphi}_{\alpha_n}(x_n(t))$, $t_n^* = t^*(x_n)$. From continuity of t^* and ϕ (see Assumption (B4)) we have that, as $n \to \infty$, $x_n(t) \to x(t)$, and, whenever $t^*(x) < \infty$, $x_n(t_n^*) \to \phi(x,t^*(x))$. From the fact that $\Re h_\alpha$ is bounded from below and ρ_α is bounded, we can apply the Fatou's lemma in (45) to obtain that

$$\begin{aligned} h_*(x) = \underline{\lim}_{n\to\infty} h_{\alpha_n}(x_n) \geq \int_{]0,+\infty[} \underline{\lim}_{n\to\infty} \Big(I_{[0,t_n^*)}(t)e^{-(K+\alpha_n)t} [\Re h_{\alpha_n}(x_n(t)) - \rho_{\alpha_n}]\Big)dt \\ + \underline{\lim}_{n\to\infty} e^{-(K+\alpha_n)t_n^*} \mathfrak{T} h_{\alpha_n}(x_n(t_n^*)). \end{aligned} \tag{48}$$

The convergence of ρ_{α_n} to ρ together with Assumption (B5) implies that, *a.s.* on $[0,\infty)$,

$$\begin{aligned} &\underline{\lim}_{n\to\infty} I_{[0,t_n^*)}(t)e^{-(K+\alpha_n)t} \Big\{\Re h_{\alpha_n}(x_n(t)) - \rho_{\alpha_n}\Big\} \\ &\qquad = I_{[0,t^*(x))}(t)e^{-Kt}\Big\{\underline{\lim}_{n\to\infty} \Re h_{\alpha_n}(x_n(t)) - \rho\Big\}, \end{aligned} \tag{49}$$

and $\underline{\lim}_{n\to\infty} e^{-(K+\alpha_n)t_n^*} \mathfrak{T} h_{\alpha_n}(x_n(t_n^*)) = e^{-Kt^*(x)} \underline{\lim}_{n\to\infty} \mathfrak{T} h_{\alpha_n}(x_n(t_n^*))$. The goal now is to show that

$$\varliminf_{n\to\infty} \Re h_{\alpha_n}(x_n(t))) \geq \Re h_*(x(t)), \tag{50}$$

and that

$$\varliminf_{n\to\infty} \mathfrak{T} h_{\alpha_n}(x_n(t_n^*)) \geq \mathfrak{T} h_\alpha(x(t^*(x))). \tag{51}$$

Let us first show (50). For a fixed $t \in (0, t^*(x))$, there is no loss of generality in assuming that $t < t_n^*$ for any $n \in \mathbb{N}$ and thus $x_n(t) \in \mathbf{X}$. Consider a subsequence $\{n_j\}$ of $\{n\}$ such that

$$\varliminf_{n\to\infty} \Re h_{\alpha_n}(x_n(t)) = \lim_{j\to\infty} \Re h_{\alpha_{n_j}}(x_{n_j}(t)).$$

From Assumptions (B1) and (C1) the multifunction Ψ^g is compact valued and upper semi-continuous so that, from the fact that $x_{n_j}(t) \to x(t)$, we can find a subsequence of $\{a_{n_j}(t)\} \in \mathbf{A}(x_{n_j}(t))$, still denoted by $\{a_{n_j}(t)\}$ such that $a_{n_j}(t) \to a \in \mathbf{A}(x(t))$ (see Theorem 17.16 in [1]) as $j \to \infty$. From (46) we have that

$$\begin{aligned}\varliminf_{n\to\infty} \Re h_{\alpha_n}(x_n(t)) = \lim_{j\to\infty} &\Big(C^g(x_{n_j}(t), a_{n_j}(t)) + q h_{\alpha_{n_j}}(x_{n_i}(t), x_{n_j}(t))\Big) \\ &+ \lim_{j\to\infty} \Big(K h_{\alpha_{n_j}}(x_{n_j}(t))\Big),\end{aligned}$$

and therefore

$$\begin{aligned}\varliminf_{n\to\infty} \Re h_{\alpha_n}(x_n(t)) \geq{}& \varliminf_{j\to\infty} C^g(x_{n_j}(t), a_{n_j}(t)) \\ &+ \varliminf_{j\to\infty} \Big(q h_{\alpha_{n_j}}(x_{n_i}(t), x_{n_j}(t)) + K h_{\alpha_{n_j}}(x_{n_j}(t))\Big).\end{aligned} \tag{52}$$

Lower semicontinuity of C^g on $\mathbf{K}^g$ yields to

$$\varliminf_{j\to\infty} C^g(x_{n_j}(t), a_{n_j}(t)) \geq C^g(x(t), a). \tag{53}$$

From Proposition 1 (i) and (iii), the fact that Q is weakly continuous on $\mathbf{K}^g$ (Assumption (B2)), and the continuity of λ (Assumption (B3)), we get that

$$\varliminf_{j\to\infty} \lambda(x_{n_j}(t), a_{n_j}(t)) Q h_{\alpha_{n_j}}(x_{n_j}(t), a_{n_j}(t)) \geq \lambda(x(t), a) Q h_*(x(t), a) \tag{54}$$

and, recalling that $K - \lambda(x_{n_j}(t), a_{n_j}(t)) \geq 0$ from Assumption (A1), we get that

$$\varliminf_{j\to\infty} [K - \lambda(x_{n_j}(t), a_{n_j}(t))] h_{\alpha_{n_j}}(x_{n_j}(t), a_{n_j}(t)) \geq [K - \lambda(x(t), a)] h_*(x(t), a). \tag{55}$$

Combining (46), (52), (53), (54), (55), we conclude that

$$\varliminf_{n\to\infty} \Re h_{\alpha_n}(x_n(t))) \geq C^g(x(t), a) + q h_*(x(t), a) + K h_*(x(t)) \geq \Re h_*(x(t)),$$

showing (50).

Let us now show (51) for $t^*(x) < \infty$. From the fact that Ψ^i is compact valued and upper semi-continuous and $x_n(t_n^*) \to x(t^*(x))$, and using similar arguments as before (in particular equation (47)), we can find a subsequence $\{a_{n_j}(t_{n_j}^*)\} \in \mathbf{A}(x_{n_j}(t_{n_j}^*))$ such that $a_{n_j}(t_{n_j}^*) \to b \in \mathbf{A}(x(t^*(x)))$ (see again Theorem 17.16 in [1]), and that

$$\underline{\lim}_{n\to\infty} \mathfrak{T}h_{\alpha_n}(x_n(t_n^*)) \geq C^i(x(t^*(x)), b) + Qh_*(x(t^*(x)), b) \geq \mathfrak{T}h_\alpha(x(t^*(x)))$$

showing (51).

Combining (48), (49), (50) and (51) we get that (27) holds, showing Lemma 5. □

Proof of the Lemma 6: From Theorem 3 we get that for any $x \in \mathbf{X}$

$$\rho \geq \mathcal{X}H(\phi(x,s)) + C^{g,\widehat{\varphi}}(\phi(x,s)) + q^{\widehat{\varphi}}H((\phi(x,s))), \tag{56}$$

and for the case $t^*(x) < \infty$,

$$H(\phi(x,t^*(x))) \geq C^{i,\widehat{\varphi}}(\phi(x,t^*(x))) + Q^{\widehat{\varphi}}H(\phi(x,t^*(x))). \tag{57}$$

Multiplying (56) by $e^{-\Lambda^{\widehat{\varphi}}(x,s)}$ and taking the integral from 0 to t we obtain that

$$\begin{aligned}\rho\int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}ds \geq \int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}(\mathcal{X}H(\phi(x,s)) - \lambda^{\widehat{\varphi}}(\phi(x,s))H(\phi(x,s)))ds \\ + \int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}C^{g,\widehat{\varphi}}(\phi(x,s))ds + \int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}\lambda^{\widehat{\varphi}}Q^{\widehat{\varphi}}H(\phi(x,s))ds.\end{aligned} \tag{58}$$

Replacing

$$\begin{aligned}\int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}(\mathcal{X}H(\phi(x,s)) - \lambda^{\widehat{\varphi}}(\phi(x,s))H(\phi(x,s)))ds \\ = e^{-\Lambda^{\widehat{\varphi}}(x,t)}H(\phi(x,t)) - H(x)\end{aligned}$$

into (58) yields to

$$\begin{aligned}H(x) \geq -\rho\int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}ds + \int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}\lambda^{\widehat{\varphi}}Q^{\widehat{\varphi}}H(\phi(x,s))ds \\ + e^{-\Lambda^{\widehat{\varphi}}(x,t)}H(\phi(x,t)) + \int_{]0,t[} e^{-\Lambda^{\widehat{\varphi}}(x,s)}C^{g,\widehat{\varphi}}(\phi(x,s))ds.\end{aligned}$$

Taking the limit as $t \to t^*(x)$ and using (57) for the case $t^*(x) < \infty$ we obtain (37). From (37) and Proposition 3.4 in [8] we obtain (38).

References

1. C.D. Aliprantis and K.C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.
2. A. Almudevar. A dynamic programming algorithm for the optimal control of piecewise deterministic Markov processes. *SIAM Journal on Control and Optimization*, 40(2):525–539, 2001.
3. N. Bauerle and U. Rieder. Optimal control of piecewise deterministic Markov processes with finite time horizon. In *Modern Trends in Controlled Stochastic Processes: Theory and Applications*, pages 123–143. Luniver Press, United Kingdom, 2010.
4. R. Cavazos-Cadena and F. Salem-Silva. The discounted method and equivalence of average criteria for risk-sensitive Markov decision processes on Borel spaces. *Appl. Math. Optim.*, 61:167–190, 2010.
5. O. L. V. Costa and F. Dufour. The vanishing discount approach for the average continuous control of piecewise deterministic Markov processes. *Journal of Applied Probability*, 46(4):1157–1183, 2009.
6. O. L. V. Costa and F. Dufour. Hamilton-Jacobi-Bellman inequality for the average control of piecewise deterministic markov processes. *Stochastics*, 91(6):817–835, 2019.
7. O. L. V. Costa, F. Dufour, and A. B. Piunovskiy. Constrained and unconstrained optimal discounted control of piecewise deterministic Markov processes. *SIAM Journal on Control and Optimization*, 54(3):1444–1474, 2016.
8. O.L.V Costa and F. Dufour. *Continuous average control of piecewise deterministic Markov processes*. Springer Briefs in Mathematics. Springer, New York, 2013.
9. M.H.A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society (B)*, 46(3):353–388, 1984.
10. M.H.A. Davis. Control of piecewise-deterministic processes via discrete-time dynamic programming. In *Stochastic differential systems (Bad Honnef, 1985)*, volume 78 of *Lecture Notes in Control and Inform. Sci.*, pages 140–150. Springer, Berlin, 1986.
11. M.H.A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
12. M.H.A Davis and M. Farid. A target recognition problem: Sequential analysis and optimal control. *SIAM Journal on Control and Optimization*, 34(6):2116–2132, 1996.
13. M.A.H. Dempster and J.J. Ye. Necessary and sufficient optimality conditions for control of piecewise deterministic processes. *Stochastics and Stochastics Reports*, 40(3-4):125–145, 1992.
14. M.A.H. Dempster and J.J. Ye. Generalized Bellman-Hamilton-Jacobi optimality conditions for a control problem with boundary conditions. *Applied Mathematics and Optimization*, 33(3):211–225, 1996.
15. L. Forwick, M. Schäl, and M. Schmitz. Piecewise deterministic Markov control processes with feedback controls and unbounded costs. *Acta Applicandae Mathematicae*, 82(3):239–267, 2004.
16. X. Guo and O. Hernández-Lerma. *Continuous-time Markov decision processes*, volume 62 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Theory and applications.
17. X. Guo, O. Hernández-Lerma, T. Prieto-Rumeau, X.R. Cao, J. Zhang, Q. Hu, M. E. Lewis, and R. Vélez. A survey of recent results on continuous-time Markov decision processes. *TOP*, 14(2):177–261, Dec 2006.
18. X. Guo and A. Piunovskiy. Discounted continuous-time Markov decision processes with constraints: unbounded transition and loss rates. *Math. Oper. Res.*, 36(1):105–132, 2011.
19. X. Guo and L. Ye. New discount and average optimality conditions for continuous-time Markov decision processes. *Advances in Applied Probability*, 42(4):953âĂŞ985, 2010.
20. X. Guo and Y. Zhang. On risk-sensitive piecewise deterministic Markov decision processes. *Applied Mathematics & Optimization*, pages 1–26, Feb 2018.

21. O. Hernández-Lerma and J.-B. Lasserre. *Discrete-time Markov control processes*, volume 30 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1996. Basic optimality criteria.
22. Y. Huang, Z. Lian, and X. Guo. Risk-sensitive finite-horizon piecewise deterministic Markov decision processes. *Operations Research Letters*, pages 1–8, 2019.
23. J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:235–253, 1974/75.
24. J. Jacod. *Calcul stochastique et problèmes de martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.
25. A. Jaśkiewicz and A.S. Nowak. Zero-sum ergodic stochastic games with feller transition probabilities. *SIAM J. Control Optim.*, 45:773–789, 2006.
26. A. Piunovskiy and Y. Zhang. The transformation method for continuous-time Markov decision processes. *Journal of Optimization Theory and Applications*, 154(2):691–712, Aug 2012.
27. M. Schäl. Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32(3):179–196, 1975.
28. M. Schäl. On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.*, 3(4):345–364, 1975.
29. M. Schäl. Average optimality in dynamic programming with general state space. *Math. Oper. Res.*, 18(1):163–172, 1993.
30. M. Schäl. On piecewise deterministic Markov control processes: control of jumps and of risk processes in insurance. *Insurance Math. Econom.*, 22(1):75–91, 1998.
31. A.A. Yushkevich. On reducing a jump controllable Markov model to a model with discrete time. *Theory of Probability and Its Applications*, 25(1):58–69, 1980.
32. A.A. Yushkevich. Bellman inequalities in Markov decision deterministic drift processes. *Stochastics*, 23(1):25–77, 1987.
33. Y. Zhang. Average optimality for continuous-time Markov decision processes under weak continuity conditions. *J. Appl. Probab.*, 51(4):954–970, 2014.
34. Q. Zhu. Average optimality for continuous-time Markov decision processes with a policy iteration approach. *Journal of Mathematical Analysis and Applications*, 339(1):691 – 704, 2008.

Pathwise Approximations for the Solution of the Non-Linear Filtering Problem

Dan Crisan, Alexander Lobbe, and Salvador Ortiz-Latorre

Abstract We consider high order approximations of the solution of the stochastic filtering problem, derive their pathwise representation in the spirit of the earlier work of Clark [2] and Davis [10, 11] and prove their robustness property. In particular, we show that the high order discretised filtering functionals can be represented by Lipschitz continuous functions defined on the observation path space. This property is important from the practical point of view as it is in fact the pathwise version of the filtering functional that is sought in numerical applications. Moreover, the pathwise viewpoint will be a stepping stone into the rigorous development of machine learning methods for the filtering problem. This work is a cotinuation of [5] where a discretisation of the solution of the filtering problem of arbitrary order has been established. We expand the work in [5] by showing that robust approximations can be derived from the discretisations therein.

1 Introduction

With the present article on non-linear filtering we wish to honor the work of Mark H. A. Davis in particular to commemorate our great colleague. The topic of filtering is an area that has seen many excellent contributions by Mark. It is remarkable that

Dan Crisan
Department of Mathematics, Imperial College London, Huxley's Building, 180 Queen's Gate, London SW7 2AZ, UK,
e-mail: dcrisan@imperial.ac.uk

Alexander Lobbe
Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway,
e-mail: alexalob@math.uio.no

Salvador Ortiz-Latorre
Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway,
e-mail: salvadoo@math.uio.no

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_4

he was able to advance the understanding of non-linear filtering from a variety of angles. He considered many aspects of the field in his work, spanning the full range from the theory of the filtering equations to the numerical solution of the filtering problem via Monte-Carlo methods.

Mark Davis' work on filtering can be traced back to his doctoral thesis where he treats stochastic control of partially observable processes. The first article specifically on the topic of filtering that was co-authored by Mark appeared back in 1975 and considered a filtering problem with discontinuous observation process [12]. There, they used the so-called innovations method to compute the evolution of the conditional density of a process that is used to modulate the rate of a counting process. This method is nowadays well-known and is a standard way also to compute the linear (Kalman) filter explicitly. Early on in his career, Mark also contributed to the dissemination of filtering in the mathematics community with his monograph *Linear Estimation and Stochastic Control* [7], published in 1977, which deals with filtering to a significant degree. Moreover, his paper *An Introduction to Nonlinear Filtering* [9], written together with S. I. Marcus in 1981, has gained the status of a standard reference in the field.

Importantly, and in connection to the theme of the present paper, Mark has worked on computation and the robust filter already in 1980 [8]. Directly after the conception of the robust filter by Clark in 1978 [2], Mark took up the role of a driving figure in the subsequent development of robust, also known as pathwise, filtering theory [10, 11]. Here, he was instrumental in the development of the pathwise solution to the filtering equations with one-dimensional observation processes. Additionally, also correlated noise was already analysed in this work.

Robust filtering remains a highly relevant and challenging problem today. Some more recent work on this topic includes the article [6] which can be seen as an extension of the work by Mark, where correlated noise and a multidimensional observation process are considered. The work [4] is also worth mentioning in this context, as it establishes the validity of the robust filter rigorously.

Non-linear filtering is an important area within stochastic analysis and has numerous applications in a variety of different fields. For example, numerical weather prediction requires the solution of a high dimensional, non-linear filtering problem. Therefore, accurate and fast numerical algorithms for the approximate solution of the filtering problem are essential. In this contribution we analyse a recently developed high order time discretisation of the solution of the filtering problem from the literature [5] and prove that the so discretised solution possesses a property known as *robustness*. Thus, the present paper is a continuation of the previous work [5] by two of the authors which gives a new high-order time discretisation for the filtering functional. We extend this result to produce the robust version, of any order, of the discretisation from [5]. The implementation of the resulting numerical method remains open and is subject of future research. In subsequent work, the authors plan to deal with suitable extensions, notably a machine learning approach to pathwise filtering.

Robustness is a property that is especially important for the numerical approximation of the filtering problem in continuous time, since numerical observations can only be made in a discrete way. Here, the robustness property ensures that despite the discrete approximation, the solution obtained from it will still be a reasonable approximation of the true, continuous filter.

The present paper is organised as follows: In Section 2 we discuss the established theory leading up to the contribution of this paper. We introduce the stochastic filtering problem in sufficient generality in Subsection 2.1 whereafter the high order discretisation from the recent paper [5] is presented in Subsection 2.2 together with all the necessary notations. The Subsection 2.2 is concluded with the Theorem 1, taken from [5], which shows the validity of the high order discretisation and is the starting point for our contribution. Then, Section 3 serves to concisely present the main result of this work, which is Theorem 2 below. Our Theorem is a general result applying to corresponding discretisations of arbitrary order and shows that all of these discretisations do indeed assume a robust version. In Section 4 we present the proof of the main result in detail. The argument proceeds along the following lines. First, we establish the robust version of the discretisations for any order by means of a *formal* application of the integration by parts formula. In Lemma 1 we then show that the new robust approximation is locally bounded over the set of observation paths. Thereafter, Lemma 2 shows that the robustly discretised filtering functionals are locally Lipschitz continuous over the set of observation paths. Based on the elementary but important auxilliary Lemma 3 we use the path properties of the typical observation in Lemma 4 to get a version of the stochastic integral appearing in the robust approximation which is product measurable on the Borel sigma-algebra of the path space and the chosen filtration. Finally, after simplifying the arguments by lifting some of the random variables to an auxilliary copy of the probability space, we can show in Lemma 5 that, up to a null-set, the lifted stochastic integral appearing in the robust approximation is a random variable on the correct space. And subsequently, in Lemma 6 that the pathwise integral almost surely coincides with the standard stochastic integral of the observation process. The argument is concluded with Theorem 3 where we show that the robustly discretised filtering functional is a version of the high-order discretisation of the filtering functional as derived in the recent paper [5].

Our result in Theorem 2 can be interpreted as a remedy for some of the shortcomings of the earlier work [5] where the discretisation of the filter is viewed as a random variable and the dependence on the observation path is not made explicit. Here, we are correcting this in the sense that we give an interpretation of said random variable as a continuous function on path space. Our approach has two main advantages. Firstly, from a practitioner's point of view, it is exactly the path dependent version of the discretised solution that we are computing in numerical applications. Thus it is natural to consider it explicitly. The second advantage lies in the fact that here we are building a foundation for the theoretical development of machine learning approaches to the filtering problem which rely on the simulation of observation paths. With Theorem 2 we offer a first theoretical justification for this approach.

2 Preliminaries

Here, we begin by introducing the theory leading up to the main part of the paper which is presented in Sections 3 and 4.

2.1 The filtering problem

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a complete and right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$. We consider a $d_X \times d_Y$-dimensional partially observed system (X, Y) satisfying the system of stochastic integral equations

$$\begin{cases} X_t = X_0 + \int_0^t f(X_s)\,ds + \int_0^t \sigma(X_s)\,dV_s, \\ Y_t = \int_0^t h(X_s)\,ds + W_t, \end{cases} \tag{1}$$

where V and W are independent $(\mathcal{F}_t)_{t\geq 0}$-adapted d_V- and d_Y-dimensional standard Brownian motions, respectively. Further, X_0 is a random variable, independent of V and W, with distribution denoted by π_0. We assume that the coefficients

$$f = (f_i)_{i=1,\ldots,d_X} : \mathrm{R}^{d_X} \to \mathrm{R}^{d_X} \text{ and } \sigma = \left(\sigma_{i,j}\right)_{i=1,\ldots,d_X, j=1,\ldots,d_V} : \mathrm{R}^{d_X} \to \mathrm{R}^{d_X \times d_V}$$

of the *signal process* X are globally Lipschitz continuous and that the *sensor function*

$$h = (h_i)_{i=1,\ldots,d_Y} : \mathrm{R}^{d_X} \to \mathrm{R}^{d_Y}$$

is Borel-measurable and has linear growth. These conditions ensure that strong solutions to the system (1) exist and are almost surely unique. A central object in filtering theory is the *observation filtration* $\{\mathcal{Y}_t\}_{t\geq 0}$ that is defined as the augmentation of the filtration generated by the *observation process* Y, so that $\mathcal{Y}_t = \sigma(Y_s, s \in [0,t]) \vee \mathcal{N}$, where $\mathcal{N}$ are all P-null sets of $\mathcal{F}$.

In this context, non-linear filtering means that we are interested in determining, for all $t > 0$, the conditional law, called *filter* and denoted by π_t, of the signal X at time t given the information accumulated from observing Y on the interval $[0, t]$. Furthermore, this is equivalent to knowing for every bounded and Borel measurable function φ and every $t > 0$, the value of

$$\pi_t(\varphi) = \mathrm{E}[\varphi(X_t) \mid \mathcal{Y}_t].$$

A common approach to the non-linear filtering problem introduced above is via a change of probability measure. This approach is explained in detail in the monograph [1]. In summary, a probability measure $\tilde{P}$ is constructed that is absolutely continuous with respect to P and such that Y becomes a $\tilde{P}$-Brownian motion independent of X. Additionally, the law of X remains unchanged under $\tilde{P}$. The

Radon-Nikodym derivative of $\tilde{P}$ with respect to P is further given by the process Z that is given, for all $t \geq 0$, by

$$Z_t = \exp\left(\sum_{i=1}^{d_Y} \int_0^t h_i(X_s)\, dY_s^i - \frac{1}{2}\sum_{i=1}^{d_Y} \int_0^t h_i^2(X_s)\, ds\right).$$

Note that Z is an $(\mathcal{F}_t)_{t\geq 0}$-adapted martingale under $\tilde{P}$. This process is used in the definition of another, measure-valued process ρ that is given, for all bounded and Borel measurable functions φ and all $t \geq 0$, by

$$\rho_t(\varphi) = \tilde{\mathrm{E}}[\varphi(X_t)Z_t \mid \mathcal{Y}_t], \tag{2}$$

where we denote by $\tilde{\mathrm{E}}$ the expectation with respect to $\tilde{P}$. We call ρ the *unnormalised filter*, because it is related to the probability measure-valued process π through the Kallianpur-Striebel formula establishing that for all bounded Borel measurable functions φ and all $t \geq 0$ we have P-almost surely that

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(\mathbf{1})} = \frac{\tilde{\mathrm{E}}\left[\varphi(X_t)Z_t|\mathcal{Y}_t\right]}{\tilde{\mathrm{E}}\left[Z_t|\mathcal{Y}_t\right]} \tag{3}$$

where $\mathbf{1}$ is the constant function. Hence, the denominator $\rho_t(\mathbf{1})$ can be viewed as the normalising factor for π_t.

2.2 High order time discretisation of the filter

As shown by the Kallianpur-Striebel formula (3), $\pi_t(\varphi)$ is a ratio of two conditional expectations. In the recent paper [5] a high order time discretisation of these conditional expectations was introduced which leads further to a high order time discretisation of $\pi_t(\varphi)$. The idea behind this discretisation is summarised as follows.

First, for the sake of compactness, we augment the observation process as $\hat{Y}_t = (\hat{Y}_t^i)_{i=0}^{d_Y} = (t, Y_t^1, \ldots, Y_t^{d_Y})$ for all $t \geq 0$ and write

$$\hat{h} = \left(-\frac{1}{2}\sum_{i=1}^{d_Y} h_i^2, h_1, \ldots, h_{d_Y}\right).$$

Then, consider the *log-likelihood* process

$$\xi_t = \log(Z_t) = \sum_{i=0}^{d_Y} \int_0^t \hat{h}_i(X_s)\, \mathrm{d}\hat{Y}_s^i, \qquad t \geq 0. \tag{4}$$

Now, given a positive integer m, the order m time discretisation is achieved by a stochastic Taylor expansion up to order m of the processes $(\hat{h}_i(X_t))_{t\geq 0}$, $i = 0, \ldots, d_Y$ in (4). Finally, we substitute the discretised log-likelihood back into the original relationships (2) and the Kallianpur-Striebel formula (3) to obtain a discretisation of the filtering functionals. However, it is important to note that for the orders $m > 2$ an additional truncation procedure is needed, which we will make precise shortly, after introducing the necessary notation for the stochastic Taylor expansion.

2.2.1 Stochastic Taylor expansions

Let $\mathcal{M} = \{\alpha \in \{0,\ldots,d_V\}^l : l = 0,1,\ldots\}$ be the set of all multi-indices with range $\{0,\ldots,d_V\}$, where $\emptyset$ denotes the multi-index of length zero. For $\alpha = (\alpha_1,\ldots,\alpha_k) \in \mathcal{M}$ we adopt the notation $|\alpha| = k$ for its length, $|\alpha|_0 = \#\{j : \alpha_j = 0\}$ for the number of zeros in α, and $\alpha- = (\alpha_1,\ldots,\alpha_{k-1})$ and $-\alpha = (\alpha_2,\ldots,\alpha_k)$, for the right and left truncations, respectively. By convention $|\emptyset| = 0$ and $-\emptyset = \emptyset- = \emptyset$. Given two multi-indices $\alpha,\beta \in \mathcal{M}$ we denote their concatenation by $\alpha * \beta$. For positive and non-zero integers n and m, we will also consider the subsets of multi-indices

$$\mathcal{M}_{n,m} = \{\alpha \in \mathcal{M} : n \leq |\alpha| \leq m\}, \text{ and}$$
$$\mathcal{M}_m = \mathcal{M}_{m,m} = \{\alpha \in \mathcal{M} : |\alpha| = m\}.$$

For brevity, and by slight abuse of notation, we augment the Brownian motion V and now write $V = \left(V^i\right)_{i=0}^{d_V} = (t, V_t^1,\ldots,V_t^{d_V})$ for all $t \geq 0$. We will consider the filtration $\{\mathcal{F}_t^{0,V}\}_{t\geq 0}$ defined to be the usual augmentation of the filtration generated by the process V and initially enlarged with the random variable X_0. Moreover, for fixed $t \geq 0$, we will also consider the filtration $\{\mathcal{H}_s^t = \mathcal{F}_s^{0,V} \vee \mathcal{Y}_t\}_{s\leq t}$. For all $\alpha \in \mathcal{M}$ and all suitably integrable $\mathcal{H}_s^t$-adapted processes $\gamma = \{\gamma_s\}_{s\leq t}$ denote by $I_\alpha(\gamma_\cdot)_{s,t}$ the ItÃť iterated integral given for all $s \leq t$ by

$$I_\alpha(\gamma_\cdot)_{s,t} = \begin{cases} \gamma_t, & \text{if } |\alpha| = 0 \\ \displaystyle\int_s^t I_{\alpha-}(\gamma_\cdot)_{s,u}\, dV_u^{\alpha_{|\alpha|}}, & \text{if } |\alpha| \geq 1. \end{cases}$$

Based on the coefficient functions of the signal X, we introduce the differential operators L^0 and L^r, $r = 1,\ldots,d_V$, defined for all twice continuously differentiable functions $g : \mathrm{R}^{d_X} \to \mathrm{R}$ by

$$L^0 g = \sum_{k=1}^{d_X} f_k \frac{\partial g}{\partial x^k} + \frac{1}{2}\sum_{k,l=1}^{d_X}\sum_{r=1}^{d_V} \sigma_{k,r}\sigma_{l,r}\frac{\partial^2 g}{\partial x^k \partial x^l} \quad \text{and}$$
$$L^r g = \sum_{k=1}^{d_X} \sigma_{k,r}\frac{\partial g}{\partial x^k}, \quad r = 1,\ldots,d_V.$$

Lastly, for $\alpha = (\alpha_1,\ldots,\alpha_k) \in \mathcal{M}$, the differential operator L^α is defined to be the composition $L^\alpha = L^{\alpha_1} \circ \cdots \circ L^{\alpha_k}$, where, by convention, $L^\emptyset g = g$.

2.2.2 Discretisation of the log-likelihood process

With the stochastic Taylor expansion at hand, we can now describe the discretisation of the log-likelihood in (4). To this end, let for all $t > 0$,

$$\Pi(t) = \{\{t_0, \dots, t_n\} \subset [0,t]^{n+1} : 0 = t_0 < t_1 < \dots < t_n = t,\ n = 1, 2, \dots\}$$

be the set of all partitions of the interval $[0,t]$. For a given partition we call the quantity $\delta = \max\{t_{j+1} - t_j : j = 0, \dots, n-1\}$ the *meshsize* of τ. Then we discretise the log-likelihood as follows. For all $t > 0$, $\tau \in \Pi(t)$ and all positive integers m we consider

$$\begin{aligned}\xi_t^{\tau,m} &= \sum_{j=0}^{n-1} \xi_t^{\tau,m}(j) = \sum_{j=0}^{n-1} \sum_{i=0}^{d_Y} \sum_{\alpha \in \mathcal{M}_{0,m-1}} L^\alpha \hat{h}_i(X_{t_j}) \int_{t_j}^{t_{j+1}} I_\alpha(\mathbf{1})_{t_j,s} d\hat{Y}_s^i \\ &= \sum_{j=0}^{n-1} \{\kappa_j^{0,m} + \int_{t_j}^{t_{j+1}} \langle \eta_j^{0,m}(s), \mathrm{d}Y_s \rangle\},\end{aligned}$$

where we define for all integers $l \le m-1$ and $j = 0, \dots, n-1$ the quantities

$$\begin{aligned}\kappa^{l,m} &= \sum_{j=0}^{n-1} \kappa_j^{l,m} = \sum_{j=0}^{n-1} \Big\{ -\frac{1}{2} \sum_{\alpha \in \mathcal{M}_{l,m-1}} L^\alpha \langle h(\cdot), h(\cdot) \rangle (X_{t_j}) \int_{t_j}^{t_{j+1}} I_\alpha(\mathbf{1})_{t_j,s}\, \mathrm{d}s \Big\} \\ \eta_j^{l,m}(s) &= \Big(\sum_{\alpha \in \mathcal{M}_{l,m-1}} L^\alpha h_i(X_{t_j}) I_\alpha(\mathbf{1})_{t_j,s} \Big)_{i=1,\dots,d_Y}.\end{aligned}$$

and $\langle \cdot, \cdot \rangle$ denotes the euclidean inner product. Note that by setting, in the case of $m > 2$,

$$\begin{aligned}\mu^{\tau,m}(j) &= \sum_{i=0}^{d_Y} \sum_{\alpha \in \mathcal{M}_{2,m-1}} L^\alpha \hat{h}_i(X_{t_j}) \int_{t_j}^{t_{j+1}} I_\alpha(\mathbf{1})_{t_j,s} d\hat{Y}_s^i \\ &= \kappa_j^{2,m} + \int_{t_j}^{t_{j+1}} \langle \eta_j^{2,m}(s), \mathrm{d}Y_s \rangle,\end{aligned}$$

we may write the above as

$$\xi_t^{\tau,m} = \xi_t^{\tau,2} + \sum_{j=0}^{n-1} \mu^{\tau,m}(j).$$

As outlined before, the discretisations $\xi^{\tau,m}$ are obtained by replacing the processes $(\hat{h}_i(X_t))_{t \ge 0}$, $i = 0, \dots, d_Y$ in (4) with the truncation of degree $m-1$ of the corresponding stochastic Taylor expansion of $\hat{h}_i(X_t)$. These discretisations are subsequently used to obtain discretisation schemes of first and second order for the filter

$\pi_t(\varphi)$. However, they cannot be used directly to produce discretisation schemes of any order $m > 2$ because they do not have finite exponential moments (required to define the discretisation schemes). More precisely, the quantities $\mu^{\tau,m}(j)$ do not have finite exponential moments because of the high order iterated integral involved. For this, we need to introduce a truncation of $\mu^{\tau,m}(j)$ resulting in a (partial) taming procedure to the stochastic Taylor expansion of $(\hat{h}_i(X_t))_{t\geq 0}$. To achieve this, we introduce for every positive integer q and all $\delta > 0$ the truncation functions $\Gamma_{q,\delta}\colon \mathbb{R} \to \mathbb{R}$ such that

$$\Gamma_{q,\delta}(z) = \frac{z}{1 + (z/\delta)^{2q}} \tag{5}$$

and set, for all $j = 0, \ldots, n-1$,

$$\bar{\xi}_t^{\tau,m}(j) = \begin{cases} \xi_t^{\tau,m}(j), & \text{if } m = 1,2 \\ \xi_t^{\tau,2}(j) + \Gamma_{m,(t_{j+1}-t_j)}(\mu^{\tau,m}(j)), & \text{if } m > 2 \end{cases}.$$

Utilising the above, the truncated discretisations of the log-likelihood finally read

$$\bar{\xi}_t^{\tau,m} = \sum_{j=0}^{n-1} \bar{\xi}_t^{\tau,m}(j). \tag{6}$$

We end this section with a remark about the properties of the truncation function before we go on to discretising the filter.

Remark 1 The following two properties of the truncation function Γ, defined in (5), are readily checked. For all positive integers q and all $\delta > 0$ we have that

i) the truncation function is bounded, specifically, for all $z \in \mathbb{R}$,

$$|\Gamma_{q,\delta}(z)| \leq \frac{\delta}{(2q-1)^{1/2q}},$$

ii) and that its derivative is bounded for all $z \in \mathbb{R}$ as

$$\frac{q(1-q)-1}{2q} \leq \frac{\mathrm{d}}{\mathrm{d}z}\Gamma_{q,\delta}(z) \leq 1.$$

In particular, the truncation function is Lipschitz continuous.

2.2.3 Discretisation of the filter

Since $\bar{\xi}_t^{\tau,m}$ in (6) is a discretisation of the log-likelihood we will now consider, for all $t > 0$, $\tau \in \Pi(t)$ and all positive integers m, the discretised likelihood

$$Z_t^{\tau,m} = \exp\left(\bar{\xi}_t^{\tau,m}\right).$$

The filter is now discretised, under the condition that the Borel measurable function φ satisfies $\tilde{\mathrm{E}}[|\varphi(X_t)Z_t^{\tau,m}|] < \infty$, to the m-th order by

$$\rho_t^{\tau,m}(\varphi) = \tilde{\mathrm{E}}\left[\varphi(X_t)Z_t^{\tau,m} \mid \mathcal{Y}_t\right]$$

and

$$\pi_t^{\tau,m}(\varphi) = \frac{\rho_t^{\tau,m}(\varphi)}{\rho_t^{\tau,m}(\mathbf{1})}. \tag{7}$$

It remains to show that the achieved discretisation is indeed of order m.

2.2.4 Order of approximation for the filtering functionals

In the framework developed thus far, we can state the main result of [5] which justifies the construction and proves the high order approximation. To this end, we consider the L^p-norms $\|\cdot\|_{L^p} = \tilde{\mathrm{E}}[|\cdot|^p]^{1/p}$, $p \geq 1$.

Theorem 1 (Theorem 2.3 in [5])
Let m be a positive integer, let $t > 0$, let φ be an $(m+1)$-times continuously differentiable function with at most polynomial growth and assume further that the coefficients of the partially observed system (X, Y) in (1) satisfy that

- *f is bounded and $\max\{2, 2m-1\}$-times continuously differentiable with bounded derivatives,*
- *σ is bounded and $2m$-times continuously differentiable with bounded derivatives,*
- *h is bounded and $(2m+1)$-times continuously differentiable with bounded derivatives, and that*
- *X_0 has moments of all orders.*

Then there exist positive constants δ_0 and C, such that for all partitions $\tau \in \Pi(t)$ with meshsize $\delta < \delta_0$ we have that

$$\|\rho_t(\varphi) - \rho_t^{\tau,m}(\varphi)\|_{L^2} \leq C\delta^m.$$

Moreover, there exist positive constants $\bar{\delta}_0$ and $\bar{C}$, such that for all partitions $\tau \in \Pi(t)$ with meshsize $\delta < \bar{\delta}_0$,

$$\mathrm{E}\left[|\pi_t(\varphi) - \pi_t^{\tau,m}(\varphi)|\right] \leq \bar{C}\delta^m.$$

Remark 2 Under the above assumption that h is bounded and φ has at most polynomial growth, the required condition from Theorem 2.4 in [5] that there exists $\varepsilon > 0$ such that $\sup_{\{\tau\in\Pi(t):\delta<\delta_0\}} \|\pi_t^{\tau,m}(\varphi)\|_{L^{2+\varepsilon}} < \infty$ holds.

3 Robustness of the approximation

The classical robustness of the filter as in Theorem 5.12 in [1] states that for every $t > 0$ and bounded Borel measurable function φ the filter $\pi_t(\varphi)$ can be represented as a function of the observation *path*

$$Y_{[0,t]}(\omega) = \{Y_s(\omega) : s \in [0,t]\}, \qquad \omega \in \Omega.$$

In particular, $Y_{[0,t]}$ is here a path-valued random variable. The precise meaning of robustness is then that there exists a unique bounded Borel measurable function $F^{t,\varphi}$ on the path space $C([0,t];\mathbb{R}^{d_Y})$, that is the space of continuous $\mathbb{R}^{d_Y}$-valued functions on $[0,t]$, with the properties that

i) P-almost surely,
$$\pi_t(\varphi) = F^{t,\varphi}(Y_{[0,t]})$$
and
ii) $F^{t,\varphi}$ is continuous with respect to the supremum norm[1].

The volume [1] contains further details on the robust representation. In the present paper, we establish the analogous result for the discretised filter $\pi_t^{\tau,m}(\varphi)$ from (7). It is formulated as follows.

Theorem 2 *Let $t > 0$, $\tau = \{t_0, \dots, t_n\} \in \Pi(t)$, let m be a positive integer and let φ be a bounded Borel measurable function. Then there exists a function $F_\varphi^{\tau,m} : C([0,t];\mathbb{R}^{d_Y}) \to \mathbb{R}$ with the properties that*

i) P-almost surely,
$$\pi_t^{\tau,m}(\varphi) = F_\varphi^{\tau,m}(Y_{[0,t]})$$
and
ii) for every two bounded paths $y_1, y_2 \in C([0,t];\mathbb{R}^{d_Y})$ there exists a positive constant C such that
$$|F_\varphi^{\tau,m}(y_1) - F_\varphi^{\tau,m}(y_2)| \le C\|\varphi\|_\infty \|y_1 - y_2\|_\infty.$$

Note that Theorem 2 implies the following statement in the total variation norm.

Corollary 1 *Let $t > 0$, $\tau = \{t_0, \dots, t_n\} \in \Pi(t)$, and let m be a positive integer. Then, for every two bounded paths $y_1, y_2 \in C([0,t];\mathbb{R}^{d_Y})$ there exists a positive constant C such that*

$$\|\pi_t^{\tau,m,y_1} - \pi_t^{\tau,m,y_2}\|_{TV} = \sup_{\varphi \in B_b,\, \|\varphi\|_\infty \le 1} |F_\varphi^{\tau,m}(y_1) - F_\varphi^{\tau,m}(y_2)| \le C\|y_1 - y_2\|_\infty,$$

where B_b is the set of bounded and Borel measurable functions.

Remark 3 A natural question that arises in this context is to seek the rate of pathwise convergence of $F_\varphi^{\tau,m}$ to F_φ (defined as the limit of $F_\varphi^{\tau,m}$ when the meshsize goes to zero) as functions on the path space. The rate of pathwise convergence is expected to be dependent on the Hölder constant of the observation path. Therefore, it is expected to be not better than $\frac{1}{2} - \epsilon$ for a semimartingale observation. The absence of high order iterated integrals of the observation process in the construction of $F_\varphi^{\tau,m}$ means

[1] For a subset $D \subseteq \mathbb{R}^l$ and a function $\psi : D \to \mathbb{R}^d$ we set $\|\psi\|_\infty = \max_{i=1,\dots,d} \|\psi_i\|_\infty = \max_{i=1,\dots,d} \sup_{x \in D} |\psi_i(x)|$

that one cannot obtain *pathwise* high order approximations based on the work in [5]. Such approximations will no longer be continuous in the supremum norm. Thus we need to consider rough path norms in this context. In a different setting, Clark showed in the earlier paper [3] that one cannot construct pathwise approximations of solutions of SDEs by using only increments of the driving Brownian motion.

In the following and final part of the paper, we exhibit the proof of Theorem 2.

4 Proof of the robustness of the approximation

We begin by constructing what will be the robust representation. Consider, for all $y \in C([0,t];\mathrm{R}^{d_Y})$,

$$\begin{aligned}
\Xi_t^{\tau,m}(y) &= \sum_{j=0}^{n-1}\{\kappa_j^{0,m} + \langle\eta_j^{0,m}(t_{j+1}), y_{t_{j+1}}\rangle - \langle\eta_j^{0,m}(t_j), y_{t_j}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{0,m}(s)\rangle\} \\
&= \sum_{j=0}^{n-1}\{\kappa_j^{0,m} + \langle\eta_j^{0,m}(t_{j+1}), y_{t_{j+1}}\rangle - \langle h(X_{t_j}), y_{t_j}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{0,m}(s)\rangle\} \\
&= \langle h(X_{t_n}), y_{t_n}\rangle - \langle h(X_{t_0}), y_{t_0}\rangle \\
&+ \sum_{j=0}^{n-1}\{\kappa_j^{0,m} + \langle\eta_j^{0,m}(t_{j+1}) - h(X_{t_{j+1}}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{0,m}(s)\rangle\}
\end{aligned}$$

and further, for $m > 2$,

$$M_j^{\tau,m}(y) = \kappa_j^{2,m} + \langle\eta_j^{2,m}(t_{j+1}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, d\eta_j^{2,m}(s)\rangle$$

so that we can define

$$\bar{\Xi}_t^{\tau,m}(y) = \begin{cases} \Xi_t^{\tau,m}(y), & \text{if } \ m = 1,2 \\ \Xi_t^{\tau,2}(y) + \sum_{j=0}^{n-1} \Gamma_{m,(t_{j+1}-t_j)}(M_j^{\tau,m}(y)), & \text{if } \ m > 2 \end{cases}.$$

Furthermore, set

$$\mathcal{Z}_t^{\tau,m}(y) = \exp(\bar{\Xi}_t^{\tau,m}(y)).$$

Example 1 The robust approximation for $m = 1$ and $m = 2$ are given as follows. First, if $m = 1$, then

$$\begin{aligned}\Xi_t^{\tau,1}(y) &= \sum_{j=0}^{n-1}\{\kappa_j^{0,1} + \langle \eta_j^{0,1}(t_{j+1}), y_{t_{j+1}}\rangle - \langle h(X_{t_j}), y_{t_j}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{0,1}(s)\rangle\}\\ &= \sum_{j=0}^{n-1}\{-\frac{1}{2}\langle h,h\rangle(X_{t_j})(t_{j+1}-t_j) + \langle h(X_{t_j}), y_{t_{j+1}} - y_{t_j}\rangle\}\end{aligned}$$

and also $\bar{\Xi}_t^{\tau,1}(y) = \Xi_t^{\tau,1}(y)$ so that $\mathcal{Z}_t^{\tau,1}(y) = \exp(\Xi_t^{\tau,1}(y))$. If $m = 2$, then

$$\begin{aligned}\Xi_t^{\tau,2}(y) &= \Xi_t^{\tau,1}(y) + \sum_{j=0}^{n-1}\{\kappa_j^{1,2} + \langle \eta_j^{1,2}(t_{j+1}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{1,2}(s)\rangle\}\\ &= \Xi_t^{\tau,1}(y) - \sum_{\alpha\in\mathcal{M}_1}\sum_{j=0}^{n-1}\frac{1}{2}L^\alpha\langle h,h\rangle(X_{t_j})\int_{t_j}^{t_{j+1}} V_s^\alpha - V_{t_j}^\alpha\,\mathrm{d}s\\ &+ \sum_{\alpha\in\mathcal{M}_1}\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\langle L^\alpha h(X_{t_j}), y_{t_{j+1}} - y_s\rangle\,\mathrm{d}V_s^\alpha.\end{aligned}$$

Therefore, also $\bar{\Xi}_t^{\tau,2}(y) = \Xi_t^{\tau,2}(y)$ so that $\mathcal{Z}_t^{\tau,2}(y) = \exp(\Xi_t^{\tau,2}(y))$. □

First, we show that the newly constructed $\mathcal{Z}_t^{\tau,m}$ is locally bounded.

Lemma 1 *Let $t > 0$, let $\tau = \{t_0,\ldots,t_n\} \in \Pi(t)$ be a partition with mesh size δ and let m be a positive integer. Then, for all $R > 0$, $p \geq 1$ there exists a positive constant $B_{p,R}$ such that*

$$\sup_{\|y\|_\infty\leq R}\|\mathcal{Z}_t^{\tau,m}(y)\|_{L^p} \leq B_{p,R}.$$

Proof Notice that, by Remark 1, in the case $m \geq 2$, we have for all $y \in C([0,t];\mathrm{R}^{d_Y})$ that

$$\bar{\Xi}_t^{\tau,m}(y) \leq \Xi_t^{\tau,2}(y) + \frac{n\delta}{(2m-1)^{1/2m}}.$$

This implies that for all $y \in C([0,t];\mathrm{R}^{d_Y})$,

$$\mathcal{Z}_t^{\tau,m}(y) = \exp(\bar{\Xi}_t^{\tau,m}(y)) \leq \exp(\Xi_t^{\tau,2}(y))\exp\Big(\frac{n\delta}{(2m-1)^{1/2m}}\Big).$$

For $m = 1$, we clearly have $\mathcal{Z}_t^{\tau,1}(y) = \exp(\Xi_t^{\tau,1}(y))$. Hence, it suffices to show the result for $m = 1,2$ only. We have

$$\Xi_t^{\tau,2}(y) = \Xi_t^{\tau,1}(y) + \sum_{j=0}^{n-1}\{\kappa_j^{1,2} + \langle \eta_j^{1,2}(t_{j+1}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}}\langle y_s, \mathrm{d}\eta_j^{1,2}(s)\rangle\}.$$

Now, by the triangle inequality, boundedness of y, and boundedness of h, we get

$$|\Xi_t^{\tau,1}(y)| = \Big|\sum_{j=0}^{n-1}\{\kappa_j^{0,1} + \langle \eta_j^{0,1}(t_{j+1}), y_{t_{j+1}}\rangle - \langle h(X_{t_j}), y_{t_j}\rangle - \int_{t_j}^{t_{j+1}}\langle y_s, \mathrm{d}\eta_j^{0,1}(s)\rangle\}\Big|$$

$$
\begin{aligned}
&= \Big|\sum_{j=0}^{n-1}\{\kappa_j^{0,1} + \langle h(X_{t_j}), y_{t_{j+1}} - y_{t_j}\rangle\}\Big| \\
&= \Big|\sum_{j=0}^{n-1}\{-\frac{1}{2}\langle h(X_{t_j}), h(X_{t_j})\rangle(t_{j+1}-t_j) + \langle h(X_{t_j}), y_{t_{j+1}} - y_{t_j}\rangle\}\Big| \\
&\leq \frac{t d_Y \|h\|_\infty^2}{2} + 2R\|h\|_\infty = C_0,
\end{aligned}
$$

where we denote the final constant by C_0. Furthermore, by the triangle inequality, boundedness of y, and boundedness of h and its derivatives,

$$
\begin{aligned}
&\Big|\sum_{j=0}^{n-1}\{\kappa_j^{1,2} + \langle \eta_j^{1,2}(t_{j+1}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{1,2}(s)\rangle\}\Big| \\
&= \sum_{\alpha\in\mathcal{M}_1}\Bigg\{\Big|\sum_{j=0}^{n-1}\frac{1}{2}L^\alpha\langle h,h\rangle(X_{t_j})\int_{t_j}^{t_{j+1}} V_s^\alpha - V_{t_j}^\alpha\,\mathrm{d}s \\
&+ \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\langle L^\alpha h(X_{t_j}), y_{t_{j+1}} - y_s\rangle\,\mathrm{d}V_s^\alpha\Big|\Bigg\} \\
&\leq \sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\Bigg\{\Big|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\frac{1}{2}L^\alpha\langle h,h\rangle(X_{t_j})(t_{j+1}-s) + \langle L^\alpha h(X_{t_j}), y_{t_{j+1}} - y_s\rangle\,\mathrm{d}V_s^\alpha\Big|\Bigg\} \\
&+ \Big|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\frac{1}{2}L^0\langle h,h\rangle(X_{t_j})(s-t_j) + \langle L^0 h(X_{t_j}), y_{t_{j+1}} - y_s\rangle\,\mathrm{d}s\Big| \\
&\leq \sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\Bigg\{\Big|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\frac{1}{2}L^\alpha\langle h,h\rangle(X_{t_j})(t_{j+1}-s) + \langle L^\alpha h(X_{t_j}), y_{t_{j+1}} - y_s\rangle\,\mathrm{d}V_s^\alpha\Big|\Bigg\} \\
&+ \frac{1}{2}\delta t\|L^0\langle h,h\rangle\|_\infty + 2d_Y R t\|L^0 h\|_\infty \\
&= C_1 + \sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\Bigg\{\Big|\int_0^t \frac{1}{2}L^\alpha\langle h,h\rangle(X_{\lfloor s\rfloor})(\lceil s\rceil - s) + \langle L^\alpha h(X_{\lfloor s\rfloor}), y_{\lceil s\rceil} - y_s\rangle\,\mathrm{d}V_s^\alpha\Big|\Bigg\}.
\end{aligned}
$$

Here, C_1 is a constant introduced for conciseness. Then,

$$
\begin{aligned}
&\|\mathcal{Z}_t^{\tau,2}(y)\|_{L^p} \\
&= \Big\|\mathcal{Z}_t^{\tau,1}(y)\exp\Big(\sum_{j=0}^{n-1}\{\kappa_j^{1,2} - \langle\eta_j^{1,2}(t_{j+1}), y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}}\langle y_s, \mathrm{d}\eta_j^{1,2}(s)\rangle\}\Big)\Big\|_{L^p} \\
&\leq \exp(C_0 + C_1) \\
&\Bigg\|\exp\Big(\sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\Bigg\{\Big|\int_0^t \frac{1}{2}L^\alpha\langle h,h\rangle(X_{\lfloor s\rfloor})(\lceil s\rceil - s) + \langle L^\alpha h(X_{\lfloor s\rfloor}), y_{\lceil s\rceil} - y_s\rangle\,\mathrm{d}V_s^\alpha\Big|\Bigg\}\Big)\Bigg\|_{L^p}
\end{aligned}
$$

$< \infty.$

The lemma is thus proved. □

In analogy to the filter, we define the functions

$$G_{\varphi}^{\tau,m}(y) = \tilde{\mathrm{E}}[\varphi(X_t)\mathcal{Z}_t^{\tau,m}(y)]$$

and

$$F_{\varphi}^{\tau,m}(y) = \frac{G_{\varphi}^{\tau,m}(y)}{G_{\mathbf{1}}^{\tau,m}(y)} = \frac{\tilde{\mathrm{E}}[\varphi(X_t)\mathcal{Z}_t^{\tau,m}(y)]}{\tilde{\mathrm{E}}[\mathcal{Z}_t^{\tau,m}(y)]}.$$

Lemma 2 *Let $\tau \in \Pi(t)$ be a partition, let m be a positive integer and let φ be a bounded Borel measurable function. Then the functions $G_{\varphi}^{\tau,m} : C([0,t];\mathrm{R}^{d_Y}) \to \mathrm{R}$ and $F_{\varphi}^{\tau,m} : C([0,t];\mathrm{R}^{d_Y}) \to \mathrm{R}$ are locally Lipschitz continuous and locally bounded. Specifically, for every two paths $y_1, y_2 \in C([0,t];\mathrm{R}^{d_Y})$ such that there exists a real number $R > 0$ with $\|y_1\|_\infty \le R$ and $\|y_2\|_\infty \le R$, there exist constants L_G, M_G, L_F, and M_F such that*

$$|G_{\varphi}^{\tau,m}(y_1) - G_{\varphi}^{\tau,m}(y_2)| \le L_G \|\varphi\|_\infty \|y_1 - y_2\|_\infty \qquad \textit{and} \qquad |G_{\varphi}^{\tau,m}(y_1)| \le M_G \|\varphi\|_\infty$$

and

$$|F_{\varphi}^{\tau,m}(y_1) - F_{\varphi}^{\tau,m}(y_2)| \le L_F \|\varphi\|_\infty \|y_1 - y_2\|_\infty \qquad \textit{and} \qquad |F_{\varphi}^{\tau,m}(y_1)| \le M_F \|\varphi\|_\infty.$$

Proof We first show the results for $G_{\varphi}^{\tau,m}$. Note that

$$|\mathcal{Z}_t^{\tau,m}(y_1) - \mathcal{Z}_t^{\tau,m}(y_2)| \le (\mathcal{Z}_t^{\tau,m}(y_1) + \mathcal{Z}_t^{\tau,m}(y_2))|\bar{\Xi}_t^{\tau,m}(y_1) - \bar{\Xi}_t^{\tau,m}(y_2)|.$$

Then, by the Cauchy-Schwarz inequality, for all $p \ge 1$ we have

$$\|\varphi(X_t)\mathcal{Z}_t^{\tau,m}(y_1) - \varphi(X_t)\mathcal{Z}_t^{\tau,m}(y_2)\|_{L^p} \le 2B_{2p,R}\|\varphi\|_\infty \|\bar{\Xi}_t^{\tau,m}(y_1) - \bar{\Xi}_t^{\tau,m}(y_2)\|_{L^{2p}}. \tag{8}$$

Thus, for $m > 2$, we can exploit the effect of the truncation function and, similarly to the proof of Lemma 1, it suffices to show the result for $m = 1, 2$. To this end, consider for all $q \ge 1$,

$$\begin{aligned}\|\Xi_t^{\tau,2}(y_1) - \Xi_t^{\tau,2}(y_2)\|_{L^q} &\le \|\Xi_t^{\tau,1}(y_1) - \Xi_t^{\tau,1}(y_2)\|_{L^q} \\ &+ \left\| \sum_{j=0}^{n-1} \{\langle \eta_j^{1,2}(t_{j+1}), y_1(t_{j+1}) - y_2(t_{j+1})\rangle - \int_{t_j}^{t_{j+1}} \langle y_1(s) - y_2(s), \mathrm{d}\eta_j^{1,2}(s)\rangle\} \right\|_{L^q}.\end{aligned}$$

First, we obtain for all $q \ge 1$,

$$\|\Xi_t^{\tau,1}(y_1)-\Xi_t^{\tau,1}(y_2)\|_{L^q}=\left\|\sum_{j=0}^{n-1}\langle h(X_{t_j}),(y_1(t_{j+1})-y_2(t_{j+1}))-(y_1(t_j)-y_2(t_j))\rangle\right\|_{L^q}$$
$$\leq 2d_Y\|h\|_\infty\|y_1-y_2\|_\infty.$$

And second we have for all $q \geq 1$ that

$$\begin{aligned}
&\left\|\sum_{j=0}^{n-1}\{\int_{t_j}^{t_{j+1}}\langle(y_1(t_{j+1})-y_1(s))-(y_2(t_{j+1})-y_2(s)),\mathrm{d}\eta_j^{1,2}(s)\rangle\}\right\|_{L^q}\\
&\leq \sum_{j=0}^{n-1}\Bigg\|\left|\langle L^0h(X_{t_j}),y_1(t_{j+1})-y_2(t_{j+1})\rangle(t_{j+1}-t_j)\right|\\
&+\left|\int_{t_j}^{t_{j+1}}\langle L^0h(X_{t_j}),y_1(s)-y_2(s)\rangle\mathrm{d}s\right|\\
&+\sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\left|\langle L^\alpha h(X_{t_j}),y_1(t_{j+1})-y_2(t_{j+1})\rangle(V_{t_{j+1}}^\alpha-V_{t_j}^\alpha)\right|\\
&+\left|\int_{t_j}^{t_{j+1}}\langle L^\alpha h(X_{t_j}),y_1(s)-y_2(s)\rangle\mathrm{d}V_s^\alpha\right|\Bigg\|_{L^q}\\
&\leq\left[\bar{C}_1+\bar{C}_2\sum_{j=0}^{n-1}\sum_{\alpha\in\mathcal{M}_1\setminus\{0\}}\|V_{t_{j+1}}^\alpha-V_{t_j}^\alpha\|_{L^q}\right]\|y_1-y_2\|_\infty\\
&\leq C\|y_1-y_2\|_\infty
\end{aligned}$$

This and Lemma 1 imply that $G_\varphi^{\tau,m}$ is locally Lipschitz and locally bounded. To show the result for $F_\varphi^{\tau,m}$ we need to establish that $1/G_{\mathbf{1}}^{\tau,m}$ is locally bounded. We have, using Jensen's inequality, that for $m \geq 2$

$$G_{\mathbf{1}}^{\tau,m}=\tilde{\mathrm{E}}[\mathcal{Z}_t^{\tau,m}]\geq\exp\Big(\tilde{\mathrm{E}}[\bar{\Xi}_t^{\tau,m}]\Big)\geq\exp\Big(\tilde{\mathrm{E}}[\Xi_t^{\tau,2}]\Big)\exp\left(-\frac{n\delta}{(2m-1)^{1/2m}}\right)$$

and for $m = 1$ clearly

$$G_{\mathbf{1}}^{\tau,1}=\tilde{\mathrm{E}}[\mathcal{Z}_t^{\tau,1}]\geq\exp\Big(\tilde{\mathrm{E}}[\Xi_t^{\tau,1}]\Big).$$

Since the quantities $\tilde{\mathrm{E}}[\Xi_t^{\tau,1}]$ and $\tilde{\mathrm{E}}[\Xi_t^{\tau,2}]$ are finite, the lemma is proved. □

In the following, given $t > 0$, we set for every $\gamma \in (0, 1/2)$,

$$\mathcal{H}_\gamma=\left\{y\in C([0,t];\mathrm{R}^{d_Y}):\sup_{s_1,s_2\in[0,t]}\frac{\|y_{s_1}-y_{s_2}\|_\infty}{|s_1-s_2|^\gamma}<\infty\right\}\subseteq C([0,t];\mathrm{R}^{d_Y})$$

and recall that $Y_{[0,t]}:\Omega\to C([0,t];\mathrm{R}^{d_Y})$ denotes the random variable in path space corresponding to the observation process Y.

Lemma 3 *For all $t > 0$ and $\gamma \in (0, 1/2)$, we have $\tilde{P}$-almost surely that $Y_{[0,t]} \in \mathcal{H}_\gamma$.*

Proof Recall that, under $\tilde{P}$, the observation process Y is a Brownian motion and, by the Brownian scaling property, it suffices to show the result for $t = 1$. Therefore, let $\gamma \in (0, 1/2)$ and note that for all $\delta \in (0, 1]$ we have

$$\sup_{s_1,s_2\in[0,1]} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{|s_1 - s_2|^\gamma} = \max\left\{ \sup_{\substack{s_1,s_2\in[0,1]\\ |s_1-s_2|\le\delta}} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{|s_1 - s_2|^\gamma}, \sup_{\substack{s_1,s_2\in[0,1]\\ |s_1-s_2|\ge\delta}} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{|s_1 - s_2|^\gamma} \right\}.$$

The second element of the maximum above is easily bounded, $\tilde{P}$-almost surely, by the sample path continuity. For the first element, note that there exists $\delta_0 \in (0, 1)$ such that for all $\delta \in (0, \delta_0]$,

$$\delta^\gamma \ge \sqrt{2\delta \log(1/\delta)}.$$

Therefore, it follows that $\tilde{P}$-almost surely,

$$\sup_{\substack{s_1,s_2\in[0,1]\\ |s_1-s_2|\le\delta_0}} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{|s_1 - s_2|^\gamma} \le \sup_{\substack{s_1,s_2\in[0,1]\\ |s_1-s_2|\le\delta_0}} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{\sqrt{2|s_1 - s_2| \log(1/|s_1 - s_2|)}}.$$

The Lévy modulus of continuity of Brownian motion further ensures that $\tilde{P}$-almost surely,

$$\limsup_{\delta\downarrow 0} \sup_{\substack{s_1,s_2\in[0,1]\\ |s_1-s_2|\le\delta}} \frac{\|Y_{s_1} - Y_{s_2}\|_\infty}{\sqrt{2\delta \log(1/\delta)}} = 1.$$

The Lemma 3 thus follows. □

Lemma 4 *Let $\tau = \{0 = t_1 < \ldots < t_n = t\} \in \Pi(t)$ be a partition, let $j \in \{0, \ldots, n-1\}$ and let c be a positive integer. Then, there exists a version of the stochastic integral*

$$C([0,t]; \mathbb{R}^{d_Y}) \times \Omega \ni (y, \omega) \mapsto \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{c,c+1}(s, \omega)\rangle \in \mathbb{R}$$

such that it is equal on $\mathcal{H}_\gamma \times \Omega$ to a $\mathcal{B}(C([0,t]; \mathbb{R}^{d_Y})) \times \mathcal{F}$-measurable mapping.

Proof For k a positive integer, define for $y \in C([0,t]; \mathbb{R}^{d_Y})$,

$$\mathcal{J}_j^{c,k}(y) = \sum_{i=0}^{k-1} \Big\langle y_{s_{i,j}}, \Big(\eta_j^{c,c+1}(s_{i+1,j}) - \eta_j^{c,c+1}(s_{i,j})\Big)\Big\rangle,$$

where $s_{i,j} = \frac{i(t_{j+1}-t_j)}{k} + t_j$, $i = 0, \ldots, k$. Furthermore, we set $\lfloor s \rfloor = s_{i,j}$ for $s \in [\frac{i(t_{j+1}-t_j)}{k} + t_j, \frac{(i+1)(t_{j+1}-t_j)}{k} + t_j)$. Then, for $y \in \mathcal{H}_\gamma$, we have

$$\tilde{\mathrm{E}}\Big[\Big(\mathcal{J}_j^{c,2^l}(y) - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{c,c+1}(s)\rangle\Big)^2\Big]$$

$$
\begin{aligned}
&= \tilde{\mathrm{E}}\Big[\Big(\int_{t_j}^{t_{j+1}} \langle y_{\lfloor s\rfloor} - y_s, \mathrm{d}\eta_j^{c,c+1}(s)\rangle\Big)^2\Big] \\
&= \tilde{\mathrm{E}}\Big[\Big(\sum_{\alpha\in\mathcal{M}_c}\sum_{i=0}^{d_Y}\int_{t_j}^{t_{j+1}} (y^i_{\lfloor s\rfloor} - y^i_s) L^\alpha h_i(X_{t_j})\,\mathrm{d}I_\alpha(\mathbf{1})_{t_j,s}\Big)^2\Big] \\
&\le (d_V+1)d_Y \sum_{i=0}^{d_Y}\sum_{\alpha\in\mathcal{M}_c} \tilde{\mathrm{E}}\Big[\Big(\int_{t_j}^{t_{j+1}} (y^i_{\lfloor s\rfloor} - y^i_s) L^\alpha h_i(X_{t_j})\,\mathrm{d}I_\alpha(\mathbf{1})_{t_j,s}\Big)^2\Big] \\
&= (d_V+1)d_Y \sum_{i=0}^{d_Y}\sum_{\substack{\alpha\in\mathcal{M}_c\\ \alpha_{|\alpha|}\neq 0}} \tilde{\mathrm{E}}\Big[\int_{t_j}^{t_{j+1}} ((y^i_{\lfloor s\rfloor} - y^i_s) L^\alpha h_i(X_{t_j}))^2\,\mathrm{d}\langle I_\alpha(\mathbf{1})_{t_j,\cdot}\rangle_s\Big] \\
&+ (d_V+1)d_Y \sum_{i=0}^{d_Y}\sum_{\substack{\alpha\in\mathcal{M}_c\\ \alpha_{|\alpha|}= 0}} \tilde{\mathrm{E}}\Big[\Big(\int_{t_j}^{t_{j+1}} (y^i_{\lfloor s\rfloor} - y^i_s) L^\alpha h_i(X_{t_j})\,\mathrm{d}\Big[\int_{t_j}^{s} I_{\alpha-}(\mathbf{1})_{t_j,r}\mathrm{d}r\Big]\Big)^2\Big] \\
&\le (d_V+1)d_Y \frac{K(t_{j+1}-t_j)^{2\gamma}}{2^{2l\gamma}} \max_{\alpha\in\mathcal{M}_c}\|L^\alpha h(X_{t_j})\|_\infty \Big\{\sum_{i=0}^{d_Y}\sum_{\substack{\alpha\in\mathcal{M}_c\\ \alpha_{|\alpha|}\neq 0}} \tilde{\mathrm{E}}\Big[\int_{t_j}^{t_{j+1}} (I_{\alpha-}(\mathbf{1})_{t_j,s})^2\,\mathrm{d}s\Big] \\
&+ \sum_{i=0}^{d_Y}\sum_{\substack{\alpha\in\mathcal{M}_c\\ \alpha_{|\alpha|}= 0}} \tilde{\mathrm{E}}\Big[\Big(\int_{t_j}^{t_{j+1}} I_{\alpha-}(\mathbf{1})_{t_j,s}\,\mathrm{d}s\Big)^2\Big]\Big\} \\
&\le \frac{(d_V+1)d_Y C K(t_{j+1}-t_j)^{2\gamma}}{2^{2l\gamma}},
\end{aligned}
$$

Where the constant C is independent of l. Thus, by Chebyshev's inequality, we get for all $\epsilon > 0$ that

$$
\tilde{P}\Big(\Big|\mathcal{J}_j^{c,2^l}(y) - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{c,c+1}(s)\rangle\Big| > \epsilon\Big) \le \frac{1}{\epsilon^2}\frac{(d_V+1)d_Y C K(t_{j+1}-t_j)^{2\gamma}}{2^{2l\gamma}}.
$$

However, the bound on the right-hand side is summable over l so that we conclude using the first Borel-Cantelli Lemma that, for all $\epsilon > 0$,

$$
\tilde{P}\Big(\limsup_{l\to\infty}\Big|\mathcal{J}_j^{c,2^l}(y) - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\eta_j^{c,c+1}(s)\rangle\Big| > \epsilon\Big) = 0.
$$

Thus, for all $y \in \mathcal{H}_\gamma$, the integral $\mathcal{J}_j^{c,k}(y)$ converges $\tilde{P}$-almost surely to the integral $\int_{t_j}^{t_{j+1}}\langle y_s, \mathrm{d}\eta_j^{c,c+1}(s)\rangle$. Hence, we can define the limit on $\mathcal{H}_\gamma \times \Omega$ to be

$$
\mathcal{J}_j^{c}(y)(\omega) = \limsup_{l\to\infty}\mathcal{J}_j^{c,l}(y)(\omega); \qquad (y,\omega)\in\mathcal{H}_\gamma\times\Omega.
$$

Since the mapping

$$C([0,T];\mathrm{R}^{d_Y}) \times \Omega \ni (y,\omega) \mapsto \limsup_{l\to\infty} \mathcal{J}_j^{c,l}(y)(\omega) \in \mathrm{R}$$

is jointly $\mathcal{B}(C([0,T];\mathrm{R}^{d_Y})) \otimes \mathcal{F}$ measurable the lemma is proved. □

It turns out that proving the robustness result is simplified by first decoupling the processes X and Y in the following manner. Let $(\mathring{\Omega}, \mathring{\mathcal{F}}, \mathring{P})$ be an indentical copy of the probability space $(\Omega, \mathcal{F}, \tilde{P})$. Then

$$\mathring{G}_{\varphi}^{\tau,m}(y) = \mathring{\mathrm{E}}[\varphi(\mathring{X}_t)\mathring{\mathcal{Z}}_t^{\tau,m}(y)]$$

is the corresponding representation of $G_{\varphi}^{\tau,m}(y)$ in the new space, where $\mathring{\mathcal{Z}}_t^{\tau,m}(y) = \exp(\mathring{\bar{\Xi}}_t^{\tau,m}(y))$ with

$$\mathring{\Xi}_t^{\tau,m}(y) = \sum_{j=0}^{n-1} \mathring{\kappa}_j^{0,m} + \langle \mathring{\eta}_j^{0,m}(t_{j+1}), y_{t_{j+1}} \rangle - \langle h(\mathring{X}_{t_j}), y_{t_j} \rangle - \int_{t_j}^{t_{j+1}} \langle y_s, \mathrm{d}\mathring{\eta}_j^{0,m}(s) \rangle$$

and, for $m > 2$,

$$\mathring{M}_j^{\tau,m}(y) = \mathring{\kappa}_j^{2,m} - \langle \mathring{\eta}_j^{2,m}(t_{j+1}), y_{t_{j+1}} \rangle - \int_{t_j}^{t_{j+1}} \langle y_s, d\mathring{\eta}_j^{2,m}(s) \rangle,$$

so that, finally,

$$\mathring{\bar{\Xi}}_t^{\tau,m}(y) = \begin{cases} \mathring{\Xi}_t^{\tau,m}(y), & \text{if } m = 1,2 \\ \mathring{\Xi}_t^{\tau,2}(y) + \sum_{j=0}^{n-1} \Gamma_{m,(t_{j+1}-t_j)}(\mathring{M}_j^{\tau,m}(y)), & \text{if } m > 2. \end{cases}$$

Moreover, with $\mathring{\mathcal{J}}_j^c(y)$ corresponding to Lemma 4 we can write for $y \in \mathcal{H}_\gamma$,

$$\begin{aligned} \mathring{\Xi}_t^{\tau,m}(y) = & \sum_{j=0}^{n-1} \mathring{\kappa}_j^{0,m} + \langle \mathring{\eta}_j^{0,m}(t_{j+1}), y_{t_{j+1}} \rangle - \langle h(\mathring{X}_{t_j}), y_{t_j} \rangle \\ & - \sum_{c=0}^{m-1} \sum_{j=0}^{n-1} \mathring{\mathcal{J}}_j^c(y). \end{aligned}$$

In the same way we get, *mutatis mutandis*, the expression for $\mathring{\bar{\Xi}}_t^{\tau,m}(y)$ on $\mathcal{H}_\gamma$. Now, we denote by

$$(\check{\Omega}, \check{\mathcal{F}}, \check{P}) = (\Omega \times \mathring{\Omega}, \mathcal{F} \otimes \mathring{\mathcal{F}}, \tilde{P} \otimes \mathring{P})$$

the product probability space. In the following we lift the processes $\mathring{\eta}$ and Y from the component spaces to the product space by writing $Y(\omega, \mathring{\omega}) = Y(\omega)$ and $\mathring{\eta}_j^{c,c+1}(\omega, \mathring{\omega}) = \mathring{\eta}_j^{c,c+1}(\mathring{\omega})$ for all $(\omega, \mathring{\omega}) \in \check{\Omega}$.

Lemma 5 *Let c be a positive integer and let $j \in \{0,\dots,n\}$. Then there exists a nullset $N_0 \in \mathcal{F}$ such that the mapping $(\omega,\mathring{\omega}) \mapsto \mathring{\mathcal{J}}_j^c(Y_{[0,t]}(\omega))(\mathring{\omega})$ coincides on $(\Omega \setminus N_0)\times\mathring{\Omega}$ with an $\check{\mathcal{F}}$-measurable map.*

Proof Notice first that the set

$$N_0 = \{\omega \in \Omega\colon\ Y_{[0,t]}(\omega) \notin \mathcal{H}_\gamma\}$$

is clearly a member of $\mathcal{F}$ and we have that $\tilde{P}(N_0) = 0$. With N_0 so defined, the lemma follows from the definition and measurability of $(\omega,\mathring{\omega}) \mapsto \mathring{\mathcal{J}}_j^c(Y_{[0,t]}(\omega))(\mathring{\omega})$. □

Lemma 6 *Let c be a positive integer and $j \in \{0,\dots,n\}$. Then we have $\check{P}$-almost surely that*

$$\int_{t_j}^{t_{j+1}} \langle Y_s, \mathrm{d}\mathring{\eta}_j^{c,c+1}(s)\rangle = \mathring{\mathcal{J}}_j^c(Y_{[0,t]}).$$

Proof Note that we can assume without loss of generality that $d_Y = 1$ because the result follows componentwise. Then, let $K > 0$ and $T = \inf\{s \in [0,t]\colon |Y_s| \le K\}$ to define

$$Y_s^K = Y_s\mathbb{I}_{s\le T} + Y_T\mathbb{I}_{s>T}\ ; \qquad s \in [0,t].$$

Then Fubini's theorem and Lemma 5 imply that

$$\check{\mathrm{E}}\left[\left(\sum_{i=0}^{k-1} Y_{s_{i,j}}^K\left(\mathring{\eta}_j^{c,c+1}(s_{i+1,j}) - \mathring{\eta}_j^{c,c+1}(s_{i,j})\right) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega))\right)^2\right]$$
$$= \int_{\Omega\setminus N_0} \mathring{\mathrm{E}}[(\mathring{\mathcal{J}}_j^{c,k}(Y_{[0,t]}^K(\omega)) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega)))^2]\,\mathrm{d}\tilde{P}(\omega)$$

Now, since the function $s \mapsto Y_s^K(\omega)$ is continuous and $\mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega))$ is a version of the integral $\int_{t_j}^{t_{j+1}} Y_s^K(\omega)\,\mathrm{d}\mathring{\eta}_j^{c,c+1}(s)$ we have for every $\omega \in \Omega \setminus N_0$ that

$$\lim_{k\to\infty} \mathring{\mathrm{E}}[(\mathring{\mathcal{J}}_j^{c,k}(Y_{[0,t]}^K(\omega)) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega)))^2] = 0.$$

Moreover, clearly,

$$\mathring{\mathrm{E}}[(\mathring{\mathcal{J}}_j^{c,k}(Y_{[0,t]}^K(\omega)) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega)))^2] \le 4K^2\mathring{\mathrm{E}}[\mathring{\eta}_t^2] < \infty$$

So that we can conclude by the dominated convergence theorem that

$$\lim_{k\to\infty} \check{\mathrm{E}}\left[\left(\sum_{i=0}^{k-1} Y_{s_{i,j}}^K\left(\eta_j^{c,c+1}(s_{i+1,j}) - \eta_j^{c,c+1}(s_{i,j})\right) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega))\right)^2\right]$$
$$= \int_{\Omega\setminus N_0} \lim_{k\to\infty} \mathring{\mathrm{E}}[(\mathring{\mathcal{J}}_j^{c,k}(Y_{[0,t]}^K(\omega)) - \mathring{\mathcal{J}}_j^c(Y_{[0,t]}^K(\omega)))^2]\,\mathrm{d}\tilde{P}(\omega) = 0$$

As K is arbitrary, the lemma is proved. □

Finally, we are ready to show the main result, Theorem 2. We restate it here again, in a slightly different manner which reflects the current line of argument.

Theorem 3 *The random variable* $F_\varphi^{\tau,m}(Y_{[0,t]})$ *is a version of* $\pi_t^{\tau,m}(\varphi)$.

Proof By the Kallianpur-Striebel formula it suffices to show that for all bounded and Borel measurable functions φ we have $\tilde{P}$-almost surely

$$\rho_t^{\tau,m}(\varphi) = G_\varphi^{\tau,m}(Y_{[0,t]}).$$

Furthermore, this is equivalent to showing that for all continuous and bounded functions $b\colon C([0,t];\mathrm{R}^{d_Y}) \to \mathrm{R}$ the equality

$$\tilde{\mathrm{E}}[\rho_t^{\tau,m}(\varphi)b(Y_{[0,t]})] = \tilde{\mathrm{E}}[G_\varphi^{\tau,m}(Y_{[0,t]})b(Y_{[0,t]})].$$

holds. As for the left-hand side we can write

$$\begin{aligned}
&\tilde{\mathrm{E}}[\rho_t^{\tau,m}(\varphi)b(Y_{[0,t]})]\\
&= \tilde{\mathrm{E}}[\varphi(X_t)Z_t^{\tau,m}b(Y_{[0,t]})]\\
&= \tilde{\mathrm{E}}[\varphi(X_t)\exp(\bar{\xi}_t^{\tau,m})b(Y_{[0,t]})]\\
&= \check{\mathrm{E}}[\varphi(\mathring{X}_t)\exp(\mathring{\bar{\xi}}_t^{\tau,m})b(Y_{[0,t]})]\\
&= \check{\mathrm{E}}[\varphi(\mathring{X}_t)\exp(\mathtt{IBP}(\mathring{\bar{\xi}}_t^{\tau,m}))b(Y_{[0,t]})]
\end{aligned}$$

where $\mathtt{IBP}(\mathring{\bar{\xi}}_t^{\tau,m})$ is given by the application of the integration by parts formula for semimartingales as

$$\begin{aligned}
\mathtt{IBP}(\mathring{\xi}_t^{\tau,m}) &= \sum_{j=0}^{n-1} \mathtt{IBP}(\mathring{\xi}_t^{\tau,m})(j)\\
&= \sum_{j=0}^{n-1} \{\mathring{\kappa}_j^{0,m} + \langle \mathring{\eta}_j^{0,m}(t_{j+1}), Y_{t_{j+1}}\rangle - \langle h(\mathring{X}_{t_j}), Y_{t_j}\rangle - \int_{t_j}^{t_{j+1}} \langle Y_s, \mathrm{d}\mathring{\eta}_j^{0,m}(s)\rangle\}
\end{aligned}$$

$$\mathtt{IBP}(\mathring{\mu}^{\tau,m})(j) = \mathring{\kappa}_j^{2,m} + \langle \mathring{\eta}_j^{2,m}(t_{j+1}), Y_{t_{j+1}}\rangle - \int_{t_j}^{t_{j+1}} \langle Y_s, \mathrm{d}\mathring{\eta}_j^{2,m}(s)\rangle$$

$$\mathtt{IBP}(\mathring{\bar{\xi}}_t^{\tau,m})(j) = \begin{cases} \mathtt{IBP}(\mathring{\xi}_t^{\tau,m})(j), & \text{if } m = 1,2\\ \mathtt{IBP}(\mathring{\xi}_t^{\tau,m})(j) + \Gamma_{m,(t_{j+1}-t_j)}\left(\mathtt{IBP}(\mathring{\mu}^{\tau,m})(j)\right), & \text{if } m > 2 \end{cases}.$$

And, on the other hand, the right-hand side is

$$
\begin{aligned}
&\tilde{\mathrm{E}}[G_{\varphi}^{\tau,m}(Y_{[0,t]})b(Y_{[0,t]})] \\
&= \tilde{\mathrm{E}}[\varphi(X_t)\mathcal{Z}_t^{\tau,m}(Y_{[0,t]})b(Y_{[0,t]})] \\
&= \tilde{\mathrm{E}}[\varphi(X_t)\exp(\bar{\Xi}_t^{\tau,m}(Y_{[0,t]}))b(Y_{[0,t]})] \\
&= \tilde{\mathrm{E}}[\mathring{\mathrm{E}}[\varphi(\mathring{X}_t)\exp(\mathring{\bar{\Xi}}_t^{\tau,m}(Y_{[0,t]}))]b(Y_{[0,t]})] \\
&= \check{\mathrm{E}}[\varphi(\mathring{X}_t)\exp(\mathring{\bar{\Xi}}_t^{\tau,m}(Y_{[0,t]}))b(Y_{[0,t]})],
\end{aligned}
$$

where the last equality follows from Fubini's theorem. As the representations coincide, the theorem is thus proved. □

Acknowledgements Part of this research was funded within the project STORM: Stochastics for Time-Space Risk Models, from the Research Council of Norway (RCN). Project number: 274410.

References

1. Bain, A., Crisan, D.: Fundamentals of stochastic filtering. Springer, Stochastic modelling and applied probability 60 (2009)
2. Clark, J. M. C.: The design of robust approximations to the stochastic differential equations of nonlinear filtering. In J.K. Skwirzynski, editor, Communication systems and random processes theory, volume 25 of Proc. 2nd NATO Advanced Study Inst. Ser. E, Appl. Sci., pages 721-734. Sijthoff & Noordhoff, Alphen aan den Rijn (1978)
3. Clark, J.M.C., Cameron, R.J.: The maximum rate of convergence of discrete approximations for stochastic differential equations. In Stochastic differential systems, Lecture Notes in Control and Information Science 25, pages 162–171, Springer, Berlin-New York (1980)
4. Clark, J., Crisan, D.: On a robust version of the integral representation formula of nonlinear filtering. In Probab. Theory Relat. Fields 133, 43âĂŞ56 (2005). https://doi.org/10.1007/s00440-004-0412-5
5. Crisan, D., Ortiz-Latorre, S.: A high order time discretization of the solution of the non-linear filtering problem. Stoch PDE: Anal Comp (2019) doi: 10.1007/s40072-019-00157-3
6. Crisan, D., Diehl, J., Friz, P. K., Oberhauser, H.: Robust filtering: Correlated noise and multidimensional observation. In Ann. Appl. Probab. 23 (2013), no. 5, 2139–2160. doi:10.1214/12-AAP896.
7. Davis, M.H.A.: Linear estimation and stochastic control. Chapman and Hall, London (1977)
8. Davis, M.H.A., Wellings, P.H.: Computational problems in nonlinear filtering. In Analysis and Optimization of Systems, Springer (1980)
9. Davis, M.H.A., Marcus S.I.: An Introduction to Nonlinear Filtering. In Hazewinkel M., Willems J.C. (eds) Stochastic Systems - The Mathematics of Filtering and Identification and Applications. NATO Advanced Study Institutes Series (Series C - Mathematical and Physical Sciences), vol 78. Springer, Dordrecht. doi: 10.1007/978-94-009-8546-9_4 (1981)
10. Davis, M.H.A.: A Pathwise Solution of the Equations of Nonlinear Filtering. In Theory of Probability and its Applications, 27:1, pages 167-175 (1982)
11. Davis, M.H.A., Spathopoulos, M. P.: Pathwise Nonlinear Filtering for Nondegenerate Diffusions with Noise Correlation. In SIAM Journal on Control and Optimization, 25:2, pages 260-278 (1987)
12. Segall, A., Davis, M.H.A., Kailath, T.: Nonlinear filtering with counting observations. In IEEE Transactions on Information Theory, volume 21, no. 2, pages 143-149. doi: 10.1109/TIT.1975.1055360. (1975)

Discrete-Time Portfolio Optimization under Maximum Drawdown Constraint with Partial Information and Deep Learning Resolution

Carmine de Franco, Johann Nicolle, and Huyên Pham

In Memory of Mark H Davis

Abstract We study a discrete-time portfolio selection problem with partial information and maximum drawdown constraint. Drift uncertainty in the multidimensional framework is modeled by a prior probability distribution. In this Bayesian framework, we derive the dynamic programming equation using an appropriate change of measure, and obtain semi-explicit results in the Gaussian case. The latter case, with a CRRA utility function is completely solved numerically using recent deep learning techniques for stochastic optimal control problems. We emphasize the informative value of the learning strategy versus the non-learning one by providing empirical performance and sensitivity analysis with respect to the uncertainty of the drift. Furthermore, we show numerical evidence of the close relationship between the non-learning strategy and a no short-sale constrained Merton problem, by illustrating the convergence of the former towards the latter as the maximum drawdown constraint vanishes.

1 Introduction

This paper is devoted to the study of a constrained allocation problem in discrete time with partial information. We consider an investor who is willing to maximize the expected utility of her terminal wealth over a given investment horizon. The

Carmine de Franco
OSSIAM, 80, Avenue de la Grande Armée 75017 Paris, France,
e-mail: carmine.de-franco@ossiam.com

Johann Nicolle
LPSM-OSSIAM, 80, Avenue de la Grande Armée 75017 Paris, France,
e-mail: johann.nicolle@ossiam.com

Huyên Pham
Université de Paris, Bâtiment Sophie Germain, Case courrier 7012, 75205 Paris Cedex 13, France,
e-mail: pham@lpsm.paris

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_5

risk-averse investor is looking for the optimal portfolio in financial assets under a maximum drawdown constraint. The maximum drawdown is a common metric in finance and represents the largest drop in the portfolio value. Our framework incorporates this constraint by setting a threshold representing the proportion of the current maximum of the wealth process that the investor is willing to keep.

The expected rate of assets' return (drift) is unknown, but information can be learnt by progressive observation of the financial asset prices. The uncertainty about the rate of return is modeled by a probability distribution, i.e., a prior belief on the drift. To take into account the information conveyed by the prices, this prior will be updated using a Bayesian learning approach.

An extensive literature exists on parameters uncertainty and especially on filtering and learning techniques in a partial information framework. To cite just a few, see [18], [20], [5], [16], [2], and [6]. Somme articles deal with risk constraints in a portfolio allocation framework. For instance, paper [19] tackles dynamic risk constraints and compares the continuous and discrete time trading while some papers especially focus on drawdown constraints, see in particular seminal paper [11] or [4]. More recently, the authors of [8] study infinite-horizon optimal consumption-investment problem in continuous-time, and in paper [3], authors use forecasts of the mean and covariance of financial returns from a multivariate hidden Markov model with time-varying parameters to build the optimal controls.

As it is not possible to solve analytically our constrained optimal allocation problem, we have applied a machine learning algorithm developed in [13] and [1]. This algorithm, called *Hybrid-Now*, is particularly suited for solving stochastic control problems in high dimension using deep neural networks.

Our main contributions to the literature is twofold: a detailed theoretical study of a discrete-time portfolio selection problem including both drift uncertainty and maximum drawdown constraint, and a numerical resolution using a deep learning approach for an application to a model of three risky assets, leading to a five-dimensional problem. We derive the dynamic programming equation (DPE), which is in general of infinite-dimensional nature, following the change of measure suggested in [9]. In the Gaussian case, the DPE is reduced to a finite-dimensional equation by exploiting the Kalman filter. In the particular case of constant relative risk aversion (CRRA) utility function, we reduce furthermore the dimensionality of the problem. Then, we solve numerically the problem in the Gaussian case with CRRA utility functions using the deep learning *Hybrid-Now* algorithm. Such numerical results allow us to provide a detailed analysis of the performance and allocations of both the learning and non-learning strategies benchmarked with a comparable equally-weighted strategy. Finally, we assess the performance of the learning compared to the non-learning strategy with respect to the sensitivity of the uncertainty of the drift. Additionally, we provide empirical evidence of convergence of the non-learning strategy to the solution of the classical Merton problem when the parameter controlling the maximum drawdown vanishes.

The paper is organized as follows: Section 2 sets up the financial market model and the associated optimization problem. Section 3 describes, in the general case, the change of measure and the Bayesian filtering, the derivation of the dynamic programming equation and details some properties of the value function. Section 4 focuses on the Gaussian case. Finally, Section 5 presents the neural network techniques used, and shows the numerical results.

2 Problem setup

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a discrete filtration $(\mathcal{F}_k)_{k=0,\ldots,N}$ satisfying the usual conditions, we consider a financial market model with one riskless asset assumed normalized to one, and d risky assets. The price process $(S^i_k)_{k=0,\ldots,N}$ of asset $i \in [\![1, d]\!]$ is governed by the dynamics

$$S^i_{k+1} = S^i_k e^{R^i_{k+1}}, \quad k = 0, \ldots, N-1, \tag{1}$$

where $R_{k+1} = (R^1_{k+1}, \ldots, R^N_{k+1})$ is the vector of the assets log-return between time k and $k+1$, and modeled as:

$$R_{k+1} = B + \epsilon_{k+1}. \tag{2}$$

The drift vector B is a d-dimensional random variable with probability distribution (prior) μ_0 of known mean $b_0 = \mathbb{E}[B]$ and finite second order moment. Note that the case of known drift B means that μ_0 is a Dirac distribution. The noise $\epsilon = (\epsilon_k)_k$ is a sequence of centered i.i.d. random vector variables with covariance matrix $\Gamma = \mathbb{E}[\epsilon_k \epsilon'_k]$, and assumed to be independent of B. We also assume the fundamental assumption that the probability distribution ν of ϵ_k admits a strictly positive density function g on $\mathbb{R}^d$ with respect to the Lebesgue measure.

The price process S is observable, and notice by relation (1) that R can be deduced from S, and vice-versa. We will then denote by $\mathbb{F}^o = \left\{\mathcal{F}^o_k\right\}_{k=0,\ldots,N}$ the observation filtration generated by the process S (hence equivalently by R) augmented by the null sets of $\mathcal{F}$, with the convention that for $k=0$, $\mathcal{F}^o_0$ is the trivial algebra.

An investment strategy is an $\mathbb{F}^o$-progressively measurable process $\alpha = (\alpha_k)_{k=0,\ldots,N-1}$, valued in $\mathbb{R}^d$, and representing the proportion of the current wealth invested in each of the d risky assets at each time $k = 0, \ldots, N-1$. Given an investment strategy α and an initial wealth $x_0 > 0$, the (self-financed) wealth process X^α evolves according to

$$\begin{cases} X^\alpha_{k+1} = X^\alpha_k \left(1 + \alpha'_k \left(e^{R_{k+1}} - \mathbb{1}_d\right)\right), & k = 0, \ldots, N-1, \\ X^\alpha_0 = x_0. \end{cases} \tag{3}$$

where $e^{R_{k+1}}$ is the d-dimensional random variable with components $\left[e^{R_{k+1}}\right]_i = e^{R^i_{k+1}}$ for $i \in [\![1, d]\!]$, and $\mathbb{1}_d$ is the vector in $\mathbb{R}^d$ with all components equal to 1.

Let us introduce the process Z^α_k, as the maximum up to time k of the wealth process X^α, i.e.,

$$Z_k^\alpha := \max_{0 \le \ell \le k} X_\ell^\alpha, \qquad k = 0, \ldots, N.$$

The maximum drawdown constraints the wealth X_k^α to remain above a fraction $q \in (0,1)$ of the current historical maximum Z_k^α. We then define the set of *admissible* investment strategies $\mathcal{A}_0^q$ as the set of investment strategies α such that

$$X_k^\alpha \ge q Z_k^\alpha, \quad \text{a.s.,} \quad k = 0, \ldots, N.$$

In this framework, the portfolio selection problem is formulated as

$$V_0 := \sup_{\alpha \in \mathcal{A}_0^q} \mathbb{E}\left[U\left(X_N^\alpha\right)\right], \tag{4}$$

where U is a utility function on $(0, \infty)$ satisfying the standard Inada conditions: continuously differentiable, strictly increasing, concave on $(0, \infty)$ with $U'(0) = \infty$ and $U'(\infty) = 0$.

3 Dynamic programming system

In this section, we show how Problem (4) can be characterized from dynamic programming in terms of a backward system of equations amenable for algorithms. In a first step, we will update the prior on the drift uncertainty, and take advantage of the newest available information by adopting a Bayesian filtering approach. This relies on a suitable change of probability measure.

3.1 Change of measure and Bayesian filtering

We start by introducing a change of measure under which R_1,..., R_N are mutually independent, identically distributed random variables and independent from the drift B, hence behaving like a noise. Following the methodology detailed in [9] we define the σ-algebras

$$\mathcal{G}_k^0 := \sigma(B, R_1, \ldots, R_k), \quad k = 0, \ldots, N,$$

and $\mathbb{G} = (\mathcal{G}_k)_k$ the corresponding complete filtration. We then define a new probability measure $\overline{\mathbb{P}}$ on $(\Omega, \bigvee_{k=1}^N \mathcal{G}_k)$ by

$$\left.\frac{\mathrm{d}\overline{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{G}_k} := \Lambda_k, \quad k = 0, \ldots, N,$$

with

$$\Lambda_k := \prod_{\ell=1}^{k} \frac{g(R_\ell)}{g(\epsilon_\ell)}, \quad k = 1, \ldots, N, \quad \Lambda_0 \;=\; 1.$$

The existence of $\overline{\mathbb{P}}$ comes from the Kolmogorov's theorem since Λ_k is a strictly positive martingale with expectation equal to one. Indeed, for all $k = 1, \ldots, N$,

- $\Lambda_k > 0$ since the probability density function g is strictly positive
- Λ_k is $\mathcal{G}_k$-adapted,
- As $\epsilon_k \perp\!\!\!\perp \mathcal{G}_{k-1}$, we have

$$\begin{aligned}
\mathbb{E}[\Lambda_k|\mathcal{G}_{k-1}] &= \Lambda_{k-1}\mathbb{E}\Big[\frac{g(B+\epsilon_k)}{g(\epsilon_k)}|\mathcal{G}_{k-1}\Big] \\
&= \Lambda_{k-1}\int_{\mathbb{R}^d} \frac{g(B+e)}{g(e)} g(e)de \;=\; \Lambda_{k-1}\int_{\mathbb{R}^d} g(z)dz \;=\; \Lambda_{k-1}.
\end{aligned}$$

Proposition Under $\overline{\mathbb{P}}$, $(R_k)_{k=1,\ldots,N}$, is a sequence of i.i.d. random variables, independent from B, having the same probability distribution ν as ϵ_k. □

Proof. See Appendix 6.1. □

Conversely, we recover the initial measure $\mathbb{P}$ under which $(\epsilon_k)_{k=1,\ldots,N}$ is a sequence of independent and identically distributed random variables having probability density function g where $\epsilon_k = R_k - B$. Denoting by $\overline{\Lambda_k}$ the Radon-Nikodym derivative $\mathrm{d}\mathbb{P}/\mathrm{d}\overline{\mathbb{P}}$ restricted to the σ-algebra $\mathcal{G}_k$:

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\overline{\mathbb{P}}}\Big|_{\mathcal{G}_k} = \overline{\Lambda}_k,$$

we have

$$\overline{\Lambda}_k \;=\; \prod_{i=1}^{k} \frac{g(R_i - B)}{g(R_i)}.$$

It is clear that, under $\mathbb{P}$, the return and wealth processes have the form stated in equations (2) and (3). Moreover, from Bayes formula, the posterior distribution of the drift, i.e. the conditional law of B given the asset price observation, is

$$\mu_k(db) \;:=\; \mathbb{P}[B \in db|\mathcal{F}_k^o] \;=\; \frac{\pi_k(db)}{\pi_k(\mathbb{R}^d)}, \quad k = 0, \ldots, N, \tag{5}$$

where π_k is the so-called unnormalized conditional law

$$\pi_k(db) := \overline{\mathbb{E}}[\overline{\Lambda}_k \mathbb{1}_{\{B\in db\}}|\mathcal{F}_k^o], \quad k = 0, \ldots, N.$$

We then have the key recurrence linear relation on the unnormalized conditional law.

Proposition We have the recursive linear relation

$$\pi_\ell = \bar{g}(R_\ell - \cdot)\pi_{\ell-1}, \quad \ell = 1, \dots, N, \tag{6}$$

with initial condition $\pi_0 = \mu_0$, where

$$\bar{g}(R_\ell - b) = \frac{g(R_\ell - b)}{g(R_\ell)}, \quad b \in \mathbb{R}^d,$$

and we recall that g is the probability density function of the identically distributed ϵ_k under $\mathbb{P}$. □

Proof. See Appendix 6.2 . □

3.2 The static set of admissible controls

In this subsection, we derive some useful characteristics of the space of controls which will turn out to be crucial in the derivation of the dynamic programming system.

Given time $k \in [\![0, N]\!]$, a current wealth $x = X_k^\alpha > 0$, and current maximum wealth $z = Z_k^\alpha \geq x$ that satisfies the drawdown constraint $qz \leq x$ at time k for an admissible investment strategy $\alpha \in \mathcal{A}_0^q$, we denote by $A_k^q(x, z) \subset \mathbb{R}^d$ the set of static controls $a = \alpha_k$ such that the drawdown constraint is satisfied at next time $k+1$, i.e. $X_{k+1}^\alpha \geq qZ_{k+1}^\alpha$. From the relation (3), and noting that $Z_{k+1}^\alpha = \max[Z_k^\alpha, X_{k+1}^\alpha]$, this yields

$$A_k^q(x,z) = \left\{a \in \mathbb{R}^d : 1 + a'(e^{R_{k+1}} - \mathbb{1}_d) \;\geq\; q \max\left[\frac{z}{x}, 1 + a'(e^{R_{k+1}} - \mathbb{1}_d)\right] \text{ a.s.}\right\}. \tag{7}$$

Recalling from Proposition 1, that the random variables $R_1, \dots, R_N$ are i.i.d. under $\overline{\mathbb{P}}$, we notice that the set $A_k^q(x,z)$ does not depend on the current time k, and we will drop the subscript k in the sequel, and simply denote by $A^q(x,z)$.

Remembering that the support of ν, the probability distribution of ϵ_k, is $\mathbb{R}^d$, the following lemma characterizes more precisely the set $A^q(x,z)$.

Lemma 1 *For any* $(x,z) \in \mathcal{S}^q := \{(x,z) \in (0,\infty)^2 : qz \leq x \leq z\}$, *we have*

$$A^q(x,z) = \left\{a \in \mathbb{R}_+^d : |a|_1 \leq 1 - q\frac{z}{x}\right\},$$

where $|a|_1 = \sum_{i=1}^d |a_i|$ *for* $a = (a_1, \dots, a_d) \in \mathbb{R}_+^d$.

Proof. See Appendix 6.3. □

Let us prove some properties on the admissible set $A^q(x,z)$.

Lemma 2 *For any* $(x,z) \in \mathcal{S}^q$, *the set* $A^q(x,z)$ *satisfies the following properties:*

1. *It is decreasing in* q: $\forall q_1 \leq q_2$, $A^{q_2}(x,z) \subseteq A^{q_1}(x,z)$,

2. *It is continuous in q,*
3. *It is increasing in x:* $\forall x_1 \leq x_2$, $A^q(x_1, z) \subseteq A^q(x_2, z)$,
4. *It is a convex set,*
5. *It is homogeneous:* $a \in A^q(x, z) \Leftrightarrow a \in A^q(\lambda x, \lambda z)$, *for any* $\lambda > 0$.

Proof. See Appendix 6.4. □

3.3 Derivation of the dynamic programming equation

The change of probability detailed in Subsection 3.1 allows us to turn the initial partial information Problem (4) into a full observation problem as

$$\begin{aligned} V_0 \;:=\; \sup_{\alpha \in \mathcal{A}_0^q} \mathbb{E}[U(X_N^\alpha)] &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}[\overline{\Lambda}_N U(X_N^\alpha)] \\ &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}\Big[\overline{\mathbb{E}}[\overline{\Lambda}_N U(X_N^\alpha) | \mathcal{F}_N^o]\Big] \\ &= \sup_{\alpha \in \mathcal{A}_0^q} \overline{\mathbb{E}}\Big[U(X_N^\alpha)\pi_N(\mathbb{R}^d)\Big], \end{aligned} \tag{8}$$

from Bayes formula, the law of conditional expectations, and the definition of the unnormalized filter π_N valued in $\mathcal{M}_+$, the set of nonnegative measures on $\mathbb{R}^d$. In view of Equation (3), Proposition 1, and Proposition 2, we then introduce the dynamic value function associated to Problem (8) as

$$v_k(x, z, \mu) = \sup_{\alpha \in \mathcal{A}_k^q(x,z)} J_k(x, z, \mu, \alpha), \quad k \in [\![0, N]\!],\ (x, z) \in \mathcal{S}^q,\ \mu \in \mathcal{M}_+,$$

with

$$J_k(x, z, \mu, \alpha) = \overline{\mathbb{E}}\Big[U(X_N^{k,x,\alpha})\pi_N^{k,\mu}(\mathbb{R}^d)\Big],$$

where $X^{k,x,\alpha}$ is the solution to Equation (3) on $[\![k, N]\!]$, starting at $X_k^{k,x,\alpha} = x$ at time k, controlled by $\alpha \in \mathcal{A}_k^q(x, z)$, and $(\pi_\ell^{k,\mu})_{\ell=k,\dots,N}$ is the solution to (6) on $\mathcal{M}_+$, starting from $\pi_k^{k,\mu} = \mu$, so that $V_0 = v_0(x_0, x_0, \mu_0)$. Here, $\mathcal{A}_k^q(x, z)$ is the set of admissible investment strategies embedding the drawdown constraint: $X_\ell^{k,x,\alpha} \geq qZ_\ell^{k,x,z,\alpha}$, $\ell = k, \dots, N$, where the maximum wealth process $Z^{k,x,z,\alpha}$ follows the dynamics: $Z_{\ell+1}^{k,x,z,\alpha} = \max[Z_\ell^{k,x,z,\alpha}, X_{\ell+1}^{k,x,\alpha}]$, $\ell = k, \dots, N-1$, starting from $Z_k^{k,x,z,\alpha} = z$ at time k. The dependence of the value function upon the unnormalized filter μ means that the probability distribution on the drift is updated at each time step from Bayesian learning by observing assets price.

The dynamic programming equation associated to (8) is then written in backward induction as

$$\begin{cases} v_N(x,z,\mu) = U(x)\mu(\mathbb{R}^d), \\ v_k(x,z,\mu) = \sup_{\alpha \in \mathcal{A}_k^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}(X_{k+1}^{k,x,\alpha}, Z_{k+1}^{k,x,z,\alpha}, \pi_{k+1}^{k,\mu})\Big], \quad k = 0,\ldots,N-1. \end{cases}$$

Recalling Proposition 2 and Lemma 1, this dynamic programming system is written more explicitly as

$$\begin{cases} v_N(x,z,\mu) = U(x)\mu(\mathbb{R}^d), \quad (x,z) \in \mathcal{S}^q, \ \mu \in \mathcal{M}_+, \\ v_k(x,z,\mu) = \sup_{a \in A^q(x,z)} \overline{\mathbb{E}}\Big[v_{k+1}\Big(x(1 + a'(e^{R_{k+1}} - \mathbb{1}_d)), \\ \qquad \max\left[z, x(1 + a'(e^{R_{k+1}} - \mathbb{1}_d))\right], \bar{g}(R_{k+1} - \cdot)\mu\Big)\Big], \end{cases} \tag{9}$$

for $k = 0,\ldots,N-1$. Notice from Proposition 1 that the expectation in the above formula is only taken with respect to the noise R_{k+1}, which is distributed under $\overline{\mathbb{P}}$ according to the probability distribution ν with density g on $\mathbb{R}^d$.

3.4 Special case: CRRA utility function

In the case where the utility function is of CRRA (Constant Relative Risk Aversion) type, i.e.,

$$U(x) = \frac{x^p}{p}, \quad x > 0, \ \text{for some } 0 < p < 1, \tag{10}$$

one can reduce the dimensionality of the problem. For this purpose, we introduce the process $\rho = (\rho_k)_k$ defined as the ratio of the wealth over its maximum up to current as:

$$\rho_k^\alpha = \frac{X_k^\alpha}{Z_k^\alpha}, \quad k = 0,\ldots,N.$$

This ratio process lies in the interval $[q,1]$ due to the maximum drawdown constraint. Moreover, recalling (3), and observing that $Z_{k+1}^\alpha = \max[Z_k^\alpha, X_{k+1}^\alpha]$, together with the fact that $\frac{1}{\max[z,x]} = \min[\frac{1}{z}, \frac{1}{x}]$, we notice that the ratio process ρ can be written in inductive form as

$$\rho_{k+1}^\alpha = \min\Big[1, \rho_k^\alpha(1 + \alpha_k'(e^{R_{k+1}} - \mathbb{1}_d))\Big], \quad k = 0,\ldots,N-1.$$

The following result states that the value function inherits the homogeneity property of the utility function.

Lemma 3 *For a utility function U as in* (10)*, we have for all* $(x,z) \in \mathcal{S}^q$*,* $\mu \in \mathcal{M}_+$*,* $k \in [\![0,N]\!]$*,*

$$v_k(\lambda x, \lambda z, \mu) = \lambda^p v_k(x,z,\mu), \quad \lambda > 0.$$

Proof. See Appendix 6.5. □

In view of the above Lemma, we consider the sequence of functions w_k, $k \in [\![0, N]\!]$, defined by

$$w_k(r, \mu) = v_k(r, 1, \mu), \quad r \in [q, 1], \ \mu \in \mathcal{M}_+,$$

so that $v_k(x, z, \mu) = z^p w_k(\frac{x}{z}, \mu)$, and we call w_k the reduced value function. From the dynamic programming system satisfied by v_k, we immediately obtain the backward system for $(w_k)_k$ as

$$\begin{cases} w_N(r, \mu) = \frac{r^p}{p} \mu(\mathbb{R}^d), \quad r \in [q, 1], \ \mu \in \mathcal{M}_+, \\ w_k(r, \mu) = \sup\limits_{a \in A^q(r)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min\left[1, r(1 + a'(e^{R_{k+1}} - \mathbb{1}_d))\right], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big], \end{cases} \tag{11}$$

for $k = 0, \ldots, N - 1$, where

$$A^q(r) = \Big\{a \in \mathbb{R}^d_+ : a'\mathbb{1}_d \leq 1 - \frac{q}{r}\Big\}.$$

We end this section by stating some properties on the reduced value function.

Lemma 4 *For any $k \in [\![0, N]\!]$, the reduced value function w_k is nondecreasing and concave in $r \in [q, 1]$.*

Proof. See proof in Appendix 6.6. □

4 The Gaussian case

We consider in this section the Gaussian framework where the noise and the prior belief on the drift are modeled according to a Gaussian distribution. In this special case, the Bayesian filtering is simplified into the Kalman filtering, and the dynamic programming system is reduced to a finite-dimensional problem that will be solved numerically. It is convenient to deal directly with the posterior distribution of the drift, i.e. the conditional law of the drift B given the assets price observation, also called normalized filter. From (5) and Proposition 2, it is given by the inductive relation

$$\mu_k(db) = \frac{g(R_k - b)\mu_{k-1}(db)}{\int_{\mathbb{R}^d} g(R_k - b)\mu_{k-1}(db)}, \quad k = 1, \ldots, N. \tag{12}$$

4.1 Bayesian Kalman filtering

We assume that the probability law ν of the noise ϵ_k is Gaussian: $\mathcal{N}(0, \Gamma)$, and so with density function

$$g(r) = (2\pi)^{-\frac{d}{2}} |\Gamma|^{-\frac{1}{2}} e^{-\frac{1}{2} r' \Gamma^{-1} r}, \quad r \in \mathbb{R}^d. \tag{13}$$

Assuming also that the prior distribution μ_0 on the drift B is Gaussian with mean b_0, and invertible covariance matrix Σ_0, we deduce by induction from (12) that the posterior distribution μ_k is also Gaussian: $\mu_k \sim \mathcal{N}(\hat{B}_k, \Sigma_k)$, where $\hat{B}_k = \mathbb{E}[B|\mathcal{F}_k^o]$ and Σ_k satisfy the well-known inductive relations:

$$\hat{B}_{k+1} = \hat{B}_k + K_{k+1}(R_{k+1} - \hat{B}_k), \quad k = 0, \ldots, N-1 \tag{14}$$

$$\Sigma_{k+1} = \Sigma_k - \Sigma_k(\Sigma_k + \Gamma)^{-1}\Sigma_k, \tag{15}$$

where K_{k+1} is the so-called Kalman gain given by

$$K_{k+1} = \Sigma_k(\Sigma_k + \Gamma)^{-1}, \quad k = 0, \ldots, N-1. \tag{16}$$

We have the initialization $\hat{B}_0 = b_0$, and the notation for Σ_k is coherent at time $k = 0$ as it corresponds to the covariance matrice of B. While the Bayesian estimation $\hat{B}_k$ of B is updated from the current observation of the log-return R_k, notice that Σ_k (as well as K_k) is deterministic, and is then equal to the covariance matrix of the error between B and its Bayesian estimation, i.e. $\Sigma_k = \mathbb{E}[(B - \hat{B}_k)(B - \hat{B}_k)']$. Actually, we can explicitly compute Σ_k by noting from Equation (12) with g as in (13) and $\mu_0 \sim \mathcal{N}(b_0, \Sigma_0)$ that

$$\mu_k \sim \frac{e^{-\frac{1}{2}\left(b - \left(\Sigma_0^{-1} + \Gamma^{-1}k\right)^{-1}\left(\Gamma^{-1}\sum_{j=1}^k R_j + \Sigma_0^{-1} b_0\right)\right)\left(\Sigma_0^{-1} + \Gamma^{-1}k\right)\left(b - \left(\Sigma_0^{-1} + \Gamma^{-1}k\right)^{-1}\left(\Gamma^{-1}\sum_{j=1}^k R_j + \Sigma_0^{-1} b_0\right)\right)}}{(2\pi)^{\frac{d}{2}} |(\Sigma_0^{-1} + \Gamma^{-1}k)^{-1}|^{\frac{1}{2}}}.$$

By identification, we then get

$$\Sigma_k = (\Sigma_0^{-1} + \Gamma^{-1}k)^{-1} = \Sigma_0(\Gamma + \Sigma_0 k)^{-1}\Gamma. \tag{17}$$

Moreover, the innovation process $(\tilde{\epsilon}_k)_k$, defined as

$$\tilde{\epsilon}_{k+1} = R_{k+1} - \mathbb{E}[R_{k+1}|\mathcal{F}_k^o] = R_{k+1} - \hat{B}_k, \quad k = 0, \ldots, N-1, \tag{18}$$

is a $\mathbb{F}^o$-adapted Gaussian process. Each $\tilde{\epsilon}_{k+1}$ is independent of $\mathcal{F}_k^0$ (hence $\tilde{\epsilon}_k$, $k = 1, \ldots, N$ are mutually independent), and is a centered Gaussian vector with covariance matrix:

$$\tilde{\epsilon}_{k+1} \sim \mathcal{N}(0, \tilde{\Gamma}_{k+1}), \quad \text{with } \tilde{\Gamma}_{k+1} = \Sigma_k + \Gamma.$$

We refer to [15] and [14] for these classical properties about the Kalman filtering and the innovation process.

Remark 1 From (14), and (18), we see that the Bayesian estimator $\hat{B}_k$ follows the dynamics

$$\begin{cases} \hat{B}_{k+1} = \hat{B}_k + K_{k+1}\tilde{\epsilon}_{k+1}, & k = 0, \ldots, N-1 \\ \hat{B}_0 = b_0, \end{cases}$$

which implies in particular that $\hat{B}_k$ has a Gaussian distribution with mean b_0, and covariance matrix satisfying

$$\mathrm{Var}(\hat{B}_{k+1}) = \mathrm{Var}(\hat{B}_k) + K_{k+1}(\Sigma_k + \Gamma)K'_{k+1} = \mathrm{Var}(\hat{B}_k) + \Sigma_k(\Sigma_k + \Gamma)^{-1}\Sigma_k.$$

Recalling the inductive relation (15) on Σ_k, this shows that $\mathrm{Var}(\hat{B}_k) = \Sigma_0 - \Sigma_k$. Note that, from Equation (15), $(\Sigma_k)_k$ is a decreasing sequence which ensures that $\mathrm{Var}(\hat{B}_k)$ is positive semi-definite and is nondecreasing with time k. ◊

4.2 Finite-dimensional dynamic programming equation

From (18), we see that our initial portfolio selection Problem (4) can be reformulated as a full observation problem with state dynamics given by

$$\begin{cases} X_{k+1}^{\alpha} = X_k^{\alpha}\Big(1 + \alpha'_k\big(e^{\hat{B}_k + \tilde{\epsilon}_{k+1}} - \mathbb{1}_d\big)\Big), \\ \hat{B}_{k+1} = \hat{B}_k + K_{k+1}\tilde{\epsilon}_{k+1}, \quad k = 0, \dots, N-1. \end{cases} \tag{19}$$

We then define the value function on $[\![0, N]\!] \times \mathcal{S}^q \times \mathbb{R}^d$ by

$$\tilde{v}_k(x, z, b) = \sup_{\alpha \in \mathcal{A}_k^q(x,z)} \mathbb{E}\big[U(X_N^{k,x,b,\alpha})\big], \quad k \in [\![0, N]\!],\ (x, z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d,$$

where the pair $(X^{k,x,b,\alpha}, \hat{B}^{k,b})$ is the process solution to (19) on $[\![k, N]\!]$, starting from (x, b) at time k, so that $V_0 = \tilde{v}_0(x_0, x_0, b_0)$. The associated dynamic programming system satisfied by the sequence $(\tilde{v}_k)_k$ is

$$\begin{cases} \tilde{v}_N(x, z, b) = U(x), \quad (x, z) \in \mathcal{S}^q,\ b \in \mathbb{R}^d, \\ \tilde{v}_k(x, z, b) = \sup_{a \in A^q(x,z)} \mathbb{E}\Big[\tilde{v}_{k+1}\Big(x(1 + a'(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d)), \\ \qquad\qquad \max\big[z, x(1 + a'(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d))\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \dots, N-1$. Notice that in the above formula, the expectation is taken with respect to the innovation vector $\tilde{\epsilon}_{k+1}$, which is distributed according to $\mathcal{N}(0, \tilde{\Gamma}_{k+1})$.

Moreover, in the case of CRRA utility functions $U(x) = x^p/p$, and similarly as in Section 3.4, we have the dimension reduction with

$$\tilde{w}_k(r, b) = \tilde{v}_k(r, 1, b), \quad r \in [q, 1],\ b \in \mathbb{R}^d,$$

so that $\tilde{v}_k(x, z, b) = z^p \tilde{w}_k(\frac{x}{z}, b)$, and this reduced value function satisfies the backward system on $[q, 1] \times \mathbb{R}^d$:

$$\begin{cases} \tilde{w}_N(r, b) = \frac{r^p}{p}, \quad r \in [q, 1],\ b \in \mathbb{R}^d, \\ \tilde{w}_k(r, b) = \sup_{a \in A^q(r)} \mathbb{E}\Big[\tilde{w}_{k+1}\Big(\min\big[1, r(1 + a'(e^{b+\tilde{\epsilon}_{k+1}} - \mathbb{1}_d))\big], b + K_{k+1}\tilde{\epsilon}_{k+1}\Big)\Big], \end{cases}$$

for $k = 0, \ldots, N-1$.

Remark 2 (No short-sale constrained Merton problem) In the limiting case when $q = 0$, the drawdown constraint is reduced to a non-negativity constraint on the wealth process, and by Lemma 1, this means a no-short selling and no borrowing constraint on the portfolio strategies. When the drift B is also known, equal to b_0, and for a CRRA utility function, let us then consider the corresponding constrained Merton problem with value function denoted by v_k^M, $k = 0, \ldots, N$, which satisfies the standard backward recursion from dynamic programming:

$$\begin{cases} v_N^M(x) = \dfrac{x^p}{p}, \quad x > 0, \\ v_k^M(x) = \sup\limits_{\substack{a'\mathbb{1}_d \le 1 \\ a \in [0,1]^d}} \mathbb{E}\Big[v_{k+1}^M\big(x(1 + a'(e^{b_0+\epsilon_{k+1}} - \mathbb{1}_d))\big)\Big], \quad k = 0, \ldots, N-1. \end{cases} \tag{20}$$

Searching for a solution of the form $v_k^M(x) = K_k x^p/p$, with $K_k \ge 0$ for all $k \in [\![0, N]\!]$, we see that the sequence $(K_k)_k$ satisfies the recursive relation:

$$K_k = SK_{k+1}, \quad k = 0, \ldots, N-1,$$

starting from $K_N = 1$, where

$$S := \sup_{\substack{a'\mathbb{1}_d \le 1 \\ a \in [0,1]^d}} \mathbb{E}\Big[\Big(1 + a'(e^{b_0+\epsilon_1} - \mathbb{1}_d)\Big)^p\Big],$$

by recalling that $\epsilon_1, \ldots, \epsilon_N$ are i.i.d. random variables. It follows that the value function of the constrained Merton problem, unique solution to the dynamic programming system (20), is equal to

$$v_k^M(x) = S^{N-k}\frac{x^p}{p}, \quad k = 0, \ldots, N,$$

and the constant optimal control is given by

$$a_k^M = \underset{\substack{a'\mathbb{1} \le 1 \\ a \in [0,1]^d}}{\operatorname{argmax}} \mathbb{E}\left[\left(1 + a'\left(e^{R_1} - \mathbb{1}_d\right)\right)^p\right] \quad k = 0, \ldots, N-1.$$

◊

5 Deep learning numerical resolution

In this section, we exhibit numerical results to promote the benefits of learning from new information. To this end, we compare the learning strategy (Learning) to the non-learning one (Non-Learning) in the case of the CRRA utility function and the Gaussian distribution for the noise. The prior probability distribution of B is the Gaussian distribution $\mathcal{N}(b_0, \Sigma_0)$ for Learning while it is the Dirac distribution concentrated at b_0 for Non-Learning.

We use deep neural network techniques to compute numerically the optimal solutions for both Learning and Non-Learning. To broaden the analysis, in addition to the learning and non-learning strategies, we have computed an "admissible" equally weighted (EW) strategy. More precisely, this EW strategy will share the quantity $X_k - qZ_k$ equally among the d assets. Eventually, we show numerical evidence that the Non-Learning converges to the optimal strategy of the constrained Merton problem, when the loss aversion parameter q vanishes.

5.1 Architectures of the deep neural networks

Neural networks (NN) are able to approximate nonlinear continuous functions, typically the value function and controls of our problem. The principle is to use a large amount of data to train the NN so that it progressively comes close to the target function. It is an iterative process in which the NN is tuned on a training set, then tested on a validation set to avoid over-fitting. For more details, see for instance [12] and [10].

The algorithm we use, relies on two dense neural networks: the first one is dedicated to the controls (A_{NN}) and the second one to the value function (VF_{NN}). Each NN is composed of four layers: an input layer, two hidden layers and an output layer:

(i) The input layer is $d+1$-dimensional since it embeds the conditional expectations of each of the d assets and the ratio of the current wealth to the current historical maximum ρ.
(ii) The two hidden layers give the NN the flexibility to adjust its weights and biases to approximate the solution. From numerical experiments, we see that, given the complexity of our problem, a first hidden layer with $d + 20$ neurons and a second one with $d + 10$ are a good compromise between speed and accuracy.
(iii) The output layer is d-dimensional for the controls, one for each asset representing the weight of the instrument, and is one-dimensional for the value function. See Figures 1 and 2 for an overview of the NN architectures in the case of $d = 3$ assets.

Parameter	$\mathbf{A_{NN}}$	$\mathbf{VF_{NN}}$
Initializer	uniform(0, 1)	He_uniform
Regularizers	L2 norm	L2 norm
Activation functions	Elu and Sigmoid for output layer	Elu and Sigmoid for output layer
Optimizer	Adam	Adam
Learning rates: step N-1	5e-3	1e-3
steps k = 0,...,N-2	6.25e-4	5e-4
Scale	1e-3	1e-3
Number of elements in a training batch	3e2	3e2
Number of training batches	1e2	1e2
Size of the validation batches	1e3	1e3
Penalty constant	3e-1	NA
Number of epochs: step N-1	2e3	2e3
steps k = 0,...,N-2	5e2	5e2
Size of the training set: step N-1	6e7	6e7
steps k = 0,...,N-2	1.5e7	1.5e7
Size of the validation set: step N-1	2e6	2e6
steps k = 0,..., N-2	5e5	5e5

Table 1 Parameters for the neural networks of the controls $\mathbf{A_{NN}}$ and the value function $\mathbf{VF_{NN}}$.

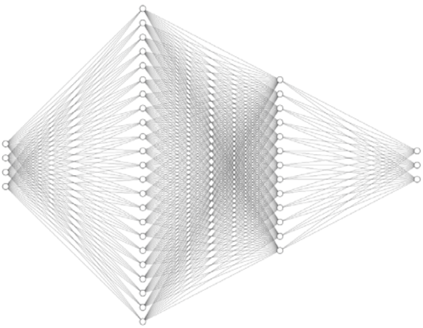

Fig. 1 A_{NN} architecture with $d = 3$ assets

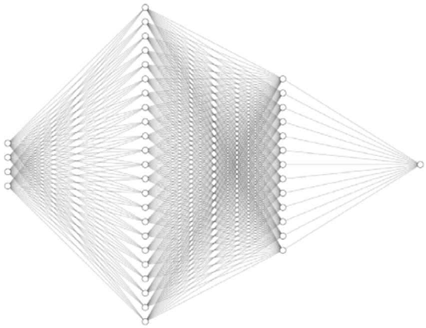

Fig. 2 VF_{NN} architecture with $d = 3$ assets

We follow the indications in [10] to setup and define the values of the various inputs of the neural networks which are listed in Table 1.

To train the NN, we simulate the input data. For the conditional expectation $\hat{B}_k$, we use its time-dependent Gaussian distribution (see Remark 1): $\hat{B}_k \sim \mathcal{N}(b_0, \Sigma_0 - \Sigma_k)$, with Σ_k as in Equation (17). On the other hand, the training of ρ is drawn from the uniform distribution between q and 1, the interval where it lies according to the maximum drawdown constraint.

5.2 *Hybrid-Now* algorithm

We use the *Hybrid-Now* algorithm developped in [1] in order to solve numerically our problem. This algorithm combines optimal policy estimation by neural networks and dynamic programming principle which suits the approach we have developped in Section 4.

With the same notations as in Algorithm 1 detailed in the next insert, at time k, the algorithm computes the proxy of the optimal control $\hat{\alpha}_k$ with A_{NN}, using the known function $\hat{V}_{k+1}$ calculated the step before, and uses V_{NN} to obtain a proxy of the value function $\hat{V}_k$. Starting from the known function $\hat{V}_N := U$ at terminal time N, the algorithm computes sequentially $\hat{\alpha}_k$ and $\hat{V}_k$ with backward iteration until time 0. This way, the algorithm loops to build the optimal controls and the value function pointwise and gives as output the optimal strategy, namely the optimal controls from 0 to $N-1$ and the value function at each of the N time steps.

The maximum drawdown constraint is a time-dependent constraint on the maximal proportion of wealth to invest (recall Lemma 1). In practice, it is a constraint on the sum of weights of each asset or equivalently on the output of A_{NN}. For that reason, we have implemented an appropriate penalty function that will reject undesirable values:

$$G_{Penalty}(A, r) = K \max\left(|A|_1 \leq 1 - \frac{q}{r}, 0\right), \quad A \in [0,1]^d, \quad r \in [q, 1].$$

This penalty function ensures that the strategy respects the maximum drawdown constraint at each time step, when the parameter K is chosen sufficiently large.

Algorithm 1: *Hybrid-Now*

Input: the training distributions μ_{Unif} and μ^k_{Gauss};

▷ $\mu_{Unif} = \mathcal{U}(q, 1)$

▷ $\mu^k_{Gauss} = \mathcal{N}(b_0, \Sigma_0 - \Sigma_k)$

Output:

- estimate of the optimal strategy $(\hat{a}_k)_{k=0}^{N-1}$;
- estimate of the value function $\left(\hat{V}_k\right)_{k=0}^{N-1}$;

Set $\hat{V}_N = U$;

for $k = N-1, \ldots, 0$ **do**

Compute:

$$\hat{\beta}_k \in \underset{\beta \in \mathbb{R}^{2d^2+56d+283}}{\operatorname{argmin}} \mathbb{E}\left[G_{Penalty}(A_{NN}(\rho_k, \hat{B}_k; \beta), \rho_k) - \hat{V}_{k+1}\left(\rho^{\beta}_{k+1}, \hat{B}_{k+1}\right)\right]$$

where $\rho_k \sim \mu_{Unif}$, $\hat{B}_k \sim \mu^k_{Gauss}$,

$\hat{B}_{k+1} = \tilde{H}_k(\hat{B}_k, \tilde{\epsilon}_{k+1})$ and $\rho^{\beta}_{k+1} = F\left(\rho_k, \hat{B}_k, A_{NN}\left(\rho_k, \hat{B}_k; \beta\right), \tilde{\epsilon}_{k+1}\right)$;

▷ $F(\rho, b, a, \epsilon) = \min\left(1, \rho\left(1 + \sum_{i=1}^d a^i\left(e^{b^i+\epsilon^i} - 1\right)\right)\right)$

▷ $\tilde{H}_k(b, \epsilon) = b + \Sigma_0(\Gamma + \Sigma_0 k)^{-1}\epsilon$

Set $\hat{a}_k = A_{NN}\left(.; \hat{\beta}_k\right)$;

▷ $\hat{a}_k$ `is the estimate of the optimal control at time` k.

Compute:

$$\hat{\theta}_k \in \underset{\theta \in \mathbb{R}^{2d^2+54d+261}}{\operatorname{argmin}} \mathbb{E}\left[\left(\hat{V}_{k+1}\left(\rho^{\hat{\beta}_k}_{k+1}, \hat{B}_{k+1}\right) - VF_{NN}\left(\rho_k, \hat{B}_k; \theta\right)\right)^2\right]$$

Set $\hat{V}_k = VF_{NN}\left(., \hat{\theta}_k\right)$;

▷ $\hat{V}_k$ `is the estimate of the value function at time` k.

A major argument behind the choice of this algorithm is that, it is particularly relevant for problems in which the neural network approximation of the controls and value function at time k, are close to the ones at time $k+1$. This is what we expect in our case. We can then take a small learning rate for the Adam optimizer which enforces the stability of the parameters' update during the gradient-descent based learning procedure.

5.3 Numerical results

In this section, we explain the setup of the simulation and exhibit the main results. We have used Tensorflow 2 and deep learning techniques for Python developped in [10]. We consider $d = 3$ risky assets and a riskless asset whose return is assumed

Parameter	Value
Number of risky assets d	3
Investment horizon in years T	1
Number of steps/rebalancing N	24
Number of simulations/trajectories $\tilde{N}$	1000
Degree of the CRRA utility function p	0.8
Parameter of risk aversion q	0.7
Annualized expectation of the drift B	$\begin{bmatrix} 0.05 & 0.025 & 0.12 \end{bmatrix}$
Annualized covariance matrix of the drift B	$\begin{bmatrix} 0.2^2 & 0 & 0 \\ 0 & 0.15^2 & 0 \\ 0 & 0 & 0.1^2 \end{bmatrix}$
Annualized volatility of ϵ	$\begin{bmatrix} 0.08 & 0.04 & 0.22 \end{bmatrix}$
Correlation matrix of ϵ	$\begin{bmatrix} 1 & -0.1 & 0.2 \\ -0.1 & 1 & -0.25 \\ 0.2 & -0.25 & 1 \end{bmatrix}$
Annualized covariance matrix of the noise ϵ	$\begin{bmatrix} 0.0064 & -0.00032 & 0.00352 \\ -0.00032 & 0.0016 & -0.0022 \\ 0.00352 & -0.0022 & 0.0484 \end{bmatrix}$

Table 2 Values of the parameters used in the simulation.

to be 0, on a 1-year investment horizon for the sake of simplicity. We consider 24 portfolio rebalancing during the 1-year period, i.e., one every two weeks. This means that we have $N = 24$ steps in the training of our neural networks. The parameters used in the simulation are detailed in Table 2.

First, we show the numerical results for the learning and the non-learning strategies by presenting a performance and an allocation analysis in Subsection 5.3.1. Then, we add the admissible constrained EW to the two previous ones and use this neutral strategy as a benchmark in Subsection 5.3.2. Ultimately, in Subsection 5.3.3, we illustrate numerically the convergence of the non-learning strategy to the constrained Merton problem when the loss aversion parameter q vanishes.

5.3.1 Learning and non-learning strategies

We simulate $\tilde{N} = 1000$ trajectories for each strategy and exhibit the performance results with an initial wealth $x_0 = 1$. Figures 3 illustrates the average historical level of the learning and non-learning strategies with a 95% confidence interval. Learning outperforms significantly Non-Learning with a narrower confidence interval revealing that less uncertainty surrounds Learning performance, thus yielding less risk.

An interesting phenomenon, visible in Fig. 3, is the nearly flat curve for Learning between time 0 and time 1. Indeed, whereas Non-Learning starts investing immediately, Learning adopts a safer approach and needs a first time step before allocating a significant proportion of wealth. Given the level of uncertainty surrounding b_0, this first step allows Learning to fine-tune its allocation by updating the prior belief with the first return available at time 1. On the contrary, Non-Learning, which cannot update its prior, starts investing at time 0.

Fig. 4 shows the ratio of Learning over Non-Learning. A ratio greater than one means that Learning outperforms Non-Learning and underperforms when less than one. It shows the significant outperformance of Learning over Non-Learning except during the first period where Learning was not significantly invested and Non-Learning had a positive return. Moreover, this graph reveals the typical increasing concave curve of the value of information described in [17], in the context of investment decisions and costs of data analytics, and in [6] in the resolution of the Markowitz portfolio selection problem using a Bayesian learning approach.

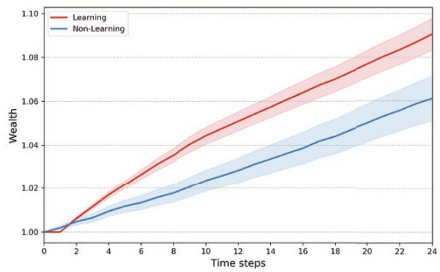

Fig. 3 Historical Learning and Non-Learning levels with a 95% confidence interval.

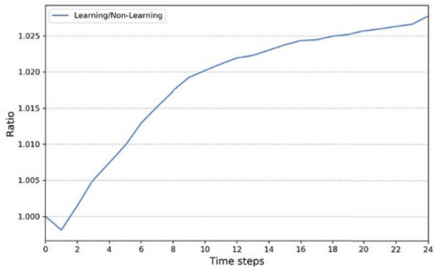

Fig. 4 Historical ratio of Learning over Non-Learning levels.

Table 3 gathers relevant statistics for both Learning and Non-Learning such as: average total performance, standard deviation of the terminal wealth X_T, Sharpe ratio computed as average total performance over standard deviation of terminal wealth. The maximum drawdown (MD) is examined through two statistics: noting $MD_\ell^{\tilde{s}}$ the maximum drawdown of the ℓ-th trajectory of a strategy $\tilde{s}$, the average MD is defined as,

$$\text{Avg MD}^{\tilde{s}} = \frac{1}{\tilde{N}} \sum_{\ell=1}^{\tilde{N}} \text{MD}_\ell^{\tilde{s}},$$

for $\tilde{N}$ trajectories of the strategy $\tilde{s}$, and the worst MD is defined as,

$$\text{Worst MD}^{\tilde{s}} = \min\left(MD_1^{\tilde{s}}, \ldots, MD_{\tilde{N}}^{\tilde{s}}\right).$$

Finally, the Calmar ratio, computed as the ratio of the average total performance over the average maximum drawdown, is the last statistic exhibited.

With the simulated dataset, Learning delivered, on average, a total performance of 9.34% while Non-Learning only 6.40%. Integrating the most recent information yielded a 2.94% excess return. Moreover, risk metrics are significantly better for Learning than for Non-Learning. Learning exhibits a lower standard deviation of

Statistic	Learning	Non-Learning	Difference
Avg total performance	9.34%	6.40%	2.94%
Std dev. of X_T	11.88%	16.67%	-4.79%
Sharpe ratio	0.79	0.38	104.95%
Avg MD	-1.53%	-6.54%	5.01%
Worst MD	-11.74%	-27.18%	15.44%
Calmar ratio	6.12	0.98	525.26%

Table 3 Performance metrics: Learning and Non-Learning. The difference for ratios are computed as relative improvement.

terminal wealth than Non-Learning (11.88% versus 16.67%), with a difference of 4.79%. More interestingly, the maximum drawdown is notably better controlled by Learning than by Non-Learning, on average (−1.53% versus −6.54%) and in the worst case (−11.74% versus −27.18%). This result suggests that learning from new observations, helps the strategy to better handle the dual objective of maximizing total wealth while controlling the maximum drawdown. We also note that learning improves the Sharpe ratio by 104.95% and the Calmar ratio by 525.26%.

Fig. 5 and 6 focus more precisely on the portfolio allocation. The graphs of Fig. 5 show the historical average allocation for each of the three risky assets. First, none of the strategies invests in Asset 2 since it has the lowest expected return according to the prior, see Table 2. Whereas Non-Learning focuses on Asset 3, the one with the highest expected return, Learning performs an optimal allocation between Asset 1 and Asset 3 since this strategy is not stuck with the initial estimate given by the prior. Therefore, Learning invests little at time 0, then balances nearly equally both Assets 1 and 3, and then invests only in Asset 3 after time step 12. Instead, Non-Learning is investing only in Asset 3, from time 0 until the end of the investment horizon.

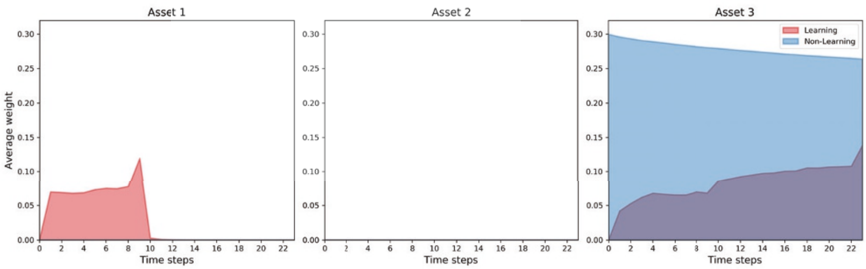

Fig. 5 Historical Learning and Non-Learning asset allocations.

The curves in Fig. 6 recall each asset's optimal weight, but the main features are the colored areas that represent the average historical total percentage of wealth invested by each strategy. The dotted line represents the total allocation constraint they should satisfy to be admissible. To satisfy the maximum drawdown constraint, admissible strategies can only invest in risky assets the proportion of wealth that, in theory, could be totally lost. This explains why the non-learning strategy invests at full capacity on the asset that has the maximum expected return according to the prior distribution.

We clearly see that both strategies satisfy their respective constraints. Indeed, looking at the left panel, Learning is far from saturating the constraint. It has invested, on average, roughly 10% of its wealth while its constraint was set around 30%. Non-learning invests at full capacity saturating its allocation constraint. Remark that this constraint is not a straight line since it depends on the value of the ratio: current wealth over current historical maximum, and evolves according to time.

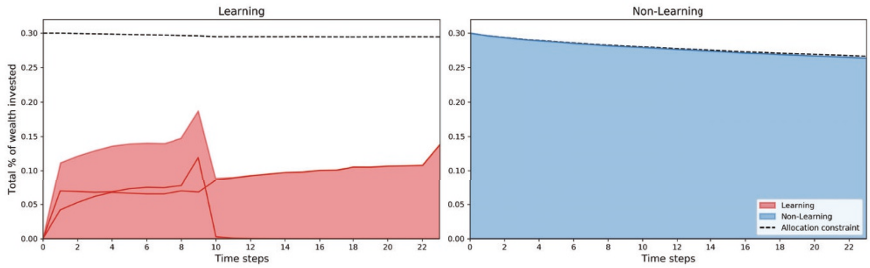

Fig. 6 Historical Learning and Non-Learning total allocations.

5.3.2 Learning, non-learning and constrained equally-weighted strategies

In this section, we add a simple constrained equally-weighted (EW) strategy to serve as a benchmark for both Learning and Non-Learning. At each time step, the constrained EW strategy invests, equally across the three assets, the proportion of wealth above the threshold q.

Fig. 7 shows the average historical levels of the three strategies: Learning, Non-Learning and constrained EW. We notice Non-Learning outperforms constrained EW and both have similar confidence intervals. It is not surprising to see that Non-Learning outperforms constrained EW since Non-Learning always bets on Asset 3, the most performing, while constrained EW diversifies the risks equally among the three assets.

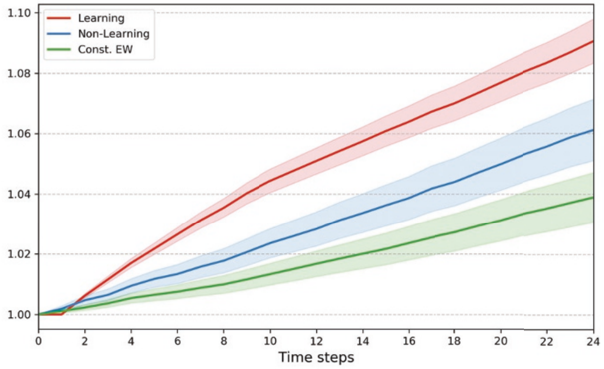

Fig. 7 Historical Learning, Non-Learning and constrained EW (Const. EW) levels with a 95% confidence interval.

Fig. 8 shows the ratio of Learning over constrained EW: it depicts the same concave shape as Fig. 4. The outperformance of Non-Learning with respect to constrained EW is plot in Fig. 9 and confirms, on average, the similarity of the two strategies.

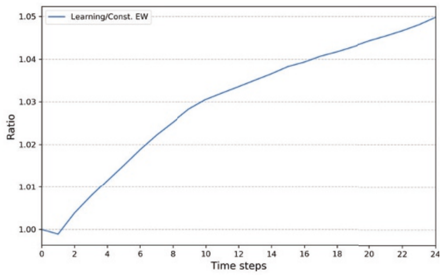

Fig. 8 Ratio Learning over constrained EW (Const. EW) according to time.

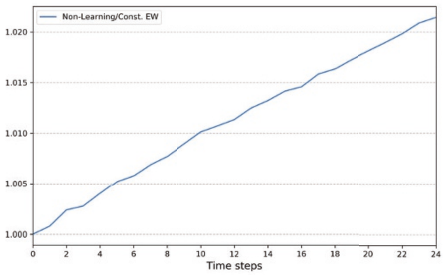

Fig. 9 Ratio Non-Learning over constrained EW (Const. EW) according to time.

Table 4 collects relevant statistics for the three strategies. Learning clearly surpasses constrained EW: it outperforms by 5.49% while reducing uncertainty on terminal wealth by 1.92% resulting in an improvement of 182.08% of the Sharpe ratio. Moreover, it better handles maximum drawdown regarding both the average and the worst case, exhibiting an improvement of 3.17% and 10.09% respectively, enhancing the Calmar ratio by 647.56%.

The Non-Learning and the constrained EW have similar profiles. Even if Non-Learning outperforms constrained EW by 2.5%, it has a higher uncertainty in terminal wealth (+2.87%). This results in similar Sharpe ratios. Maximum drawdown, both on average and considering the worst case are better handled by constrained EW (−4.70% and −21.83% respectively) than by Non-Learning (−6.54% and −27.18% respectively) thanks to the diversification capacity of constrained EW. The better per-

Statistic	Const. EW	L	NL	L - Const. EW	NL - Const. EW
Avg total performance	3.85%	9.34%	6.40%	5.49%	2.55%
Std dev. of X_T	13.80%	11.88%	16.67%	-1.92%	2.87%
Sharpe ratio	0.28	0.79	0.38	182.08%	37.63%
Avg MD	-4.70%	-1.53%	-6.54%	3.17%	-1.84%
Worst MD	-21.83%	-11.74%	-27.18%	10.09%	-5.34%
Calmar ratio	0.82	6.12	0.98	647.56%	-19.56%

Table 4 Performance metrics: Constrained EW (Const. EW) vs Learning (L) and Non-Learning (NL). The difference for ratios are computed as relative improvement.

formance of Non-Learning compensates the better maximum drawdown handling of constrained EW, entailing a better Calmar ratio for Non-Learning 0.98 versus 0.82 for constrained EW.

5.3.3 Non-learning and Merton strategies

We numerically analyze the impact of the drawdown parameter q, and compare the non-learning strategies (assuming that the drift is equal to b_0), with the constrained Merton strategy as described in Remark 2. Fig. 10 confirms that when the loss aversion parameter q goes to zero, the non-learning strategy approaches the Merton strategy.

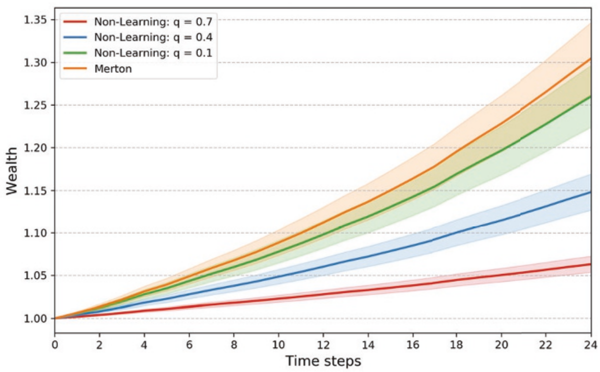

Fig. 10 Wealth curves resulting from the Merton strategy and the non-learning strategy for different values of q.

In terms of assets' allocation, the Merton strategy saturates the constraint only by investing in the asset with the highest expected return, Asset 3, while the non-learning strategy adopts a similar approach and invests at full capacity in the same asset. To illustrate this point, we easily see that the areas at the top and bottom-left corner converge to the area at the bottom-right corner of Fig. 11.

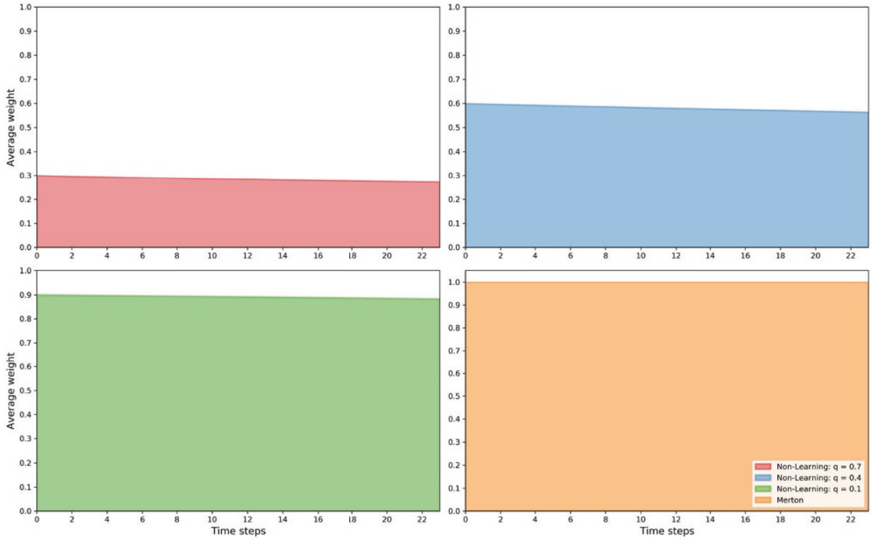

Fig. 11 Asset 3 average weights of the non-learning strategies with $q \in \{0.7, 0.4, 0.1\}$ and the Merton strategy.

As q vanishes, we observe evidence of the convergence of the Merton and the non-learning strategies, materialized by a converging allocation pattern and resulting wealth trajectories. It should not be surprising since both have in common not to learn from incoming information conveyed by the prices.

5.4 Sensitivities analysis

In this subsection, we study the effect of changes in the uncertainty about the beliefs of B. These beliefs take the form of an estimate b_0 of B, and a degree of uncertainty about this estimate, the covariance of Σ_0 of B. For the sake of simplicity, we design Σ_0 as a diagonal matrix whose diagonal entries are variances representing the confidence the investor has in her beliefs about the drift. To easily model a change in Σ_0, we define the modified covariance matrix $\tilde{\Sigma}$ as

$$\tilde{\Sigma}_{unc} := unc * \Sigma_0,$$

where $unc > 0$. From now on, the prior of B is $\mathcal{N}(b_0, \tilde{\Sigma}_{unc})$.

A higher value of unc means a higher uncertainty materialized by a lower confidence in the prior estimate of the expected return of B, b_0. We consider learning strategies with values of $unc \in \{1/6,\ 1,\ 3,\ 6,\ 12\}$. The value $unc = 1$ was used for Learning in Subsection 5.3.

Equation (2) implies that the returns' probability distribution depends upon *unc*. It implies that for each value of *unc*, we need to compute both Learning and Non-Learning on the returns sample drawn from the same probability law to make relevant comparisons.
Therefore, from a sample of a thousand returns paths' draws, we plot in Fig. 12 the average curves of the excess return of Learning over its associated Non-Learning, for different values of the uncertainty parameter *unc*.

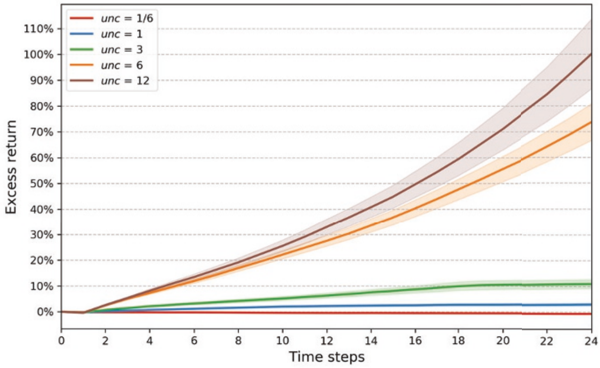

Fig. 12 Excess return of Learning over Non-Learning with a 95% confidence interval for different levels of uncertainty.

Looking at Fig. 12, we notice that when uncertainty about b_0 is low, i.e. $unc = 1/6$, Learning is close to Non-Learning and unsurprisingly the associated excess return is small. Then, as we increase the value of *unc* the curves steepen increasingly showing the effect of learning in generating excess return.

Table 5 summarises key statistics for the ten strategies computed in this section. When $unc = 1/6$, Learning underperforms Non-Learning. This is explained by the fact that Non-Learning has no doubt about b_0 and knows Asset 3 is the best performing asset acoording to its prior, whereas Learning, even with low uncertainty, needs to learn it generating a lag which explains the underperformance on average. For values of $unc \geq 1$ Learning outperforms Non-learning increasingly, as can be seen on Fig. 13, at the cost of a growing standard deviation of terminal wealth.

The Sharpe ratio of terminal wealth is higher for Learning than for Non-Learning for any value of *unc*. Nevertheless, an interesting fact is that the ratio rises from $unc = 1/6$ to $unc = 1$, then reaches a level close to 0.8 for values of $unc = 1, 3, 6$ then decreases when $unc = 12$.
This phenomenon is more visible on Fig. 14 that displays the Sharpe ratio of terminal wealth of Learning and Non-Learning according to the values of *unc*, and the associated relative improvement. Clearly, looking at Figures 13 and 14, we remark

	unc = 1/6		*unc* = 1		*unc* = 3		*unc* = 6		*unc* = 12	
Statistic	L	NL	L	NL	L	NL	L	NL	L	NL
Avg total performance	3.87%	4.35%	9.45%	6.00%	19.96%	10.25%	90.03%	16.22%	130.07%	30.44%
Std dev. of X_T	5.81%	9.22%	12.10%	17.28%	25.01%	28.18 %	113.69%	41.24%	222.77%	70.84%
Sharpe ratio	0.67	0.47	0.78	0.35	0.80	0.36	0.79	0.39	0.58	0.43
Avg MD	-2.51%	-5.21%	-1.40%	-6.78%	-1.90%	-8.40%	-2.68%	-10.14%	-3.58%	-11.35%
Worst MD	-7.64%	-17.88%	-5.46%	-24.01%	-7.99%	-26.68%	-15.62%	-29.22%	-16.98%	-29.47%
Calmar ratio	1.54	0.83	6.77	0.89	10.49	1.22	33.65	1.60	36.32	2.68

Table 5 Performance and risk metrics: Learning (L) vs Non-Learning (NL) for different values of uncertainty *unc*.

that while increasing *unc* gives more excess return, too high values of *unc* in the model turn out to be a drag as far as Sharpe ratio improvement is concerned.

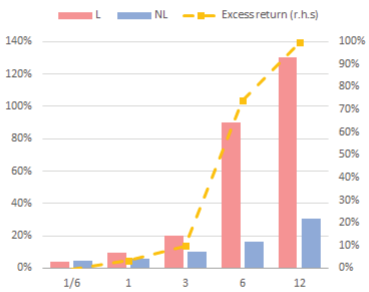

Fig. 13 Average total performance of Learning (L) and Non-Learning (NL), and excess return, for $unc \in \{1/6, 1, 3, 6, 12\}$.

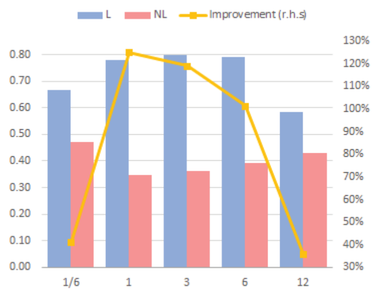

Fig. 14 Sharpe ratio of terminal wealth of Learning (L) and Non-Learning (NL), and relative improvement, for $unc \in \{1/6, 1, 3, 6, 12\}$.

For any value of *unc*, Learning handles maximum drawdown significantly better than Non-Learning whatever it is the average or the worst. This results in a better performance per unit of average maximum drawdown (Calmar ratio), for Learning. We also see that the maximum drawdown constraint is satisfied for every strategies of the sample and for any value of *unc* since the worst maximum drawdown is always above −30%, the lowest admissible value with a loss aversion parameter q set at 0.7. Fig. 15 reveals how the average maximum drawdown behaves regarding the level of uncertainty. Non-Learning maximum drawdown behaves linearly with uncertainty: the wider the range of possible values of B the higher the maximum drawdown is on average. It emphasizes its inability to adapt to an environment in which the returns have different behaviors compared to their expectations. Learning instead, manages to keep a low maximum drawdown for any value of *unc*. Given the previous remarks, it is obvious that the gain in maximum drawdown from learning grows with the level of uncertainty.

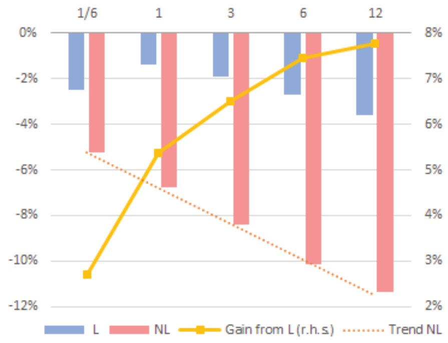

Fig. 15 Average maximum drawdown of Learning (L) and Non-Learning (NL) and the gain from learning for *unc* $\in \{1/6, 1, 3, 6, 12\}$.

Figures 16-20 represent portfolio allocations averaged over the simulations. They depict, for each value of the uncertainty parameter *unc*, the average proportion of wealth invested, in each of the three assets, by Learning and Non-Learning. The purpose is not to compare the graphs with different values of *unc* since the allocation is not performed on the same sample of returns. Rather, we can identify trends that are typically differentiating Learning from Non-Learning allocations.

Since the maximum drawdown constraint is satisfied by the capped sum of total weights that can be invested, the allocations of both Learning and Non-Learning are mainly based on the expected returns of the assets.
Non-Learning, by definition, does not depend on the value of the uncertainty parameter. Hence, no matter the value of *unc*, its allocation is easy to characterize since it saturates its constraint investing in the asset that has the best expected return according to the prior. In our setup, Asset 3 has the highest expected return, so Non-Learning invests only in it and saturates its constraint of roughly 30% during all the investment period. The slight change of the average weight in Asset 3 comes from ρ, the ratio wealth over maximum wealth, changing over time.

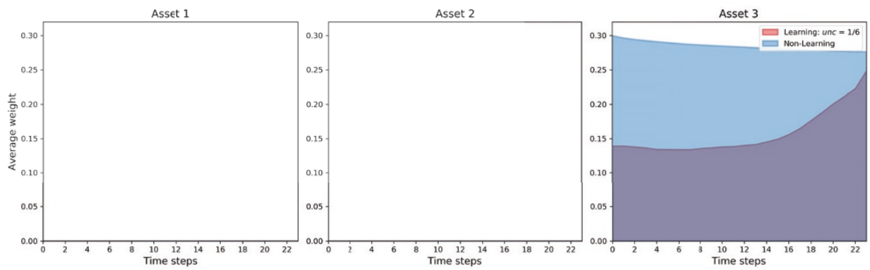

Fig. 16 Learning and Non-Learning historical assets' allocations with $unc = 1/6$.

Unlike Non-Learning, depending of the value of *unc*, Learning can perform more sophisticated allocations because it can adjust the weights according to the incoming information. Nonetheless, in Fig. 16, when *unc* is low, Learning and Non-Learning look similar regarding their weights allocation since both strategies invest, as of time 0, a significant proportion of their wealth only in Asset 3.
On the right panel of Fig. 16, the progressive increase in the weight of Asset 3 illustrates the learning process. As time goes by, Learning progressively increases the weight in Asset 3 since it has the highest expected return. It also explains why Learning underperforms Non-Learning for low values of *unc*; contrary to Non-Learning which invests at full capacity in Asset 3, Learning needs to learn that Asset 3 is the optimal choice.

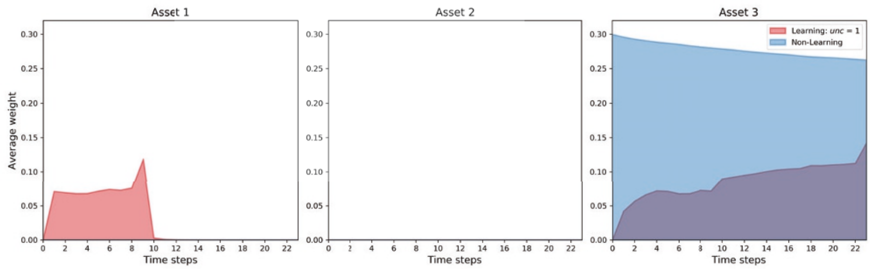

Fig. 17 Learning and Non-Learning historical assets' allocations with $unc = 1$.

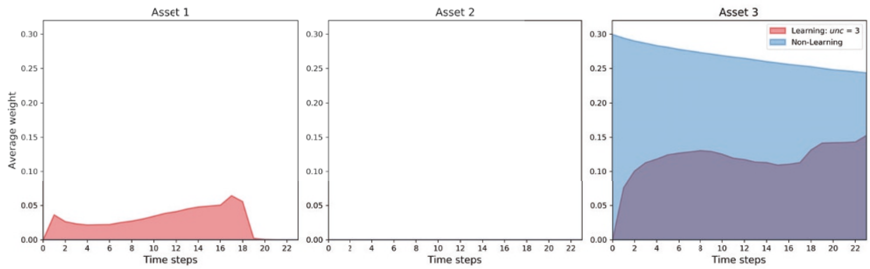

Fig. 18 Learning and Non-Learning historical assets' allocations with $unc = 3$.

However, as uncertainty increases, Learning and Non-Learning strategies start differentiating. When $unc \geq 1$, Learning invests little, if any, at time 0. In addition, an increase in *unc* allows the inital drift to lie in a wider range and generates investment opportunities for Learning. This explains why Learning invests in Asset 1 when $unc = 1, 3, 6, 12$ although the estimate b_0 for this asset is lower than for Asset 3. In Fig. 19, we see that Learning even invests in Asset 2 which has the lowest expected drift.

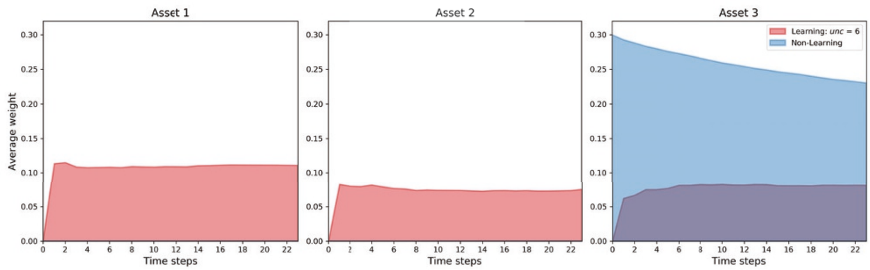

Fig. 19 Learning and Non-Learning historical assets' allocations with $unc = 6$.

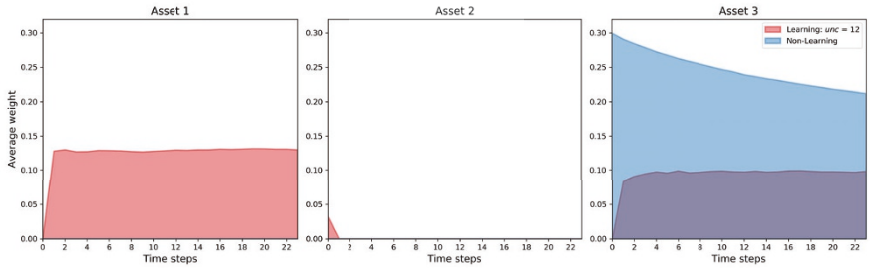

Fig. 20 Learning and Non-Learning historical assets' allocations with $unc = 12$.

Figures 21-25 illustrate the historical total percentage of wealth allocated for Learning and Non-Learning with different levels of uncertainty. As seen previously, Non-Learning has fully invested in Asset 3 for any value of unc.

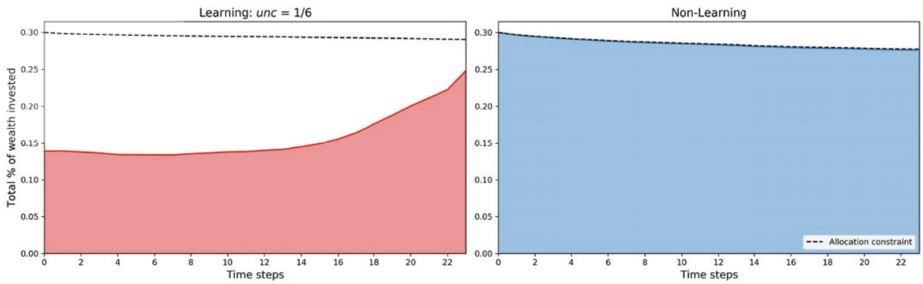

Fig. 21 Historical total allocations of Learning and Non-Learning with $unc = 1/6$.

Moreover, Learning has always less investment that Non-Learning for any level of uncertainty. It suggests that Learning yields a more cautious strategy than Non-Learning. This fact, in addition to its wait-and-see approach at time 0 and its ability to better handle maximum drawdown, makes Learning a safer and more conservative strategy than Non-Learning. This can be seen in Fig. 21, where both Learning and Non-Learning have invested in Asset 3, but not at the same pace. Non-Learning goes fully in Asset 3 at time 0, whereas Learning increments slowly its weight in Asset 3 reaching 25% at the final step. When *unc* is low, there is no value added to choose Learning over Non-Learning from a performance perspective. Nevertheless, Learning allows for a better management of risk as Table 5 exhibits.

As *unc* increases, in addition to being cautious, Learning mixes allocation in different assets, see Figures 22-25, while Non-Learning is stuck with the highest expected return asset.

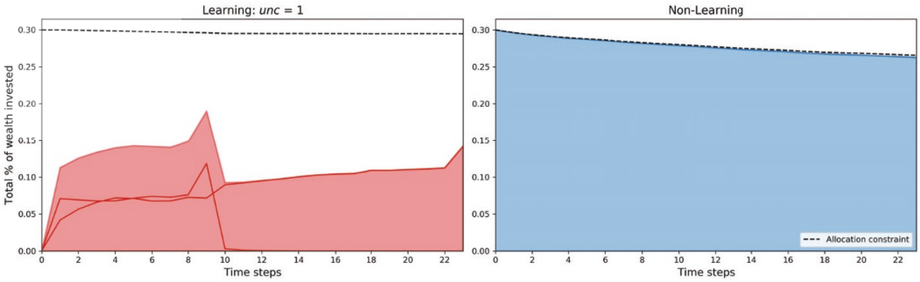

Fig. 22 Historical total allocations of Learning and Non-Learning with $unc = 1$.

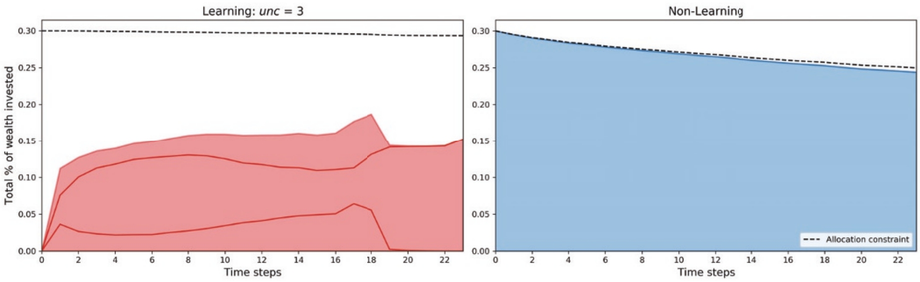

Fig. 23 Historical total allocations of Learning and Non-Learning with $unc = 3$.

Learning is able to be opportunistic and changes its allocation given the prices observed. For example in Fig. 22, Learning starts investing in Asset 1 and 3 at time 1 and stops at time 12 to weigh Asset 1 while keeping Asset 3. Similar remarks can be made for Fig. 23, where Learning puts non negligeable weights in all three risky assets for $unc = 6$ in Fig. 24.

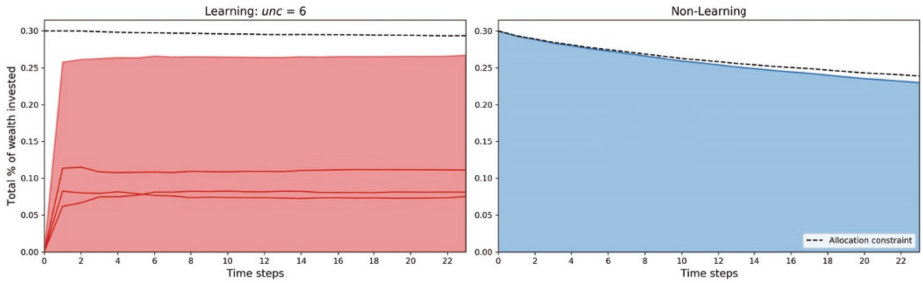

Fig. 24 Historical total allocations of Learning and Non-Learning with $unc = 6$.

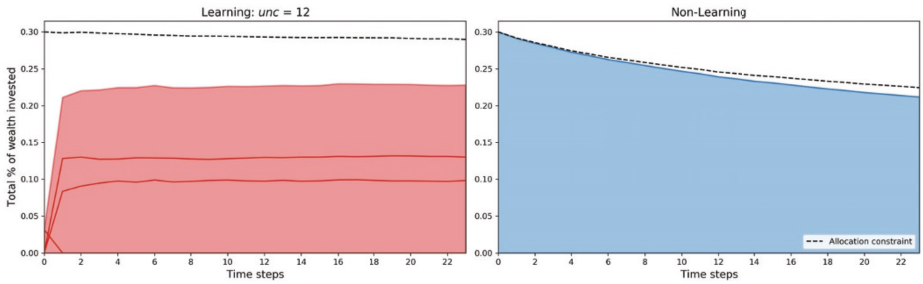

Fig. 25 Historical total allocations of Learning and Non-Learning with $unc = 12$.

6 Conclusion

We have studied a discrete-time portfolio selection problem by taking into account both drift uncertainty and maximum drawdown constraint. The dynamic programming equation has been derived in the general case thanks to a specific change of measure. More explicit results have been provided in the Gaussian case using the Kalman filter. Moreover, a change of variable has reduced the dimensionality of the problem in the case of CRRA utility functions. Next, we have provided extensive numerical results in the Gaussian case with CRRA utility functions using recent deep neural network techniques. Our numerical analysis has clearly shown and quantified the better risk-return profile of the learning strategy versus the non-learning one. Indeed, besides outperforming the non-learning strategy, the learning one provides a significantly lower standard deviation of terminal wealth and a better controlled maximum drawdown. Confirming the results established in [7], this study exhibits the benefits of learning in providing optimal portfolio allocations.

Appendix

6.1 Proof of Proposition 1

For all $k = 1, \ldots, N$, the law under $\overline{\mathbb{P}}$, of R_k given the filtration $\mathcal{G}_{k-1}$ yields the unconditional law under $\mathbb{P}$ of ϵ_k. Indeed, since $(\Lambda_k)_k$ is a $(\mathbb{P}, \mathbb{G})$-martingale, we have from Bayes formula, for all Borelian $F \subset \mathbb{R}^d$,

$$\begin{aligned}
\overline{\mathbb{P}}[R_k \in F|\mathcal{G}_{k-1}] = \overline{\mathbb{E}}[\mathbb{1}_{\{R_k \in F\}}|\mathcal{G}_{k-1}] &= \frac{\mathbb{E}[\Lambda_k \mathbb{1}_{\{R_k \in F\}}|\mathcal{G}_{k-1}]}{\mathbb{E}[\Lambda_k|\mathcal{G}_{k-1}]} \\
&= \mathbb{E}[\frac{\Lambda_k}{\Lambda_{k-1}} \mathbb{1}_{\{R_k \in F\}}|\mathcal{G}_{k-1}] = \mathbb{E}\left[\frac{g(B+\epsilon_k)}{g(\epsilon_k)} \mathbb{1}_{\{R_k \in F\}}|\mathcal{G}_{k-1}\right] \\
&= \int_{\mathbb{R}^d} \frac{g(B+e)}{g(e)} \mathbb{1}_{\{B+e \in F\}} g(e) de = \int_{\mathbb{R}^d} g(z) \mathbb{1}_{\{z \in F\}} dz \\
&= \mathbb{P}[\epsilon_k \in F].
\end{aligned}$$

This means that, under $\overline{\mathbb{P}}$, R_k is independent from B and from $R_1, .., R_{k-1}$ and that R_k has the same probability distribution as ϵ_k. □

6.2 Proof of Proposition 2

For any borelian function $f : \mathbb{R}^d \mapsto \mathbb{R}$ we have, on one hand, by definition of π_{k+1}:

$$\overline{\mathbb{E}}[\overline{\Lambda}_{k+1} f(B)|\mathcal{F}^o_{k+1}] = \int_{\mathbb{R}^d} f(b)\pi_{k+1}(db),$$

and, on the other hand, by definition of $\overline{\Lambda}_k$:

$$\begin{aligned}
\overline{\mathbb{E}}[\overline{\Lambda}_{k+1} f(B)|\mathcal{F}^o_{k+1}] &= \overline{\mathbb{E}}\left[\overline{\Lambda}_k f(B) \frac{g(R_{k+1}-B)}{g(R_{k+1})} \middle| \mathcal{F}^o_{k+1}\right] \\
&= \overline{\mathbb{E}}\left[\overline{\Lambda}_k f(B) g(R_{k+1}-B) \middle| \mathcal{F}^o_{k+1}\right] (g(R_{k+1}))^{-1} \\
&= \int_{\mathbb{R}^d} f(b) \frac{g(R_{k+1}-b)}{g(R_{k+1})} \pi_k(db),
\end{aligned}$$

where we use in the last equality the fact that R_{k+1} is independent of B under $\overline{P}$ (recall Proposition 1). By identification, we obtain the expected relation. □

6.3 Proof of Lemma 1

Since the support of the probability distribution ν of ϵ_k is $\mathbb{R}^d$, we notice that the law of the random vector $Y_k := e^{R_k} - \mathbb{1}_d$ has support equal to $(-1, \infty)^d$. Recall from (7) that $a \in A_k^q(x, z)$ iff

$$1 + a'Y_{k+1} \geq q \max\Big[\frac{z}{x}, 1 + a'Y_{k+1}\Big], \quad a.s. \tag{21}$$

(i) Take some $a \in A_k^q(x, z)$, and assume that $a^i < 0$ for some $i \in [\![1, d]\!]$. Let us then define the event $\Omega_M^i = \{Y_{k+1}^i \geq M, Y_{k+1}^M \in [0, 1], j \neq i\}$, for $M > 0$, and observe that $\mathbb{P}[\Omega_M^i] > 0$. It follows from (21) that

$$1 + a_i M + \max_{j \neq i} |a_j| \geq q\frac{z}{x}, \quad \text{on } \Omega_M^i,$$

which leads to a contradiction for M large enough. This shows that $a^i \geq 0$ for all $i \in [\![1, d]\!]$, i.e. $A_k^q(x, z) \subset \mathbb{R}_+^d$.

(ii) For $\varepsilon \in (0, 1)$, let us define the event $\Omega_\varepsilon = \{Y_{k+1}^i \leq -1 + \varepsilon, i = 1, \ldots, d\}$, which satisfies $\mathbb{P}[\Omega_\varepsilon] > 0$. For $a \in A^q(x, z)$, we get from (21), and since $a \in \mathbb{R}_+^d$ by Step (i):

$$1 - (1 - \varepsilon)a'\mathbb{1}_d \geq q\frac{z}{x}, \quad \text{on } \Omega_\varepsilon.$$

By taking ε small enough, this shows by a contradiction argument that

$$A_k^q(x, z) \subset \Big\{a \in \mathbb{R}_+^d : 1 - a'\mathbb{1}_d \geq q\frac{z}{x}\Big\}. =: \tilde{A}^q(x, z). \tag{22}$$

(iii) Let us finally check the equality in (22). Fix some $a \in \tilde{A}^q(x, z)$. Since the random vector Y_{k+1} is valued in $(-1, \infty)^d$, it is clear that

$$1 + a'Y_{k+1} \geq 1 - a'\mathbb{1}_d \ \geq \ q\frac{z}{x} \ \geq \ 0, \quad a.s.,$$

and thus

$$1 + a'Y_{k+1} \geq q\big[1 + a'Y_{k+1}\big], \quad a.s.,$$

which proves (21), hence the equality $A^q(x, z) = \tilde{A}(x, z)$. □

6.4 Proof of Lemma 2

1. Fix $q_1 \leq q_2$ and $(x, z) \in \mathcal{S}^{q_2} \subset \mathcal{S}^{q_1}$. We then have

$$a \in A^{q_2}(x,z) \;\Rightarrow\; a \in \mathbb{R}_+^d \text{ and } a'\mathbb{1}_d \;\le\; 1 - q_2\frac{z}{x} \;\le\; 1 - q_1\frac{z}{x} \;\Longrightarrow\; a \in A^{q_1}(x,z),$$

which means that $A^{q_2}(x,z) \subseteq A^{q_1}(x,z)$.

2. Fix $q \in (0,1)$, and consider the decreasing sequence $q_n = q + \frac{1}{n}$, $n \in \mathbb{N}^*$. For any $(x,z) \in \mathcal{S}^{q_n}$, we then have $A^{q_n}(x,z) \subseteq A^{q_{n+1}}(x,z) \subset A^a(x,z)$, which implies that the sequence of increasing sets $A^{q_n}(x,z)$ admits a limit equal to

$$\lim_{n\to\infty} A^{q_n}(x,z) = \bigcup_{n\ge 1} A^{q_n}(x,z) \;=\; A^q(x,z),$$

since $\lim_{n\to\infty} q_n = q$. This shows the right continuity of $q \mapsto A^q(x,z)$. Similarly, by considering the increasing sequence $q_n = q - \frac{1}{n}$, $n \in \mathbb{N}^*$, we see that for any $(x,z) \in A^q(x,z)$, the sequence of decreasing sets $A^{q_n}(x,z)$ admits a limit equal to

$$\lim_{n\to\infty} A^{q_n}(x,z) = \bigcap_{n\ge 1} A^{q_n}(x,z) \;=\; A^q(x,z),$$

since $\lim_{n\to\infty} q_n = q$. This proves the continuity in q of the set $A^q(x,z)$.

3. Fix $q \in (0,1)$, and $(x_1,z), (x_2,z) \in \mathcal{S}^q$ s.t. $x_1 \le x_2$. Then,

$$a \in A^q(x_1,z) \;\Longrightarrow\; a \in \mathbb{R}_+^d \text{ and } a'\mathbb{1}_d \le 1 - q\frac{z}{x_1} \le 1 - q\frac{z}{x_2} \;\Longrightarrow\; a \in A^q(x_2,z),$$

which shows that $A^q(x_1,z) \subseteq A^q(x_2,z)$.

4. Fix $q \in (0,1)$, $(x,z) \in A^a(x,z)$. Then, for any a_1, a_2 of the set $A^q(x,z)$, and $\beta \in (0,1)]$, and denoting by $a_3 = \beta a_1 + (1-\beta)a_2 \in \mathbb{R}_+^d$, we have

$$a_3'\mathbb{1}_d \;=\; \beta a_1'\mathbb{1}_d + (1-\beta)a_2'\mathbb{1}_d \le \beta(1 - q\frac{z}{x}) + (1-\beta)(1 - q\frac{z}{x}) \;=\; 1 - q\frac{z}{x}.$$

This proves the convexity of the set $A^q(x,z)$.

4. The homogeneity property of $A^q(x,z)$ is obvious from its very definition. □

6.5 Proof of Lemma 3

We prove the result by backward induction on time k from the dynamic programming equation for the value function.

• At time N, we have for all $\lambda > 0$,

$$v_N(\lambda x, \lambda z, \mu) = \frac{(\lambda x)^p}{p} = \lambda^p v_N(x,z,\mu),$$

which shows the required homogeneity property.

• Now, assume that the homogeneity property holds at time $k+1$, i.e $v_{k+1}(\lambda x, \lambda z, \mu)$ $= \lambda^p v_{k+1}(x,z,\mu)$ for any $\lambda > 0$. Then, from the backward relation (9), and the homogeneity property of $A^q(x,z)$ in Lemma 2, it is clear that v_k inherits from v_{k+1} the homogeneity property. □

6.6 Proof of Lemma 4

1. We first show by backward induction that $r \mapsto w_k(r,\cdot)$ is nondecreasing in on $[q,1]$ for all $k \in [\![0,N]\!]$.

• For any $r_1, r_2 \in [q,1]$, with $r_1 \le r_2$, and $\mu \in \mathcal{M}_+$, we have at time N

$$w_N(r_1,\mu) \;=\; U(r_1)\mu(\mathbb{R}^d) \le U(r_2)\mu(\mathbb{R}^d) \;=\; w_N(r_2,\mu).$$

This shows that $w_N(r,\cdot)$ is nondecreasing on $[q,1]$.

• Now, suppose by induction hypothesis that $r \mapsto w_{k+1}(r,\cdot)$ is nondecreasing. Denoting by $Y_k := e^{R_k} - \mathbb{1}_d$ the random vector valued in $(-1,\infty)^d$, we see that for all $a \in A^q(r_1)$

$$\min\Big[1, r_1(1 + a'Y_{k+1})\Big] \le \min\Big[1, r_2(1 + a'Y_{k+1})\Big], \quad a.s.$$

since $1 + a'Y_{k+1} \ge 1 - a'\mathbb{1}_d \ge q\frac{1}{r_1} \ge 0$. Therefore, from backward dynamic programming Equation (11), and noting that $A^q(r_1) \subset A^q(r_2)$, we have

$$\begin{aligned} w_k(r_1,\mu) &= \sup_{a \in A^q_(r_1)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min[1, r_1(1 + a'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\ &\le \sup_{a \in A^q(r_2)} \overline{\mathbb{E}}\Big[w_{k+1}\big(\min[1, r_2(1 + a'Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \;=\; w_k(r_2,\mu), \end{aligned}$$

which shows the required nondecreasing property at time k.

2. We prove the concavity of $r \in [q,1] \mapsto w_k(r,\cdot)$ by backward induction for all $k \in [\![0,N]\!]$. For $r_1, r_2 \in [q,1]$, and $\lambda \in (0,1)$, we set $r = \lambda r_1 + (1-\lambda)r_2$, and for $a_1 \in A^q(r_1)$, $a_2 \in A^q(r_2)$, we set $a = (\lambda r_1 a_1 + (1-\lambda)r_2 a_2)/r$ which belongs to $A^q(r)$. Indeed, since $a_1, a_2 \in \mathbb{R}^d_+$, we have $a \in \mathbb{R}^d_+$, and

$$a \;=\; \Big(\frac{\lambda r_1 a_1 + (1-\lambda)r_2 a_2}{r}\Big)'\mathbb{1}_d \le \frac{\lambda r_1}{r}\big(1 - \frac{q}{r_1}\big) + \frac{(1-\lambda)r_2}{r}\big(1 - \frac{q}{r_2}\big) \;=\; 1 - \frac{q}{r}.$$

• At time N, for fixed $\mu \in \mathcal{M}_+$, we have

$$\begin{aligned} w_N(\lambda r_1 + (1-\lambda)r_2,\mu) &= U(\lambda r_1 + (1-\lambda)r_2) \\ &\ge \lambda U(r_1) + (1-\lambda)U(r_2) \;=\; \lambda w_N(r_1,\mu) + (1-\lambda)w_N(r_2,\mu), \end{aligned}$$

since U is concave. This shows that $w_N(r,\cdot)$ is concave on $[q,1]$.

• Suppose now the induction hypothesis holds true at time $k+1$: $w_{k+1}(r,\cdot)$ is concave on $[q,1]$. From the backward dynamic programming relation (11), we then have

$$
\begin{aligned}
& \lambda w_k(r_1, \mu) + (1-\lambda) w_k(r_2, \mu) \\
\le\; & \lambda \mathbb{E}\Big[w_{k+1}\big(\min[1, r_1(1 + a_1' Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\
& \qquad\qquad + (1-\lambda)\mathbb{E}\Big[w_{k+1}\big(\min[1, r_2(1 + a_2' Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\
\le\; & \mathbb{E}\Big[w_{k+1}\big(\lambda \min[1, r_1(1 + a_1' Y_{k+1})] + (1-\lambda) \min[1, r_2(1 + a_2' Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \\
=\; & \mathbb{E}\Big[w_{k+1}\big(\min[1, r(1 + a' Y_{k+1})], \bar{g}(R_{k+1} - \cdot)\mu\big)\Big] \;\le\; w_k(r, \mu),
\end{aligned}
$$

where we used for the second inequality, the induction hypothesis joint with the concavity of $x \mapsto \min(1, x)$, and the nondecreasing monotonicity of $r \mapsto w_{k+1}(r, \cdot)$. This shows the required inductive concavity property of $r \mapsto w_k(r, \cdot)$ on $[q, 1]$. □

References

[1] Achref Bachouch, Côme Huré, Nicolas Langrené, and Huyên Pham. Deep neural networks algorithms for stochastic control problems on finite horizon, part 2: numerical applications. *Methodology and Computing in Applied Probability*, https://doi.org/10.1007/s11009-019-09767-9, 2021.

[2] Alexis Bismuth, Olivier Guéant, and Jiang Pu. Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty. *Mathematics and Financial Economics*, 13(4):661–719, 2019.

[3] Stephen Boyd, Erik Lindström, Henrik Madsen, and Peter Nystrup. Multi-period portfolio selection with drawdown control. *Annals of Operations Research*, 282(1-2):245–271, 2019.

[4] Jaksa Cvitanić and Ioannis Karatzas. On portfolio optimization under "drawdown" constraints. *Constraints, IMA Lecture Notes in Mathematics & Applications 65*, 1994.

[5] Jakša Cvitanić, Ali Lazrak, Lionel Martellini, and Fernando Zapatero. Dynamic portfolio choice with parameter uncertainty and the economic value of analysts? recommendations. *The Review of Financial Studies*, 19(4):1113–1156, 2006.

[6] Carmine De Franco, Johann Nicolle, and Huyên Pham. Bayesian learning for the markowitz portfolio selection problem. *International Journal of Theoretical and Applied Finance*, 22(07), 2019.

[7] Carmine De Franco, Johann Nicolle, and Huyên Pham. Dealing with drift uncertainty: a bayesian learning approach. *Risks*, 7(1):5, 2019.

[8] Romuald Elie and Nizar Touzi. Optimal lifetime consumption and investment under a drawdown constraint. *Finance and Stochastics*, 12(3):299, 2008.

[9] Robert J Elliott, Lakhdar Aggoun, and John B Moore. *Hidden Markov Models: Estimation and Control*. Springer, 2008.

[10] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2019.

[11] Sanford J Grossman and Zhongquan Zhou. Optimal investment strategies for controlling drawdowns. *Mathematical finance*, 3(3):241–276, 1993.

[12] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, 1991.

[13] Côme Huré, Huyên Pham, and Nicolas Bachouch, Achref a Langrené. Deep neural networks algorithms for stochastic control problems on finite horizon, part 1: convergence analysis. *SIAM Journal of Numerical Analysis*, 59(1):525–557, 2021.

[14] R. E. Kalman and R. S. Bucy. New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(1):95–108, 03 1961.

[15] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82:35–45, 1960.

[16] Ioannis Karatzas and X. Zhao. Bayesian Adaptative Portfolio Optimization. In *Option Pricing, Interest Rates and Risk Management*. Cambridge University Press, 2001.

[17] Jussi Keppo, Hong Ming Tan, and Chao Zhou. Investment decisions and falling cost of data analytics. 2018.

[18] Peter Lakner. Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications*, 76(1):77–97, 1998.

[19] Imke Redeker and Ralf Wunderlich. Portfolio optimization under dynamic risk constraints: Continuous vs. discrete time trading. *Statistics & Risk Modeling*, 35(1-2):1–21, 2018.

[20] L Chris G Rogers. The relaxed investor and parameter uncertainty. *Finance and Stochastics*, 5(2):131–154, 2001.

Estimating the Matthew Effects: Switching Pareto Dynamics

Robert J. Elliott

Abstract Pareto distributions can describe the clustering of observations and give rise to sayings such as 'The rich gets richer and the poor gets poorer'. They are sometimes generated by counting processes whose rate depends on external factors. In turn, these factors are modelled by a finite state Markov chain Z. New filters are derived which estimate Z together with other parameters of the model.

1 Introduction

The Matthew effect is paraphrased by saying 'The rich get richer and the poor get poorer'. A probability describing such a distribution can be given by a power law of which an example is the Pareto distribution.

If X is a real random variable with a Pareto distribution then there is a (positive) value x_m and a parameter $\alpha > 0$ such that

$$P(X > x) = \begin{cases} (\frac{x_m}{x})^\alpha & \text{if} \quad x \geq x_m \\ 1 & \text{if} \quad x < x_m \,. \end{cases}$$

It is immediate that if X is Pareto with parameter α then

$$Y = \log\left(\frac{X}{x_m}\right)$$

is exponentially distributed with parameter α. In fact

Robert J. Elliott
School of Business, University of South Australia, Haskayne School of Business, University of Calgary,
e-mail: relliott@ucalgary.ca

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_6

$$\begin{aligned} P(Y < y) &= P\Big(\log\big(\frac{X}{x_m}\big) < y\Big) \\ &= P(X < x_m e^y) \\ &= \Big(1 - \big(\frac{x_m}{x_m e^y}\big)^\alpha\Big) = 1 - e^{-\alpha y}. \end{aligned}$$

Conversely, if Y is exponentially distributed then $x_m e^y$ is Pareto distributed with minimum x_m and index α.

2 Generating Pareto Random Variables

Our processes are defined on $(\Omega, \mathcal{F}, \overline{P})$. Consider a (counting) point process

$$Y = \{Y_t, t \geq 0\}$$

with jump times $\tau_1, \tau_2, \tau_3, \dots$.

Write $\tau_0 = 0$.

Suppose the compensator of Y is λt, that is λ is the rate of jumping. Write

$$\begin{aligned} \mathcal{Y}_t &= \sigma\{Y_s : s \leq t\} \quad \text{and} \\ \mathcal{Y} &= \{\mathcal{Y}_t\} \end{aligned}$$

for the right continuous, complete filtration generated by Y. Then with

$$\begin{aligned} Y_t &:= \sum_n I_{t \geq \tau_n} \\ \overline{Q}_t &:= Y_t - \lambda t \quad \text{is a } (\mathcal{Y}, \overline{P}) \quad \text{martingale.} \end{aligned}$$

The times between the jumps of Y are independent and identically exponentially distributed with parameter λ.

That is, for each n, $\Delta_{n+1} := \tau_{n+1} - \tau_n \sim \exp \lambda \Delta_{n+1}$. This generates a family of i.i.d. exponential random variables $\Delta_1, \Delta_2, \dots$.

Consequently, $x_m \exp \Delta_1, x_m \exp \Delta_2, \dots$ is a family of Pareto random variables with minimum x_m and parameter λ.

3 Switching Parameter Values

Suppose the parameter λ is not constant but can switch between values $\alpha_1, \alpha_2, \dots, \alpha_N$. In a simple case perhaps $N = 2$ so there are just two values α_1, α_2. The value α_i is determined by some 'state' of the market, or the property of some Reddit post.

Suppose there is a finite, N, state Markov chain $Z = \{Z_t, t \geq 0\}$ which represents the (hidden) state loss of generality the state space of Z can be identified with unit vectors

$$S = \{e_1, e_2, \ldots, e_N\}, \text{ where } e_i = (0, 0, \ldots, 1, 0, \ldots, 0)' \in R^N.$$

That is for each $t > 0$, $Z_t \in S$. Suppose the rate matrix of Z is given by the matrix $A = (a_{ji}, 1 \le i, j \le N)$. Here a_{ji}, $j \ne i$, is the rate of jumping from e_i to e_j. Then, (see [2]), Z has the semimartingale representation

$$Z_t = Z_0 + \int_0^t AZ_s ds + M_t \in R^N$$

where M is a (vector) martingale. That is, if $\mathcal{F}_t = \sigma\{Z_s, s \le t\}$ and $\mathcal{F} = \{\mathcal{F}_t\}$ is the right continuous complete filtration generated by Z, then for $s \le t$

$$\overline{E}[M_t | \mathcal{F}_s] = M_s \in R^N.$$

For our counting process we now suppose that at time t the rate is

$$\alpha_t = \langle \alpha, Z_t \rangle.$$

Then, with $\mathcal{G} = \{\mathcal{G}_t\}$ has filtration generated by Y and Z, the process

$$Q_t := Y_t - \int_0^t \langle \alpha, Z_s \rangle ds$$

is a martingale.

4 Estimation

The problem now is: suppose the counting process Y, or equivalently the jump times $\tau_1, \tau_2, \ldots$, are observed. We wish to estimate the state of Z and the parameters in $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)'$.

A Filter

We shall use a 'reference probability' $\overline{P}$. Suppose that under $\overline{P}$

1) Z is a Markov chain with rate matrix A
2) Y is a counting process with compensation λt.

Then as in Section 1, $\overline{Q}_t = Y_t - \lambda t$ is a $\overline{P}$ martingale.

Definition 1 Write

$$\Lambda_t = 1 + \int_0^t \Lambda_{s-}\Big(\frac{\langle \alpha, Z_{s-} \rangle}{\lambda} - 1\Big)(dY_s - \lambda ds) \tag{1}$$

so:

$$\Lambda_t = \exp\Big(-\int_0^t \big(\frac{\langle \alpha, Z_{s-}\rangle}{\lambda} - 1\big)\lambda ds + \int_0^t \log \frac{\langle \alpha, Z_{s-}\rangle}{\lambda} dY_s\Big).$$

Note that Λ is a $(\overline{P}, \mathcal{G})$ martingale. Define a probability measure P by

$$\frac{dP}{d\overline{P}}\Big|_{\mathcal{G}_t} = \Lambda_t\,.$$

Theorem 1 *Write*

$$Q_t = Y_t - \int_0^t \langle \alpha, Z_s\rangle ds.$$

Then $Q = \{Q_t\}$ is a $(P, \mathcal{G})$ martingale.

Proof From [1], Lemma 15.2.1 this is so if and only if $\Lambda_t Q_t$ is a $(P, \mathcal{G})$ martingale. Now

$$\begin{aligned}\Lambda_t Q_t &= \int_0^t \Lambda_{s-} dQ_s + \int_0^t Q_{s-} d\Lambda_s + \sum_{0\le s\le t} \Delta\Lambda_s \Delta Q_s \\ &= \int_0^t Q_{s-} d\Lambda_s + \int_0^t \Lambda_{s-} \frac{\langle \alpha, Z_{s-}\rangle}{\lambda}(dY_s - \lambda ds).\end{aligned}$$

We wish to obtain a recursive estimate for

$$E[Z_t|\mathcal{Y}_t] \in R^N.$$

By Bayes' Theorem, (see [2], Theorem 3.2),

$$E[Z_t|Y_t] = \frac{\overline{E}[\Lambda_t Z_t|\mathcal{Y}_t]}{\overline{E}[\lambda_t|\mathcal{Y}_t]}.$$

Write

$$q_t = \overline{E}[\Lambda_t Z_t|Y_t] \in R^N.$$

This is an unnormalized conditional expected value of Z_t given $\mathcal{Y}_t$.

If $1 = (1, 1, \ldots, 1)' \in R^N$ is a vector of 1s then

$$\langle Z_t, 1\rangle = 1.$$

Consequently

$$\begin{aligned}\langle q_t, 1\rangle &= \overline{E}[\Lambda_t \langle Z_t, 1\rangle\,|\mathcal{Y}_t] \\ &= \overline{E}[\Lambda_t\,|\mathcal{Y}_t]\end{aligned}$$

which gives the normalizing.

Notation Write $\mathrm{diag}(\frac{\alpha_i}{\lambda})$ for the $N \times N$ diagonal matrix with $(\frac{\alpha_1}{\lambda}, \frac{\alpha_2}{\lambda}, \ldots, \frac{\alpha_N}{\lambda})$ down the diagonal. □

Theorem 2 *q satisfies the recursion*

$$q_t = q_0 + \int_0^t Aq_s ds + \int_0^t \Big(\text{diag}(\frac{\alpha_i}{\lambda}) - I\Big) q_{s-}(dY_s - \lambda ds).$$

Proof Recall $dZ_t = AZ_t dt + dM_t \in R^N$. Then

$$\Lambda_t Z_t = \Lambda_0 Z_0 + \int_0^t \Lambda_{s-} dZ_s + \int_0^t (d\Lambda_s) Z_{s-}.$$

(Note there are no common jump terms.)

From (1) this is:

$$= \Lambda_0 Z_0 + \int_0^t \Lambda_{s-}(AZ_s ds + dM_s) + \int_0^t Z_{s-}\Lambda_{s-}\Big(\frac{\langle \alpha, Z_{s-}\rangle}{\lambda} - 1\Big)(dY_S - \lambda ds)$$
$$= \Lambda_0 Z_0 + \int_0^t A\Lambda_s Z_s ds + \int_0^t \Lambda_{s-} dM_s + \int_0^t \Lambda_{s-} Z_{s-}\Big(\frac{\langle \alpha, Z_{s-}\rangle}{\lambda} - 1\Big)(dY_s - \lambda ds).$$

We now take a conditional expectation under $\overline{P}$ given $\mathcal{Y}_t$ and obtain

$$q_t = q_0 + \int_0^t Aq_s ds + \int_0^t \Big(\text{diag}(\frac{\alpha_i}{\lambda}) - I\Big) q_s(dY_s - \lambda ds).$$

5 Parameter Estimation

We have seen that to change the rate from λ to $\langle \alpha, Z_t\rangle$ the Girsanov density given by (1) is used.

Write this density as Λ_t^α and the related probability P^α.

Suppose there is a second possible set of parameter values

$$\alpha' = (\alpha'_1, \alpha_2^0, \dots, \alpha'_N)' \in R^N$$

giving a related probability $P^{\alpha'}$.

Then the Girsanov density $\frac{\Lambda_t^\alpha}{\Lambda_t^{\alpha'}}$ will change the probability $P^{\alpha'}$ to P^α and the compensator of Y from $\langle \alpha', Z_t\rangle$ to $\langle \alpha, Z_t\rangle$.

Suppose the model has been implemented with a parameter set $\{A = (a_{ji}), \alpha' = (\alpha'_1, \dots, \alpha'_N)'\}$.

Given the observations of Y we wish to re-estimate the parameters in α'. The conditional expectation of the log-likelihood to change parameters α' to α is

$$E\Big[\log \frac{\Lambda_t^\alpha}{\Lambda_t^{\alpha'}} \big| \mathcal{Y}_t\Big] = E\Big[-\int_0^t \Big(\frac{\langle \alpha, Z_s\rangle}{\lambda} - 1\Big)\lambda ds + \int_0^t \log \frac{\langle \alpha, Z_{s-}\rangle}{\lambda} dY_s | \mathcal{Y}_t\Big] + R$$

where R represents terms which do not depend on α.

In turn, this is

$$= E\Big[- \int_0^t \langle \alpha, Z_s \rangle + \int_0^t \log \langle \alpha, Z_{s-} \rangle, dY_s | \mathcal{Y}_t \Big]$$
$$+ \text{ terms which do not depend on } \alpha.$$

Write $J_t^i = \int_0^t \langle e_i, Z_s \rangle ds$ for the amount of time Z has spent in state e_i upto time t. Also,note

$$\int_0^t \log \langle \alpha, Z_{s-} \rangle \, dY_s = \sum_{i=1}^N \log \alpha_i \int_0^t \langle e_i, Z_{s-} \rangle \, dY_s$$

so

$$E\Big[\log \frac{\Lambda_t^{\alpha}}{\lambda_t^{\alpha'}} \Big| \mathcal{Y}_t \Big] = E\,[- \sum_{i=1}^N \alpha_i J_t^i + \sum_{i=1}^N \log \, \alpha_i \int_0^t \langle e_i, Z_{s-} \rangle \, dY_s \, | \mathcal{Y}_t \Big]$$
$$+ \text{ terms which do not depend on } \alpha.$$

To find the α_i which maximizes this conditional expected log-likelihood the first order condition gives

$$E[-J_t^i | \mathcal{Y}_t] + \frac{1}{\alpha_i} E\Big[\int_0^t \langle e_i, Z_{s-} \rangle \, dY_s \, | \mathcal{Y}_t \Big] = 0$$

so the maximizing α_i is given as

$$\alpha_i = E\Big[\int_0^t \langle e_i, Z_{s-} \rangle \, dY_s | \mathcal{Y}_t \Big] \Big/ E[J_t^i | \mathcal{Y}_t].$$

Our final task is to find expressions for these quantities. Write Λ_t for $\Lambda_t^{\alpha'}$. Recall

$$E[J_t^i | \mathcal{Y}_t] = \frac{\overline{E}[\Lambda_t J_t^i | \mathcal{Y}_t]}{\overline{E}[\Lambda_t | \mathcal{Y}_t]}.$$

Also from Section 4

$$\langle q_t, \mathbf{1} \rangle = \overline{E}[\Lambda_t | \mathcal{Y}_t].$$

Notation For any process $H = \{H_t, t \geq 0\}$ write

$$\sigma(H)_t = \overline{E}[\Lambda_t H_t | \mathcal{Y}_t].$$

Then

$$E[J_t^i | \mathcal{Y}_t] = \sigma(J)_t \Big/ \langle q_t, 1 \rangle.$$

6 Recursive Estimates

Rather than $\sigma(J)_t$ we shall initially obtain a recursive estimate for

$$\sigma(JZ)_t = \overline{E}\Big[\Lambda_t J_t Z_t | \mathcal{Y}_t] \in R^N .$$

As $\langle Z_t, 1\rangle = 1$ we then would have

$$\begin{aligned}\langle\sigma(JZ)_t, 1\rangle &= \overline{E}[\Lambda_t J_t \langle Z_t, 1\rangle | \mathcal{Y}_t] \\ &= \overline{E}[\Lambda_t J_t | \mathcal{Y}_t] \\ &= \sigma(J)_t .\end{aligned}$$

Theorem 3 *A recursive estimate for $\sigma(J^i Z)_t$ is given by*

$$\begin{aligned}\sigma(J^i Z)_t = \int_0^t A\sigma(J^i Z)_s ds + \int_0^t \Big(\mathrm{diag}(\frac{\alpha}{\lambda}) - I\Big)\sigma(J^i Z)_{s-}(dY_s - \lambda ds) \\ + \int_0^t \langle q_s, e_i\rangle\, ds\, e_i .\end{aligned}$$

Proof Recall

$$dZ_t = AZ_t dt + dM_t \in R^N \tag{2}$$

$$dJ_t^i = \int_0^t \langle e_i Z_s\rangle ds \tag{3}$$

and

$$d\Lambda_t = \Lambda_{t-}\Big(\frac{\langle\alpha, Z_{t-}\rangle}{\lambda} - 1\Big)(dY_t - \lambda dt). \tag{4}$$

Then

$$\begin{aligned}d(\Lambda_t J_t^i) &= \Lambda_t dJ_t^i + J_t^i d\Lambda_t \\ &= \Lambda_t\langle e_i, Z_t\rangle dt + J_t^i \Lambda_{t-}\Big(\frac{\langle\alpha, Z_{t-}\rangle}{\lambda} - 1\Big)(dY_t - \lambda dt)\end{aligned}$$

and

$$\begin{aligned}d(\Lambda_t J_t^i Z_t) &= (\Lambda_{t-} J_t^i) dZ_t + d(\Lambda_t J_t^i) Z_t \\ &= \Lambda_{t-} J_t^i (AZ_t dt + dM_t) + \Lambda_t \langle e_i, Z_t\rangle Z_t dt \\ &\quad + \Lambda_{t-} J_t^i \Big(\frac{\langle\alpha, Z_{t-}\rangle}{\lambda} - 1\Big) Z_{t-}(dY_t - \lambda dt) \\ &= A(\Lambda_t J_t^i Z_t) dt + (\Lambda_{t-} J_t^i) dM_t + \langle e_i, \Lambda_t Z_t\rangle dt\, e_i \\ &\quad + \sum_{i=1}^N J_t^i \Lambda_{t-}\langle e_i, Z_{t-}\rangle \Big(\frac{\alpha_i}{\lambda} - 1\Big)(dY_t - \lambda dt) e_i .\end{aligned}$$

That is:

$$\Lambda_t J_t^i Z_{t-} = A\int_0^t \Lambda_s J_s^i Z_s\,ds + \int_0^t \Lambda_{s-} J_{s-}^i\,dM_s$$
$$+ \int_0^t \langle e_i, \Lambda_s Z_s\rangle\,ds\,e_i$$
$$+ \sum_{i=1}^N \int_0^t J_s^i \Lambda_{s-}\langle e_i, Z_{s-}\rangle\Big(\frac{\alpha_i}{\lambda} - 1\Big)(dY_s - \lambda ds)e_i\,.$$

Taking the conditional expectation under $\overline{P}$ given $\mathcal{Y}_t$ gives

$$\sigma(J^i Z) = \int_0^t A\sigma(J^i Z)_s + \int_0^t \langle e_i, q_s\rangle\,ds\,e_i$$
$$+ \int_0^t \Big(\operatorname{diag}\big(\frac{\sigma}{\lambda}\big) - I\Big)\sigma(J^i Z)_{s-}(dY_s - \lambda ds).$$

Write

$$H_t^i := \int_0^t \langle e_i, Z_{s-}\rangle\,dY_s\,.$$

Theorem 4 *A recursive estimate for $\sigma(H^i Z)_t$ is given by*

$$\sigma(H^i Z)_t = \int_0^t A\sigma(H^i Z)_s ds$$
$$+ \int_0^t \Big(\operatorname{diag}\big(\frac{\alpha}{\lambda}\big) - I\Big)\sigma(H^i Z_{s-})(\lambda Y_{s-}\lambda ds) + \int_0^t \langle e_i, q_{s-}\rangle\,dY_s e_i\,.$$

Proof

$$dH_t^i = \langle e_i, Z_{t-}\rangle\,dY_t \tag{5}$$

so from (2) and (4)

$$d(\Lambda_t H_t^i) = \Lambda_{t-}\langle e_i, Z_{t-}\rangle dY_t + H_{t-}^i \Lambda_{t-}\Big(\frac{\langle \alpha, Z_{t-}\rangle}{\lambda} - 1\Big)(dY_t - \lambda dt)$$
$$+ \langle e_i, Z_{t-}\rangle \Lambda_{t-}\Big(\frac{\langle \alpha, Z_{t-}\rangle}{\lambda} - 1\Big)dY_t\,.$$

Then

$$d(\Lambda_t H_t^i Z_t) = \Lambda_{t-} H_{t-}^i (AZ_t dt + dM_t)$$
$$+ \Lambda_{t-}\langle e_i, Z_{t-}\rangle Z_{t-} dY_t + H_{t-}^i \Lambda_{t-} Z_{t-}\Big(\frac{\langle \alpha, Z_{t-}\rangle}{\lambda} - 1\Big)(dY_t - \lambda dt)$$
$$+ \langle e_i, Z_{t-}\rangle \Lambda_{t-} Z_{t-}\Big(\frac{\langle \alpha, Z_{t-}\rangle}{\lambda} - 1\Big)dY_t\,.$$

That is:

$$\Lambda_t H_t^i Z_t = \int_0^t A\Lambda_s H_s^i Z_s ds + \int_0^t \Lambda_{s-} H_{s-}^i dM_s$$
$$+ \int_0^t \langle e_i, \Lambda_{s-} Z_{s-}\rangle dY_s e_i + \int_0^t \langle e_i, \Lambda_{s-} Z_{s-}\rangle \Big(\frac{\alpha_i}{\lambda} - 1\Big) dY_s e_i$$
$$+ \int_0^t \Big(\mathrm{diag}(\frac{\alpha}{\lambda}) - I\Big)\Lambda_{s-} H_{s-}^i Z_{s-}(dY_s - \lambda ds).$$

Taking the conditional expectation under $\overline{P}$ given $\mathcal{Y}_t$ gives

$$\sigma(H^i Z)_t = \int_0^t A\sigma(H^i Z)_s + \int_0^t \langle e_i, q_{s-}\rangle dY_s e_i$$
$$+ \int_0^t \Big(\mathrm{diag}(\frac{\alpha}{\lambda}) - I\Big)\sigma(H^i Z)_{s-}(dY_s - \lambda ds).$$

Remark 1 Again $\sigma(H^i)_t = \langle \sigma(H^i Z)_t, \mathbf{1}\rangle$ and

$$E[H_t^i | \mathcal{Y}_t] = \sigma(H^i)_t \Big/ \langle q_t, \mathbf{1}\rangle .$$

7 Implemention

Suppose we receive a sequence of data values $d_1, d_2, \ldots$ which we believe to be Pareto distributed with parameter α and minimum x_m. Then

$$\Delta_1 = \log \frac{d_1}{x_m},\ \Delta_2 = \log \frac{d_2}{x_m}, \ldots$$

are exponentially distributed with parameter α.

The Δ_i can be considered as inter-arrival times of a point process Y as above, that is

$$\Delta_i = \tau_i - \tau_{i-1} .$$

Between arrivals $dY_t = \Delta Y_t = 0$.

At an arrival time $dY_t = \Delta Y_t = 1$. The above theory can be used to estimate possible values of α.

Note for example

$$\int_{\tau_{i-1}}^{\tau_i} \gamma_{s-} dY_s = \gamma_{\tau_i -}$$

and in between jumps $dY_s = 0$.

References

1. Cohen, S. and Elliott, R.J.: Stochastic Calculus and Applications, 2^{nd} edition. New York-London: Birkhäuser-Springer, xxiii + 666 pp. (2015).

2. Elliott, R.J., Aggoun, L. and Moore, J.B.: (1994) Hidden Markov Models: Estimation and Control, (Applications of Mathematics **29**). Berlin-Heidelberg-New York: Springer Verlag, 350 pp. (1994).

Optimal Couplings on Wiener Space and An Extension of Talagrand's Transport Inequality

Hans Föllmer

Abstract For a probability measure Q on Wiener space, Talagrand's transport inequality takes the form $W_{\mathscr{H}}(Q,P)^2 \leq 2H(Q|P)$, where the Wasserstein distance $W_{\mathscr{H}}$ is defined in terms of the Cameron-Martin norm, and where $H(Q|P)$ denotes the relative entropy with respect to Wiener measure P. Talagrand's original proof takes a bottom-up approach, using finite-dimensional approximations. As shown by Feyel and Üstünel in [3] and Lehec in [10], the inequality can also be proved directly on Wiener space, using a suitable coupling of Q and P. We show how this top-down approach can be extended beyond the absolutely continuous case $Q \ll P$. Here the Wasserstein distance is defined in terms of quadratic variation, and $H(Q|P)$ is replaced by the specific relative entropy $h(Q|P)$ on Wiener space that was introduced by N. Gantert in [7].

1 Introduction

There are many ways of quantifying the extent to which a probability measure Q on the path space $C[0,1]$ deviates from Wiener measure P. In this paper we discuss the following two approaches and the relation between them. One involves the notion of entropy, the other uses a Wasserstein distance, that is, the solution of an optimal transport problem on Wiener space. We will do this in two stages.

In the first stage, the measure Q will be absolutely continuous with respect to Wiener measure P, and we consider the relative entropy $H(Q|P)$ of Q with respect to P. On the other hand, we use the Wasserstein distance

$$W_{\mathscr{H}}(Q,P) = \inf\Big(\int ||\omega - \eta)||_{\mathscr{H}} P(d\omega)R(\omega,d\eta)\Big)^{1/2}, \tag{1}$$

Humboldt-Universität zu Berlin, Institut für Mathematik, Unter den Linden 6, 10099 Berlin, Germany, e-mail: foellmer@math.hu-berlin.de

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_7

where the infimum is taken over all transition kernels R on Wiener space which transport P into Q, and where the transportation cost is defined by the Cameron-Martin norm. Talagrand's transport inequality

$$W_{\mathscr{H}}(Q,P) \leq \sqrt{2H(Q|P)} \tag{2}$$

on Wiener space shows that these two measures of deviation are closely related. In fact, inequality (2) becomes an identity as soon as we introduce the additional constraint that the transport should be adapted to the natural filtration on Wiener space; this was first shown by R. Lassalle in [9].

On Wiener space, inequality (2) was first studied by Feyel and Üstünel [3]. In Talagrand's original version [13], the inequality is formulated on Euclidean space $\mathbb{R}^n$, including the case $n=\infty$; the Wasserstein distance is defined in terms of the Euclidean norm, and the reference measure P is the product of standard normal distributions. But the Lévy-Ciesielski construction of Brownian motion in terms of the Schauder functions shows that inequality (2) on Wiener space can be viewed as a direct translation of the Euclidean case for $n=\infty$, as explained in Section 3.

Talagrand's original proof in [13] takes a bottom-up approach, using finite-dimensional approximations. Instead, as shown by D. Feyel and A. S. Üstünel in [3] and by J. Lehec in [10], Talagrand's inequality can be proved directly on Wiener space, using a suitable coupling of Q and P. This top-down approach involves the computation of relative entropy in terms of the *intrinsic drift* of Q that was used in [4] and [5] for the analysis of time reversal and large deviations on Wiener space. The intrinsic drift b^Q is such that Q can be viewed as a weak solution of the stochastic differential equation $dW = dW^Q + b^Q(W)dt$, that is, W^Q is a Wiener process under Q. Coupling W^Q with the coordinate process W under Q immediately yields inequality (2), and it solves the optimal transport problem for the Cameron-Martin norm if the coupling is required to be adapted.

Clearly, inequality (2) is of interest only if the relative entropy is finite, and so Q should be absolutely continuous with respect to Wiener measure. In the second stage, we go beyond this restriction. Here we replace $H(Q|P)$ by the *specific relative entropy*

$$h(Q|P) := \liminf_{N\uparrow\infty} 2^{-N} H_N(Q|P),$$

where $H_N(Q|P)$ denotes the relative entropy of Q with respect to P on the σ-field generated by observing the path along the N-th dyadic partition of the unit interval. The notion of specific relative entropy on Wiener space was introduced by N. Gantert in her thesis [7], where it serves as a rate function for large deviations of the quadratic variation from its ergodic behaviour; cf. also [8]. In our context, the specific relative entropy appears if we rewrite the finite-dimensional Talagrand inequality for $n=2^N$ in the form

$$W_N^2(Q,P) \leq 2 \cdot 2^{-N} H_N(Q|P), \tag{3}$$

where the Wasserstein metric W_N is defined in terms of the discrete quadratic variation along the N-th dyadic partition. This suggests that a passage to the limit should yield an extension of Talagrand's inequality, where $H(Q|P)$ is replaced by $h(Q|P)$, and where $W_{\mathcal{H}}$ is replaced by a Wasserstein metric $W_{\mathcal{S}}$ that is defined in terms of quadratic variation. Here again, we take a top-down approach. Instead of analyzing the convergence on the left-hand side of (3), we argue directly on Wiener space, assuming that the coordinate process W is a special semimartingale under Q. We show that $h(Q|P) < \infty$ implies that Q admits the construction of an intrinsic Wiener process W^Q such that the pair (W, W^Q) defines a coupling of P and Q. This coupling solves the optimal transport problem defined by $W_{\mathcal{S}}$, and for a martingale measure Q it yields the inequality

$$W_{\mathcal{S}}(Q,P) \leq \sqrt{2h(Q|P)}. \tag{4}$$

If, more generally, Q is a semimartingale measure that admits a unique equivalent martingale measure Q^*, then we obtain the following extension of Talagrand's inequality on Wiener space:

$$W_{\mathcal{S}}(Q|P)^2 \leq 2\big(h(Q|P) + H(Q|Q^*)\big). \tag{5}$$

In this form, inequality (5) includes both (4) and Talagrand's inequality (2) as special cases.

The paper is organized as follows. In Section 2 we introduce the basic concepts of relative entropy and of a Wasserstein distance. Section 3 describes the top-down approach to inequality (2) in the absolutely continuous case; the exposition will be reasonably self-contained because we repeatedly refer to it in the sequel. In the second stage, we consider measures Q on $C[0,1]$ such that the coordinate process W is a semimartingale under Q. Section 4 shows how the semimartingale structure of Q is reflected in the specific relative entropy $h(Q|P)$; this extends Theorem 1.2 in [7] for martingale measures to the general case. In section 5 we show that the condition $h(Q|P) < \infty$ implies that Q admits the construction of an intrinsic Wiener process W^Q. Coupling W^Q with the coordinate process W under Q, we obtain the solution of an optimal transport problem on Wiener space that yields inequalities (4) and (5).

2 Preliminaries

In this section we recall some basic notions, in particular the definitions of relative entropy and of the Wasserstein distances that we are going to use.

For two probability measures μ and ν on some measurable space $(S, \mathcal{S})$, the *relative entropy* of ν with respect to μ is defined as

$$H(\nu|\mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

For $\nu \ll \mu$ we can write

$$H(\nu|\mu) = \int h(\frac{d\nu}{d\mu})d\mu,$$

denoting by h the strictly convex function $h(x) = x\log x$ on $[0,\infty)$, and Jensen's inequality implies $H(\nu|\mu) \geq 0$, with equality if and only if $\mu = \nu$. Sometimes we will deal with different σ-fields $\mathscr{S}$ on the same space S, and then we will also use the notation $H_{\mathscr{S}}(\nu|\mu)$. We are going to use repeatedly the fact that

$$\lim_{n\uparrow\infty} H_{\mathscr{S}_n}(\nu|\mu) = H_{\mathscr{S}}(\nu|\mu) \tag{6}$$

if $(\mathscr{S}_n)_{n=1,2...}$ is a sequence of σ-fields increasing to $\mathscr{S}$.

Consider a measurable cost function $c(\cdot,\cdot)$ on $S\times S$ with values in $[0,\infty]$; typically, $c(\cdot,\cdot)$ will be a metric on S. We define the corresponding *Wasserstein distance* between ν and μ as

$$W(\nu,\mu) = \inf_{\gamma\in\Gamma(\mu,\nu)} (\int c^2(x,y)\gamma(dx,dy))^{1/2},$$

where $\Gamma(\mu,\nu)$ denotes the class of all probability measures γ on the product space $S\times S$ with marginals μ and ν. Equivalently, we can write

$$W(\nu,\mu) = \inf \tilde{E}[c^2(\tilde{X},\tilde{Y})]^{1/2},$$

where the infimum is taken over all couples $(\tilde{X},\tilde{Y})$ of S-valued random variables on some probability space $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{P})$ such that $\tilde{X}$ and $\tilde{Y}$ have distributions μ and ν, respectively. Such a couple, and also any measure $\gamma \in \Gamma(\mu,\nu)$, will be called a *coupling of* μ *and* ν. We refer to [15] for a thorough discussion of Wasserstein distances in various contexts.

In the sequel, the space S will be either a Euclidean space $\mathbb{R}^n$, including the infinite-dimensional case $n=\infty$, or the space

$$\Omega = C_0[0,1]$$

of all continuous functions ω on $[0,1]$ with initial value $\omega(0)=0$.

For $S=\mathbb{R}^n$ with $n\in\{1,\ldots,\infty\}$ we are going to use the cost function $c(x,y) = ||x-y||_n$, defined by the Euclidean norm $||x||_n = (\sum_{i=1}^n x_i^2)^{1/2}$. Thus, the corresponding Wasserstein distance is given by

$$W_n(\nu,\mu) = \inf_{\gamma\in\Gamma(\mu,\nu)} (\int ||x-y||_n^2)\gamma(dx,dy))^{1/2}.$$

Taking as reference measure the Gaussian measure

$$\mu_n = \prod_{i=1}^n N(0,1),$$

Talagrand's inequality on Euclidean space can now be stated as follows:

Theorem 1 *For any $n \in \{0,\dots,\infty\}$ and for any probability measure ν on $\mathbb{R}^n$,*

$$W_n(\nu,\mu_n) \leq \sqrt{2H(\nu|\mu_n)}. \tag{7}$$

Talagrand's proof in [13] takes a bottom-up approach. First the inequality is proved in the one-dimensional case, using Vallender's expression

$$W_1(\nu,\mu) = \Big(\int_0^1 \big(q_\nu(\alpha) - q_\mu(\alpha)\big)^2 d\alpha\Big)^{1/2} \tag{8}$$

in [14] of the Wasserstein distance on $\mathbb{R}^1$ in terms of the quantile functions q_ν and q_μ, followed by an integration by parts that involves the special form of the normal distribution. The finite-dimensional case is shown by induction, applying the one-dimensional inequality to the conditional distributions $\nu(dx_{n+1}|x_1,\dots,x_n)$ of ν. The infinite-dimensional case $n=\infty$ follows by applying (7) to the finite-dimensional marginals and taking the limit $n\uparrow\infty$, using a standard martingale argument to obtain convergence of the relative entropies on the right-hand side.

Let us now turn to the case $S=\Omega=C_0[0,1]$. We denote by $(\mathscr{F}_t)_{0\leq t\leq 1}$ the right-continuous filtration on Ω generated by the coordinate process

$$W = (W_t)_{0\leq t\leq 1}$$

defined by $W_t(\omega)=\omega(t)$. We set $\mathscr{F}=\mathscr{F}_1$ and denote by P the *Wiener measure* on $(\Omega,\mathscr{F})$. Let $\mathscr{H}$ denote the *Cameron-Martin space* of all absolutely continuous functions $\omega\in\Omega$ such that the derivative $\dot\omega$ is square integrable on $[0,1]$. First we will consider the cost function $c(\omega,\eta)=||\omega-\eta||_{\mathscr{H}}$, where

$$||\omega||_{\mathscr{H}} = \begin{cases} (\int_0^1 \dot\omega^2(t)dt)^{1/2} & \text{if } \omega\in\mathscr{H} \\ +\infty & \text{otherwise.} \end{cases}$$

The corresponding Wasserstein distance will be denoted by $W_{\mathscr{H}}$, that is,

$$W_{\mathscr{H}}(Q,P) = \inf_{\gamma\in\Gamma(P,Q)} \int ||\omega-\eta||^2_{\mathscr{H}}\gamma(d\omega,d\eta)^{1/2},$$

for any probability measure Q on $(\Omega,\mathscr{F})$. In this setting, Talagrand's inequality takes the following form, first stated by D. Feyel and A. S. Ustunel in [3].

Theorem 2 *For any probability measure Q on $(\Omega,\mathscr{F})$,*

$$W_{\mathscr{H}}(Q,P) \leq \sqrt{2H(Q|P}. \tag{9}$$

In fact, inequality (9) can be viewed as a direct translation of Talagrand's inequality on $\mathbb{R}^\infty$. To see this, recall the Lévy-Ciesielski representation

$$W_t(\omega) = \sum_{i \in I} X_i(\omega) e_i(t)$$

of Brownian motion in terms of the Schauder basis $(e_i)_{i \in I}$ of $C_0[0,1]$. Under Wiener measure P, the coordinates X_i are independent with distribution $N(0,1)$. Thus, the random vector $(X_i(\omega))_{i \in I}$, viewed as a measurable map T from Ω to R^∞, has distribution μ_∞ under P. Relative entropy is invariant under T, and so we get

$$H(\nu|\mu_\infty) = H(Q|P),$$

where ν denotes the image of Q under T. On the other hand we have $||\omega||_{\mathscr{H}} = ||(X_i(\omega))_{i \in I}||_\infty$, and this implies

$$W_{\mathscr{H}}(Q,P) = W_\infty(\nu, \mu_\infty).$$

Thus, Talagrand's inequality (7) for $n = \infty$ translates into inequality (9) on Wiener space.

Having scetched the bottom-up approach to Talagrand's inequality on Wiener space, we are now going to focus on the top-down approach. It consists in proving Talagrand's inequality (9) directly on Wiener space, using a suitable coupling of Q and P.

3 Intrinsic drift and optimal coupling in the absolutely continuous case

Take any probability measure Q on $(\Omega, \mathscr{F})$ that is absolutely continuous with respect to Wiener measure P. Let us first recall the following computation of the relative entroopy $H(Q|P)$ in terms of the *intrinsic drift* of Q; cf. [4], [5] or, for the first two parts, Th. 7.11 in [11].

Proposition 1 *There exists a predictable process $b^Q = (b_t^Q(\omega))_{0 \le t \le 1}$ with the following properties:*

1*)*

$$\int_0^1 \left(b_t^Q(\omega)\right)^2 dt < \infty \quad Q\text{-a.s.}, \tag{10}$$

that is, the process B^Q defined by $B_t^Q(\omega) = \int_0^t b_s^Q(\omega) ds$ satisfies

$$B^Q(\omega) \in \mathscr{H} \quad Q\text{-a.s.}$$

2*) $W^Q := W - B^Q$ is a Wiener process under Q, that is, W is a special semimartingale under Q with canonical decomposition*

$$W = W^Q + B^Q.$$

3) The relative entropy of Q with respect to P is given by

$$H(Q|P) = \frac{1}{2}E_Q\Big[\int_0^1 (b_t^Q)^2 dt\Big] = \frac{1}{2}E_Q\big[||B^Q||^2_{\mathscr{H}}\big]. \tag{11}$$

The process b^Q will be called the intrinsic drift *of Q.*

Proof For the convenience of the reader we scetch the argument; cf., e.g., [5] for details.

1) By Itô's representation theorem, the density $\phi = \frac{dQ}{dP}$ can be represented as a stochastic integral of the Brownian motion W, that is, there exists a predictable process $(\xi_t)_{0\le t\le 1}$ such that $\int_0^1 \xi_t(\omega)dt < \infty \quad P-a.s.$ and

$$\phi = 1 + \int_0^1 \xi_t dW_t \qquad P\text{-a.s.}$$

Moreover, the process

$$\phi_t := E_P[\phi|\mathscr{F}_t] = 1 + \int_0^t \xi_s dW_s, \qquad 0 \le t \le 1,$$

is a continuous martingale with quadratic variation

$$\langle\phi\rangle_t = \int_0^t \xi_s^2 ds \qquad P\text{-a.s.}$$

and

$$\inf_{0\le t\le 1} \phi_t > 0 \quad P\text{-a.s. on } \{\phi > 0\},$$

hence Q-a.s.. Thus, the predictable process b^Q defined by

$$b_t^Q := \frac{\xi_t}{\phi_t} I_{\{\phi_t > 0\}}, \qquad 0 \le t \le 1,$$

satisfies the integrability condition (10).

2) Applying Itô's formula to $\log\phi_t$, we get

$$\begin{aligned}\log\phi_t &= \int_0^t \frac{1}{\phi_s} d\phi_s - \frac{1}{2}\int_0^t (\frac{1}{\phi_s})^2 d\langle\phi\rangle_s \\ &= \int_0^t b_s^Q dW_s - \frac{1}{2}\int_0^t (b_s^Q)^2 ds \\ &= \int_0^t b_s^Q dW_s^Q + \frac{1}{2}\int_0^t (b_s^Q)^2 ds\end{aligned}$$

The second part now follows from Girsanov's theorem.

3) Equation (11) for $H(Q|P) = E_Q[\log\phi_1]$ follows from the preceding equation for $t = 1$. Indeed, if $E_Q\big[\int_0^1 (b_s^Q)^2 ds\big] < \infty$ then we get

$$E_Q\Big[\int_0^1 b_s^Q dW_s^Q\Big] = 0,$$

and this implies (11). In the general case, the same argument applies up to each stopping time $T_n = \inf\{t \mid \int_0^t (b_s^Q)^2 ds > n\} \wedge 1$, and for $n \uparrow \infty$ we obtain (11). □

Remark 1 Apart from our present purpose, the intrinsic drift of Q is also an efficient tool in proving a number of inequalities, including logarithmic Sobolev and Shannon-Stam inequalities; see [10] and [2].

As observed by J. Lehec in [10], proposition 1 can be rephrased as follows in terms of coupling, and in this form it yields an immediate proof of Talagrand's inequality on Wiener space.

Proposition 2 *The processes $W^Q = W - B^Q$ and W, defined on the probability space $(\Omega, \mathscr{F}, Q)$, form a coupling of P and Q such that*

$$E_Q\big[||W - W^Q||^2_{\mathscr{H}}\big] = 2H(Q|P). \tag{12}$$

Corollary 1 *Any probability measure Q on $(\Omega, \mathscr{F})$ satisfies Talagrand's inequality*

$$W_{\mathscr{H}}(Q,P) \leq \sqrt{2H(Q|P)}. \tag{13}$$

Proof If Q is not absolutely continuous with respect to Wiener measure P then we have $H(Q|P) = \infty$, and (13) holds trivially. In the absolutely continuous case, inequality (13) follows immediately from equation (12) and the definition of the Wasserstein distance $W_{\mathscr{H}}$. □

Note that the coupling (W^Q, W) of P and Q, which is defined on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leq t\leq 1}, Q)$, is *adaptive* in the following sense.

Definition 1 A coupling $(\tilde{X}, \tilde{Y})$ of P and Q will be called an adaptive coupling, if it is defined on a filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\leq t\leq 1}, \tilde{P})$ such that

1. $\tilde{Y} = (\tilde{Y}_t)$ is adapted with respect to $\tilde{P}$ and $(\tilde{\mathscr{F}}_t)_{0\leq t\leq 1}$,
2. $\tilde{X}$ is a Wiener process with respect to $\tilde{P}$ and $(\tilde{\mathscr{F}}_t)_{0\leq t\leq 1}$. that is, each increment $\tilde{X}_t - \tilde{X}_s$ is independent of $\tilde{\mathscr{F}}_s$ with law $N(0, t-s)$.

Theorem 3 *The* optimal adaptive coupling *of P and Q is given by (W^Q, W), that is,*

$$E_Q\big[||W - W^Q||^2_{\mathscr{H}}\big] \leq \tilde{E}\big[||\tilde{Y} - \tilde{X}||^2_{\mathscr{H}}\big], \tag{14}$$

for any adaptive coupling $(\tilde{X}, \tilde{Y})$ of P and Q, and equality holds iff

$$\tilde{Y} = W^Q(\tilde{Y}) + B^Q(\tilde{Y}), \qquad \tilde{P}-\text{a.s.}. \tag{15}$$

Proof Take any adapted coupling $(\tilde{X}, \tilde{Y})$ of P and Q, defined on a filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\leq t\leq 1}, \tilde{P})$, such that

$$\tilde{E}\big[||\tilde{Y} - \tilde{X}||^2_{\mathscr{H}}\big] < \infty.$$

Since $\tilde{Y}$ is adapted with continuous paths, $\tilde{B} := \tilde{Y} - \tilde{X}$ is an adapted continuous process such that $\tilde{E}\left[||B||^2_{\mathscr{H}}\right] < \infty$. This implies $\tilde{B}_t = \int_0^t \tilde{b}_s ds$ for some predictable process $\tilde{b} = (\tilde{b}_s)_{0 \le s \le 1}$ such that $\tilde{E}\left[\int_0^1 \tilde{b}_s^2 ds\right] < \infty$. Since $\tilde{X}$ is a Brownian motion with respect to the filtration $(\tilde{\mathscr{F}}_t)$, the process $\tilde{Y}$ is a special semimartingale with canonical decomposition

$$\tilde{Y}_t = \tilde{X}_t + \int_0^t \tilde{b}_s ds \tag{16}$$

under $\tilde{P}$ with respect to $(\tilde{\mathscr{F}}_t)$. On the other hand, since $\tilde{Y}$ has law Q under $\tilde{P}$ and W^Q is a Brownian motion under Q, the process $W^Q(\tilde{Y})$ is a Brownian motion under $\tilde{P}$ with respect to the smaller filtration $(\tilde{\mathscr{F}}_t^0)$ generated by the adapted process $\tilde{Y}$. Thus, $\tilde{Y}$ has the canonical decomposition

$$\tilde{Y}_t = W_t^Q(\tilde{Y}) + \int_0^t b_s^Q(\tilde{Y}) ds \tag{17}$$

under $\tilde{P}$ with respect to $(\tilde{\mathscr{F}}_t^0)$. This implies

$$b_t^Q(\tilde{Y}) = \tilde{E}\left[\tilde{b}_t \,|\, \tilde{\mathscr{G}}_t\right] \qquad \tilde{P} \otimes dt - \text{a.s.}; \tag{18}$$

cf., for example, Th. 8.1 in [11] or the proof of equation 68 in the general context of Proposition 4 below. Applying Jensen's inequality, we obtain

$$\begin{aligned} \tilde{E}\left[||\tilde{Y} - \tilde{X}||^2_{\mathscr{H}}\right] &= \tilde{E}\left[\int_0^1 \tilde{b}_t^2 dt\right] \\ &\ge \tilde{E}\left[\int_0^1 (b_t^Q(\tilde{Y}))^2 dt\right] = E_Q\left[\int_0^1 (b_t^Q(W))^2 dt\right] \\ &= 2H(Q|P). \end{aligned}$$

Equality holds iff

$$\tilde{b}_t = b_t^Q(\tilde{Y}) \qquad \tilde{P} \otimes dt - \text{a.s.},$$

and in this case (16) and (17) imply $\tilde{X} = W^Q(\tilde{Y})$ $\tilde{P}$-a.s.. □

Let us define $W_{\mathscr{H},ad}(Q,P)$ as the infimum of the right hand side in (14), taken only over the *adaptive* couplings of P and Q. Clearly we have

$$W_{\mathscr{H}}(Q,P) \le W_{\mathscr{H},ad}(Q,P), \tag{19}$$

and Theorem 3 shows that the following identity holds, first proved by R. Lassalle in [9].

Corollary 2 *For any probability measure Q on $(\Omega, \mathscr{F})$ we have*

$$W_{\mathscr{H},ad}(Q,P) = \sqrt{2H(Q|P)}. \tag{20}$$

Remark 2 For a thorough discussion of optimal transport problems on Wiener space under various constraints, with special emphasis on the effects of an enlargement of filtration, we refer to [1].

The following example illustrates the difference between $W_{\mathcal{H}}$ and $W_{\mathcal{H},ad}$. It also shows how the finite-dimensional inequalities in (7) can be derived from Talagrand's inequality on Wiener space, thus completing the top-down approach.

For a probability measure ν on $\mathbb{R}^1$ we introduce the probability measure

$$Q^\nu = \int P^x \nu(dx)$$

on $(\Omega, \mathcal{F})$, where P^x denotes the law of the Brownian bridge from 0 to $x \in \mathbb{R}^1$. If $\nu \ll \mu := N(0,1)$, then Q^ν is absolutely continuous with respect to P with density

$$\frac{dQ^\nu}{dP} = \frac{d\nu}{d\mu}(W_1),$$

and the relative entropy is given by

$$H(Q^\nu|P) = \int \log \frac{d\nu}{d\mu}(W_1) dQ^\nu = \int \log \frac{d\nu}{d\mu} d\nu = H(\nu|\mu). \tag{21}$$

Corollary 3 *We have*

$$W_{\mathcal{H}}(Q^\nu, P) = W_1(\nu, \mu) \quad \text{and} \quad W_{\mathcal{H},ad}(Q^\nu, P) = \sqrt{2H(\nu|\mu)}. \tag{22}$$

Thus, inequality (19) *implies*

$$W_1(\nu, \mu) \le \sqrt{2H(\nu|\mu)}. \tag{23}$$

Inequality (23) *is strict except for the case where* $\nu = N(m,1)$ *for some* $m \in \mathbb{R}^1$.

Proof 1) The second identity in (22) follows from Corollary 2 and equation (21).
2) To prove the first identity, take any coupling $(\tilde{X}, \tilde{Y})$ of P and Q, defined on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, such that $Z := \tilde{Y} - \tilde{X} \in \mathcal{H}$. Then the endpoints $\tilde{X}_1$ and $\tilde{Y}_1$ form a coupling of μ and ν. Since

$$(\tilde{Y}_1 - \tilde{X}_1)^2 = Z_1^2 = (\int_0^1 \dot{Z}_s ds)^2 \le \int_0^1 \dot{Z}_s^2 ds = ||\tilde{Y} - \tilde{X}||^2_{\mathcal{H}},$$

we obtain

$$W_1^2(\nu, \mu) \le \tilde{E}\left[(\tilde{Y}_1 - \tilde{X}_1)^2\right] \le \tilde{E}\left[||\tilde{Y} - \tilde{X}||^2_{\mathcal{H}}\right],$$

hence

$$W_1^2(\nu, \mu) \le W_{\mathcal{H}}^2(Q, P). \tag{24}$$

We now show that the lower bound $W_1^2(\nu, \mu)$ is attained by the following coupling (W, Y) of P and Q^ν, defined on the Wiener space $(\Omega, \mathcal{F}, P)$. The process Y is given by

$$Y_t = W_t + t\big(f_\nu(W_1) - W_1\big), \quad 0 \le t \le 1,$$

where $f_\nu(x) = q_\nu(\Phi(x))$ and Φ denotes the distribution function of $\mu = N(0,1)$. The endpoint $Y_1 = f_\nu(W_1)$ has distribution ν under P, and the conditional distribution of Y given the endpoint $Y_1 = y$ coincides with the Brownian bridge P^y. Thus Y has distribution Q^ν under P, and (W,Y) is a coupling of P and Q^ν, defined on $(\Omega, \mathscr{F}, P)$. Note that this coupling is not adaptive with respect to the filtration $(\mathscr{F}_t)$, since Y anticipates the endpoint W_1 of the Brownian path. Since $||Y-W||^2_{\mathscr{H}} = (f_\nu(W_1) - W_1)^2$, we get

$$\begin{aligned} E_P\big[||Y-W||^2_{\mathscr{H}}\big] &= \int \big(f_\nu(x) - x\big)^2 \mu(dx) \\ &= \int_0^1 \big(q_\nu(\alpha) - \Phi^{-1}(\alpha)\big)^2 d\alpha = W_1^2(\nu,\mu), \end{aligned}$$

using equation (8) in the last step. This completes the proof of the first identity in (22)

3) Let us write $Q = Q^\nu$. Theorem 3 shows that the optimal adapted coupling of Q and P is given by (W, W^Q) under Q. Since

$$(W_1 - W_1^Q)^2 = \Big(\int_0^1 b_t^Q dt\Big)^2 \le \int_0^1 (b_t^Q)^2 dt = ||B^Q||^2_{\mathscr{H}}$$

and

$$W_1^2(\nu,\mu) \le E_Q\big[(W_1 - W_1^Q)^2\big] \le E_Q\big[||B^Q||^2_{\mathscr{H}}\big] = 2H(\nu|\mu),$$

equality in (23) implies, Q-a.s., that $b_t^Q(\cdot)$ is almost everywhere constant in t, hence equal to $m(\cdot) := W_1 - W_1^Q$. Since the process b^Q is adapted to the filtration $(\mathscr{F}_t)$, $m(\cdot)$ is measurable with respect to $\mathscr{F}_0 = \bigcap_{t>0} \mathscr{F}_t$. But P is 0-1 on $\mathscr{F}_0$, and the same is true for $Q \ll P$. This implies $m(\cdot) = m \quad Q$-a.s. for some $m \in R^1$, that is, $W_1 = W_1^Q + m$ and $\nu = N(m,1)$. □

Talagrand's inequality in any finite dimension $n > 1$ follows in the same manner. For our purpose it is convenient to use the following equivalent version, where the reference measure is taken to be

$$\tilde{\mu}_n = \prod_{i=1}^n N\Big(0, \frac{1}{n}\Big)$$

instead of $\mu_n = \prod_{i=1}^n N(0,1)$ as in (7).

Corollary 4 *For any probability measure ν on $\mathbb{R}^n$,*

$$nW_n^2(\nu, \tilde{\mu}_n) \le 2H(\nu|\tilde{\mu}_n). \tag{25}$$

Proof We may assume $\nu \ll \tilde{\mu}_n$. Let $T_n : \Omega \to \mathbb{R}^n$ denote the map that associates to each path ω the vector of its increments $\omega(i/n) - \omega((i-1)/n) \quad (i = 1, \dots, n)$. Under Wiener measure P, the distribution of T_n is given by $\tilde{\mu}_n$. Define Q^ν on $(\Omega, \mathscr{F})$ by

$$\frac{dQ^\nu}{dP} = \frac{d\nu}{d\mu}(T_n).$$

For any coupling $(\tilde{X}, \tilde{Y})$ of P and Q^ν such that $Z := \tilde{Y} - \tilde{X} \in \mathscr{H}$, the vectors $X_n := T_n(\tilde{X})$ and $Y_n := T_n(\tilde{Y})$ form a coupling of ν and $\tilde{\mu}_n$. Since

$$||X_n - Y_n||^2 = \sum_{i=1}^{n} \Big(\int_{(i-1)/n}^{i/n} \dot{Z}_s ds\Big)^2 \le \sum_{i=1}^{n} \frac{1}{n} \int_{(i-1)/n}^{i/n} \dot{Z}_s^2 ds = \frac{1}{n} ||\tilde{Y} - \tilde{X}||^2_{\mathscr{H}},$$

we obtain

$$W_n^2(\nu, \tilde{\mu}_n) \le \tilde{E}\left[||Y_n - X_n||^2\right] \le \frac{1}{n} \tilde{E}\left[(||\tilde{Y} - \tilde{X}||^2_{\mathscr{H}}\right],$$

hence

$$W_n^2(\nu, \tilde{\mu}_n) \le \frac{1}{n} W^2_{\mathscr{H}}(Q, P) \le \frac{2}{n} H(Q^\nu | P).$$

due to Corollary 1. Since $H(Q^\nu|P) = H(\nu|\tilde{\mu}_n)$, we have proved (25). □

4 Specific Relative Entropy

The following concept of specific relative entropy on Wiener space was introduced by N. Gantert in her thesis [7], where it plays the role of a rate function for large deviations of the quadratic variation from its ergodic behaviour; cf. also [8]. In our context, it will allow us to extend Talagrand's inequality on Wiener space beyond the absolutely continuous case $Q \ll P$.

From now on, the index N will refer to the N-th dyadic partition of the unit interval, that is, $D_N = \{k2^{-N} | k = 1, \dots, 2^N\}$. In particular we introduce the discretized filtration

$$\mathscr{F}_{N,t} = \sigma(\{W_s | s \in D_N, s \le t\}), \quad 0 \le t \le 1$$

on $\Omega = C_0[0,1]$, and we set $\mathscr{F}_N = \mathscr{F}_{N,1} = \sigma(\{W_s | s \in D_N\})$.

Definition 2 For any probability measure Q on $(\Omega, \mathscr{F})$, the specific relative entropy of Q with respect to Wiener measure P is defined as

$$h(Q|P) = \liminf_{N \uparrow \infty} 2^{-N} H_N(Q|P), \tag{26}$$

where $H_N(Q|P)$ denotes the relative entropy of Q with respect to P on the σ-field $\mathscr{F}_N$.

Since $H(Q|P) = \lim_N H_N(Q|P)$, we get $h(Q|P) = 0$ for any Q such that $H(Q|P) < \infty$. Thus, the notion of specific relative entropy is of interest only if we look beyond the cases that we have considered so far.

Remark 3 Note that $\mathscr{F}_N = \sigma(T_n)$ for $n = 2^N$, where $T_n : \Omega \to \mathbb{R}^n$ maps a path ω to the vector of its increments along the N-th dyadic partition; cf. the proof of Corollary 4. Identifying the restrictions of Q and P to $\mathscr{F}_N$ with their images ν and $\tilde{\mu}_n$ under T_n, Talagrand's finite-dimensional inequality (25) can be written in the form

$$2^N W_N^2(Q,P) \leq 2H_N(Q|P), \tag{27}$$

with

$$W_N(Q,P) := \inf\left(\tilde{E}_{\tilde{P}}\left[\left\langle \tilde{Y} - \tilde{X}\right\rangle_N\right]\right)^{1/2},$$

where the infimum is taken over all couplings of Q and P and $\langle \cdot \rangle_N$ denotes the discrete quadratic variation along the N-th dyadic partition, that is, $\langle \omega \rangle_N = ||T_n(\omega)||_n^2$ for any continuous function $\omega \in \Omega = C_0[0,1]$. For $N \uparrow \infty$, the right hand side of (27) increases to $2H(Q|P)$. Thus, an alternative version of the bottom-up approach to Talagrand's inequality on Wiener space consists in showing that, in the limit $N \uparrow \infty$, the left hand side of (27) can be replaced by $W_{\mathscr{H}}(Q,P)$ if $H(Q|P < \infty$.

In order to go beyond the absolutely continuous case $Q \ll P$, let us rewrite the finite-dimensional inequality (27) as

$$W_N^2(Q,P) \leq 2 \cdot 2^{-N} H_N(Q|P). \tag{28}$$

Taking the limit $N \uparrow \infty$, the specific relative entropy $h(Q|P)$ appears on the right hand side of (28), while the left hand side suggests to define a new Wasserstein distance on Wiener space in terms of quadratic variation. The resulting extension of Talagrand's inequality is contained in Theorems 6 and 7 below. Instead of analyzing the limit behaviour of the left hand side of (28), we are going to use again a top-down approach, arguing directly in terms of couplings on Wiener space. As a first step in that direction, we now show how the specific relative entropy $h(Q|P)$ reflects the special structure of a semimartingale measure Q on $C_0[0,1]$.

Definition 3 Let $\mathscr{Q}_{\mathscr{S}}$ denote the class of all probability measures Q on $\Omega = C_0[0,1]$ such that the coordinate process W is a special semimartingale of the form

$$W = M^Q + A^Q \tag{29}$$

under Q with respect to the filtration $(\mathscr{F}_t)$, where

1. $M^Q = (M^Q)_{0 \leq t \leq 1}$ is a square-integrable martingale under Q
2. $A^Q = (A^Q)_{0 \leq t \leq 1}$ is an adapted process with continuous paths of bounded variation such that its total variation $|A|^Q$ satisfies $|A|_1^Q \in L^2(Q)$.

A probability measure $Q \in \mathscr{Q}_{\mathscr{S}}$ will be called a martingale measure if $A^Q = 0$, that is, if W is a square-integrable martingale under Q. The class of all such martingale measures will be denoted by $\mathscr{Q}_{\mathscr{M}}$.

Remark 4 Proposition 1 shows that any probability measure Q on $(\Omega, \mathscr{F})$ with finite relative entropy $H(Q|P) < \infty$ belongs to the class $\mathscr{Q}_{\mathscr{S}}$, with $M^Q = W^Q$ and $A^Q = B^Q$.

Let us now fix a measure $Q \in \mathscr{Q}_{\mathscr{S}}$. We denote by

$$\langle W \rangle = (\langle W \rangle_t)_{0 \le t \le 1}$$

the continuous quadratic variation process defined, Q-a.s., by the decomposition

$$W^2 = \int W dW + \langle W \rangle$$

of the continuous semimartingale W^2 under Q. Our assumptions for $Q \in \mathscr{Q}_{\mathscr{S}}$ imply that

$$\langle W \rangle_t = \lim_{N \uparrow \infty} \sum_{t \in D_N} \left(W_t - W_{t-2^{-N}} \right)^2 \qquad \text{in } L^1(Q) \tag{30}$$

and that

$$\lim_{N \uparrow \infty} \sum_{t \in D_N} \left(A_t - A_{t-2^{-N}} \right)^2 = 0 \qquad \text{in } L^1(Q) \tag{31}$$

cf., e.g., Ch. VI in [12].

Let us introduce the finite measure $q(\omega, dt)$ on $[0,1]$ with distribution function $\langle W \rangle(\omega)$, defined Q-a.s., and denote by

$$q(\omega, dt) = q_s(\omega, dt) + \sigma^2(\omega, t) dt \tag{32}$$

its Lebesgue decomposition into a singular and an absolutely continuous part with respect to Lebesgue measure λ on $[0,1]$; an explicit construction will be given in the second part of the following proof.

Our next aim is to derive, for a large class of probability measures $Q \in \mathscr{Q}_{\mathscr{S}}$, a lower bound for the specific relative entropy $h(Q|P)$ in terms of the quadratic variation of W under Q, that is, in terms of the random measure $q(\cdot,\cdot)$. In a first step we focus on the case $Q \in \mathscr{Q}_{\mathscr{M}}$. The following theorem for martingale measures is essentially due to N. Gantert in [7]; here we extend it to the case where the quadration variation may have a singular component.

Theorem 4 *For any martingale measure $Q \in \mathscr{Q}_{\mathscr{M}}$, the specific relative entropy of Q with respect to Wiener measure P satisfies*

$$\begin{aligned} h(Q|P) &\ge \frac{1}{2} E_Q\left[q(\omega, [0,1]) - 1 + H(\lambda | q(\omega, \cdot)) \right] \\ &= \frac{1}{2} E_Q\left[q_s(\omega, [0,1]) \right] + E_Q\Big[\int_0^1 f\big(\sigma^2(\omega, t)\big) dt \Big], \end{aligned} \tag{33}$$

where f is the convex function on $[0,\infty)$ defined by $f(x) = \frac{1}{2}(x - 1 - \log x) \ge 0$. In particular,

$$h(Q|P) < \infty \quad \Longrightarrow \quad \sigma^2(\cdot,\cdot) > 0 \quad Q \otimes \lambda - a.s. \tag{34}$$

Proof 1) First we look at the general case $Q \in \mathscr{Q}_{\mathscr{S}}$. Thus we can write $W = M + A$, where M is a square-integrable Q-martingale and A is an adapted process with continuous paths of bounded variation such that $E_Q\big[|A|_1^2\big] < \infty$.

For $N \geq 1$ and $i = 1, \dots, 2^N$ we write $t_i = i2^{-N}$ and denote by $\nu_{N,i}(\omega, \cdot)$ the conditional distribution of the increment $W_{t_i} - W_{t_{i-1}}$ under Q given the σ-field $\mathscr{F}_{N,t_{i-1}}$, by

$$m_{N,i} = E_Q\big[W_{t_i} - W_{t_{i-1}}|\mathscr{F}_{N,t_{i-1}}\big] = E_Q\big[A_{t_i} - A_{t_{i-1}}|\mathscr{F}_{N,t_{i-1}}\big]$$

its conditional mean, by

$$\tilde{\sigma}^2_{N,i} = E_Q\big[(W_{t_i} - W_{t_{i-1}})^2|\mathscr{F}_{N,t_{i-1}}\big] - m^2_{N,i}$$

its conditional variance, and by

$$\sigma^2_{N,i} = E_Q\big[(M_{t_i} - M_{t_{i-1}})^2|\mathscr{F}_{N,t_{i-1}}\big] = E_Q\big[\langle W\rangle_{t_i} - \langle W\rangle_{t_{i-1}}|\mathscr{F}_{N,t_{i-1}}\big] \tag{35}$$

the conditional variance of the martingale increment $M_{t_i} - M_{t_{i-1}}$. We can write

$$H_N(Q|P) = \sum_{i=1}^{2^N} E_Q\big[H\big(\nu_{N,i}(\omega,\cdot)|N(0,2^{-N})\big)\big].$$

Since

$$H\big(N(m,\alpha)|N(0,\beta)\big) = f(\frac{\alpha}{\beta}) + \frac{m^2}{2\beta}$$

for $\alpha, \beta > 0$ and $m \in R^1$, we get

$$\begin{aligned}
&H\big(\nu_{N,i}|N(0,2^{-N})\big)\\
&\quad = H\big(\nu_{N,i}|N(m_{N,i},\tilde{\sigma}^2_{N,i})\big) + H\big(N(m_{N,i},\tilde{\sigma}^2_{N,i})|N(0,2^{-N})\big)\\
&\quad = H\big(\nu_{N,i}|N(m_{N,i},\tilde{\sigma}^2_{N,i})\big) + f(2^N\tilde{\sigma}^2_{N,i}) + \frac{1}{2}2^N m^2_{N,i},
\end{aligned}$$

hence

$$H_N(Q|P) = H_N(Q|Q^*_N) + E_Q\big[\sum_{i=1}^{2^N} f(2^N\tilde{\sigma}^2_{N,i})\big] + \frac{1}{2}2^N I_N, \tag{36}$$

where we define

$$I_N := E_Q\big[\sum_{i=1}^{2^N} m^2_{N,i}\big], \tag{37}$$

and where Q^*_N denotes the probability measure on $(\Omega, \mathscr{F}_N)$ such that the increments $W_{t_i} - W_{t_{i-1}}$ have conditional distribution $N(m_{N,i}, \tilde{\sigma}^2_{N,i})$ given the σ-field $\mathscr{F}_{N,t_{i-1}}$. Note that Jensen's inequality yields

$$I_N \le E_Q\Big[\sum_{i=1}^{2^N}(A_{t_i} - A_{t_{i-1}})^2\Big],$$

hence

$$\lim_{N\uparrow\infty} I_N = 0, \tag{38}$$

due to (31). Note also that $H_N(Q|P) < \infty$ implies $\tilde{\sigma}^2_{N,i}(\omega) > 0 \quad Q$-a.s., since $f(0) = \infty$.

2) Let $Q \otimes q$ denote the finite measure on $\bar{\Omega} = \Omega \times [0,1]$ defined by $Q \otimes q(d\omega, dt) = Q(d\omega)q(\omega, dt)$. On the σ-field

$$\mathscr{P}_N := \sigma(\{A_t \times (t,1] \mid t \in D_N, A_t \in \mathscr{F}_{N,t}\}),$$

the measure $Q \otimes q$ is absolutely continuous with respect to the product measure $Q \otimes \lambda$, where λ denotes the Lebesgue measure on $(0,1]$, and the density is given by

$$\sigma^2_N(\omega,t) := \sum_{i=1}^{2^N} 2^N \sigma^2_{N,i}(\omega) I_{(t_{i-1},t_i]}(t).$$

The σ-fields $\mathscr{P}_N$ increase to the predictable σ-field $\mathscr{P}$ on $\bar{\Omega}$, generated by the sets $A_t \times (t,1]$ with $t \in [0,1]$ and $A_t \in \mathscr{F}_t$. Applying the first part of Lemma 1 with $\mu = Q \otimes \lambda$ and $\nu = Q \otimes q$, we see that the limit

$$\sigma^2(\omega,t) = \lim_{N\uparrow\infty} \sigma^2_N(\omega,t)$$

exists both $Q \otimes q$ -a.s. and $Q \otimes \lambda$ -a.s., with

$$\sigma^2(\omega,t) \in [0,\infty) \qquad Q \otimes \lambda - a.s.$$

and

$$\sigma^2(\omega,t) \in (0,\infty] \qquad Q \otimes q - a.s..$$

Moreover, the Lebesgue decomposition of $Q \otimes q$ with respect to $Q \otimes \lambda$ on the predictable σ-field $\mathscr{P}$ takes the form

$$Q \otimes q\,[\bar{A}] = Q \otimes q\,[\bar{A} \cap \{\sigma^2 = \infty\}] + E_{Q\otimes\lambda}\,[\sigma^2; \bar{A}],$$

for $\bar{A} \in \mathscr{P}$. This implies, Q-a.s., the Lebesgue decomposition

$$q(\omega,dt) = q_s(\omega,dt) + \sigma^2(\omega,t)\lambda(dt),$$

of $q(\omega,\cdot)$ with respect to Lebesgue measure λ, where the singular part $q_s(\omega,\cdot)$ is given by the restriction of $q(\omega,\cdot)$ to the λ-null set

$$N(\omega) := \{t \mid \sigma^2(\omega,t) = \infty\}. \tag{39}$$

3) Let us now focus on the case where Q is a martingale measure. For $Q \in \mathscr{Q}_{\mathscr{M}}$, we have $\tilde{\sigma}^2_{N,i} = \sigma^2_{N,i}$ and $A = 0$, hence $I_N = 0$. Thus, equation (36) can be written as

$$2^{-N} H_N(Q|P) = 2^{-N} H_N(Q|Q_N^*) + E_Q\Big[\int_0^1 f(\sigma_N^2(\cdot,t))dt\Big]. \tag{40}$$

Since $H_N(Q|Q_N^*) \geq 0$, we obtain

$$\begin{aligned} h(Q|P) &\geq \lim_{N\uparrow\infty} E_Q\Big[\int_0^1 f(\sigma_N^2(\cdot,t))dt\Big] \\ &= \frac{1}{2}E_Q\big[q_s(\omega,(0,1])\big] + E_Q\Big[\int_0^1 f(\sigma^2(\cdot,t))dt\Big]. \end{aligned} \tag{41}$$

where we apply the second part of Lemma 1 below, with $\mu = Q \otimes \lambda$ and $\nu = Q \otimes q$. Since $f(0) = \infty$, we see that $h(Q|P) < \infty$ implies that $\sigma^2(\cdot,\cdot)$ is strictly positive $Q \otimes \lambda$-a.s.. □

Remark 5 The proof of Theorem 4 shows that we obtain existence of the limit

$$h(Q|P) = \lim_{N\uparrow\infty} 2^{-N} H_N(Q|P) \tag{42}$$

together with the equality

$$h(Q|P) = \frac{1}{2}E_Q\big[q_s(\omega,[0,1])\big] + E_Q\Big[\int_0^1 f\big(\sigma^2(\omega,t)\big)dt\Big], \tag{43}$$

if and only if Q is "almost locally Gaussian" in the sense that the measures Q_N^* appearing in (36) satisfy

$$\lim_{N\uparrow\infty} 2^{-N} H_N(Q|Q_N^*) = 0. \tag{44}$$

This was already observed by N. Gantert in [7].

In the proof of Theorem 4 we have used the following general lemma.

Lemma 1 *Consider two probability measures μ and ν on a measurable space $(S,\mathscr{S})$ and a sequence of $(\mathscr{S}_n)_{n=1,2,...}$ of sub-σ-fields increasing to $\mathscr{S}_\infty$. Suppose that ν is equivalent to μ on $\mathscr{S}_n$ with density ϕ_n.*
1) The limit $\phi_\infty = \lim_n \phi_n$ exists both μ-a.s. and ν-a,s,, with

$$\phi_\infty \in [0,\infty) \ \mu - a.s. \quad \text{and} \quad \phi_\infty \in (0,\infty] \ \nu - a.s.,$$

and the Lebesgue decomposition $\nu = \nu_s + \nu_a$ of ν with respect to μ on $\mathscr{S}_\infty$ is given by

$$\nu_s(A) = \nu(A \cap \{\phi_\infty = \infty\}) \quad \text{and} \quad \nu_a(A) = \int_A \phi_\infty d\mu.$$

2) If $\sup_n \int f(\phi_n)d\mu < \infty$ for $f(x) = \frac{1}{2}(x - 1 - \log x)$ then we have

$$\lim_{n\uparrow\infty}\int f(\phi_n)d\mu = \frac{1}{2}\nu_s(S) + \int f(\phi_\infty)d\mu. \tag{45}$$

Proof The first part is well-known; the proof uses standard martingale arguments. To prove the second part, we write

$$\begin{aligned}\int 2f(\phi_n)d\mu &= \int \phi_n d\mu - 1 + \int \log(\phi_n^{-1})d\mu \\ &= \nu_s(S) + \int \phi_\infty d\mu - 1 + H_{\mathscr{S}_n}(\nu|\mu).\end{aligned}$$

Due to (6), we get

$$\lim_{n\uparrow\infty}\int f(\phi_n)d\mu = \frac{1}{2}\big(\nu_s(S) + \int \phi_\infty d\mu - 1 + H_{\mathscr{S}_\infty}(\nu|\mu)\big)$$

If the left hand side is finite, the relative entropy is finite and reduces to $\int \log(\phi_\infty^{-1})d\mu$, and this yields equation (45). □

Let us now go beyond the case of a martingale measure. Take $Q \in \mathscr{Q}_{\mathscr{S}}$ and let $W = M + A$ be the canonical decomposition of the semimartingale W under Q. As soon as the process A is non-deterministic, the conditional variances $\sigma_{N,i}^2$ of M defined in (35) do no longer coincide with the conditional variances $\tilde{\sigma}_{N,i}^2$ of W along the N-th dyadic partition. Instead we have

$$\tilde{\sigma}_{N,i}^2 = \sigma_{N,i}^2 + \delta_{N,i},$$

where

$$\delta_{N,i} = \alpha_{N,i}^2 + 2E_Q\big[(M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})|\mathscr{F}_{N,t_{i-1}}\big],$$

and where we denote by

$$\alpha_{N,i}^2 = E_Q\big[(A_{t_i} - A_{t_{i-1}})^2|\mathscr{F}_{N,t_{i-1}}\big] - m_{N,i}^2$$

the conditional variances of A along the N-th dyadic partition.

Lemma 2 *The differences $\delta_{N,i}$ and the conditional variances $\alpha_{N,i}^2$ satisfy*

$$\lim_{n\uparrow\infty} E_Q\Big[\sum_{i=1}^{2^N}|\delta_{N,i}|\Big] = \lim_{n\uparrow\infty} E_Q\Big[\sum_{i=1}^{2^N}\alpha_{N,i}^2\Big] = 0.$$

Proof Since

$$J_N := E_Q\Big[\sum_{i=1}^{2^N}\alpha_{N,i}^2\Big] \le E_Q\Big[\sum_{i=1}^{2^N}(A_{t_i} - A_{t_{i-1}})^2\Big],$$

we obtain

$$\lim_{n\uparrow\infty} J_N = 0 \tag{46}$$

due to (31). On the other hand, since

$$|\delta_{N,i}| \leq \alpha_{N,i}^2 + 2\sigma_{N,i}\alpha_{N,i}, \tag{47}$$

we get

$$\begin{aligned} E_Q\Big[\sum_{i=1}^{2^N} |\delta_{N,i}|\Big] &\leq E_Q\Big[\sum_{i=1}^{2^N} \alpha_{N,i}^2\Big] + 2\sum_{i=1}^{2^N} E_Q\big[\sigma_{N,i}^2\big]^{1/2} E_Q\big[\alpha_{N,i}^2\big]^{1/2} \\ &\leq J_N + 2E_Q\big[M_1^2\big]^{1/2} J_N^{1/2}, \end{aligned}$$

hence

$$\lim_{N\uparrow\infty} E_Q\Big[\sum_{i=1}^{2^N} |\delta_{N,i}| = 0, \tag{48}$$

due to (46). □

To prove our extended version of Theorem 4, we use an additional assumption.

Definition 4 We denote by $\mathscr{Q}_{\mathscr{S}}^0$ the class of all probability measures $Q \in \mathscr{Q}_{\mathscr{S}}$ such that

$$\lim_{n\uparrow\infty} E_Q\big[2^{-N} \sum_{i=1}^{2^N} \alpha_{N,i}^2 \sigma_{N,i}^{-2}\big] = 0. \tag{49}$$

Remark 6 Condition (49) is satisfied if $\sigma^2(\cdot,\cdot)$ is bounded away from 0. Indeed, if $\sigma^2(\cdot,\cdot) \geq c \quad Q\otimes\lambda$-a.s. for some $c > 0$ then

$$\sum_{i=1}^{2^N} 2^N \sigma_{N,i}^2(\omega) I_{(t_{i-1},t_i]}(t) = \sigma_N^2(\omega,t) \geq E_{Q\otimes\lambda}\left[\sigma^2|\mathscr{P}_N\right] \geq c \quad Q\otimes\lambda - a.s.;$$

cf. the second part of the proof of Theorem 4. Thus, (49) follows from Lemma 2.

Theorem 5 *For any $Q \in \mathscr{Q}_{\mathscr{S}}^0$,*

$$h(Q|P) \geq \frac{1}{2} E_Q\big[q_s(\omega,[0,1])\big] + E_Q\Big[\int_0^1 f\big(\sigma^2(\omega,t)\big)dt\Big]. \tag{50}$$

Proof 1) Let us return to the first part of the proof of Theorem 4. Since $H_N(Q|Q_N^*) \geq 0$, equation (36) yields

$$2^{-N} H_N(Q|P) \geq E_Q\Big[\sum_{i=1}^{2^N} f(2^N \tilde{\sigma}_{N,i}^2) 2^{-N}\Big] + \frac{1}{2} I_N.$$

Since f is convex with $f'(x) = \frac{1}{2}(1-x^{-1})$, we obtain

$$f(2^N \tilde{\sigma}_{N,i}^2) \geq f(2^N \sigma_{N,i}^2) + \frac{1}{2}(1 - 2^{-N}\sigma_{N,i}^{-2}) 2^N \delta_{N,i}.$$

Due to (38), this implies

$$h(Q|P) \geq \liminf_{N\uparrow\infty} E_Q\Big[\int_0^1 f\big(\sigma_N^2(\omega,t)\big)dt + \frac{1}{2}\Delta_N\Big],$$

where

$$\Delta_N = \sum_{i=1}^{2^N} \big(\delta_{N,i} - 2^{-N}\sigma_{N,i}^{-2}\,\delta_{N,i}\big).$$

Applying the second part of Lemma 1 as in the proof of Theorem 4, we see that inequality (50) holds as soon as we show that

$$\lim_{N\uparrow\infty} \Delta_N = 0 \quad \text{in} \quad L^1(Q). \tag{51}$$

2) In view of Lemma 2 it is enough to show convergence to 0 for

$$\begin{aligned} &E_Q\Big[\sum_{i=1}^{2^N} 2^{-N}\sigma_{N,i}^{-2}|\delta_{N,i}|\Big] \\ &\leq E_Q\Big[\sum_{i=1}^{2^N} 2^{-N}\sigma_{N,i}^{-2}\alpha_{N,i}^2\Big] + 2E_Q\Big[2^{-N}\sum_{i=1}^{2^N}\alpha_{N,i}\sigma_{N,i}^{-1}\Big] \\ &\leq E_Q\Big[\sum_{i=1}^{2^N} 2^{-N}\sigma_{N,i}^{-2}\alpha_{N,i}^2\Big] + 2E_Q\Big[2^{-N}\sum_{i=1}^{2^N}\sigma_{N,i}^{-2}\alpha_{N,i}^2\Big]^{1/2}. \end{aligned}$$

But the last two terms converge to 0 due to our assumption (49), and this completes the proof of (51). □

Corollary 5 *Let $Q \in \mathcal{Q}_{\mathcal{S}}$ be such that $||A^Q||_{\mathcal{H}} \in L^2(Q)$. Then we have*

$$h(Q|P) = 0 \quad \Longleftrightarrow \quad H(Q|P) < \infty,$$

and in this case the canonical decomposition (29) *of W under Q takes the form $M^Q = W^Q$ and $A^Q = B^Q$.*

Proof Let us assume $h(Q|P) = 0$. Inequality (33) implies $q_s(\omega,\cdot) = 0$ Q-a.s and $f\big(\sigma^2(\omega,t)\big) = 0$ $Q\otimes\lambda$-a.s, hence $\sigma^2(\omega,t) = 1$ $Q\otimes\lambda$-a.s. Thus, W has quadratic variation

$$\langle W\rangle_t = \langle M^Q\rangle_t = t$$

under Q, and so M^Q is a Wiener process under Q. Uniqueness of the canonical decomposition of W under Q yields $M^Q = W^Q$ and $A^Q = B^Q$, hence

$$H(Q|P) = \frac{1}{2}E_Q\big[||A^Q||_{\mathcal{H}}^2\big] < \infty$$

due to Proposition 1. Conversely, $H(Q|P) < \infty$ implies $h(Q|P) = 0$, as we have already observed above, following the definition of $h(Q|P)$. □

5 Intrinsic Wiener Process and Optimal Coupling for Semimartingale Measures

We fix a probability measure $Q \in \mathcal{Q}_{\mathscr{S}}$ and denote by

$$W = M + A \tag{52}$$

the canonical decomposition of the coordinate process W under Q. Recall the Lebesgue decomposition

$$q(\omega, dt) = q(\omega, dt) + \sigma^2(\omega, t)dt$$

of the random measure $q(\omega, \cdot)$ on $[0,1]$ with distribution function $\langle W \rangle(\omega)$, and put

$$A(\omega) := \{t \in [0,1] \mid \sigma^2(\omega, t) < \infty\}.$$

The following construction of an intrinsic Wiener process W^Q for Q extends the definition in Proposition 1 beyond the absolutely continuous case $Q \ll P$.

Lemma 3 *If $h(Q|P) < \infty$ then the process $W^Q = (W_t^Q)_{0 \le t \le 1}$, defined Q-a.s. by*

$$W_t^Q := \int_0^t \sigma(\cdot, s)^{-1} I_{A(\cdot)}(s) dM_s, \tag{53}$$

is a Wiener process under Q.

Proof By Theorem 4, our assumption $h(Q|P) < \infty$ implies

$$E_Q\Big[\int_0^1 f(\sigma^2(\omega, t))dt\Big] < \infty,$$

where $f(x) = \frac{1}{2}(x - 1 - \log x)$, and in particular

$$0 < \sigma^2(\cdot, \cdot) < \infty \quad Q \otimes \lambda - a.s..$$

since $f(0) = \infty$. Since $\langle M \rangle = \langle W \rangle$ and $\lambda\left(A(\cdot)\right) = 1$ Q-a.s., the predictable integrand $\phi_s = \sigma(\cdot, s)^{-1} I_{A(\cdot)}(s)$ in (53) satisfies

$$\int_0^t \phi_s^2 d\langle M \rangle_s = \int_0^t \sigma_s^{-2} I_{A(\cdot)}(s) \sigma_s^2 ds = \int_0^t I_{A(\cdot)}(s) ds = t.$$

Thus, the stochastic integrals in (53) are well defined, and they define a continuous martingale W^Q under Q with quadratic variation $\langle W^Q \rangle_t = t$. This implies that W^Q is a Wiener process under Q. □

For the rest of this section we assume that $Q \in \mathcal{Q}_{\mathscr{S}}$ satisfies the condition

$$h(Q|P) < \infty, \tag{54}$$

and so W^Q will be a Wiener process under Q.

Definition 5 W^Q will be called the intrinsic Wiener process of Q.

Remark 7 If $H(Q|P) < \infty$ then the intrinsic Wiener process coincides with the Wiener process $W^Q := W - B^Q$ defined in Proposition 1; cf. the proof of corollary 5.

Definition 6 An adaptive coupling $(\tilde{X}, \tilde{Y})$ of P and Q on a filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\le t\le 1}, \tilde{P})$ will be called a semimartingale coupling if $\tilde{Y}$ is a special semimartingale with respect to $\tilde{P}$ and $(\tilde{\mathscr{F}}_t)_{0\le t\le 1}$, and if the canonical decomposition $\tilde{Y} = \tilde{M} + \tilde{A}$ is such that

1. $\tilde{M}$ is a square-integrable martingale,
2. $\tilde{A}$ is an adapted process with continuous paths of bounded variation such that its total variation $|\tilde{A}|$ satisfies $|\tilde{A}|_1 \in L^2(\tilde{P})$.

Clearly, the pair (W^Q, W) is a semimartingale coupling of P and Q, defined on the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\le t\le 1}, Q)$. In fact, we are going to show that (W^Q, W) is the *optimal* semimartingale coupling for the Wasserstein distance $W_{\mathscr{S}}(Q,P)$ defined below.

Proposition 3 *For any semimartingale coupling $(\tilde{X}, \tilde{Y})$ of P and Q on some filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\le t\le 1}, \tilde{P})$ we have*

$$\tilde{E}\big[\langle \tilde{Y} - \tilde{X}\rangle_1\big] \ge E_Q\big[\langle W - W^Q\rangle_1\big], \tag{55}$$

and equality holds if and only if $\tilde{X} = W^Q(\tilde{Y}) \quad \tilde{P}$-a.s.. Moreover,

$$E_Q\big[\langle W - W^Q\rangle_1\big] = E_Q\Big[\int_0^1 \big(\sigma(\cdot,s) - 1\big)^2 ds + q_s(\cdot,(0,1])\Big]. \tag{56}$$

.

Proof 1) First we show that the last equality holds. Recall from the proof of Theorem 4 that $q_s(\omega,\cdot)$ is given, Q-a.s., by the restriction of $q(\omega,\cdot)$ to the λ-nullset $N(\omega)$ defined in (39). Since $A(\cdot)\cup N(\cdot) = [0,1]$, we have

$$\begin{aligned} W_t &= \int_0^t I_{A(\cdot)}(s)dW_s + \int_0^1 I_{N(\cdot)}(s)dW_s \\ &= \int_0^t \sigma(\cdot,s)dW_s^Q + \int_0^1 I_{N(\cdot)}(s)dW_s, \end{aligned}$$

hence

$$(W - W^Q)_t = \int_0^t \big(\sigma(\cdot,s) - 1\big)dW_s^Q + \int_0^1 I_{N(\cdot)}(s)dW_s$$

and

$$\langle W - W^Q \rangle_t = \int_0^t \big(\sigma(\cdot,s) - 1\big)^2 ds + \int_0^1 I_{N(\cdot)}(s) d\langle W \rangle_s \\ + 2\int_0^t \big(\sigma(\cdot,s) - 1\big) I_{N(\cdot)}(s) d\langle W^Q, W \rangle_s.$$

The last term vanishes since, Q-a.s., $N(\omega)$ is a nullset with respect to $d\langle W^Q, W\rangle(\omega) \ll d\langle W^Q \rangle(\omega) = dt$. This implies

$$E_Q\big[\langle W - W^Q \rangle_1\big] = E_Q\Big[\int_0^1 \big(\sigma(\cdot,s) - 1\big)^2 ds + q_s(\cdot,(0,1])\Big].$$

2) Consider any semimartingale coupling $(\tilde{X}, \tilde{Y})$ of P and Q, defined on some filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\le t\le 1}, \tilde{P})$. Both $\tilde{X}$ and the process $\tilde{W} := W^Q(\tilde{Y})$, defined by

$$\tilde{W}_t := \int_0^t \sigma(\tilde{Y}, s)^{-1} I_{A(\tilde{Y})}(s) d\tilde{Y}_s\,,$$

are Wiener processes under $\tilde{P}$ with respect to the filtration $(\tilde{\mathscr{F}}_t)$. Projecting the first on the second, we can write

$$\tilde{X}_t = \int_0^t \rho_s d\tilde{W}_s + \tilde{L}_t\,,$$

where $\tilde{L} = (\tilde{L}_t)_{0\le t\le 1}$ is a martingale orthogonal to $\tilde{W}$. Since

$$t = \langle \tilde{X} \rangle_t = \int_0^t \rho_s^2 ds + \langle \tilde{L} \rangle_t,$$

we get $\rho_t^2 \le 1$ and $d\langle \tilde{L} \rangle_t = (1 - \rho_t^2)dt$. This implies

$$\begin{aligned} d\langle \tilde{X}, \tilde{Y} \rangle &= \rho_t d\langle \tilde{W}, \tilde{Y} \rangle \\ &= \rho_t \sigma^{-1}(\tilde{Y}, t) I_{A(\tilde{Y})}(t) \sigma^2(\tilde{Y}, t) dt \\ &\le \sigma(\tilde{Y}, t) dt, \end{aligned}$$

hence

$$\begin{aligned} \langle \tilde{Y} - \tilde{X} \rangle_1 &= \langle \tilde{Y} \rangle_1 + \langle \tilde{X} \rangle_1 - 2\langle \tilde{X}, \tilde{Y} \rangle_1 \\ &\ge \int_0^1 \sigma^2(\tilde{Y}, t) dt + q_s(\tilde{Y}, (0,1]) + 1 - 2\int_0^1 \sigma(\tilde{Y}, t) dt \\ &= \int_0^1 \big(\sigma(\tilde{Y}, t) - 1\big)^2 dt + q_s(\tilde{Y}, (0,1]). \end{aligned}$$

Thus,

$$\begin{aligned}\tilde{E}\big[\langle\tilde{Y}-\tilde{X}\rangle_1\big] &\geq \tilde{E}\Big[\int_0^1 \big(\sigma\big(\tilde{Y},t\big)-1\big)^2 dt + q_s(\tilde{Y},(0,1])\Big] \\ &= E_Q\Big[\int_0^1 \big(\sigma(\cdot,t)-1\big)^2 dt + q_s(\cdot,(0,1])\Big] \\ . &= E_Q\big[\langle W-W^Q\rangle_1\big],\end{aligned}$$

and equality holds iff $\rho_t(\cdot)=1$ $\tilde{P}\otimes dt$ -a.s., that is, iff $\tilde{X}=\tilde{W}=W^Q(\tilde{Y})$ $\tilde{P}$-a.s.. □

Now consider the following Wasserstein distance $W_{\mathscr{S}}(Q,P)$, where the cost function is defined in terms of quadratic variation.

Definition 7 The Wasserstein distance $W_{\mathscr{S}}(Q,P)$ between Q and Wiener measure P is defined as

$$W_{\mathscr{S}}(Q,P)=\inf\big(\tilde{E}\big[\langle\tilde{Y}-\tilde{X}\rangle_1+||\tilde{A}||^2_{\mathscr{S}}\big]\big)^{\frac{1}{2}}, \tag{57}$$

where the infimum is taken over all semimartingale couplings $(\tilde{Y},\tilde{X})$ of Q and P on some filtered probability space, where $\tilde{M}+\tilde{A}$ is the canonical decomposition of $\tilde{Y}$, and where we set

$$||\tilde{A}||_{\mathscr{S}}=\Big(\int_0^1 \tilde{a}_t^2 d\langle\tilde{Y}\rangle_t\Big)^{1/2}$$

if $\tilde{A}$ is absolutely continuous with respect to $\langle\tilde{Y}\rangle$ with density process $\tilde{a}$, and $||\tilde{A}||_{\mathscr{S}}=\infty$ otherwise.

Remark 8 In the absolutely continuous case $Q\ll P$ we have

$$d\langle\tilde{Y}\rangle=d\langle\tilde{X}\rangle=dt \qquad Q\text{–a.s.},$$

and so the norm $||\tilde{A}||_{\mathscr{S}}$ reduces to the Cameron-Martin norm $||\tilde{A}||_{\mathscr{H}}$.

As an immediate corollary to the preceding proposition we obtain the following inequality for martingale measures. It provides a first extension of Talagrand's inequality (13) on Wiener space beyond the absolutely continuous case.

Theorem 6 *For a martingale measure $Q\in\mathscr{Q}_{\mathscr{M}}$,*

$$W^2_{\mathscr{S}}(Q,P)=E_Q\big[\langle W-W^Q\rangle_1\big]\leq 2h(Q|P), \tag{58}$$

and equality holds iff $Q=P$.

Proof 1) For $Q\in\mathscr{Q}_{\mathscr{M}}$, the pair (W,W^Q) is a semimartingale coupling of Q and P, defined on $(\Omega,\mathscr{F},(\mathscr{F}_t)_{0\leq t\leq 1},Q)$, such that $W-W^Q=M-W^Q$ is a martingale under Q. Thus, the expected cost in (57) only involves the quadratic variation component, and Proposition 3 implies

$$W^2_{\mathscr{S}}(Q,P)=E_Q\big[\langle W-W^Q\rangle_1\big]=E_Q\Big[\int_0^1 \big(\sigma(\cdot,s)-1\big)^2 ds+q_s(\cdot,(0,1])\Big]. \tag{59}$$

Note that

$$(\sigma-1)^2 \le \sigma^2 - 1 - \log\sigma^2 = 2f(\sigma^2),$$

with equality iff $\sigma^2 = 1$. Thus,

$$\begin{aligned} E_Q\big[\langle W - W^Q\rangle_1\big] &\le E_Q\big[2\int_0^1 f\big(\sigma^2(\cdot,s)\big)dt + q_s(\cdot,(0,1])\big] \\ &\le 2h(Q|P), \end{aligned} \tag{60}$$

where the second inequality follows from Theorem 4.

2) Equality in (58) implies equality in (60). It follows from part 1) that $\sigma^2(\cdot,\cdot) = 1$ $Q\otimes\lambda$-a.s.. This implies $W = M = W^Q$ under Q, hence $Q = P$. The converse is obvious. □

Definition 8 We write $Q \in \mathscr{Q}^*_{\mathscr{S}}$ if the canonical decomposition $W = M + A$ of the coordinate process W under $Q \in \mathscr{Q}_{\mathscr{S}}$ is such that

$$E_Q\big[||A||^2_{\mathscr{S}}\big] < \infty, \tag{61}$$

that is, $dA_t = a_t d\langle W\rangle_t$ with $\int_0^1 a_t^2 d\langle W\rangle_t \in L^1(Q)$, and if

$$G^* := \exp\big(-\int_0^1 a_t dM - \frac{1}{2}\int_0^1 a_t^2 d\langle M\rangle_t\big)$$

satisfies

$$G^* \in L^2(Q) \quad \text{and} \quad E_Q[G^*] = 1. \tag{62}$$

Remark 9 For $Q \in \mathscr{Q}^*_{\mathscr{S}}$, the probability measure Q^* defined by

$$dQ^* = G^* dQ \tag{63}$$

is an equivalent martingale measure for Q; cf., for example, [6]. Note that $\mathscr{Q}_{\mathscr{M}} \subset \mathscr{Q}^*_{\mathscr{S}}$, and that $Q^* = Q$ for $Q \in \mathscr{Q}_{\mathscr{M}}$.

Proposition 4 *For $Q \in \mathscr{Q}^*_{\mathscr{S}}$, the coupling (W, W^Q) of Q and P is optimal for $W_{\mathscr{S}}$, that is,*

$$W^2_{\mathscr{S}}(Q,P) = E_Q\big[\langle W - W^Q\rangle_1 + ||A||^2_{\mathscr{S}}\big]. \tag{64}$$

Proof For $Q \in \mathscr{Q}^*_{\mathscr{S}}$, the right-hand side in (64) is finite, and so we have $W_{\mathscr{S}}(Q,P) < \infty$. Now take any semimartingale coupling $(\tilde{Y},\tilde{X})$ of Q and P, defined on some filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{0\le t\le 1}, \tilde{P})$, such that

$$\tilde{E}\big[\langle \tilde{Y} - \tilde{X}\rangle_1 + ||\tilde{A}||^2_{\mathscr{S}}\big] < \infty.$$

Since

$$\tilde{E}\big[\langle \tilde{Y} - \tilde{X}\rangle_1\big] \ge E_Q\big[\langle W - W^Q\rangle_1\big] \tag{65}$$

by Proposition 3, it only remains to show that

$$\tilde{E}\big[||\tilde{A}||^2_{\mathscr{S}}\big] \ge E_Q\big[||A||^2_{\mathscr{S}}\big],$$

that is,

$$\tilde{E}\Big[\int_0^1 \tilde{a}_t^2 d\langle\tilde{Y}\rangle_t\Big] \geq E_Q\Big[\int_0^1 a_t^2 d\langle W\rangle_t\Big]. \tag{66}$$

We denote by $\tilde{\mathscr{P}}$ the predictable σ-field on $\tilde{\Omega}\times(0,1]$ corresponding to the filtration $(\tilde{\mathscr{F}}_t)$, and by $\mathscr{P}^0 \subseteq \mathscr{P}$ the predictable σ-field corresponding to the smaller filtration $(\tilde{\mathscr{F}}_t^0)$ generated by $(\tilde{Y}_t)$. Since $\tilde{E}\big[||\tilde{A}||^2_{\mathscr{S}}\big] < \infty$, we have

$$d\tilde{A}_t = \tilde{a}_t d\langle\tilde{Y}\rangle_t = \tilde{a}_t dq(\tilde{Y},t),$$

where $\tilde{a} = (\tilde{a}_t)$ is $\mathscr{P}$-measurable and square-integrable with respect to the finite measure $\tilde{P}\otimes q(\tilde{Y},\cdot)$ on $\tilde{\mathscr{P}}$. Let $\tilde{a}^0 = (\tilde{a}_t^0)$ denote the process defined by the conditional expectation

$$\tilde{a}^0 := E_{\tilde{P}\otimes q(\tilde{Y},\cdot)}\big[\tilde{a}\,\big|\,\mathscr{P}^0\big],$$

and note that Jensen's inequality implies

$$E_{\tilde{P}\otimes q(\tilde{Y},\cdot)}\big[(\tilde{a}^0)^2\big] \leq E_{\tilde{P}\otimes q(\tilde{Y},\cdot)}\big[\tilde{a}^2\big]. \tag{67}$$

For any $A_t^0 \in \mathscr{F}_t^0$ we can write

$$\begin{aligned}
\tilde{E}\big[\tilde{Y}_{t+h} - \tilde{Y}_t; A_t^0\big] &= \tilde{E}\big[\tilde{M}_{t+h} - \tilde{M}_t; A_t^0\big] + \tilde{E}\big[\tilde{A}_{t+h} - \tilde{A}_t; A_t^0\big] \\
&= \tilde{E}\Big[\int_t^{t+h} \tilde{a}_s d\langle\tilde{Y}\rangle_s; A_t^0\Big] = E_{\tilde{P}\otimes q(\tilde{Y},\cdot)}\big[\tilde{a}; A_t^0 \times (t,t+h]\big] \\
&= E_{\tilde{P}\otimes q(\tilde{Y},\cdot)}\big[\tilde{a}^0; A_t^0 \times (t,t+h]\big] = \tilde{E}\Big[\int_t^{t+h} \tilde{a}_s^0 d\langle\tilde{Y}\rangle_s; A_t^0\Big].
\end{aligned}$$

This implies that the canonical decomposition of the semimartingale $\tilde{Y}$ in the smaller filtration $(\tilde{\mathscr{F}}_t^0)$ is of the form

$$\tilde{Y}_t = \tilde{M}_t^0 + \int_0^t \tilde{a}_s^0 d\langle\tilde{Y}\rangle_s.$$

where $\tilde{M}^0$ is a martingale with respect to $(\tilde{\mathscr{F}}_t^0)$. On the other hand, since the law of $\tilde{Y}$ under $\tilde{P}$ is given by Q, we have

$$\tilde{Y}_t = M_t(\tilde{Y}) + \int_0^t a_s(\tilde{Y}) d\langle\tilde{Y}\rangle_s.$$

Uniqueness of the canonical decomposition implies

$$\tilde{a}^0 = a(\tilde{Y}) \qquad \tilde{P}\otimes q(\tilde{Y},\cdot) - a.s. \tag{68}$$

Thus, inequality (67) yields

$$\tilde{E}\Big[\int_0^1 \tilde{a}_t^2 d\langle\tilde{Y}\rangle_t\Big] \geq \tilde{E}\Big[\int_0^1 a_t^2(\tilde{Y}) d\langle\tilde{Y}\rangle_t\Big] = E_Q\Big[\int_0^1 a_t^2(W) d\langle W\rangle_t\Big],$$

and so we have shown inequality (66). □

The following inequality extends Theorem 6 beyond the case of a martingale measure. As explained in Remark 10 below, it contains inequality (58) for $Q \in \mathscr{Q}_{\mathscr{M}}$, Talagrand's inequality (9) for $Q \ll P$, and Corollary 2 for $W_{\mathscr{H},ad}$ as special cases.

Theorem 7 *For* $Q \in \mathscr{Q}^*_{\mathscr{S}}$,

$$W^2_{\mathscr{S}}(Q,P) \leq 2\big(h(Q|P) + H(Q|Q^*)\big), \tag{69}$$

where Q^* *is the equivalent martingale measure for* Q *defined by* (63). *Equality holds iff* $H(Q|P) < \infty$.

Proof 1) Proposition 4 combined with inequality (60) shows that

$$\begin{aligned} W^2_{\mathscr{S}}(Q,P) &= E_Q\big[\langle W - W^Q\rangle_1 + ||A||^2_{\mathscr{S}}\big] \\ &\leq 2h(Q|P) + E_Q\Big[\int_0^1 a_t^2 d\langle W\rangle_t\Big]. \end{aligned} \tag{70}$$

Since Q^* is equivalent to Q, we have

$$\begin{aligned} H(Q|Q^*) &= E_Q\big[\log\big(dQ^*/dQ\big)^{-1}\big] \\ &= E_Q\Big[\int_0^1 a_t dM_t + \frac{1}{2}\int_0^1 a_t^2 d\langle M\rangle_t\Big]. \end{aligned}$$

But M is a square-integrable martingale under Q and $E_Q[\int_0^1 a_t^2 d\langle M\rangle_t] < \infty$ for $Q \in \mathscr{Q}^*_{\mathscr{S}}$. This implies $E_Q\big[\int_0^1 a_t dM_t\big] = 0$, hence

$$H(Q|Q^*) = \frac{1}{2}E_Q\Big[\int_0^1 a_t^2 d\langle M\rangle_t\Big].$$

Thus,

$$W^2_{\mathscr{S}}(Q,P) \leq E_Q\big[\langle W - W^Q\rangle_1 + ||A||^2_{\mathscr{S}}\big] \leq 2h(Q|P) + 2H(Q|Q^*).$$

and so we have shown inequality (69).

2) Equality in (69) implies equality in (70), hence

$$E_Q\big[\langle W - W^Q\rangle_1\big] = 2h(Q|P).$$

Recall that the left-hand side satisfies equation (56). As in the proof of Theorem 6, it follows that $M = W^Q$. This implies $W = W^Q + A$ and $||A||_{\mathscr{H}} = ||A||_{\mathscr{S}} \in L^2(Q)$, hence

$$H(Q|P) = \frac{1}{2}E_Q\big[||A||^2_{\mathscr{H}}\big] < \infty,$$

due to Proposition 1.

Conversely, $H(Q|P) < \infty$ implies $h(Q|P) = 0$ and $Q \in \mathscr{Q}^*_{\mathscr{S}}$ with $\mathscr{Q}^* = P$, hence $H(Q|Q^*) = H(Q|P)$. Thus, the right-hand side of (69) reduces to $2H(Q|P) = E_Q\left[||B^Q||^2_{\mathscr{H}}\right]$. Moreover, since $W = W^Q + B^Q$ and $\langle W \rangle_t = t$ under Q, we get $A = B^Q$, and the left-hand side becomes $W^2_{\mathscr{S}}(Q,P) = W^2_{\mathscr{H},ad}(Q,P) = E_Q\left[||B^Q||^2_{\mathscr{H}}\right]$. Thus, equality holds in (69). □

Remark 10 Inequality (69) includes inequality (58) for martingale measures as a special case. Indeed, for $Q \in \mathscr{Q}_{\mathscr{M}} \subset \mathscr{Q}^*_{\mathscr{S}}$ we have $Q = Q^*$, hence $H(Q|Q^*) = 0$ and

$$W^2_{\mathscr{S}}(Q,P) \le 2h(Q|P).$$

Part 2) of the preceding proof shows how Talagrand's inequality (9) and the identity (20) for $W_{\mathscr{H},ad}$ follow from Theorem 7.

Acknowledgments: This paper is dedicated to the memory of Mark Davis, a cherished fellow explorer of the rich landscape of Stochastic Analysis, including a long excursion to the uncharted territory of Mathematical Finance.

Work on this paper began in the fall of 2018 while I was giving lectures on "Optimal Transport on Wiener Space" at Jiao Tong University, Shanghai. It is a pleasure to express my thanks to Samuel Drapeau and his colleagues in the School of Mathematical Sciences and in the China Academy of Finance Research for their kind hospitality.

References

1. B. Acciaio, J. Backhoff Veraguas, and A. Zalashko. Causal optimal transport and its links to enlargement of filtrations and continuous-time stochastic optimization. *Stochastic Processes and their Applications*, 9(3):203–228, 2019.
2. R. Eldan and D. Mikulincer. Stability of the Shannon-Stam inequality via the Föllmer process. *Probab. Theor. Relat. Fields*, 177:891–922, 2020.
3. D. Feyel and A.S. Üstünel. Monge-Kantorovich measure transportation and Monge-Ampere equation on Wiener Space. *Probab. Theor. Relat. Fields*, 128(3):347–385, 2004.
4. H. Föllmer. Time reversal on Wiener Space. In *Stochastic Processes - Mathematics and Physics (Bielefeld 1984)*, volume 1158 of *Lecture Notes in Mathematics*, pages 119–129. Springer, 1986.
5. H. Föllmer. Random Fields and Diffusion Processes. In *Ecole d' été de Probabilités de St. Flour XV- XVII, 1985-87*, volume 1362 of *Lecture Notes in Mathematics*, pages 101–203. Springer, Berlin, 1988.
6. H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M.H.A. Davis and R.J. Elliott, editors, *Applied Stochastic Analysis*, volume 5 of *Stochastic Monographs*, pages 389–414. Gordon and Breach, London, New York, 1986.
7. N. Gantert. *Einige grosse Abweichungen der Brownschen Bewegung*. PhD thesis, Universität Bonn, 1991.
8. N. Gantert. Self-similarity of Brownian motion and a large deviation principle for random fields on a binary tree. *Probability Theory and Related Fields*, 98:7–20, 1994.
9. R. Lassalle. Causal transference plans and their Monge-Kantorovich problems. *Stochastic Analysis and Applications*, 36(1):1–33, 2018.

10. J. Lehec. Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré - Probabilités et Statistiques*, 49(3):885–899, 2013.
11. R.S. Liptser and A.N. Shiryayev. *Statistics of Random Processes I, General Theory*. Applications of Mathematics 5. Springer, 1977.
12. P. A. Meyer. Un Cours sur les Intégrales Stochastiques. In *Séminaire de Probabilités X*, volume 511 of *Lecture Notes in Mathematics*, pages 246–400. Springer, 1976.
13. M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.
14. S.S. Vallender. Calculation of the Wasserstein distance between probability distributions on the line. *Theory Probab. Appl.*, 18:784–788, 1973.
15. C. Villani. *Optimal Transport: Old and New*. Grundlehren der Mathematischen Wissenschaften 338. Springer, 2009.

Who Are I: Time Inconsistency and Intrapersonal Conflict and Reconciliation*

Xue Dong He and Xun Yu Zhou

Abstract Time inconsistency is prevalent in dynamic choice problems: a plan of actions to be taken in the future that is optimal for an agent today may not be optimal for the same agent in the future. If the agent is aware of this intra-personal conflict but unable to commit herself in the future to following the optimal plan today, the rational strategy for her today is to reconcile with her future selves, namely to correctly anticipate her actions in the future and then act today accordingly. Such a strategy is named intra-personal equilibrium and has been studied since as early as in the 1950s. A rigorous treatment in continuous-time settings, however, had not been available until a decade ago. Since then, the study on intra-personal equilibrium for time-inconsistent problems in continuous time has grown rapidly. In this chapter, we review the classical results and some recent development in this literature.

1 Introduction

When making dynamic decisions, the decision criteria of an agent at different times may not align with each other, leading to time-inconsistent behavior: an action that is optimal under the decision criterion today may no longer be optimal under the decision criterion at certain future time. A variety of preference models can lead to time inconsistent behaviors, such as those involving present-bias, mean-variance criterion, and probability weighting.

Xue Dong He
The Chinese University of Hong Kong,
e-mail: xdhe@se.cuhk.edu.hk

Xun Yu Zhou
Columbia University,
e-mail: xz2574@columbia.edu

* There is no grammatical error in this title. The word "I" is *plural*; it refers to many different selves over time, which is the premise of this article.

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_8

In his seminal paper, Strotz (1955-1956) describes three types of agents when facing time inconsistency. Type 1, a "spendthrift" (or a naiveté as in the more recent literature), does not recognize the time-inconsistency and at any given time seeks an optimal solution from the vantage point of that moment only. As a result, his strategies are always myopic and change all the times. The next two types are aware of the time inconsistency but act differently. Type 2 is a "precommitter" who solves the optimization problem only once at time 0 and then commits to the resulting strategy throughout, even though she knows that the original solution may no longer be optimal at later times. Type 3 is a "thrift" (or a sophisticated agent) who is unable to precommit and realizes that her future selves may disobey whatever plans she makes now. Her resolution is to compromise and choose *consistent planning* in the sense that she optimizes taking the future disobedience as a *constraint*. In this resolution, the agent's selves at different times are considered to be the players of a game, and a consistent plan chosen by the agent becomes an equilibrium of the game from which no selves are willing to deviate. Such a plan or strategy is referred to as an *intra-personal equilibrium*.

To illustrate the above three types of behavior under time inconsistency, consider an agent who has a planning horizon with a finite end date T and makes decisions at *discrete* times $t \in \{0,1,\ldots,T-1\}$. The agent's decision drives a Markov state process and the agent's decision criterion at time t is to maximize an objective function $J(t,x;\mathbf{u})$, where x stands for the Markovian state at that time and $\mathbf{u}$ represents the agent's strategy. The agent considers Markovian strategies, so $\mathbf{u}$ is a function of time $s \in \{0,1,\ldots,T-1\}$ and the Markovian state at that time. If the agent, at certain time t with state x, is a "pre-committer", she is committed to implementing throughout the remaining horizon the strategy $\mathbf{u}^{\mathrm{pc}}_{(t,x)} = \{\mathbf{u}^{\mathrm{pc}}_{(t,x)}(s,\cdot)|s=t,t+1,\ldots,T-1\}$ that maximizes $J(t,x;\mathbf{u})$, and this strategy is referred to as the *pre-committed strategy* of the agent at time t with state x. If the agent is a "spendthrift", at *every* time t with state x, she is able to implement the pre-committed strategy at that moment only and will change at the next moment; so the strategy that is actually implemented by the agent throughout the horizon is $\mathbf{u}^{\mathrm{n}} = \{\mathbf{u}^{\mathrm{pc}}_{(s,X^{\mathbf{u}^{\mathrm{n}}}(s))}(s,X^{\mathbf{u}^{\mathrm{n}}}(s))|s=0,1,\ldots,T-1\}$, where $X^{\mathbf{u}^{\mathrm{n}}}$ denotes the state process under $\mathbf{u}^{\mathrm{n}}$. This strategy is referred to as the *naïve strategy*. If the agent is a "thrift", she chooses an intra-personal equilibrium strategy $\hat{\mathbf{u}}$: At any time $t \in \{0,1,\ldots,T-1\}$ with any state x at that time, $\hat{\mathbf{u}}(t,x)$ is the optimal action of the agent given that her future selves follow $\hat{\mathbf{u}}$; i.e.,

$$\hat{\mathbf{u}}(t,x) \in \arg\max_{u} J(t,x;\mathbf{u}_{t,u}), \tag{1}$$

where $\mathbf{u}_{t,u}(t,x) := u$ and $\mathbf{u}_{t,u}(s,\cdot) := \hat{\mathbf{u}}(s,\cdot)$ for $s=t+1,\ldots,T-1$.

All the three types of behavior are important from an economic perspective. First, field and experimental studies reveal the popularity of commitment devices to help individuals to fulfill plans that would otherwise be difficult to implement due to lack of self control; see for instance Bryan et al. (2010). The demand for commitment devices implies that some individuals seek for pre-committed strategies in the presence of time inconsistency. Second, empirically observed decision-making behavior implies that some individuals are naivetés. For example, Barberis (2012) shows that a naïve agent would take on a series independent, unfavorable bets and take a gain-exit strategy, and this gambling behavior is commonly observed in casinos. Finally, when an agent foresees the time-inconsistency and a commitment device is either unhelpful or unavailable, the intra-personal equilibrium strategy becomes a rational choice of the agent.

It is important to note that it is hard or perhaps not meaningful to determine which type is superior than the others, simply because there is no uniform criteria to evaluate and compare them. So a naïve strategy, despite its name, is not necessarily inferior to an intra-personal equilibrium in terms of an agent's long-run utility. Indeed, O'Donoghue and Rabin (1999) show that in an optimal stopping problem with an immediate reward and present-biased preferences, a sophisticate agent has a larger tendency to preproperate than a naiveté and thus leads to a lower long-run utility. In this sense, studying the different behaviors under time inconsistency sometimes falls into the realm of being "descriptive" as in behavioral science, rather than being "normative" as in classical decision-making theory.

In this survey article, we focus on reviewing the studies on intra-personal equilibrium of a sophisticated agent in *continuous time*. Intra-personal equilibrium for time-inconsistent problems in discrete time, which is defined through the equilibrium condition (1), has been extensively studied in the literature and generated various economic implications. The extension to the continuous-time setting, however, is nontrivial because in this setting, taking a different action from a given strategy at only one time instant does not change the state process and thus has no impact on the objective function value. As a result, it becomes meaningless to examine whether the agent is willing to deviate from a given strategy at a particular moment by just comparing the objective function values before and after the deviation. To address this issue and to formalize the idea of Strotz (1955-1956), Ekeland and Pirvu (2008), Ekeland and Lazrak (2006), and Björk and Murgoci (2010) assume that the agent's self at each time can implement her strategy in an infinitesimally small, but positive, time period; consequently, her action has an impact on the state process and thus on the objective function. In Section 2 below, we follow the framework of Björk and Murgoci (2014) to define intra-personal equilibria, show a sufficient and necessary condition for an equilibrium, and present the so-called *extended HJB equation* that characterizes the intra-personal equilibrium strategy and the value under this strategy. In Section 3, we further discuss various issues related to intra-personal equilibria.

A close-loop strategy for a control system is a mapping from the historical path of the system state and control to the space of controls. So at each time, the control taken by the agent is obtained by plugging the historical path into this mapping. For example, a Markovian strategy for a Markovian control system is a closed-loop strategy. An open-loop strategy is a collection of controls across time (and across scenarios in case of stochastic control), and at each time the control in this collection is taken, regardless of the historical path of the system state and control. For a classical, time-consistent controlled Markov decision problem, the optimal close-loop strategy and the optimal open-loop strategy yield the same state-control path. For time-inconsistent problems, however, closed-loop and open-loop intra-personal equilibria can be vastly different. In Section 4, we review the study of open-loop intra-personal equilibrium and discuss its connection with closed-loop intra-personal equilibrium.

Optimal stopping problems can be viewed as a special case of control problems, so intra-personal equilibria can be defined similarly for time-inconsistent stopping problems. These problems, however, have very special structures, and by exploiting these structures new notions of intra-personal equilibria have been proposed in the literature. We discuss these in Section 5.

If we discretize a continuous horizon of time and assume that the agent has full self control in each subperiod under the discretization, we can define and derive intra-personal equilibria as in the discrete-time setting. The limits of the intra-personal equilibria as discretization becomes infinitely finer are used by some authors to define intra-personal equilibria for continuous-time problems. In Section 6, we review this thread of research.

Time-inconsistency arises in various economic problems, and for many of them, intra-personal equilibria have been studied and their implications discussed in the literature. In Section 7, we review this literature.

Finally, in Section 8, we review the studies on dynamic consistency preferences. In these studies, starting from a preference model for an agent at certain initial time, the authors attempt to find certain preference models for the agent's future selves such that the pre-committed strategy for the agent at the initial time is also optimal for the agent at any future time and thus can be implemented consistently over time.

2 Extended HJB Equation

Strotz (1955-1956) is the first to study the behavior of a sophisticated agent in the presence of time-inconsistency in a continuous-time model. Without formally defining the notion of intra-personal equilibrium, the author derives a consistent plan of the sophisticated agent. Barro (1999) and Luttmer and Mariotti (2003) also investigate, for certain continuous-time models, consistent plans of sophisticated agents, again without their formal definitions. In a series of papers, Ekeland and Lazrak (2006), Ekeland and Lazrak (2008), and Ekeland and Lazrak (2010) study the classical Ramsey model with a nonexponential discount function and propose for the first time a formal notion of intra-personal equilibrium for deterministic control problems in continuous time. Such a notion is proposed in a stochastic context by Björk and Murgoci (2010), which is later split into two papers, Björk and Murgoci (2014) and Björk et al. (2017), discussing the discrete-time and continuous-time settings, respectively. In this section, we follow the framework of Björk et al. (2017) to define an intra-personal equilibrium strategy and present a sufficient and necessary condition for such a strategy.

2.1 Notations

We first introduce some notations. By convention, $x \in \mathbb{R}^n$ is always a column vector. When a vector x is a row vector, we write it as $x \in \mathbb{R}^{1\times n}$. Denote by $A^\top$ the transpose of a matrix A, and by $\mathrm{tr}(A)$ the trace of a square matrix A. For a differentiable function ξ that maps $x \in \mathbb{R}^m$ to $\xi(x) \in \mathbb{R}^n$, its derivative, denoted as $\xi_x(x)$, is an $n \times m$ matrix with the entry in the i-th row and j-th column denoting the derivative of the i-th component of ξ with respect to the j-th component of x. In particular, for a mapping ξ from $\mathbb{R}^m$ to $\mathbb{R}$, $\xi_x(x)$ is an m-dimensional row vector, and we further denote by ξ_{xx} the Hessian matrix.

Consider ξ that maps $(z,x) \in \mathbb{Z} \times \mathbb{X}$ to $\xi(z,x) \in \mathbb{R}^l$, where $\mathbb{Z}$ is a certain set and $\mathbb{X}$, which represents the state space throughout, is either $\mathbb{R}^n$ or $(0,+\infty)$. ξ is *locally Lipschitz in $x \in \mathbb{X}$, uniformly in $z \in \mathbb{Z}$* if there exists a sequence of compact sets $\{\mathbb{X}_k\}_{k\geq 1}$ with $\cup_{k\geq 1}\mathbb{X}_k = \mathbb{X}$ and a sequence of positive numbers $\{L_k\}_{k\geq 1}$ such that for any $k \geq 1$, $\|\xi(z,x) - \xi(z,x')\| \leq L_k\|x-x'\|, \forall z \in \mathbb{Z}, x, x' \in \mathbb{X}_k$. ξ is *global Lipschitz in $x \in \mathbb{X}$, uniformly in $z \in \mathbb{Z}$* if there exists constant $L > 0$ such that $\|\xi(z,x) - \xi(z,x')\| \leq L\|x-x'\|, \forall z \in \mathbb{Z}, x, x' \in \mathbb{X}$. In the case $\mathbb{X} = \mathbb{R}^n$, ξ is of *linear growth in $x \in \mathbb{X}$, uniformly in $z \in \mathbb{Z}$* if there exists $L > 0$ such that $\|\xi(z,x)\| \leq L(1+\|x\|), \forall z \in \mathbb{Z}, x \in \mathbb{X}$. In the case $\mathbb{X} = (0,+\infty)$, ξ *has a bounded norm in $x \in \mathbb{X}$, uniformly in $z \in \mathbb{Z}$*, if there exists $L > 0$ such that $\|\xi(z,x)\| \leq Lx, \forall z \in \mathbb{Z}, x \in \mathbb{X}$. ξ is of *polynomial growth in $x \in \mathbb{X}$, uniformly in $z \in \mathbb{Z}$* if there exists $L > 0$ and integer $\gamma \geq 1$ such that $\|\xi(z,x)\| \leq L\left(1+\varphi_{2\gamma}(x)\right), \forall z \in \mathbb{Z}, x \in \mathbb{X}$, where $\varphi_{2\gamma}(x) = \|x\|^{2\gamma}$ when $\mathbb{X} = \mathbb{R}^n$ and $\varphi_{2\gamma}(x) = x^{2\gamma} + x^{-2\gamma}$ when $\mathbb{X} = (0,+\infty)$.

Fix integers $r \geq 0$, $q \geq 2r$, and real numbers $a < b$. Consider ξ that maps $(t,x) \in [a,b] \times \mathbb{X}$ to $\xi(t,x) \in \mathbb{R}^l$. We say $\xi \in \mathfrak{C}^{r,q}([a,b]\times\mathbb{X})$ if for any derivative index α with $|\alpha| \leq q-2j$ and $j = 0,\ldots,r$, the partial derivative $\frac{\partial^{j+\alpha}\xi(t,x)}{\partial t^j \partial x^\alpha} := \frac{\partial^{j+\alpha_1+\cdots+\alpha_n}\xi(t,x)}{\partial t^j \partial x_1^{\alpha_1}\ldots\partial x_n^{\alpha_n}}$ exists for any $(t,x) \in (a,b)\times\mathbb{X}$ and can be extended to and continuous on $[a,b]\times\mathbb{X}$. We say $\xi \in \bar{\mathfrak{C}}^{r,q}([a,b]\times\mathbb{X})$ if $\xi \in \mathfrak{C}^{r,q}([a,b]\times\mathbb{X})$ and $\frac{\partial^{j+\alpha}\xi(t,x)}{\partial t^j\partial x^\alpha}$ is of polynomial growth in $x \in \mathbb{X}$, uniformly in $t \in [a,b]$, for any derivative index α with $|\alpha| \leq q-2j$ and $j = 0,\ldots,r$.

2.2 Time-Inconsistent Stochastic Control Problems

Let be given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a standard d-dimensional Brownian motion $W(t) := \left(W_1(t),\ldots,W_d(t)\right)^\top, t \geq 0$, on the space, along with the filtration $(\mathscr{F}_t)_{t\geq 0}$ generated by the Brownian motion and augmented by the $\mathbb{P}$-null sets. Consider an agent who makes dynamic decisions in a given period $[0,T]$, and for any $(t,x) \in [0,T)\times\mathbb{X}$, the agent faces the following stochastic control problem:

$$\begin{cases} \max_{\mathbf{u}} & J(t,x;\mathbf{u}) \\ \text{subject to} & dX^{\mathbf{u}}(s) = \mu(s,X^{\mathbf{u}}(s),\mathbf{u}(s,X^{\mathbf{u}}(s)))ds \\ & \quad +\sigma(s,X^{\mathbf{u}}(s),\mathbf{u}(s,X^{\mathbf{u}}(s)))dW(s),\ s\in[t,T] \\ & X^{\mathbf{u}}(t) = x. \end{cases} \tag{2}$$

The agent's dynamic decisions are represented by a Markov strategy $\mathbf{u}$, which maps $(s,y)\in[0,T)\times\mathbb{X}$ to $\mathbf{u}(s,y)\in\mathbb{U}\subset\mathbb{R}^m$. The controlled diffusion process $X^{\mathbf{u}}$ under $\mathbf{u}$ takes values in $\mathbb{X}$, which as aforementioned is set to be either $(0,+\infty)$ or $\mathbb{R}^n$. μ and σ are measurable mappings from $[0,T]\times\mathbb{X}\times\mathbb{U}$ to $\mathbb{R}^n$ and to $\mathbb{R}^{n\times d}$, respectively, where n stands for the dimension of $\mathbb{X}$.

The agent's goal at $(t,x)\in[0,T]\times\mathbb{X}$ is to maximize the following objective function:

$$\begin{aligned} J(t,x;\mathbf{u}) = \mathbb{E}_{t,x}\left[\int_t^T C\big(t,x,s,X^{\mathbf{u}}(s),\mathbf{u}(s,X^{\mathbf{u}}(s))\big)ds + F\big(t,x,X^{\mathbf{u}}(T)\big)\right] \\ + G\big(t,x,\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]\big), \end{aligned} \tag{3}$$

where C is a measurable mapping from $[0,T)\times\mathbb{X}\times[0,T]\times\mathbb{X}\times\mathbb{U}$ to $\mathbb{R}$, and F and G are measurable mappings from $[0,T)\times\mathbb{X}\times\mathbb{X}$ to $\mathbb{R}$. Here and hereafter, $\mathbb{E}_{t,x}[Z]$ denotes the expectation of Z conditional on $X^{\mathbf{u}}(t)=x$. If C, F, and G are independent of (t,x) and $G\big(t,x,\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]\big)$ is linear in $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$, then $J(t,x;\mathbf{u})$ becomes a standard objective function in classical stochastic control where time consistency holds. Thus, with objective function (3), time inconsistency arises from the dependence of C, F, and G on (t,x) as well as from the nonlinearity of $G\big(t,x,\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]\big)$ in $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$.

For any feedback strategy $\mathbf{u}$, denote

$$\begin{aligned} &\mu^{\mathbf{u}}(t,x) := \mu(t,x,\mathbf{u}(t,x)),\ \sigma^{\mathbf{u}}(t,x) := \sigma(t,x,\mathbf{u}(t,x)), \\ &\Upsilon^{\mathbf{u}}(t,x) := \sigma(t,x,\mathbf{u}(t,x))\sigma(t,x,\mathbf{u}(t,x))^\top,\ C^{\tau,y,\mathbf{u}}(t,x) := C\big(\tau,y,t,x,\mathbf{u}(t,x)\big). \end{aligned}$$

With a slight abuse of notation, $u\in\mathbb{U}$ also denotes the feedback strategy $\mathbf{u}$ such that $\mathbf{u}(t,x)=u, \forall(t,x)\in[0,T]\times\mathbb{X}$; so $\mathbb{U}$ also stands for the set of all *constant* strategies when no ambiguity arises.

We need to impose conditions on a strategy $\mathbf{u}$ to ensure the existence and uniqueness of the SDE in (2) and the well-posedness of the objective function $J(t,x;\mathbf{u})$. This consideration leads to the following definition of feasibility:

Definition 1 A feedback strategy $\mathbf{u}$ is *feasible* if the following hold:

(i) $\mu^{\mathbf{u}}$, $\sigma^{\mathbf{u}}$ are locally Lipschitz in $x\in\mathbb{X}$, uniformly in $t\in[0,T]$.
(ii) $\mu^{\mathbf{u}}$ and $\sigma^{\mathbf{u}}$ are of linear growth in $x\in\mathbb{X}$, uniformly in $t\in[0,T]$, when $\mathbb{X}=\mathbb{R}^n$ and have bounded norm in $x\in\mathbb{X}$, uniformly in $t\in[0,T]$, when $\mathbb{X}=(0,+\infty)$.
(iii) For each fixed $(\tau,y)\in[0,T)\times\mathbb{X}$, $C^{\tau,y,\mathbf{u}}(t,x)$ and $F(\tau,y,x)$ are of polynomial growth in $x\in\mathbb{X}$, uniformly in $t\in[0,T]$.
(iv) For each fixed $(\tau,y)\in[0,T)\times\mathbb{X}$ and $x\in\mathbb{X}$, $\mu^{\mathbf{u}}(t,x)$ and $\sigma^{\mathbf{u}}(t,x)$ are right-continuous in $t\in[0,T)$ and $\lim_{t'\geq t,(t',x')\to(t,x)} C^{\tau,y,\mathbf{u}}(t',x') = C^{\tau,y,\mathbf{u}}(t,x)$ for any $t\in[0,T)$.

Denote the set of feasible strategies as $\mathbf{U}$.

We impose the following assumption:

Assumption 1 Any $u \in \mathbb{U}$ is feasible.

2.3 Intra-Personal Equilibrium

Here and hereafter, $\hat{\mathbf{u}} \in \mathbf{U}$ denotes a given strategy and we examine whether it is an equilibrium strategy. For given $t \in [0,T)$, $\varepsilon \in (0, T-t)$ and $\mathbf{a} \in \mathbf{U}$, define

$$\mathbf{u}_{t,\varepsilon,\mathbf{a}}(s,y) := \begin{cases} \mathbf{a}(s,y), & s \in [t, t+\varepsilon), y \in \mathbb{X} \\ \hat{\mathbf{u}}(s,y), & s \notin [t, t+\varepsilon), y \in \mathbb{X}. \end{cases} \tag{4}$$

Imagine that the agent at time t chooses strategy $\mathbf{a}$ and is able to commit herself to this strategy in the period $[t, t+\varepsilon)$. The agent, however, is unable to control her future selves beyond this small time period, namely in the period $[t+\varepsilon, T)$ and believes that her future selves will take strategy $\hat{\mathbf{u}}$. Then, $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$ is the strategy that the agent at time t expects herself to implement throughout the entire horizon. Note that $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$ is feasible because both $\hat{\mathbf{u}}$ and $\mathbf{a}$ are feasible.

Definition 2 (Intra-Personal Equilibrium)

$\hat{\mathbf{u}} \in \mathbf{U}$ is an *intra-personal equilibrium* if for any $x \in \mathbb{X}$, $t \in [0,T)$, and $\mathbf{a} \in \mathbf{U}$, we have

$$\limsup_{\varepsilon \downarrow 0} \frac{J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}})}{\varepsilon} \leq 0. \tag{5}$$

For each positive ε, $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$ leads to a possibly different state process and thus to a different objective function value from those of $\hat{\mathbf{u}}$, so it is meaningful to compare the objective function values of $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$ and $\hat{\mathbf{u}}$ to examine whether the agent is willing to deviate from $\hat{\mathbf{u}}$ to $\mathbf{a}$ in the period of time $[t, t+\varepsilon)$. Due to the continuous-time nature, the length of the period, ε, during which the agent at t exerts full self control, must be set to be infinitesimally small. Then, $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})$ and $J(t,x;\hat{\mathbf{u}})$ become arbitrarily close to each other; so instead of evaluating their difference, we consider the *rate of increment* in the objective function value, i.e., the limit on the left-hand side of (5). Thus, under Definition 2, a strategy $\hat{\mathbf{u}}$ is an intra-personal equilibrium if at any given time and state, the rate of increment in the objective value when the agent deviates from $\hat{\mathbf{u}}$ to any alternative strategy is nonpositive. As a result, the agent has little incentive to deviate from $\hat{\mathbf{u}}$.

2.4 Sufficient and Necessary Condition

We first introduce the generator of the controlled state process. Given $\mathbf{u} \in \mathbf{U}$ and interval $[a,b] \subseteq [0,T]$, consider ξ that maps $(t,x) \in [a,b] \times \mathbb{X}$ to $\xi(t,x) \in \mathbb{R}$. Suppose $\xi \in \mathfrak{C}^{1,2}([a,b] \times \mathbb{X})$, and denote by ξ_t, ξ_x, and ξ_{xx} respectively its first-order partial derivative in t, first-order partial derivative in x, and second-order partial derivative in x. Define the following generator:

$$\mathscr{A}^{\mathbf{u}}\xi(t,x) = \xi_t(t,x) + \xi_x(t,x)\mu^{\mathbf{u}}(t,x) + \frac{1}{2}\mathrm{tr}\left(\xi_{xx}(t,x)^\top \Upsilon^{\mathbf{u}}(t,x)\right),$$
$$t \in [a,b], x \in \mathbb{X}. \tag{6}$$

For each fixed $(\tau,y) \in [0,T) \times \mathbb{X}$, denote

$$f^{\tau,y}(t,x) := \mathbb{E}_{t,x}[F(\tau,y,X^{\hat{\mathbf{u}}}(T))], \tag{7}$$
$$g(t,x) := \mathbb{E}_{t,x}[X^{\hat{\mathbf{u}}}(T)],\ t \in [0,T], x \in \mathbb{X}. \tag{8}$$

In addition, for fixed $(\tau,y) \in [0,T) \times \mathbb{X}$ and $s \in [0,T]$, denote

$$c^{\tau,y,s}(t,x) := \mathbb{E}_{t,x}[C^{\tau,y,\hat{\mathbf{u}}}(s,X^{\hat{\mathbf{u}}}(s))],\ t \in [0,s], x \in \mathbb{X}. \tag{9}$$

In the following, $\mathscr{A}^{\mathbf{u}} f^{\tau,y}$ denotes the function that is obtained by applying the operator $\mathscr{A}^{\mathbf{u}}$ to $f^{\tau,y}(t,x)$ as a function of (t,x) while fixing (τ,y). Then, $\mathscr{A}^{\mathbf{u}} f^{t,x}(t,x)$ denotes the value of $\mathscr{A}^{\mathbf{u}} f^{\tau,y}$ at (t,x) while (τ,y) is also set at (t,x).The above notations also apply to $C^{\tau,y,\mathbf{u}}$ and $c^{\tau,y,s}$.

To illustrate how to evaluate $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}})$ and thus the rate of increment, let us consider the second term in the objective function (3). An informal calculation yields

$$\begin{aligned}
&\mathbb{E}_{t,x}[F(t,x,X^{\mathbf{u}_{t,\varepsilon,\mathbf{a}}}(T))] - \mathbb{E}_{t,x}\left[F(t,x,X^{\hat{\mathbf{u}}}(T))\right] \\
&= \mathbb{E}_{t,x}\left[\mathbb{E}_{t+\varepsilon,X^{\mathbf{u}_{t,\varepsilon,\mathbf{a}}}(t+\varepsilon)}[F(t,x,X^{\mathbf{u}_{t,\varepsilon,\mathbf{a}}}(T))]\right] - \mathbb{E}_{t,x}\left[F(t,x,X^{\hat{\mathbf{u}}}(T))\right] \\
&= \mathbb{E}_{t,x}\left[f^{(t,x)}(t+\varepsilon,X^{\mathbf{a}}(t+\varepsilon))\right] - f^{(t,x)}(t,x) \\
&\approx \mathscr{A}^{\mathbf{a}} f^{t,x}(t,x)\varepsilon,
\end{aligned}$$

where the second equality holds because $\mathbf{u}_{t,\varepsilon,\mathbf{a}}(s,\cdot) = \mathbf{a}(s,\cdot)$ for $s \in [t,t+\varepsilon)$ and $\mathbf{u}_{t,\varepsilon,\mathbf{a}}(s,\cdot) = \hat{\mathbf{u}}(s,\cdot)$ for $s \in [t+\varepsilon,T)$ in addition to the definition of $f^{\tau,y}$ in (7). The change of the other terms in the objective function when the agent deviates from $\hat{\mathbf{u}}$ to $\mathbf{a}$ in the period $[t,t+\varepsilon)$ can be evaluated similarly. As a result, we can derive the rate of increment in the objective value, namely the limit on the left-hand side of (5), which in turn enables us to derive a sufficient and necessary condition for $\hat{\mathbf{u}}$ to be an intra-personal equilibrium.

To formalize the above heuristic argument, we need to impose the following assumption:

Assumption 2 For any fixed $(\tau,y)\in[0,T)\times\mathbb{X}$ and $t\in[0,T)$, there exists $\tilde{t}\in(t,T]$ such that (i) $f^{\tau,y},g\in\bar{\mathfrak{C}}^{1,2}([t,\tilde{t}]\times\mathbb{X})$; (ii) $c^{\tau,y,s}\in\mathfrak{C}^{1,2}([t,\tilde{t}\wedge s]\times\mathbb{X})$ for each fixed $s\in(t,T]$ and $\frac{\partial^{j+\alpha}c^{\tau,y,s}(t',x')}{\partial t^j\partial x^\alpha}$ is of polynomial growth in $x'\in\mathbb{X}$, uniformly in $t'\in[t,\tilde{t}\wedge s]$ and $s\in(t,T]$, for any α with $|\alpha|\leq 2-2j$ and $j=0,1$; and (iii) $G(\tau,y,z)$ is continuously differentiable in z, with the partial derivative denoted as $G_z(\tau,y,z)$.

Theorem 1 *Suppose Assumptions 1 and 2 hold. Then, for any* $(t,x)\in[0,T)\times\mathbb{X}$ *and* $\mathbf{a}\in\mathbf{U}$, *we have*

$$\lim_{\varepsilon\downarrow 0}\frac{J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})-J(t,x;\hat{\mathbf{u}})}{\varepsilon}=\Gamma^{t,x,\hat{\mathbf{u}}}(t,x;\mathbf{a}),\tag{10}$$

where for any $(\tau,y)\in[0,T)\times\mathbb{X}$,

$$\begin{aligned}\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a}):=&\,C^{\tau,y,\mathbf{a}}(t,x)-C^{\tau,y,\hat{\mathbf{u}}}(t,x)+\int_t^T\mathscr{A}^{\mathbf{a}}c^{\tau,y,s}(t,x)ds\\&+\mathscr{A}^{\mathbf{a}}f^{\tau,y}(t,x)+G_z(\tau,y,g(t,x))\mathscr{A}^{\mathbf{a}}g(t,x).\end{aligned}\tag{11}$$

Moreover, $\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a})=\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\tilde{\mathbf{a}})$ *for any* $\mathbf{a},\tilde{\mathbf{a}}\in\mathbf{U}$ *with* $\mathbf{a}(t,x)=\tilde{\mathbf{a}}(t,x)$ *and* $\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a})=0$ *if* $\mathbf{a}(t,x)=\hat{\mathbf{u}}(t,x)$. *Consequently,* $\hat{\mathbf{u}}$ *is an intra-personal equilibrium if and only if*

$$\Gamma^{t,x,\hat{\mathbf{u}}}(t,x;u)\leq 0,\ \forall u\in\mathbb{U},x\in\mathbb{X},t\in[0,T).\tag{12}$$

Theorem 1 presents a sufficient and necessary condition (12) for an intra-personal equilibrium $\hat{\mathbf{u}}$. Because $\Gamma^{\tau,y,\hat{\mathbf{u}}}(\tau,y;\hat{\mathbf{u}}(t,x))=0$, we have

$$\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a})=\Pi^{\tau,y}(t,x;\mathbf{a})-\Pi^{\tau,y}(t,x;\hat{\mathbf{u}}),$$

where

$$\begin{aligned}\Pi^{\tau,y}(t,x;\mathbf{a}):=&\,C^{\tau,y,\mathbf{a}}(t,x)+\int_t^T\mathscr{A}^{\mathbf{a}}c^{\tau,y,s}(t,x)ds+\mathscr{A}^{\mathbf{a}}f^{\tau,y}(t,x)\\&+G_z(\tau,y,g(t,x))\mathscr{A}^{\mathbf{a}}g(t,x).\end{aligned}\tag{13}$$

As a result, condition (12) is equivalent to

$$\max_{u\in\mathbb{U}}\Gamma^{t,x,\hat{\mathbf{u}}}(t,x;u)=0,x\in\mathbb{X},t\in[0,T)\tag{14}$$

or

$$\hat{\mathbf{u}}(t,x)\in\arg\max_{u\in\mathbb{U}}\Pi^{t,x}(t,x;u),x\in\mathbb{X},t\in[0,T).\tag{15}$$

This can be regarded as a time-inconsistent version of the verification theorem in (classical) stochastic control.

The proof of Theorem 1 can be found in Björk et al. (2017) and He and Jiang (2019). Assumption 1 is easy to verify because it involves only the model parameters, i.e., μ, σ, C, F, and G. Assumption 2 imposes some regularity conditions on $\mathbf{\hat{u}}$, which usually requires $\mathbf{\hat{u}}$ to be smooth to a certain degree; see He and Jiang (2019) for a sufficient condition for this assumption. As a result, the sufficient and necessary condition (12) cannot tell us whether there exists any intra-personal equilibrium among the strategies that do not satisfy Assumption 2. This condition, however, is still very useful for us to find intra-personal equilibria for specific problems. Indeed, in most time-inconsistent problems in the literature, intra-personal equilibrium can be found and verified using (12); see Section 7.

2.5 Extended HJB

Define the *continuation value* of a strategy $\mathbf{\hat{u}}$, denoted as $V^{\mathbf{\hat{u}}}(t,x),(t,x)\in[0,T]\times\mathbb{X}$, to be the objective value over time and state under this strategy, i.e.,

$$V^{\mathbf{\hat{u}}}(t,x) := J(t,x;\mathbf{\hat{u}}) = H^{t,x}(t,x) + G\big(t,x,g(t,x)\big), \tag{16}$$

where

$$\begin{aligned} H^{\tau,y}(t,x) &:= \mathbb{E}_{t,x}\left[\int_t^T C^{\tau,y,\mathbf{\hat{u}}}(s,X^{\mathbf{\hat{u}}}(s))ds + F(\tau,y,X^{\mathbf{\hat{u}}}(T))\right] \\ &= \int_t^T c^{\tau,y,s}(t,x)ds + f^{\tau,y}(t,x). \end{aligned} \tag{17}$$

Assuming certain regularity conditions and applying the operator $\mathscr{A}^u$ to $V^{\mathbf{\hat{u}}}(t,x)$, we derive

$$\begin{aligned} \mathscr{A}^u V^{\mathbf{\hat{u}}}(t,x) = &-C^{t,x,\mathbf{\hat{u}}}(t,x) + \int_t^T \mathscr{A}^u c^{t,x,s}(t,x)ds + \mathscr{A}^u f^{t,x}(t,x) \\ &+ G_z\big(t,x,g(t,x)\big)\mathscr{A}^u g(t,x) + \mathscr{A}^u_{\tau,y}H^{t,x}(t,x) + \mathscr{A}^u_{\tau,y}G(t,x,g(t,x)) \\ &+ \mathrm{tr}\left(\Big(H^{t,x}_{xy}(t,x) + G_{zy}(t,x,g(t,x))^\top g_x(t,x)\Big)^\top \Upsilon^u(t,x)\right) \\ &+ \frac{1}{2}G_{zz}\big(t,x,g(t,x)\big)\mathrm{tr}\Big(g_x(t,x)g_x(t,x)^\top \Upsilon^u(t,x)\Big) \end{aligned}$$

where $H^{\tau,y}_{xy}(t,x)$ denotes the cross partial derivative of $H^{\tau,y}(t,x)$ in x and y, $G_{zy}(\tau,y,z)$ the cross partial derivative of $G(\tau,y,z)$ in z and y, and $G_{zz}(\tau,y,z)$ the second-order derivative of $G(\tau,y,z)$ in z. For each *fixed* (t,x), $\mathscr{A}^u_{\tau,y}H^{\tau,y}(t,x)$ denotes the generator of $\mathscr{A}^u$ applied to $H^{\tau,y}(t,x)$ as a function of (τ,y), i.e., $\mathscr{A}^u_{\tau,y}H^{\tau,y}(t,x) := \mathscr{A}^u\ell(\tau,y)$, where $\ell(\tau,y) := H^{\tau,y}(t,x),(\tau,y)\in[0,T)\times\mathbb{X}$, and $\mathscr{A}^u_{\tau,y}G(\tau,y,g(t,x))$ is defined similarly.

Now, suppose $\mathbf{\hat{u}}$ is an intra-personal equilibrium. Recalling (11) and the sufficient and necessary condition (14), we derive the following equation satisfied by the continuation value of an intra-personal equilibrium $\mathbf{\hat{u}}$:

$$\begin{aligned}
&\max_{u\in\mathbb{U}}\Big[\mathscr{A}^u V^{\hat{\mathbf{u}}}(t,x)+C^{t,x,u}(t,x)-\big(\mathscr{A}^u_{\tau,y}H^{t,x}(t,x)+\mathscr{A}^u_{\tau,y}G(t,x,g(t,x))\big)\\
&-\operatorname{tr}\bigg(\Big(H^{t,x}_{xy}(tx)+G_{zy}(t,x,g(t,x))^\top g_x(t,x)\Big)^\top \Upsilon^u(t,x)\bigg)\\
&-\frac{1}{2}G_{zz}\big(t,x,g(t,x)\big)\operatorname{tr}\Big(g_x(t,x)g_x(t,x)^\top\Upsilon^u(t,x)\Big)\Big]=0,(t,x)\in[0,T)\times\mathbb{X},\\
&V^{\hat{\mathbf{u}}}(T,x)=F(T,x,x)+G(T,x,x),\ x\in\mathbb{X}.
\end{aligned}\tag{18}$$

By (17), the definitions of $c^{\tau,y,s}(t,x)$ and $f^{\tau,y}(t,x)$, and the Feymann-Kac formula, we derive the following equation for $H^{\tau,y}(t,x)$:

$$\begin{aligned}
&\mathscr{A}^{\hat{\mathbf{u}}}H^{\tau,y}(t,x)+C^{\tau,y,\hat{\mathbf{u}}}(t,x)=0,\quad (t,x)\in[0,T)\times\mathbb{X},(\tau,y)\in[0,T)\times\mathbb{X},\\
&H^{\tau,y}(T,x)=F(\tau,y,x),\quad x\in\mathbb{X},(\tau,y)\in[0,T)\times\mathbb{X}.
\end{aligned}\tag{19}$$

Similarly, we derive the following equation for g:

$$\begin{aligned}
&\mathscr{A}^{\hat{\mathbf{u}}}g(t,x)=0,\quad (t,x)\in[0,T)\times\mathbb{X},\\
&g(T,x)=x,\quad x\in\mathbb{X}.
\end{aligned}\tag{20}$$

Some remarks are in order. First, instead of a single equation for the value function of a time-consistent problem, the intra-personal equilibrium and its continuation value satisfy a system of equations (18)–(20), which is referred to as the *extended HJB equation* by Björk et al. (2017).

Second, compared to the HJB equation for a time-consistent problem, which takes the form $\max_{u\in\mathbb{U}}\big[\mathscr{A}^u V^{\hat{\mathbf{u}}}(t,x)+C^u(t,x)\big]=0$, equation (18) has three additional terms in the first, second, and third lines of the equation, respectively. Here and hereafter, when $C^{\tau,y,u}(t,x)$ does not depend on (τ,y), we simply drop the superscript (τ,y). Similar notations apply to $H^{\tau,y}(t,x)$ and to the case when there is no dependence on y. Now, recall that for the objective function (3), time inconsistency arises from (i) the dependence of C, F, and G on (t,x) and (ii) the nonlinear dependence of $G\big(t,x,\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]\big)$ on $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$. If source (i) of time inconsistency is absent, the first and second additional terms in (18) will vanish. If source (ii) of time inconsistency is absent, the third additional term in (18) will disappear. In particular, without time inconsistency, the extended HJB equation (18) reduces to the classical HJB equation.

Third, consider the case in which $G(t,x,\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)])$ is linear in $\mathbb{E}_{t,x}[X^{\mathbf{u}}(T)]$ and C, F, and G do not depend on x. In this case, the second and third lines of (18) vanish and we can assume $G\equiv 0$ without loss of generality because G can be combined with F. As a result, the extended HJB equation (18) specializes to

$$\begin{aligned}
&\max_{u\in\mathbb{U}}\Big[\mathscr{A}^u V^{\hat{\mathbf{u}}}(t,x)+C^{t,u}(t,x)\Big]=h^t(t,x),(t,x)\in[0,T)\times\mathbb{X},\\
&V^{\hat{\mathbf{u}}}(T,x)=F(T,x),\ x\in\mathbb{X}
\end{aligned}\tag{21}$$

where $h^\tau(t,x) := H^\tau_\tau(t,x)$ (with the subscript τ denoting the partial derivative with respect to τ) and thus satisfies

$$\begin{aligned}&\mathscr{A}^{\hat{\mathbf{u}}} h^\tau(t,x) + C^{\tau,\hat{\mathbf{u}}}_\tau(t,x) = 0, \quad (t,x) \in [0,T) \times \mathbb{X}, \tau \in [0,T),\\ &h^\tau(T,x) = F_\tau(\tau,x), \quad x \in \mathbb{X}, \tau \in [0,T).\end{aligned} \tag{22}$$

3 Discussions

3.1 Intra-Personal Equilibria with Fixed Initial Data

Consider an agent at time 0 with a fixed state x_0 who correctly anticipates that her self at each future time t faces the problem (2) and who has no control of future selves at any time. A strategy $\hat{\mathbf{u}}$ can be consistently implemented by the agent throughout the entire horizon $[0,T]$ if the agent has no incentive to deviate from it at any time *along the state path*. Actions that the agent might be taking were she not on the state path are irrelevant. To be more precise, for any fixed initial data $(0,x_0)$, we define $\hat{\mathbf{u}}$ to be an *intra-personal equilibrium starting from* $(0,x_0)$ if (5) holds for any $\mathbf{a} \in \mathbf{U}$, $t \in [0,T)$, and $x \in \mathbb{X}^{0,x_0,\hat{\mathbf{u}}}_t$, where $\mathbb{X}^{0,x_0,\hat{\mathbf{u}}}_t$ denotes the set of all possible states at time t along the state path starting from x_0 at the initial time and under the strategy $\hat{\mathbf{u}}$.

It is evident that the intra-personal equilibrium defined in Definition 2 is *universal* in that it is an equilibrium starting from *any* initial data $(0,x_0)$. On the other hand, starting from a fixed state x_0 at time 0, the state process in the future might not be able to visit the whole state space; so an equilibrium starting from $(0,x_0)$ is not necessarily universal, i.e., it is not necessarily an equilibrium when the agent starts from other initial data. For example, He et al. (2020) consider a continuous-time portfolio selection problem in which an agent maximizes the median of her terminal wealth. With a fixed initial wealth of the agent, the authors derive a set of intra-personal equilibrium strategies starting from this particular initial wealth level. They show that these strategies are no longer equilibria if the agent starts from some other initial wealth levels, and in particular not universal equilibria in the sense of Definition 2.

The first study of intra-personal equilibria starting from a fixed initial data dates back to Peleg and Yaari (1973). In a discrete-time setting, the authors propose that a strategy $(s^*_0, s^*_1, \dots)$, where s^*_t stands for the agent's closed-loop strategy at time t, is an equilibrium strategy if for any t, $(s^*_0, \dots, s^*_{t-1}, s_t, s^*_{t+1}, \dots)$ is dominated by $(s^*_0, \dots, s^*_{t-1}, s^*_t, s^*_{t+1}, \dots)$ for any s_t. They argue that the above definition is more desirable than the following one, which is based on a model in Pollak (1968): $(s^*_0, s^*_1, \dots)$ is an equilibrium strategy if for any time t, $(s_0, \dots, s_{t-1}, s_t, s^*_{t+1}, \dots)$ is dominated by $(s_0, \dots, s_{t-1}, s^*_t, s^*_{t+1}, \dots)$ for *any* $(s_0, \dots, s_t)$. It is clear that the equilibrium strategies considered by Peleg and Yaari (1973) are the ones starting from a fixed initial data while those studied by Pollak (1968) are universal. Recently, He and Jiang (2019), Han and Wong (2020), and Hernández and Possamaï (2020) also consider intra-personal equilibria with fixed initial data. Moreover, He and Jiang (2019) propose a formal definition of $\mathbb{X}_t^{0,x_0,\hat{\mathbf{u}}}$, calling it the set of reachable states.

Finally, let us comment that the sufficient and necessary condition in Theorem 1 is still valid for intra-personal equilibria starting from fixed initial data $(0,x_0)$, provided that we replace $\mathbb{X}$ in this condition with the set of reachable states $\mathbb{X}_t^{0,x_0,\hat{\mathbf{u}}}$; see He and Jiang (2019) for details. The extended HJB equation in Section 2.5 can be revised and applied similarly.

3.2 Set of Alternative Strategies

In Definition 2, the set of strategies that the agent can choose at time t to implement for the period $[t,t+\varepsilon)$, denoted as $\mathbf{D}$, is set to be the entire set of feasible strategies $\mathbf{U}$. This definition is used in Björk et al. (2017), Ekeland and Pirvu (2008), and Ekeland et al. (2012). In some other works, however, $\mathbf{D}$ is set to be the set of constant strategies $\mathbb{U}$; see for instance Ekeland and Lazrak (2006, 2008, 2010), Björk and Murgoci (2010), and Basak and Chabakauri (2010). He and Jiang (2019) show that the choice of $\mathbf{D}$ is irrelevant as long as it at least contains $\mathbb{U}$. Indeed, this can be seen from the observation in Theorem 1 that $\Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a}) = \Gamma^{\tau,y,\hat{\mathbf{u}}}(t,x;\mathbf{a}(t,x))$ for any $\mathbf{a}\in\mathbf{U}$. He and Jiang (2019) also show that for strong intra-personal equilibrium, which will be introduced momentarily, the choice of $\mathbf{D}$ is relevant.

3.3 Regular and Strong Intra-Personal Equilibrium

As noted in Remark 3.5 of Björk et al. (2017), condition (5) does not necessarily imply that $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})$ is less than or equal to $J(t,x;\hat{\mathbf{u}})$ however small $\varepsilon>0$ might be and thus disincentivizes the agent from deviating from $\hat{\mathbf{u}}$. For example, if $J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})-J(t,x;\hat{\mathbf{u}})=\varepsilon^2$, then (5) holds, but the agent can achieve a strictly larger objective value if she deviates from $\hat{\mathbf{u}}$ to $\mathbf{a}$ and thus is willing to do so.

To address the above issue, Huang and Zhou (2019) and He and Jiang (2019) propose the notion of *strong intra-personal equilibrium*:

Definition 3 (Strong Intra-personal Equilibrium)

$\hat{\mathbf{u}}\in\mathbf{U}$ is a *strong intra-personal equilibrium strategy* if for any $x\in\mathbb{X}$, $t\in[0,T)$, and $\mathbf{a}\in\mathbf{D}$, there exists $\varepsilon_0\in(0,T-t)$ such that

$$J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}})-J(t,x;\hat{\mathbf{u}})\le 0,\quad \forall\varepsilon\in(0,\varepsilon_0]. \tag{23}$$

It is straightforward to see that a strong intra-personal equilibrium implies the one in Definition 2, which we refer to as a *weak intra-personal equilibrium* in this subsection.

Huang and Zhou (2019) consider a stochastic control problem in which an agent can control the generator of a time-homogeneous, continuous-time, finite-state Markov chain at each time to maximize expected running reward in an infinite time horizon. Assuming that at each time the agent can implement a time-homogeneous strategy only, the authors provide a characterization of a strong intra-personal equilibrium and prove its existence under certain conditions.

He and Jiang (2019) follow the framework in (2) and derive two necessary conditions for a strategy to be strong intra-personal equilibrium. Using these conditions, the authors show that strong intra-personal equilibrium does not exist for the portfolio selection and consumption problems studied in Ekeland and Pirvu (2008), Basak and Chabakauri (2010), and Björk et al. (2014). Motivated by this non-existence result, the authors propose the so-called *regular intra-personal equilibrium* and show that it exists for the above three problems and is stronger than the weak intra-personal equilibrium and weaker than the strong intra-personal equilibrium in general.

3.4 Existence and Uniqueness

In most studies on time-inconsistent problems in the literature, a *closed-form* strategy is constructed and verified to satisfy the sufficient and necessary condition (12) or the extended HJB equation (18)–(20). The existence of intra-personal equilibrium in general is difficult to prove because it essentially relies on a fixed point argument: For each guess of intra-personal equilibrium $\hat{\mathbf{u}}$, we first calculate $\Gamma^{\tau,y,\hat{\mathbf{u}}}$ in (12) and $H^{\tau,y}(t,x)$ and g in (19) and (20), respectively, and then derive an updated intra-personal equilibrium, denoted as $\mathbb{T}\hat{\mathbf{u}}$, from the condition (12) or from the equation (18). The existence of an intra-personal equilibrium then boils down to the existence of the fixed point of $\mathbb{T}$. The mapping $\mathbb{T}$ is highly nonlinear; so the existence of its fixed point is hard to establish. Additional difficulty is caused by the regularity conditions that we need to pose on $\hat{\mathbf{u}}$ to validate the sufficient and necessary condition (12) or the extended HJB equation (18)–(20).

We are only aware of very few works on the existence of intra-personal equilibria in continuous time. Yong (2012) proposes an alternative approach to defining the strategy of a sophisticated agent, which will be discussed in detail in Section 6. Assuming $G \equiv 0$, C and F to be independent of x in the objective function (3), and $\sigma(t,x,u)$ in the controlled diffusion process (2) to be independent of control u and nondegenerate, Yong (2012) proves the existence of the sophisticated agent's strategy, which is used to imply the existence of an intra-personal equilibrium under Definition 2. Wei et al. (2017) and Wang and Yong (2019) extend the result of Yong (2012) by generalizing the objective function; however for the existence of intra-personal equilibria, they need to assume the volatility σ to be independent of control and nondegenerate. Hernández and Possamaï (2020) study intra-personal equilibria in a non-Markovian setting, where they consider a non-Markovian version of the objective function in Yong (2012) and assume the drift μ of the controlled process to be in the range of the volatility matrix at each time. The authors prove the existence of intra-personal equilibria when the volatility σ is independent of control.

Intra-personal equilibria can be non-unique; see Ekeland and Lazrak (2010), Cao and Werning (2016), and He et al. (2020). For some problems, however, uniqueness has been established in the literature. Indeed, Yong (2012), Wei et al. (2017), Wang and Yong (2019), and Hernández and Possamaï (2020) prove the uniqueness in various settings with the common assumption that the volatility σ is independent of control.

3.5 Non-Markovian Strategies

In most studies on time-inconsistent problems, where the controlled state processes are Markovian, the search for intra-personal equilibrium is restricted to the set of Markovian strategies, i.e., strategies that are functions of time t and the *current* state value x. Motivated by some practical problems such as rough volatility models and principle-agent problems, Han and Wong (2020) and Hernández and Possamaï (2020) define and search intra-personal equilibria in the class of non-Markovian or path-dependent strategies, i.e., ones that depend on time t and the whole path of the controlled state up to t.

4 Closed-Loop versus Open-Loop Intra-Personal Equilibria

A *closed-loop* or *feedback* control strategy is a function $\mathbf{u}$ that maps time t and the controlled state path $(x_s)_{s\le t}$ up to t to the space of actions. As a result, the action taken by an agent under such a strategy is $\mathbf{u}(t,(x_s)_{s\le t})$. An *open-loop* control is a collection of actions over time and state of the nature, $(u(t,\omega))_{t\ge 0}$, where $u(t,\omega)$ is the action to be taken at time t and in scenario ω, *regardless* of the state path $(x_s)_{s\le t}$. For classical time-consistent control problems and under some technical assumptions, the state-control paths under the optimal open-loop control and under the optimal closed-loop control strategy are the same if the controlled system starts from the same initial time and state; see for instance Yong and Zhou (1999).

In Section 2, intra-personal equilibrium is defined for closed-loop control strategies, which is also the approach taken by most studies on time-inconsistent problems in the literature. In some other works, intra-personal equilibrium is defined for open-loop controls; see for instance Hu et al. (2012), Hu et al. (2017), Li et al. (2019), and Hu et al. (2021).

Formally, under the same probabilistic framework in Section 2.2, we represent an open-loop strategy by a progressively measurable process $(u(t))_{t\ge 0}$ that takes values in $\mathbb{U}$. The controlled state process X^u takes the form

$$dX^u(s)=\mu(s,X^u(s),u(s))ds+\sigma(s,X^u(s),u(s))dW(s),\ s\in[t,T];\ X^u(t)=x.$$

Denote by $\mathscr{U}$ the set of feasible open-loop controls, i.e., the set of progressively measurable processes on $[0,T]$ satisfying certain integrability conditions. At time t with state x, suppose the agent's objective is to maximize $J(t,x;u(\cdot))$ by choosing $u(\cdot)\in\mathscr{U}$. Given $\hat{u}(\cdot)\in\mathscr{U}$, for any $t\in[0,T)$, $x\in\mathbb{X}$, $\varepsilon\in(0,T-t)$, and $a(\cdot)\in\mathscr{U}$, define

$$u_{t,\varepsilon,a}(s):=\begin{cases}a(s), & s\in[t,t+\varepsilon)\\ \hat{u}(s), & s\notin[t,t+\varepsilon).\end{cases} \tag{24}$$

Suppose that at time t with state x, the agent chooses an open-loop control $a(\cdot)$, but is only able to implement it in the period $[t,t+\varepsilon)$. Anticipating that her future selves will take the given control $\hat{u}(\cdot)$, the agent expects herself to follow $u_{t,\varepsilon,a}$ in the period $[t,T]$.

Definition 4 (Open-Loop Intra-Personal Equilibrium)

$\hat{u}(\cdot)\in\mathscr{U}$ is an *open-loop intra-personal equilibrium* if for any $x\in\mathbb{X}$, $t\in[0,T)$, and $a\in\mathscr{U}$ that is constant in a small period after t, we have

$$\limsup_{\varepsilon\downarrow 0}\frac{J(t,x;u_{t,\varepsilon,a}(\cdot))-J(t,x;\hat{u}(\cdot))}{\varepsilon}\le 0. \tag{25}$$

The above is analogous to the definition of an intra-personal equilibrium for closed-loop strategies. However, there is a subtle yet crucial difference between the two definitions. For the one for open-loop controls, the perturbed control $u_{t,\varepsilon,a}(s)$ defined by (24) and the original one $\hat{u}$ are identical on $[t+\varepsilon,T]$ as two stochastic processes. In other words, the perturbation in the small time period $[t,t+\varepsilon)$ will not affect the control process beyond this period. This is not the case for the closed-loop counterpart, because the perturbation (4) on $[t,t+\varepsilon)$ changes the control in the period, which will alter the state process in $[t,t+\varepsilon)$ and in particular the state at time $t+\varepsilon$. This in turn will change the control *process* on $[t+\varepsilon,T]$ upon substituting the state process into the feedback strategy.

To characterize open-loop intra-personal equilibria, we only need to compute the limit on the left-hand side of (25). This limit can be evaluated by applying the spike variation technique that is used to derive Pontryagin's maximum principle for time-consistent control problems in continuous time (Yong and Zhou, 1999). As a result, open-loop intra-personal equilibrium can be characterized by a flow of forward-backward stochastic differential equations (SDEs); see Hu et al. (2012) for more details. In contrast, the spike variation technique no longer works for closed-loop equilibria because the perturbed control process is different from the original one beyond the small time period for perturbation, as discussed above.

This discussion suggests that closed-loop and open-loop equilibria are likely different. This is confirmed by Hu et al. (2012). The authors consider a mean-variance portfolio selection problem, where an agent decides the dollar amount invested in a stock at each time, and derive an open-loop equilibrium; see Section 5.4.1 therein. They then compare this equilibrium with the closed-loop equilibrium derived by

Björk et al. (2014) for the same portfolio selection problem, and find that the state-control path under these two equilibria are different.

It can be argued that closed-loop strategies are preferred to the open-loop ones for three reasons. First, in many problems, agents' actions naturally depend on some state variables. For example, in a consumption problem, an agent's consumption at any time is more likely to depend directly on her wealth at that time. If her wealth suddenly increases, she would probably consume more.

Second, closed-loop intra-personal equilibrium is invariant to the choice of control variables while open-loop intra-personal equilibrium might not. For example, in a portfolio selection problem where an agent decides the allocation of her wealth between a risk-free asset and a risky stock, the decision variable can be set to be the dollar amount invested in the stock or the percentage of wealth invested in the stock. Suppose $\hat{\mathbf{u}}$ is a closed-loop intra-personal equilibrium representing the percentage of wealth invested in the stock. Then, we have

$$\limsup_{\varepsilon \downarrow 0} \frac{J(t,x;\mathbf{u}_{t,\varepsilon,\mathbf{a}}) - J(t,x;\hat{\mathbf{u}})}{\varepsilon} \le 0. \tag{26}$$

for all $t \in [0,T)$, $x \in \mathbb{X}$, and $\mathbf{a} \in \mathbf{U}$, where the state variable x represents the agent's wealth. Now, suppose we represent the agent's decision by the dollar amount invested in the risky stock, and denote a control strategy as π. Then, the agent's objective function is $\tilde{J}(t,x;\pi) = J(t,x;\mathbf{u})$ with $\mathbf{u}(s,y) = \pi(s,y)/y$. Condition (26) implies that

$$\limsup_{\varepsilon \downarrow 0} \frac{\tilde{J}(t,x;\pi_{t,\varepsilon,\tilde{\mathbf{a}}}) - \tilde{J}(t,x;\hat{\pi})}{\varepsilon} \le 0,$$

for any $t \in [0,T)$, $x \in \mathbb{X}$, and strategy $\tilde{\mathbf{a}}$ that represents the dollar amount invested in the stock, where $\hat{\pi}(s,y) := y\hat{\mathbf{u}}(s,y)$ and $\pi_{t,\varepsilon,\tilde{\mathbf{a}}}$ is defined similarly to $\mathbf{u}_{t,\varepsilon,\mathbf{a}}$. Thus, $\hat{\pi}$, which is the dollar amount investment strategy corresponding to the percentage investment strategy $\hat{\mathbf{u}}$, is also an intra-personal equilibrium. By contrast, for the mean-variance portfolio selection problem studied by Hu et al. (2012), where the agent's decision is the dollar amount invested in the stock, the open-loop intra-personal equilibrium yields a different control-state path from the one yielded by its closed-loop counterpart derived by Björk et al. (2014). If we change the agent's decision variable to the percentage of wealth invested in the stock, the open-loop intra-personal equilibrium and the closed-loop intra-personal equilibrium in Björk et al. (2014) yield the same control-state path. This implies that open-loop equilibria depend on the choice of control variables.

Third, open-loop intra-personal equilibrium may not be well-posed for some problems. Consider the discrete-time version of the consumption problem studied in Strotz (1955-1956): An agent decides the amount of consumption C_t at each time $t = 0,1,\ldots,T$ with the total budget x_0, i.e., $\sum_{t=0}^{T} C_t = x_0$. For this problem, any consumption plan $(\hat{C}_t)_{t\ge 0}$ is an open-loop intra-personal equilibrium. Indeed, at each time t, anticipating her future selves will consume $\hat{C}_s, s = t+1,\ldots,T$, the only amount of consumption C_t that the agent can choose at time t is $\hat{C}_t$ due to the budget constraint $(\sum_{s=0}^{t-1}\hat{C}_s)+C_t+(\sum_{s=t+1}^{T}\hat{C}_s)=x_0$. This leads to a trivial definition of intra-personal equilibrium. The above issue can be rectified if we use closed-loop strategies. To see this, we set x_t to be the agent's remaining budget at time t before the consumption at that time. For closed-loop intra-personal equilibrium, we consider a mapping from time t and the remaining budget x_t to the consumption amount. As a result, if the agent consumes more at time t, her future selves will consume less because the remaining budget in the future becomes smaller; consequently, the budget constraint is still satisfied. To elaborate, suppose the agent's future selves' strategies are to consume $\hat{k}_s$ fractional of wealth at time s, $s=t+1,\dots,T$ with $\hat{k}_s\in[0,1], s=t+1,\dots,T-1$ and $\hat{k}_T=1$. Then, given that the agent at time t consumes any amount $C_t\in[0,x_t]$, the agent's consumption in the future is $C_s=\hat{k}_s x_s$, $s=t+1,\dots,T$, where $x_s=x_{s-1}-C_{s-1}$, $s=t+1,\dots,T$. As a result, the aggregate consumption from time t to the end is $\sum_{s=t}^{T}C_s=x_t$. Recall that the aggregate consumption strictly prior to time t is x_0-x_t; so the aggregate consumption throughout the entire horizon is x_0 satisfying the budget constraint. Thus, at each time t, the agent can consume any amount up to his wealth level at that time and her future selves will adjust their consumption according to a given strategy so that the budget constraint is still satisfied.

Finally, we establish a connection between closed-loop and open-loop intra-personal equilibria. If a closed-loop equilibrium $\hat{\mathbf{u}}$ is independent of the state variable x, then it follows from the definition that it is also an open-loop equilibrium. For a general closed-loop equilibrium $\hat{\mathbf{u}}$, we can consider the following controlled state process:

$$d\hat{X}^v(s)=\hat{\mu}(s,\hat{X}^v(s),v(s))ds+\hat{\sigma}(s,\hat{X}^v(s),v(s))dW(s),\ s\in[t,T];\ X^v(t)=x,$$

where $\hat{\mu}(s,y,v):=\mu(s,y,\hat{\mathbf{u}}(s,y)+v)$, $\hat{\sigma}(s,y,v):=\sigma(s,y,\hat{\mathbf{u}}(s,y)+v)$, and $v(\cdot)$ is a progressively measurable control process. We further consider the following objective function:

$$\hat{J}(t,x;v(\cdot)):=\mathbb{E}_{t,x}\left[\int_t^T\hat{C}\big(t,x,s,\hat{X}^v(s),v(s)\big)ds+F\big(t,x,\hat{X}^v(T)\big)\right]$$
$$+G\big(t,x,\mathbb{E}_{t,x}[\hat{X}^v(T)]\big),$$

where $\hat{C}(t,x,s,y,v):=C(t,x,s,y,\hat{\mathbf{u}}(s,y)+v)$. Then, by definition, $\hat{\mathbf{u}}$ is a closed-loop equilibrium if and only if $\hat{v}(\cdot)\equiv 0$ is an open-loop equilibrium for the problem of maximizing $\hat{J}(t,x;v(\cdot))$ in $v(\cdot)$ with the controlled state process $\hat{X}^v$. In particular, we can characterize $\hat{\mathbf{u}}$ by a flow of forward-backward SDEs by applying the spike variation technique. In order to apply this technique, however, we need to assume that $\hat{\mu}(s,y,v)$ and $\hat{\sigma}(s,y,v)$ to be twice differentiable in y, which in turn requires $\hat{\mathbf{u}}$ to be twice differentiable; see Yong and Zhou (1999) for the detailed regularity conditions needed for the spike variation technique. Thus, the spike variation technique does not seem to be advantageous over the approached reviewed in Section 2.

5 Optimal Stopping

An optimal stopping problem is one to search an optimal random time τ to stop a given, *uncontrollable* process $(X_t)_{t\geq 0}$ (taking values in a state space $\mathbb{X}$) in the set of stopping times with respect to the filtration generated by the process. It is well known that if the objective function of the optimal stopping problem depends on the path of $(X_t)_{t\geq 0}$ up to the stopping time only, this problem can be "embedded" into a general control problem with (i) a closed-loop control strategy $\mathbf{u}$ taking binary values 0 and 1 representing the action of stopping and not stopping $(X_t)_{t\geq 0}$ respectively; and (ii) a controlled state process $(\tilde{X}^{\mathbf{u}})_{t\geq 0}$ that is set to be $(X_t)_{t\geq 0}$ until the first time the control path under $\mathbf{u}$ takes value 0 and is set to be an *absorbing state* afterwards; see for instance Section 3.4 of Bertsekas (2017). We call the control strategy $\mathbf{u}$ associated with a stopping time τ in the above embedding a *stopping rule*, which maps each pair of time t and a path of the process X up to time t to $\{0,1\}$. A stopping time τ is *Markovian* if the associated stopping rule is Markovian, i.e., it is a mapping from the time–state space to $\{0,1\}$. With a Markovian stopping time, at each time t, given that the process has not yet been stopped, whether to stop at t depends on the value of the process at t only.

In view of the above embedding, intra-personal equilibrium stopping rules can be defined naturally for time-inconsistent stopping problems; see for instance Tan et al. (2018), Christensen and Lindensjö (2018), Ebert et al. (2020), and Christensen and Lindensjö (2020). In particular, Tan et al. (2018) show that the smooth pasting principle, which is the main approach used to construct explicit solutions for classical time-consistent optimal stopping, may fail to find an equilibrium when one changes merely the exponential discounting to non-exponential one while keeping everything else the same. The authors also construct an explicit example in which no equilibrium exists. These results caution blindly extending the classical approach for time-consistent stopping to their time-inconsistent counterpart.

By exploiting special structures of stopping problems in continuous time, Huang and Nguyen-Huu (2018) propose an alternative approach to defining the optimal stopping rule for a sophisticated agen; see also applications of this approach in Huang et al. (2020), Ebert and Strack (2017), and Huang and Yu (2021). Precisely, consider a Markov state process

$$dX_t = \mu(t,X_t)dt + \sigma(t,X_t)dW_t$$

in $\mathbb{R}^n$, where $(W_t)_{t\geq 0}$ is an d-dimensional standard Brownian motion and μ and σ are functions of time t and state x taking values in $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}$, respectively. Following the settings in the above papers, we consider Markovian stopping times only in the following presentation, but the case of non-Markovian stopping times can be investigated similarly. At each time t with state x, give that the state process has not been stopped, the agent's goal is to choose a Markovian stopping time $\tau \in [t,T]$ to maximize an objective value $J(t,x;\tau)$. Here, $J(t,x;\tau)$ can be of the form $\mathbb{E}_{t,x}\left[\int_t^\tau g(t,x,s,X_s)ds + h(t,x,\tau,X_\tau)\right]$ for some functions g and h, or be a functional of the distribution of X_τ conditional on $X_t = x$.

Recall the embedding of optimal stopping problems into a general control framework and the stopping rule associated with each stopping time as discussed at the beginning of the present subsection. With a slight abuse of notation, we use τ to denote both a stopping time and a stopping rule. Let us now consider a given stopping rule τ and the current time–state pair (t,x). If the agent decides to stop, then she has the immediate reward $J(t,x;t)$. If the agent decides not to stop at t but expects her future selves will still follow the original rule τ, then she will stop at time $\mathscr{L}^*\tau$, the first time $s>t$ at which τ would stop the process. In this case the objective value is $J(t,x;\mathscr{L}^*\tau)$. Then, the optimal action of the agent at time t with state x is to stop if $J(t,x;t) > J(t,x;\mathscr{L}^*\tau)$, to continue if $J(t,x;t) < J(t,x;\mathscr{L}^*\tau)$, and to follow the originally assigned stopping rule τ in the break-even case $J(t,x;t) = J(t,x;\mathscr{L}^*\tau)$. The above plan across all time t and state x constitutes a *new* stopping rule, denoted as $\Theta\tau$, which can be proved to be feasible in the sense that it can generate stopping times; see Huang and Nguyen-Huu (2018) and Huang et al. (2020).

The above game-theoretic thinking shows that for any arbitrarily given stopping rule τ, at any time t with any state x, the agent finds $\Theta\tau$ to be always no worse than τ, *assuming* that her future selves will follow τ. Hence, an equilibrium stopping rule τ can be defined as one that can not be strictly improved by taking $\Theta\tau$ instead. Following Bayraktar et al. (2021), we name it as a *mild intra-personal equilibrium* stopping rule:

Definition 5 A stopping rule τ is a mild intra-personal equilibrium if $\Theta\tau = \tau$.

So a mild intra-personal equilibrium is a fix-point of the operator Θ. If τ is to stop the process at any time and with any state, then it is straightforward to see that $\mathscr{L}^*\tau = \tau$. Consequently, by definition $\Theta\tau = \tau$ and thus τ is a mild intra-personal equilibrium. In other words, following Definition 5, immediate stop is *automatically* a (trivial) mild intra-personal equilibrium.

For a general stopping rule τ, consider any time t and state x in the interior of the stopping region of τ, where the stoping region refers to the set of time-state pairs at which the stopping rule τ would stop the process. Then, it is also easy to see that $\mathscr{L}^*\tau = \tau$ at time t and state x, so one should immediately stop under $\Theta\tau$ as well. As a result, the stopping region of $\Theta\tau$ is at least as large as that of τ, if we ignore the time-state pairs that are on the boundary of the stopping region of τ. Therefore, we expect the iterative sequence $\Theta^n\tau$ to converge as $n\to\infty$, and the convergent point τ^* satisfies $\tau^* = \Theta\tau^*$ and thus is a mild intra-personal equilibrium. It is, however mathematically challenging to formalize the above heuristic derivation. Rigorous proofs have been established in various settings by Huang and Nguyen-Huu (2018), Huang et al. (2020), and Huang and Yu (2021). The above iterative algorithm, which generates a sequence $\Theta^n\tau$, $n = 0,1,\ldots$, not only yields a mild intra-personal equilibrium as the limit of the sequence, but also has a clear economic interpretation: each application of Θ corresponds to an additional level of strategic reasoning; see Huang and Nguyen-Huu (2018) and Huang et al. (2020) for elaborations.

As discussed in the above, immediate stop is always a mild equilibrium; so it is expected that there exist multiple mild intra-personal equilibrium stopping rules; see Huang and Nguyen-Huu (2018) and Huang et al. (2020). To address the issue of multiplicity, Huang and Zhou (2020) and Huang and Wang (2020) consider, in the setting of an infinite-horizon, continuous-time optimal stopping under nonexponential discounting, the "optimal" mild intra-personal equilibrium stopping rule τ^* which achieves the maximum of $J(t,x;\tau)$ over $\tau \in \mathscr{E}$ for all $t \in [0,T)$, $x \in \mathbb{X}$, where $\mathscr{E}$ is the set of all mild intra-personal equilibrium stopping rules.

Bayraktar et al. (2021) compare mild intra-personal equilibrium stopping rules with weak (respectively strong) intra-personal equilibrium stopping rules obtained by embedding optimal stopping into stochastic control and then applying Definition 2 (respectively Definition 3). Assuming the objective function to be a multiplication of a discount function and a Markov process taking values in a finite or countably infinite state space, the authors prove that the optimal mild intra-personal equilibrium is a strong intra-personal equilibrium.

6 Discretization Approach

In the discrete-time setting, an intra-personal equilibrium strategy of a sophisticated agent can be easily defined and derived in a backward manner starting from the last period. Thus, for a continuous-time problem, it is natural to discretize and then pass to the limit. Specifically, one partitions the continuous-time period $[0,T]$ into a finite number of subperiods, assumes the agent is able to commit in each subperiod but not beyond it, and computes the strategy chosen by the agent. Sending the length of the longest subperiod in the partition to zero, the limit of the above strategy, if it exists, can be regarded as the strategy of a sophisticated agent for the continuous-time problem. This ideas was first employed by Pollak (1968) to study the consumption problem of Strotz (1955-1956) and has recently been revisited and extensively studied by a series of papers; see for instance Yong (2012), Wei et al. (2017), Mei and Yong (2019), and Wang and Yong (2019).

Specifically, consider the control problem in Section 2 and assume that in the objective function in (3), C and F do not depend on x and $G \equiv 0$. For a partition Π of $[0,T]$: $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$, we denote $\|\Pi\| := \max_{k=1,\ldots,N} |t_k - t_{k-1}|$. A control strategy $\hat{\mathbf{u}}^{\Pi}$ is an intra-personal equilibrium with respect to the partition Π if

$$J(t_k, x_k; \hat{\mathbf{u}}^{\Pi}) \geq J(t_k, x_k; \mathbf{u}^{\Pi}_{k,\mathbf{a}}) \tag{27}$$

for any $k = 0, 1, \ldots, N-1$, reachable state x_k at time k under $\hat{\mathbf{u}}^{\Pi}$, and strategy $\mathbf{a}$, where $\mathbf{u}^{\Pi}_{k,\mathbf{a}}(s,\cdot) := \mathbf{a}(s,\cdot)$ for $s \in [t_k, t_{k+1})$ and $\mathbf{u}^{\Pi}_{k,\mathbf{a}}(s,\cdot) = \hat{\mathbf{u}}^{\Pi}(s,\cdot)$ for $s \in [t_{k+1}, T)$. In other words, $\hat{\mathbf{u}}(s,\cdot), s \in [t_k, t_{k+1})$, is optimal for an agent who can commit in the period $[t_k, t_{k+1})$ and anticipates that her future selves will take strategy $\hat{\mathbf{u}}$ beyond time t_{k+1}. In the aforementioned literature, the authors define a strategy $\hat{\mathbf{u}}$ to be a *limiting intra-personal equilibrium* if there exists a sequence of partition $(\Pi_m)_{m\in\mathbb{N}}$ with $\lim_{m\to\infty} \|\Pi_m\| = 0$ such that the state process, control process, and continuation value process under certain intra-personal equilibrium with respect to Π_m converge to those under $\hat{\mathbf{u}}$, respectively, as $m \to \infty$. Assuming that the diffusion coefficient of the controlled state process is independent of control and non-degenerate and that some other conditions hold, Wei et al. (2017) prove the above convergence for any sequence of partitions with mesh size going to zero, and the limit of the continuation value function satisfies a flow of PDEs. Moreover, this flow of PDEs admits a unique solution, so the limiting intra-personal equilibrium uniquely exists. Furthermore, the limiting equilibrium is also an equilibrium under Definition 2.

Whether the equilibrium with respect to Π converges when $\|\Pi\| \to 0$ for a general time-inconsistent problem, however, is still unknown. Moreover, the definition of this equilibrium relies on the assumptions that C and F do not depend on x and $G \equiv 0$. Otherwise, for a given partition Π, the optimal strategy the agent at time t_k implements in the subperiod $[t_k, t_{k+1})$ is *semi-Markovian*: the agent's action at time $s \in [t_k, t_{k+1})$ is a function of s, the state at s, and the state at t_k. As a result, the intra-personal equilibrium with respect to Π is non-Markovian; so we cannot restrict limiting equilibria to be Markov strategies.

7 Applications

7.1 Present-bias Preferences

Present-biased preferences, also known as hyperbolic discounting, refer to the following observation in intertemporal choice: when considering time preferences between two moments, individuals become more impatient when the two moments are closer to the present time. Thaler (1981) provides an illustrative example of present-biased preferences: some people may prefer an apple today to two apples tomorrow, but very few people would prefer an apple in a year to two apples in a year plus one day. Noted as early as in Strotz (1955-1956), present-biased preferences lead to time inconsistency. For example, consider an agent whose time preferences for having apples are as described in the above illustrative example by Thaler (1981). At time 0, faced with Option A of having one apple at time $t = 365$ (days) and Option B of having two apples at time $s = 366$ (days), the agent chooses Option B. When time $t = 365$ arrives, however, if the agent gets to choose again, she would choose Option A. This shows that the agent in the future will change her actions planned today; hence time-inconsistency is present. For a review of the literature on present-biased preferences, see Frederick et al. (2002).

In a time-separable discounted utility model, present-biased preferences can be modeled by a non-exponential discount function. For example, consider an intertemporal consumption model in continuous time for an agent. The agent's preference value of a random consumption stream $(C_s)_{s\in[t,T]}$ can be represented as

$$\mathbb{E}_t\left[\int_t^T h(s-t)u(C_s)ds\right], \tag{28}$$

where u is the agent's utility function, h is the agent's discount function, and $\mathbb{E}_t$ denotes the expectation conditional on all the information available at time t. To model present-biased preferences, we assume $h(s+\Delta)/h(s)$ to be *strictly* increasing in $s \geq 0$ for any fixed $\Delta > 0$; hence it excludes the standard exponential discount function. An example is the generalized hyperbolic discount function proposed by Loewenstein and Prelec (1992): $h(s) = (1+\alpha s)^{-\beta/\alpha}, s \geq 0$, where $\alpha > 0$ and $\beta > 0$ are two parameters. Ebert et al. (2020) introduce a class of weighted discount functions that is broad enough to include most commonly used non-exponential discount functions in finance and economics.

In various continuous-time settings, Barro (1999), Ekeland and Lazrak (2006), Ekeland and Lazrak (2008), Ekeland and Lazrak (2010), Ekeland and Pirvu (2008), Marín-Solano and Navas (2010), and Ekeland et al. (2012) study intra-personal equilibria for portfolio selection and consumption problems with present-biased preferences. Ebert et al. (2020) and Tan et al. (2018) study real option problems for agents with general weighted discount functions and derive equilibrium investment strategies. Harris and Laibson (2013) and Grenadier and Wang (2007) apply a stochastic, piece-wise step discount function to a consumption problem and a real option problem, respectively, and derive intra-personal equilibrium strategies. Asset pricing for sophisticated agents with present-biased preferences and without commitment has been studied by Luttmer and Mariotti (2003) and Björk et al. (2017).

7.2 Mean-Variance

A popular decision criterion in finance is mean–variance, with which an agent minimizes the variance and maximizes the mean of certain random quantity, e.g., the wealth of a portfolio at the end of a period. Any mean–variance model is inherently time inconsistent due to the variance part. To see this, consider a two-period decision problem with dates 0, 1, and 2 for an agent. The agent is offered various options at time 1 that will yield certain payoffs at time 2. The set of options offered to the agent at time 1 depends on the outcome of a fair coin that is tossed between time 0 and 1. If the toss yields a head, the agent is offered two options at time 1: Option H1 that yields \$0 and \$200 with equal probabilities and Option H2 that yields \$50 and \$150 with equal probabilities. If the toss yields a tail, the agent is offered another two options at time 1: Option T1 that yields \$0 and \$200 with equal probabilities and Option T2 that yields \$1050 and \$1150 with equal probabilities. Suppose that at both time 0 and 1, the agent's decision criterion is to minimize the variance of the terminal payoff at time 2. At time 0, the agent has not yet observed the outcome of the toss; so she will need to make choices contingent on this outcome, i.e., she chooses between the following four plans: (H1,T1), (H1,T2), (H2,T1), and (H2,T2), where the first and second components of each of the above four plans stand for the agent's planned choice when the toss yields a head and a tail, respectively. Straightforward calculation shows that the plan (H2,T1) yields the smallest variance of the terminal payoff; so at time 0 the agent plans to choose H2 when the toss yields a head and choose T1 when the toss yields a tail. At time 1, after having observed the outcome of the toss, if the agent can choose again with the objective of minimizing the variance of the terminal payoff, she would choose H2 if the outcome is a head and T2 is the outcome is a tail. Consequently, what the agent plans at time 0 is different from what is optimal for the agent at time 1, resulting in time inconsistency.

The reason of having time inconsistency above can be seen from the following conditional variance formula: $\mathrm{var}(X) = \mathbb{E}[\mathrm{var}(X|Y)] + \mathrm{var}(\mathbb{E}[X|Y])$, where X stands for the terminal payoff and Y denotes the outcome of the coin toss. At time 0, the agent's objective is to maximize $\mathrm{var}(X)$ and at time 1, her objective is to maximize $\mathrm{var}(X|Y)$. Although the plan (H2,T2) yields small variance of X given the outcome of the toss Y and thus a small value of the average conditional variance $\mathbb{E}[\mathrm{var}(X|Y)]$, it yields very different expected payoffs conditional on having a head and on having a tail, leading to a large value of $\mathrm{var}(\mathbb{E}[X|Y])$. Consequently, $\mathrm{var}(X)$ under plan (H2,T2) is larger than under plan (H2,T1), which yields a larger value of $\mathbb{E}[\mathrm{var}(X|Y)]$ than the former but a much smaller value of $\mathrm{var}(\mathbb{E}[X|Y])$. Consequently, (H2,T1) is preferred to (H2,T2) for the agent at time 0.

A lot of recent works study intra-personal equilibrium investment strategies for agents with mean-variance preferences. For continuous-time models, see for instance Basak and Chabakauri (2010), Björk et al. (2014), Pun (2018), Bensoussan et al. (2014), Cui et al. (2016), Sun et al. (2016), Landriault et al. (2018), Bensoussan et al. (2019), Kryger et al. (2020), and Han et al. (2021). In all these works, the mean-variance criterion is formulated as a weighted average of the mean and variance of wealth at a terminal time, i.e., at each time t, the agent's objective is to maximize $\mathbb{E}_t[X] - \frac{\gamma_t}{2}\mathrm{var}_t(X)$, where $\mathbb{E}_t$ and var_t stand for the conditional mean and variance of the terminal wealth X, respectively, and γ_t is a risk aversion parameter. Alternatively, He and Jiang (2020b) and He and Jiang (2020a) study intra-personal equilibria for mean-variance investors in a constrained formulation: at each time, an investor minimizes the variance of terminal wealth with a target constraint of the expected terminal wealth. Dai et al. (2021) consider a mean-variance model for log returns. Hu et al. (2012), Hu et al. (2017), Czichowsky (2013), and Yan and Wong (2020) investigate open-loop intra-personal equilibria for mean-variance portfolio selection problems. For equilibrium mean-variance insurance strategies, see for instance Zeng and Li (2011), Li et al. (2012), Zeng et al. (2013), Liang and Song (2015), and Bi and Cai (2019).

7.3 Non-EUT Preferences

There is abundant empirical and experimental evidence showing that when making choices under uncertainty, individuals do not maximize expected utility (EU); see for instance a survey by Starmer (2000). Various alternatives to the EU model, which are generally referred to as *non-EU* models, have been proposed in the literature. Some of these models employ probability weighting functions to describe the tendency of overweighing extreme outcomes that occur with small probabilities, examples being prospect theory (PT) (Kahneman and Tversky, 1979, Tversky and Kahneman, 1992) and rank-dependent utility (RDU) theory (Quiggin, 1982).

It has been noted that when applied to dynamic choice problems, non-EU models can lead to time inconsistency; see Machina (1989) for a review of early works discussing this issue. For illustration, consider a casino gambling problem studied by Barberis (2012): a gambler is offered 10 independent bets with equal probabilities of winning and losing \$1, plays these bets sequentially, and decides when to stop playing. Suppose at each time, the gambler's objective is to maximize the preference value of the payoff at end of the game and the preferences are represented by a non-EU model involving a probability weighting function. We represent the cumulative payoff of playing the bets by a binomial tree with up and down movements standing for winning and losing, respectively. At time 0, the top most state (TMS) of the tree at $t = 10$ represents the largest possible payoff achievable and the probability of reaching this state is extremely small (2^{-10}). The gambler overweighs this state due to probability weighting and aspires to reach it. Hence, at time 0, her plan is to play the 10-th bet if and when she has won all the previous 9 bets. Now, suppose she has played and indeed won the first 9 bets. If she has a chance to re-consider her decision of whether to play the 10-th bet at *that* time, she may find it no longer favorable to play because the probability of reaching the TMS at time 10 is 1/2 and thus this state is not overweighed. Consequently, when deciding whether to play the 10-th bet conditioning on she has won the first 9 bets, the gambler may choose differently when she is at time 0 and when she is at time 9, showing time inconsistency.

In a continuous-time, complete market, Hu et al. (2021) study a portfolio selection problem in which an agent maximizes the following RDU of her wealth X at a terminal time:

$$\int_{\mathbb{R}} u(x)w(1-F_X(x)), \tag{29}$$

where u is a utility function, w is a probability weighting function, and F_X is the cumulative distribution function of X. The authors derive an open-loop intra-personal equilibrium and show that it is in the same form as in the classical Merton model but with a properly scaled market price of risk. He et al. (2020) consider median and quantile maximization for portfolio selection, where the objective function, namely the quantile of the terminal wealth, can be regarded as a special case of RDU with a particular probability weighting function w. The authors study closed-loop intra-personal equilibrium and find that an affine trading strategy is an equilibrium if and only if it is a portfolio insurance strategy. Ebert and Strack (2017) consider the optimal time to stop a diffusion process with the objective to maximize the value of the process at the stopping time under a PT model. Using the notion of mild intra-personal equilibrium as previously discussed in Section 5, the authors show that under reasonable assumptions on the probability weighting functions, the only equilibrium among all two-threshold stopping rules is to immediately stop. Huang et al. (2020) study mild intra-personal equilibrium stopping rules for an agent who wants to stop a geometric Brownian motion with the objective of maximizing the RDU value at the stopping time.

Risk measures, such as value-at-risk (VaR) and conditional value-at-risk (VaR), can also be considered to be non-EU models leading to time consistency. There are, however, few studies on intra-personal equilibria for mean-risk models in continuous time. For relevant studies in discrete-time settings, see for instance Cui et al. (2019).

Models with Knightian uncertainty or ambiguity can also result in time inconsistency. For example, the α-maxmin model proposed by Ghirardato et al. (2004) is dynamically inconsistent in general; see for instance Beissner et al. (2020). Li et al. (2019) find an open-loop intra-personal equilibrium investment strategy for an agent with α-maximin preferences. Huang and Yu (2021) consider a problem of stopping a one-dimensional diffusion process with preferences represented by the α-maxmin model and study the mild intra-personal equilibrium stopping rule for the problem.

8 Dynamically Consistent Preferences

Machina (1989) notes that, in many discussions of time inconsistency in the literature, a hidden assumption is *consequentialism*: at any intermediate time t of a dynamic decision process, the agent employs the *same* preference model as used at the initial time to evaluate the choices in the *continuation* of the dynamic decision process from time t, conditional on the circumstances at time t. For example, consider a dynamic consumption problem for an agent with present-bias preferences and suppose that at the initial time 0, the agent's preference value for a consumption stream $(C_s)_{s\geq 0}$ is represented by $\mathbb{E}[\int_0^\infty h(s)u(C_s)ds]$, where the discount function h models the agent's time preferences at the initial time 0 and u is the agent's utility function. The consequentialism assumption implies that at any intermediate time t, the agent's preferences for the continuation of the consumption stream, i.e., $(C_s)_{s\geq t}$, are represented by the same preference model as at the initial time 0, conditional on the situations at time t, i.e., by $\mathbb{E}_t[\int_t^\infty h(s-t)u(C_s)ds]$, where the discount function h and u are the same as the ones in the preference model at the initial time 0. Similarly, for a dynamic choice problem with RDU preferences for the payoff at a terminal time, the consequentialism assumption stipulates that the agent uses the same utility function u and probability weighting function w at all intermediate times t when evaluating the terminal payoff at those times.

The consequentialism assumption, however, has not been broadly validated because there are few experimental or empirical studies on how individuals dynamically update their preferences. Machina (1989) consider a class of non-EU maximizers, referred to as γ-people, who adjust their preferences dynamically over time so as to remain time consistent. The idea in Machina (1989) was further developed by Karnam et al. (2017) who propose the notion of *time-consistent dynamic preference models*. The idea of considering time-consistent dynamic preferences is also central in the theory of forward performance criteria proposed and developed by Musiela and Zariphopoulou (2006, 2008, 2009, 2010a,b, 2011); see also He et al. (2021) for a related discussion.

Formally, consider a dynamic choice problem in a period $[0,T)$. A preference model at time 0 is specified for an agent, denoted as $J_0(u(\cdot))$, where $(u(s))_{s\in[0,T)}$ denotes the agent's dynamic choice. A family of dynamic preference models J_t, $t \in (0,T)$, are called time-consistent for the initial model J_0 if the optimal strategy under J_0, namely, the pre-committed strategy for the agent at time 0, is also optimal under J_t for the agent at any future time $t \in (0,T)$. Note that given the pre-committed strategy at time 0, we can always find preference models at $t > 0$ such that this strategy remains optimal. Thus, a more interesting question is whether we can find a family of time-consistent dynamic preference models that are of the same type as the initial preference model.

He et al. (2021) study portfolio selection in the Black-Scholes market for an agent whose initial preference model for wealth at a terminal time is represented by RDU. The authors show that there exists a family of time-consistent dynamic RDU models if and only if (i) the probability weighting function in the initial model belongs to a parametric class of functions proposed by Wang (1996); and (ii) the parameter of the probability weighting function, the absolute risk aversion index of the utility function, and the market price of risk must be coordinated with each other over time in a specific way. Cui et al. (2012), Karnam et al. (2017), and He and Jiang (2020a) find that mean-variance models become time consistent if the dynamic trade-off between the mean and variance over time is set properly. For mean-CVaR models, where an agent maximizes the mean and minimize the CVaR at certain confidence level, Pflug and Pichler (2016) and Strub et al. (2019) note, in discrete-time settings, that time consistency is retained as long as the tradeoff between the mean and CVaR and the confidence level evolve dynamically in a certain way.

The problem of intra-personal equilibria and that of dynamically consistent preferences can be considered primal–dual to each other: the former finds equilibrium strategies given the time-inconsistent preferences, whereas the latter identifies preferences given the problem is time-consistent. Diving deeper into this relationship may call for innovative mathematical analysis and result in profound economic insights.

Acknowledgements Xue Dong He gratefully acknowledges financial support through the General Research Fund of the Research Grants Council of Hong Kong SAR (Project No. 14200917). Xun Yu Zhou gratefully acknowledges financial support through a start-up grant at Columbia University and through the Nie Center for Intelligent Asset Management. This author also thanks the hospitality of The Chinese University of Hong Kong during the summer of 2020 when the present project started.

References

Barberis, N. (2012). A model of casino gambling, *Management Science* **58**(1): 35–51.

Barro, R. J. (1999). Ramsey meets Laibson in the neoclassical growth model, *Quarterly Journal of Economics* **114**(4): 1125–1152.

Basak, S. and Chabakauri, G. (2010). Dynamic mean-variance asset allocation, *Review of Financial Studies* **23**(8): 2970–3016.

Bayraktar, E., Zhang, J. and Zhou, Z. (2021). Equilibrium concepts for time-inconsistent stopping problems in continuous time, *Mathematical Finance* **31**: 508–530.

Beissner, P., Lin, Q. and Riedel, F. (2020). Dynamically consistent alpha-maxmin expected utility, *Mathematical Finance* **30**(3): 1073–1102.

Bensoussan, A., Wong, K. C. and Yam, S. C. P. (2019). A paradox in time-consistency in the mean–variance problem, *Finance and Stochastics* **23**(1): 173–207.

Bensoussan, A., Wong, K. C., Yam, S. C. P. and Yung, S.-P. (2014). Time-consistent portfolio selection under short-selling prohibition: From discrete to continuous setting, *SIAM Journal on Financial Mathematics* **5**(1): 153–190.

Bertsekas, D. P. (2017). *Dynamic programming and optimal control*, 4 edn, Athena Scientific, Nashua.

Bi, J. and Cai, J. (2019). Optimal investment-reinsurance strategies with state dependent risk aversion and VaR constraints in correlated markets, *Insurance: Mathematics and Economics* **85**: 1–14.

Björk, T., Khapko, M. and Murgoci, A. (2017). On time-inconsistent stochastic control in continuous time, *Finance and Stochastics* **21**(2): 331–360.

Björk, T. and Murgoci, A. (2010). A general theory of Markovian time inconsistent stochastic control problems. SSRN:1694759.

Björk, T. and Murgoci, A. (2014). A theory of Markovian time-inconsistent stochastic control in discrete time, *Finance and Stochastics* **18**(3): 545–592.

Björk, T., Murgoci, A. and Zhou, X. Y. (2014). Mean-variance portfolio optimization with state dependent risk aversion, *Mathematical Finance* **24**(1): 1–24.

Bryan, G., Karlan, D. and Nelson, S. (2010). Commitment devices, *Annual Review of Economics* **2**(1): 671–698.

Cao, D. and Werning, I. (2016). Dynamic savings choices with disagreements, *Technical report*, National Bureau of Economic Research.

Christensen, S. and Lindensjö, K. (2018). On finding equilibrium stopping times for time-inconsistent markovian problems, *SIAM Journal on Control and Optimization* **56**(6): 4228–4255.

Christensen, S. and Lindensjö, K. (2020). On time-inconsistent stopping problems and mixed strategy stopping times, *Stochastic Processes and their Applications* **130**(5): 2886–2917.

Cui, X., Gao, J., Shi, Y. and Zhu, S. (2019). Time-consistent and self-coordination strategies for multi-period mean-conditional value-at-risk portfolio selection, *European Journal of Operational Research* **276**(2): 781–789.

Cui, X., Li, D., Wang, S. and Zhu, S. (2012). Better than dynamic mean-variance: Time inconsistency and free cash flow stream, *Mathematical Finance* **22**(2): 346–378.

Cui, X., Xu, L. and Zeng, Y. (2016). Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion, *Optimization Letters* **10**(8): 1681–1691.

Czichowsky, C. (2013). Time-consistent mean-variance portfolio selection in discrete and continuous time, *Finance and Stochastics* **17**(2): 227–271.

Dai, M., Jin, H., Kou, S. and Xu, Y. (2021). A dynamic mean-variance analysis for log returns, *Management Science* **67**: 661–1328.

Ebert, S. and Strack, P. (2017). Never, ever getting started: On prospect theory without commitment. SSRN:2765550.

Ebert, S., Wei, W. and Zhou, X. Y. (2020). Weighted discounting—on group diversity, time-inconsistency, and consequences for investment, *Journal of Economic Theory* **189**: 1–40.

Ekeland, I. and Lazrak, A. (2006). Being serious about non-commitment: subgame perfect equilibrium in continuous time. arXiv:math/0604264.

Ekeland, I. and Lazrak, A. (2008). Equilibrium policies when preferences are time inconsistent. arXiv:0808.3790.

Ekeland, I. and Lazrak, A. (2010). The golden rule when preferences are time inconsistent, *Mathematics and Financial Economics* **4**(1): 29–55.

Ekeland, I., Mbodji, O. and Pirvu, T. A. (2012). Time-consistent portfolio management, *SIAM Journal on Financial Mathematics* **3**(1): 1–32.

Ekeland, I. and Pirvu, T. (2008). Investment and consumption without commitment, *Mathematics and Financial Economics* **2**(1): 57–86.

Frederick, S., Loewenstein, G. and O'donoghue, T. (2002). Time discounting and time preference: A critical review, *Journal of Economic Literature* **40**(2): 351–401.

Ghirardato, P., Maccheroni, F. and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory* **118**(2): 133–173.

Grenadier, S. R. and Wang, N. (2007). Investment under uncertainty and time-inconsistent preferences, *Journal of Financial Economics* **84**(1): 2–39.

Han, B., Pun, C. S. and Wong, H. Y. (2021). Robust mean-variance portfolio selection with state-dependent ambiguity aversion and risk aversion: A closed-loop approach, *Finance and Stochastics* **forthcoming**.

Han, B. and Wong, H. Y. (2020). Time-inconsistency with rough volatility. arXiv:1907.11378.

Harris, C. and Laibson, D. (2013). Instantaneous gratification, *Quarterly Journal of Economics* **128**(1).

He, X. D. and Jiang, Z. (2019). On the equilibrium strategies for time-inconsistent problems in continuous time. SSRN:3308274.

He, X. D. and Jiang, Z. (2020a). Dynamic mean-variance efficient fractional kelly portfolios in a stochastic volatility model. SSRN:3670621.

He, X. D. and Jiang, Z. (2020b). Mean-variance portfolio selection with dynamic targets for expected terminal wealth, *Mathematics of Operations Research* **forthcoming**. SSRN:3084657.

He, X. D., Jiang, Z. and Kou, S. (2020). Portfolio selection under median and quantile maximization. SSRN:3657661.

He, X. D., Strub, M. and Zariphopoulou, T. (2021). Forward rank-dependent performance criteria: Time-consistent investment under probability distortion, *Mathematical Finance* **31**: 683–721.

Hernández, C. and Possamaï, D. (2020). Me, myself and I: a general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. arXiv:2002.12572.

Hu, Y., Jin, H. and Zhou, X. Y. (2012). Time-inconsistent stochastic linear-quadratic control, *SIAM Journal on Control and Optimization* **50**: 1548–1572.

Hu, Y., Jin, H. and Zhou, X. Y. (2017). Time-inconsistent stochastic linear-quadratic control: Characterization and uniqueness of equilibrium, *SIAM Journal on Control and Optimization* **55**(2): 1261–1279.

Hu, Y., Jin, H. and Zhou, X. Y. (2021). Consistent investment of sophisticated rank-dependent utility agents in continuous time, *Mathematical Finance* **forthcoming**. arXiv:2006.01979.

Huang, Y.-J. and Nguyen-Huu, A. (2018). Time-consistent stopping under decreasing impatience, *Finance and Stochastics* **22**(1): 69–95.

Huang, Y.-J., Nguyen-Huu, A. and Zhou, X. Y. (2020). General stopping behaviors of naïve and non-committed sophisticated agents, with application to probability distortion, *Mathematical Finance* **30**: 310–340.

Huang, Y.-J. and Wang, Z. (2020). Optimal equilibria for multi-dimensional time-inconsistent stopping problems. arXiv:2006.00754.

Huang, Y.-J. and Yu, X. (2021). Optimal stopping under model ambiguity: a time-consistent equilibrium approach. arXiv:1906.01232.

Huang, Y.-J. and Zhou, Z. (2019). Strong and weak equilibria for time-inconsistent stochastic control in continuous time, *Mathematics of Operations Research* **forthcoming**. arXiv:1809.09243.

Huang, Y.-J. and Zhou, Z. (2020). Optimal equilibria for time-inconsistent stopping problems in continuous time, *Mathematical Finance* **30**(3): 1103–1134.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk, *Econometrica* **47**(2): 263–291.

Karnam, C., Ma, J., Zhang, J. et al. (2017). Dynamic approaches for some time-inconsistent optimization problems, *Annals of Applied Probability* **27**(6): 3435–3477.

Kryger, E., Nordfang, M.-B. and Steffensen, M. (2020). Optimal control of an objective functional with non-linearity between the conditional expectations: solutions to a class of time-inconsistent portfolio problems, *Mathematical Methods of Operations Research* **91**(3): 405–438.

Landriault, D., Li, B., Li, D. and Young, V. R. (2018). Equilibrium strategies for the mean-variance investment problem over a random horizon, *SIAM Journal on Financial Mathematics* **9**(3): 1046–1073.

Li, B., Luo, P. and Xiong, D. (2019). Equilibrium strategies for alpha-maxmin expected utility maximization, *SIAM Journal on Financial Mathematics* **10**(2): 394–429.

Li, Z., Zeng, Y. and Lai, Y. (2012). Optimal time-consistent investment and reinsurance strategies for insurers under hestons sv model, *Insurance: Mathematics and Economics* **51**(1): 191–203.

Liang, Z. and Song, M. (2015). Time-consistent reinsurance and investment strategies for mean–variance insurer under partial information, *Insurance: Mathematics and Economics* **65**: 66–76.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation, *Quarterly Journal of Economics* **107**(2): 573–597.

Luttmer, E. G. and Mariotti, T. (2003). Subjective discounting in an exchange economy, *Journal of Political Economy* **111**(5): 959–989.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty, *Journal of Economic Literature* **27**(4): 1622–1668.

Marín-Solano, J. and Navas, J. (2010). Consumption and portfolio rules for time-inconsistent investors, *European Journal of Operational Research* **201**(3): 860–872.

Mei, H. and Yong, J. (2019). Equilibrium strategies for time-inconsistent stochastic switching systems, *ESAIM: Control, Optimisation and Calculus of Variations* **25**: 64.

Musiela, M. and Zariphopoulou, T. (2006). Investments and forward utilities. https://www.ma.utexas.edu/users/zariphop/pdfs/TZ-TechnicalReport-4.pdf.

Musiela, M. and Zariphopoulou, T. (2008). Optimal asset allocation under forward exponential performance criteria, *in* S. N. Ethier, J. Feng and R. H. Stockbridge (eds), *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, Vol. 4, Institute of Mathematical Statistics, pp. 285–300.

Musiela, M. and Zariphopoulou, T. (2009). Portfolio choice under dynamic investment performance criteria, *Quantitative Finance* **9**(2): 161–170.

Musiela, M. and Zariphopoulou, T. (2010a). Portfolio choice under space-time monotone performance criteria, *SIAM Journal on Financial Mathematics* **1**(1): 326–365.

Musiela, M. and Zariphopoulou, T. (2010b). Stochastic partial differential equation and portfolio choice, *in* C. Chiarella and A. Novikov (eds), *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen*, Springer, pp. 195–216.

Musiela, M. and Zariphopoulou, T. (2011). Initial investment choice and optimal future allocations under time-monotone performance criteria, *International Journal of Theoretical and Applied Finance* **14**(1): 61–81.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later, *American Economic Review* **89**(1): 103–124.

Peleg, B. and Yaari, M. E. (1973). On the existence of a consistent course of action when tastes are changing, *Review of Economic Studies* **40**(3): 391–401.

Pflug, G. C. and Pichler, A. (2016). Time-inconsistent multistage stochastic programs: Martingale bounds, *European Journal of Operational Research* **249**(1): 155–163.

Pollak, R. A. (1968). Consistent planning, *Review of Economic Studies* **35**(2): 201–208.

Pun, C. S. (2018). Time-consistent mean-variance portfolio selection with only risky assets, *Economic Modelling* **75**: 281–292.

Quiggin, J. (1982). A theory of anticipated utility, *Journal of Economic Behavior and Organization* **3**: 323–343.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, *Journal of Economic Literature* **38**(2): 332–382.

Strotz, R. H. (1955-1956). Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* **23**(3): 165–180.

Strub, M., Li, D., Cui, X. and Gao, J. (2019). Discrete-time mean-cvar portfolio selection and time-consistency induced term structure of the cvar, *Journal of Economic Dynamics and Control* **108**(103751).

Sun, J., Li, Z. and Zeng, Y. (2016). Precommitment and equilibrium investment strategies for defined contribution pension plans under a jump-diffusion model, *Insurance: Mathematics and Economics* **67**: 158–172.

Tan, K. S., Wei, W. and Zhou, X. Y. (2018). Failure of smooth pasting principle and nonexistence of equilibrium stopping rules under time-inconsistency. arXiv:1807.01785.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency, *Economics Letters* **8**(3): 201–207.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* **5**(4): 297–323.

Wang, H. and Yong, J. (2019). Time-inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations. arXiv:1911.04995.

Wang, S. (1996). Premium calculation by transforming the layer premium density, *ASTIN Bulletin* **26**(1): 71–92.

Wei, Q., Yong, J. and Yu, Z. (2017). Time-inconsistent recursive stochastic optimal control problems, *SIAM Journal on Control and Optimization* **55**(6): 4156–4201.

Yan, T. and Wong, H. Y. (2020). Open-loop equilibrium reinsurance-investment strategy under mean–variance criterion with stochastic volatility, *Insurance: Mathematics and Economics* **90**: 105–119.

Yong, J. (2012). Time-inconsistent optimal control problems and the equilibrium HJB equation, *Mathematical Control and Related Fields* **2**(3): 271–329.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer.

Zeng, Y. and Li, Z. (2011). Optimal time-consistent investment and reinsurance policies for mean-variance insurers, *Insurance: Mathematics and Economics* **49**(1): 145–154.

Zeng, Y., Li, Z. and Lai, Y. (2013). Time-consistent investment and reinsurance strategies for mean–variance insurers with jumps, *Insurance: Mathematics and Economics* **52**(3): 498–507.

N-Player and Mean-Field Games in Itô-Diffusion Markets with Competitive or Homophilous Interaction

Ruimeng Hu and Thaleia Zariphopoulou

Abstract In Itô-diffusion environments, we introduce and analyze N-player and common-noise mean-field games in the context of optimal portfolio choice in a common market. The players invest in a finite horizon and also interact, driven either by competition or homophily. We study an incomplete market model in which the players have constant individual risk tolerance coefficients (CARA utilities). We also consider the general case of random individual risk tolerances and analyze the related games in a complete market setting. This randomness makes the problem substantially more complex as it leads to (N or a continuum of) auxiliary "individual" Itô-diffusion markets. For all cases, we derive explicit or closed-form solutions for the equilibrium stochastic processes, the optimal state processes, and the values of the games.

1 Introduction

In Itô-diffusion environments, we introduce N-player and common-noise mean-field games (MFGs) in the context of optimal portfolio choice in a common market. We build on the framework and notions of [12] (see, also, [11]) but allow for a more general market model (beyond the log-normal case) and, also, consider more complex risk preferences.

The paper consists of two parts. In the first part, we consider a common incomplete market and players with individual exponential utilities (CARA) who invest while interacting with each other, driven either by competition or homophily. We derive the equilibrium policies, which turn out to be state (wealth)-independent stochastic

Ruimeng Hu
Department of Mathematics and Department of Statistics & Applied Probability, University of California, Santa Barbara, CA, 93106-3080, USA, e-mail: rhu@ucsb.edu and Thaleia Zariphopoulou
Departments of Mathematics and IROM, The University of Texas at Austin, Austin, USA, and the Oxford-Man Institute, University of Oxford, Oxford, UK, e-mail: zariphop@math.utexas.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_9

processes. Their forms depend on the market dynamics, the risk tolerance coefficients, and the underlying minimal martingale measure. We also derive the optimal wealth and the values of both the N-player and the mean-field games, and discuss the competitive and homophilous cases.

In the second part, we assume that the common Itô-diffusion market is complete, but we generalize the model in the direction of risk preferences, allowing the risk tolerance coefficients to be random variables. For such preferences, we first analyze the single-player problem, which is interesting in its own right. Among others, we show that the randomness of the utility "distorts" the original market by inducing a "personalized" risk premium process. This effect is more pronounced in the N-player game where the common market is now replaced by "personalized" markets whose stochastic risk premia depend on the individual risk tolerances. As a result, the tractability coming from the common market assumption is lost. In the MFG setting, these auxiliary individual markets are randomly selected (depending on the type vector) and aggregate to a common market with a modified risk premium process. We characterize the optimal policies, optimal wealth processes, and game values, building on the aforementioned single-player problem.

To our knowledge, N-player games and MFGs in Itô-diffusion market settings have not been considered before except in preprint [6]. Therein, the authors used the same asset specialization framework and same CARA preferences as in [12] but allowed for Itô-diffusion price dynamics. They studied the problem using a forward-backward stochastic differential equation (FBSDE) approach. In our work, we have different model settings regarding both the measurability of the coefficients of the Itô-diffusion price processes and the individual risk tolerance inputs. We also solve the problems using a different approach, based on the analysis of portfolio optimization problems of exponential utilities in semi-martingale markets.

The theory of mean-field games was introduced by Lasry and Lions [13], who developed the fundamental elements of the mathematical theory and, independently, by Huang, Malhamé and Caines who considered a particular class [8]. Since then, the area has grown rapidly both in terms of theory and applications. Listing precise references is beyond the scope of this paper.

Our work contributes to N-player games and MFG in Itô-diffusion settings for models with controlled processes whose dynamics depend linearly on the controls and are state-independent, and, furthermore, the controls appear in both the drift and the diffusion parts. Such models are predominant in asset pricing and in optimal portfolio and consumption choice. In the context of the general MFG theory, the models considered herein are restrictive. On the other hand, their structure allows us to produce explicit/closed-form solutions for Itô-diffusion environments.

The paper is organized as follows. In Section 2, we study the incomplete market case for both the N-player game and the MFG, and for CARA utilities. In Section 3, we focus on the complete market case but allow for random risk tolerance coefficients. In analogy to Section 2, we analyze both the N-player game and the MFG. We conclude in Section 4.

2 Incomplete Itô-diffusion common market and CARA utilities

We consider an incomplete Itô-diffusion market, in which we introduce an N-player and a mean-field game for players who invest in a finite horizon while interacting among them, driven either by competition or homophily. We assume that the players (either at the finite or the continuum setting) have individual constant risk tolerance coefficients. For both the N-player and the MFG, we derive in closed form the optimal policies, optimal controlled processes, and the game values. The analysis uses the underlying minimal martingale measure, related martingales, and their decomposition.

2.1 The N-player game

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting two Brownian motions denoted as $(W_t, W_t^Y)_{t\in[0,T]}$, $T < \infty$, imperfectly correlated with the correlation coefficient $\rho \in (-1,1)$. We denote by $(\mathcal{F}_t)_{t\in[0,T]}$ the natural filtration generated by both W and W^Y, and by $(\mathcal{G}_t)_{t\in[0,T]}$ the one generated only by W^Y. We then let $(\mu_t)_{t\in[0,T]}$ and $(\sigma_t)_{t\in[0,t]}$ be $\mathcal{G}_t$-adapted processes, with $0 < c \le \sigma_t \le C$ and $|\mu_t| \le C$, $t \in [0,T]$, for some (possibly deterministic) constants c and C.

The financial market consists of a riskless bond (taken to be the numeraire and with zero interest rate) and a stock whose price process $(S_t)_{t\in[0,T]}$ satisfies

$$dS_t = \mu_t S_t\,dt + \sigma_t S_t\,dW_t, \quad S_0 = s_0 \in \mathbb{R}^+. \tag{1}$$

In this market, N players, indexed by $i \in \mathcal{I}$, $\mathcal{I} = \{1,2,\dots,N\}$, have a common investment horizon $[0,T]$ and trade between the two accounts. Each player, say player i, uses a self-financing strategy $(\pi_t^i)_{t\in[0,T]}$, representing (discounted by the numeraire) the amount invested in the stock. Then, her wealth $(X_t^i)_{t\in[0,T]}$ satisfies

$$dX_t^i = \pi_t^i\,(\mu_t\,dt + \sigma_t\,dW_t), \qquad X_0^i = x_i \in \mathbb{R}, \tag{2}$$

with π^i being an admissible policy, belonging to

$$\mathcal{A} = \Big\{ \pi : \text{self-financing, } \mathcal{F}\text{-progressively measurable} \\ \text{and } E_{\mathbb{P}}\left[\int_0^T \sigma_s^2 \pi_s^2\,ds\right] < \infty \Big\}. \tag{3}$$

As in [12] (see also [1, 4, 9, 10, 11, 20]), players optimize their expected terminal utility but are, also, concerned with the performance of their peers. For an arbitrary but *fixed* policy $(\pi_1, \dots, \pi_{i-1}, \pi_{i+1}, \dots, \pi_N)$, player i, $i \in \mathcal{I}$, seeks to optimize

$$
\begin{aligned}
&V^i(x_1,\dots,x_i,\dots,x_N)\\
&= \sup_{\pi^i\in\mathcal{A}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_i}\left(X_T^i - c_i C_T\right)\right)\middle| X_0^1 = x_1,\dots,X_0^i = x_i,\dots,X_0^N = x_N\right], \quad (4)
\end{aligned}
$$

where

$$
C_T := \frac{1}{N}\sum_{j=1}^N X_T^j \tag{5}
$$

averages all players' terminal wealth, with X_T^j, $j = 1,\dots,N$, given by (2).

The parameter $\delta_i > 0$ is the individual (absolute) risk tolerance while the constant $c_i \in (-\infty, 1]$ models the individual interaction weight towards the average wealth of all players. If $c_i > 0$, the above criterion models *competition* while when $c_i < 0$ it models *homophilous* interactions (see, for example, [14]). The optimization criterion (4) can be, then, viewed as a stochastic game among the N players, where the notion of optimality is being considered in the context of *a Nash equilibrium*, stated below (see, for example, [2]).

Definition 1 A strategy $(\pi_t^*)_{t\in[0,T]} = (\pi_t^{1,*},\dots,\pi_t^{N,*})_{t\in[0,T]} \in \mathcal{A}^{\otimes N}$ is called a Nash equilibrium if, for each $i \in \mathcal{I}$ and $\pi^i \in \mathcal{A}$,

$$
\begin{aligned}
&E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_i}\left(X_T^{i,*} - c_i C_T^*\right)\right)\middle| X_0^1 = x_1,\dots,X_0^i = x_i,\dots,X_0^N = x_N\right]\\
&\qquad \ge E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_i}\left(X_T^{i} - c_i C_T^{i,*}\right)\right)\middle| X_0^1 = x_1,\dots,X_0^i = x_i,\dots,X_0^N = x_N\right] \quad (6)
\end{aligned}
$$

with

$$
C_T^* := \frac{1}{N}\sum_{j=1}^N X_T^{j,*} \quad \text{and} \quad C_T^{i,*} := \frac{1}{N}\left(\sum_{j=1, j\neq i}^N X_T^{j,*} + X_T^i\right),
$$

where $X_T^{j,*}$, $j \in \mathcal{I}$, solve (2) with $\pi^{j,*}$ being used.

In this incomplete market, we recall the associated *minimal martingale* measure $\mathbb{Q}^{MM}$, defined on $\mathcal{F}_T$, with

$$
\frac{d\mathbb{Q}^{MM}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T \lambda_s^2\, ds - \int_0^T \lambda_s\, dW_s\right), \tag{7}
$$

where $\lambda_t := \frac{\mu_t}{\sigma_t}$, $t \in [0,T]$, is the Sharpe ratio process (see, among others, [5]). By the assumptions on the model coefficients, we have that, for $t \in [0,T]$, $\lambda_t \in \mathcal{G}_t$ and

$$
|\lambda_t| \le K, \tag{8}
$$

for some (possibly deterministic) constant K. We also consider the processes $(\widetilde{W}_t)_{t\in[0,T]}$ and $(\widetilde{W}_t^Y)_{t\in[0,T]}$ with $\widetilde{W}_t = W_t + \int_0^t \lambda_s\, ds$ and $\widetilde{W}_t^Y = W_t^Y + \rho\int_0^t \lambda_s\, ds$, which are standard Brownian motions under $\mathbb{Q}^{MM}$ with $\widetilde{W}_t \in \mathcal{F}_t$ and $\widetilde{W}_t^Y \in \mathcal{G}_t$.

Next, we introduce the $\mathbb{Q}^{MM}$-martingale $(M_t)_{t\in[0,T]}$,

$$M_t := E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_0^T \lambda_s^2\,ds}\middle|\mathcal{G}_t\right]. \tag{9}$$

From (8) and the martingale representation theorem, there exists a $\mathcal{G}_t$-adapted process $\xi \in \mathcal{L}^2(\mathbb{P})$ such that

$$dM_t = \xi_t M_t\, d\widetilde{W}_t^Y = \xi_t M_t \left(\rho\, d\widetilde{W}_t + \sqrt{1-\rho^2}\, dW_t^{\perp}\right), \tag{10}$$

where $W_t^{\perp}$ is a standard Brownian motion independent of W_t appearing in the decomposition $W_t^Y = \rho W_t + \sqrt{1-\rho^2} W_t^{\perp}$.

In the absence of interaction among the players ($c_i \equiv 0$, $i \in \mathcal{I}$), the optimization problem (4) has been analyzed by various authors (see, among others, [17, 18]). We recall its solution which will be frequently used herein.

Lemma 1 (no interaction)

Consider the optimization problem

$$v(x) = \sup_{a\in\mathcal{A}} E_{\mathbb{P}}\left[-e^{-\frac{1}{\delta}x_T}\middle|x_0 = x\right], \tag{11}$$

with $\delta > 0$ *and* $(x_t)_{t\in[0,T]}$ *solving*

$$dx_t = a_t\,(\mu_t\, dt + \sigma_t\, dW_t), \quad x_0 = x \in \mathbb{R}, \quad a \in \mathcal{A}. \tag{12}$$

Then, the optimal policy $(a_t^*)_{t\in[0,T]}$ *and the value function are given by*

$$a_t^* = \delta\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right), \tag{13}$$

and

$$v(x) = -e^{-\frac{1}{\delta}x} M_0^{\frac{1}{1-\rho^2}} = -e^{-\frac{1}{\delta}x}\left(E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_0^T \lambda_s^2\,ds}\right]\right)^{\frac{1}{1-\rho^2}}, \tag{14}$$

with $(\xi_t)_{t\in[0,T]}$ *as in* (10).

Proof We only present the key steps, showing that the process $(u_t)_{t\in[0,T]}$,

$$u_t := -e^{-\frac{1}{\delta}x_t}\left(E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_t^T \lambda_s^2\,ds}\middle|\mathcal{G}_t\right]\right)^{\frac{1}{1-\rho^2}},$$

with $u_0 = v(x)$, $x \in \mathbb{R}$, is a supermartingale for x_t solving (12) for arbitrary $\alpha \in \mathcal{A}$ and becomes a martingale for α^* as in (13). To this end, we write

$$u_t = -e^{-\frac{x_t}{\delta}} M_t^{\frac{1}{1-\rho^2}} e^{N_t} \quad \text{with} \quad N_t = \frac{1}{2}\int_0^t \lambda_u^2\, du,$$

and observe that

$$\begin{aligned} du_t &= -\frac{u_t}{\delta}\,dx_t + \frac{1}{2\delta^2}u_t\,d\langle x\rangle_t + u_t\,dN_t + \frac{1}{1-\rho^2}\frac{u_t}{M_t}\,dM_t \\ &\quad + \frac{1}{2(1-\rho^2)}\frac{\rho^2}{1-\rho^2}\frac{u_t}{M_t^2}\,d\langle M\rangle_t - \frac{1}{\delta(1-\rho^2)}\frac{u_t}{M_t}\,d\langle x, M\rangle_t \\ &= u_t\left(-\frac{1}{\delta}a_t\mu_t + \frac{1}{2}\frac{1}{\delta^2}a_t^2\sigma_t^2 + \frac{1}{2}\lambda_t^2 + \frac{\rho}{1-\rho^2}\xi_t\lambda_t + \frac{\rho^2}{2(1-\rho^2)^2}\xi_t^2 \right.\\ &\quad \left. - \frac{\rho}{\delta(1-\rho^2)}a_t\sigma_t\xi_t\right)dt + u_t\left(-\frac{1}{\delta}a_t\sigma_t\,dW_t + \frac{1}{1-\rho^2}\xi_t\,dW_t^Y\right) \\ &= \frac{1}{2}u_t\left(-\frac{1}{\delta}\sigma_t a_t + \lambda_t + \frac{\rho}{1-\rho^2}\xi_t\right)^2 dt + u_t\left(-\frac{1}{\delta}a_t\sigma_t\,dW_t + \frac{1}{1-\rho^2}\xi_t\,dW_t^Y\right). \end{aligned}$$

Because $u_t < 0$, the drift remains non-positive and vanishes for $t \in [0, T]$ if and only if the policy

$$a_t^* = \delta\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right)$$

is being used. Furthermore, $a^* \in \mathcal{A}$, as it follows from the boundedness assumption on σ, inequality (8) and that $\xi \in \mathcal{L}^2(\mathbb{P})$. The rest of the proof follows easily. □

Next, we present the first main result herein that yields the existence of a (wealth-independent) stochastic Nash equilibrium.

Proposition 1 *For $\delta_i > 0$ and $c_i \in (-\infty, 1]$, introduce the quantities*

$$\varphi_N := \frac{1}{N}\sum_{i=1}^{N}\delta_i \quad \textit{and} \quad \psi_N := \frac{1}{N}\sum_{i=1}^{N}c_i, \tag{15}$$

and

$$\bar{\delta}_i := \delta_i + \frac{\varphi_N}{1-\psi_N}c_i. \tag{16}$$

The following assertions hold:

1. *If $\psi_N < 1$, there exists a wealth-independent Nash equilibrium, $(\pi_t^*)_{t\in[0,T]} = \left(\pi_t^{1,*}, \ldots, \pi_t^{i,*}, \ldots, \pi_t^{N,*}\right)_{t\in[0,T]}$, where $\pi_t^{i,*}$, $i \in \mathcal{I}$, is given by the $\mathcal{G}_t$-adapted process*

$$\pi_t^{i,*} = \bar{\delta}_i\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right), \tag{17}$$

with $(\xi_t)_{t\in[0,T]}$ as in (10). The associated optimal wealth process $\left(X_t^{i,*}\right)_{t\in[0,T]}$ is

$$X_t^{i,*} = x_i + \bar{\delta}_i\int_0^t\left(\lambda_u + \frac{\rho}{1-\rho^2}\xi_u\right)(\lambda_u\,du + dW_u) \tag{18}$$

and the game value for player i, $i \in \mathcal{I}$, is given by

$$
\begin{aligned}
V^i(x_1, x_2, \ldots, x_N) &= -\exp\left(-\frac{1}{\delta_i}(x_i - c_i\bar{x})\right) M_0^{\frac{1}{1-\rho^2}} \\
&= -\exp\left(-\frac{1}{\delta_i}(x_i - c_i\bar{x})\right)\left(E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_0^T \lambda_s^2\, ds}\right]\right)^{\frac{1}{1-\rho^2}},
\end{aligned}
\tag{19}
$$

with $\bar{x} = \frac{1}{N}\sum_{i=1}^N x_i$.

2. *If* $\psi_N = 1$*, then it must be that* $c_i \equiv 1$*, for all* $i \in \mathcal{I}$*, and there is no such wealth-independent Nash equilibrium.*

Proof We first solve the individual optimization problem (4) for player $i \in \mathcal{I}$, taking the (arbitrary) strategies $(\pi^1, \ldots, \pi^{i-1}, \pi^{i+1}, \ldots, \pi^N)$ of all other players as given. This problem can be alternatively written as

$$
v^i(\tilde{x}_i) = \sup_{\widetilde{\pi}^i \in \mathcal{A}} \mathbb{E}_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_i}\tilde{x}_T^i\right)\middle|\, \tilde{x}_0^i = \tilde{x}_i\right], \tag{20}
$$

where $\tilde{x}_t^i := X_t^i - \frac{c_i}{N}\sum_{j=1}^N X_t^j$, $t \in [0,T]$, solves

$$
d\tilde{x}_t^i = \widetilde{\pi}_t^i(\mu_t\, dt + \sigma_t\, dW_t) \quad \text{and} \quad \tilde{x}_0^i = \tilde{x}_i := x_i - c_i\bar{x}.
$$

From Lemma 1, we deduce that its optimal policy is given by

$$
\widetilde{\pi}_t^{i,*} = \delta_i\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right),
$$

and thus the optimal policy of (4) can be written as

$$
\pi_t^{i,*} = \delta_i\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right) + \frac{c_i}{N}\left(\sum_{j\neq i}\pi_t^j + \pi_t^{i,*}\right). \tag{21}
$$

Symmetrically, all players $j \in \mathcal{I}$ follow an analogous to (21) strategy. Averaging over $j \in \mathcal{I}$ yields

$$
\frac{1}{N}\sum_{i=1}^N \pi_t^{i,*} = \psi_N \frac{1}{N}\sum_{i=1}^N \pi_t^{i,*} + \varphi_N\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right),
$$

with ψ_N and φ_N as in (15). If $\psi_N < 1$, the above equation gives

$$
\frac{1}{N}\sum_{i=1}^N \pi_t^{i,*} = \frac{\varphi_N}{1-\psi_N}\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right),
$$

and we obtain (17). The rest of the proof follows easily. □

We have stated the above result assuming that we start at $t = 0$. This is without loss of generality, as all arguments may be modified accordingly. For completeness, we present in the sequel the time-dependent case, in the context of a Markovian market.

Remark 1 As discussed in [12, Remark 2.5], problem (4) may be alternatively and equivalently represented as

$$V^i(x_1,\ldots,x_N)$$
$$= \sup_{\pi^i\in\mathcal{A}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_i'}\left(X_T^i - c_i' C_T^{-i}\right)\right)\middle| X_0^1 = x_1,\ldots,X_0^i = x_i,\ldots,X_0^N = x_N\right],$$

with $C_T^{-i} := \frac{1}{N-1}\sum_{j=1,j\neq i}^N X_T^j$, and $\delta_i = \frac{\delta_i'}{1+\frac{1}{N-1}c_i'}$ and $c_i = \frac{c_i'}{\frac{N-1}{N}+\frac{c_i'}{N}}$.

Remark 2 Instead of working with the minimal martingale measure in the incomplete Itô-diffusion market herein, one may employ the minimal entropy measure, $\mathbb{Q}^{ME}$, given by

$$\frac{d\mathbb{Q}^{ME}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T\left(\lambda_s^2+\chi_s^2\right)ds - \int_0^T \lambda_s\,dW_s - \int_0^T \chi_s\,dW_s^\perp\right), \tag{22}$$

where $\chi_t = -Z_t^\perp$ and $(y_t, Z_t, Z_t^\perp)_{t\in[0,T]}$ solves the backward stochastic differential equation (BSDE)

$$-dy_t = \left(-\frac{1}{2}\lambda_t^2 + \frac{1}{2}(Z_t^\perp)^2 - \lambda_t Z_t\right)dt - \left(Z_t\,dW_t + Z_t^\perp\,dW_t^\perp\right) \text{ and } \; y_T = 0. \tag{23}$$

The measures $\mathbb{Q}^{ME}$ and $\mathbb{Q}^{MM}$ are related through the relative entropy $\mathcal{H}$ in that $-\mathcal{H}(\mathbb{Q}^{ME}|\mathbb{P}) = \frac{1}{1-\rho^2}\ln M_0$ (cf. [17]). We choose to work with $\mathbb{Q}^{MM}$ for ease of the presentation.

From Lemma 1, we see that the Nash equilibrium process,

$$\pi_t^{i,*} = \bar{\delta}_i\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right),$$

resembles the optimal policy of an individual player of the classical optimal investment problem with exponential utility and *modified* risk tolerance, $\bar{\delta}_i$. The latter deviates from δ_i by

$$\bar{\delta}_i - \delta_i = \frac{\varphi_N}{1-\psi_N}c_i.$$

In the competitive case, $c_i > 0$, $\bar{\delta}_i > \delta_i$ and their difference increases with c_i, φ_N and ψ_N. At times t such that $\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t} > 0$ (resp. $\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t} < 0$), the competition concerns make the player invest more (resp. less) in the risky asset than without such concerns.

In the homophilous case, $c_i < 0$, we have that $\bar{\delta}_i < \delta_i$. Furthermore, direct computations show that their difference decreases with δ_i and each c_j, $j \neq i$, while it increases with c_i. In other words,

$$\partial_{\delta_j}\left(\bar{\delta}_i - \delta_i\right) < 0,\ \forall j \in \mathcal{I},\ \partial_{c_j}\left(\bar{\delta}_i - \delta_i\right) < 0,\ \forall j \in \mathcal{I} \setminus \{i\},\ \text{ and }\ \partial_{c_i}\left(\bar{\delta}_i - \delta_i\right) > 0.$$

At times t such that $\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t} > 0$, the player would invest less in the risky asset, compared to without homophilous interaction. This investment decreases if other players become more risk tolerant (their δ_j' s increase) or less homophilous (their c_j' s increase) or if the specific player i becomes more homophilous (c_i decreases). The case $\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t} < 0$ follows similarly. The comparison between the competitive and the homophilous case is described in Figure 1.

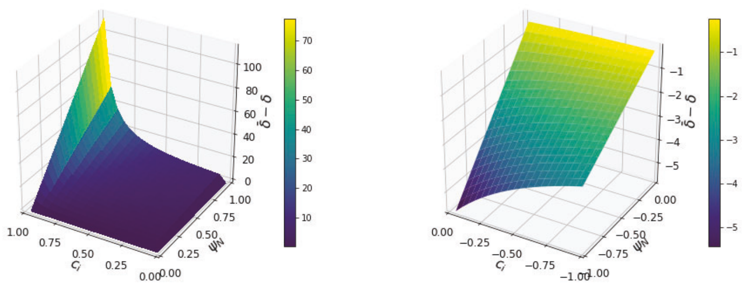

Fig. 1: The plot of $\bar{\delta}_i - \delta_i$ versus c_i and ψ_N, with $N = 25$ and $\varphi_N = 6$.

2.1.1 The Markovian case

We consider a single stochastic factor model in which the stock price process $(S_t)_{t\in[0,T]}$ solves

$$dS_t = \mu(t, Y_t) S_t\, dt + \sigma(t, Y_t) S_t\, dW_t, \tag{24}$$

$$dY_t = b(t, Y_t)\, dt + a(t, Y_t)\, dW_t^Y, \tag{25}$$

with $S_0 = S > 0$ and $Y_0 = y \in \mathbb{R}$. The market coefficients μ, σ, a and b satisfy appropriate conditions for these equations to have a unique strong solution. Further conditions, added next, are needed for the validity of the Feynman-Kac formula in Proposition 2.

Assumption 1 *The coefficients μ, σ, a and b are bounded functions, and a, b have bounded, uniformly in t, y-derivatives. It is further assumed that the Sharpe ratio function $\lambda(t, y) := \frac{\mu(t,y)}{\sigma(t,y)}$ is bounded and with bounded, uniformly in t, y-derivatives of any order.*

For $t \in [0, T]$, we consider the optimization problem

$$V^i(t,x_1,\dots,x_i,\dots,x_N,y) = \sup_{\pi^i\in\mathcal{A}} E_{\mathbb{P}}\Bigg[-\exp\left(-\frac{1}{\delta_i}\left(X_T^i - c_iC_T\right)\right)\Bigg|$$
$$X_t^1 = x_1,\dots,X_t^i = x_i,\dots,X_t^N = x_N, Y_t = y\Bigg], \quad (26)$$

with $(X_s^i)_{s\in[t,T]}$ solving $dX_s^i = \mu(s,Y_s)\pi_s^i\,ds + \sigma(s,Y_s)\pi_s^i\,dW_s$ and $\pi^i \in \mathcal{A}$, and C_T as in (5). We also consider the process $(\zeta_t)_{t\in[0,T]}$ with $\zeta_t := \zeta(t,Y_t)$, where $\zeta : [0,T]\times\mathbb{R}\to\mathbb{R}^+$ is defined as

$$\zeta(t,y) = E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_t^T \lambda^2(s,Y_s)\,ds}\Big| Y_t = y\right].$$

Under $\mathbb{Q}^{MM}$, the stochastic factor process $(Y_t)_{t\in[0,T]}$ satisfies

$$dY_t = (b(t,Y_t) - \rho\lambda(t,Y_t)a(t,Y_t))\,dt + a(t,Y_t)\,d\widetilde{W}_t^Y.$$

Thus, using the conditions on the market coefficients and the Feynman-Kac formula, we deduce that $\zeta(t,y)$ solves

$$\zeta_t + \frac{1}{2}a^2(t,y)\zeta_{yy} + (b(t,y) - \rho\lambda(t,y)a(t,y))\zeta_y = \frac{1}{2}(1-\rho^2)\lambda^2(t,y)\zeta, \quad (27)$$

with $\zeta(T,y) = 1$. In turn, the function $f(t,y) := \frac{1}{1-\rho^2}\ln\zeta(t,y)$ satisfies

$$f_t + \frac{1}{2}a^2(t,y)f_{yy} + (b(t,y) - \rho\lambda(t,y)a(t,y))f_y + \frac{1}{2}(1-\rho^2)a^2(t,y){f_y}^2 = \frac{1}{2}\lambda^2(t,y),$$
$$f(T,y) = 0. \quad (28)$$

In the absence of competitive/homophilous interaction, this problem has been examined by various authors (see, for example, [18]).

Proposition 2 *Under Assumption 1, the following assertions hold for $t\in[0,T]$.*

1. *If $\psi_N < 1$, there exists a wealth-independent Nash equilibrium $(\pi_s^*)_{s\in[t,T]} = \left(\pi_s^{1,*},\dots,\pi_s^{i,*},\dots,\pi_s^{N,*}\right)_{s\in[t,T]}$, where $\pi_s^{i,*}$, $i\in\mathcal{I}$, is given by the process*

$$\pi_s^{i,*} = \pi^{i,*}(s,Y_s), \quad (29)$$

with $(Y_t)_{t\in[0,T]}$ solving (25) and $\pi^{i,*} : [0,T]\times\mathbb{R}\to\mathbb{R}$ defined as

$$\pi^{i,*}(t,y) := \bar{\delta}_i\left(\frac{\lambda(t,y)}{\sigma(t,y)} + \rho\frac{a(t,y)}{\sigma(t,y)}f_y(t,y)\right), \quad (30)$$

with $\bar{\delta}_i$ as in (16) and $f(t,y)$ solving (28). The game value of player i, $i\in\mathcal{I}$, is given by

$$V^i(t,x_1,\dots,x_N,y) = -\exp\left(-\frac{1}{\delta_i}\left(x_i - \frac{c_i}{N}\Sigma_{i=1}^N x_i\right)\right)\zeta(t,y)^{\frac{1}{1-\rho^2}}$$
$$= -\exp\left(-\frac{1}{\delta_i}\left(x_i - \frac{c_i}{N}\Sigma_{i=1}^N x_i\right) + f(t,y)\right).$$

2. If $\psi_N = 1$, there exists no such Nash equilibrium.

Proof To ease the notation, we establish the results when $t = 0$ in (26). To this end, we first identify the process ξ in (10). For this, we rewrite the martingale in (9) as

$$M_t = \zeta(t,Y_t)e^{-\frac{1}{2}(1-\rho^2)\int_0^t \lambda^2(s,Y_s)\,ds},$$

and observe that

$$\begin{aligned} dM_t &= \Big(\zeta_t(t,Y_t) + (b(t,Y_t) - \rho a(t,Y_t)\lambda(t,Y_t))\zeta_y(t,Y_t) \\ &\quad + \frac{1}{2}a^2(t,Y_t)\zeta_{yy}(t,Y_t)\Big)\frac{M_t}{\zeta(t,Y_t)}\,dt - \frac{1}{2}(1-\rho^2)\lambda^2(t,Y_t)M_t\,dt \\ &\quad + a(t,Y_t)\frac{\zeta_y(t,Y_t)}{\zeta(t,Y_t)}M_t\left(\rho\,d\widetilde{W}_t + \sqrt{1-\rho^2}\,dW_t^{\perp}\right) \\ &= a(t,Y_t)\frac{\zeta_y(t,Y_t)}{\zeta(t,Y_t)}M_t\left(\rho\,d\widetilde{W}_t + \sqrt{1-\rho^2}\,dW_t^{\perp}\right), \end{aligned}$$

where we used that $\zeta(t,y)$ satisfies (27). Therefore, $\xi_t = a(t,Y_t)\frac{\zeta_y(t,Y_t)}{\zeta(t,Y_t)}$. In turn, using that $\zeta(t,y)^{1/(1-\rho^2)} = e^{f(t,y)}$, we obtain that

$$f_y(t,Y_t) = \frac{1}{1-\rho^2}\frac{\zeta_y(t,y)}{\zeta(t,y)} \quad \text{and} \quad \xi_t = (1-\rho^2)a(t,Y_t)f_y(t,Y_t),$$

and we easily conclude by replacing ξ_t by $(1-\rho^2)a(t,Y_t)f_y(t,Y_t)$ in (17).

It remains to show that the candidate investment process in (29) is admissible. Under Assumption 1 we deduce that $f_y(t,y)$ is a bounded function, since $\zeta(t,y)$ is bounded away from zero and $\zeta_y(t,y)$ is bounded. We easily conclude. □

Remark 3 In the Markovian model (24)–(25), the density of the minimal entropy measure $\mathbb{Q}^{ME}$ is fully specified. Indeed, the BSDE (23) admits the solution

$$y_t = f(t,Y_t), \quad Z_t = \rho a(t,Y_t)f_y(t,Y_t) \quad \text{and} \quad Z_t^{\perp} = \sqrt{1-\rho^2}a(t,Y_t)f_y(t,Y_t),$$

and, thus, the density of $\mathbb{Q}^{ME}$ is given by (22) with

$$\chi_t \equiv \chi(t,Y_t) = -\sqrt{1-\rho^2}a(t,Y_t)f_y(t,Y_t).$$

2.1.2 A fully solvable example

Consider the family of models with autonomous dynamics

$$\mu(t,y) = \mu y^{\frac{1}{2\ell}+\frac{1}{2}}, \ \ \sigma(t,y) = y^{\frac{1}{2\ell}}, \ \ b(t,y) = m - y, \quad a(t,y) = \beta\sqrt{y},$$

with $\mu > 0$, $\beta > 0$, $\ell \neq 0$ and $m > \frac{1}{2}\beta^2$. Notable cases are $\ell = 1$, which corresponds to the Heston stochastic volatility model, and $\ell = -1$ that is studied in [3].

Equation (28) depends only on $b(t,y)$, $a(t,y)$ and the Sharpe ratio $\lambda(t,y) = \mu\sqrt{y}$, and thus its solution $f(t,y)$ is independent of the parameter ℓ. Using the ansatz $f(t,y) = p(t)y + q(t)$ with $p(T) = q(T) = 0$, we deduce from (28) that $p(t)$ and $q(t)$ satisfy

$$\begin{aligned} &\dot{p}(t) - \frac{1}{2}(\mu + \rho\beta p(t))^2 - p(t) + \frac{1}{2}\beta^2 p^2(t) = 0, \\ &\dot{q}(t) + mp(t) = 0. \end{aligned} \tag{31}$$

In turn,

$$p(t) = \frac{1+\rho\mu\beta - \sqrt{\Delta}}{(1-\rho^2)\beta^2} \frac{1 - e^{-\sqrt{\Delta}(T-t)}}{1 - \frac{1+\rho\mu\beta-\sqrt{\Delta}}{1+\rho\mu\beta+\sqrt{\Delta}} e^{-\sqrt{\Delta}(T-t)}}, \qquad \Delta = 1 + \beta^2\mu^2 + 2\rho\mu\beta > 0,$$

and $q(t) = m\int_t^T p(s)ds$.

From (30), we obtain that the Nash equilibrium strategy $\left(\pi_s^{i,*}\right)_{s\in[t,T]}$, $t \in [0,T]$, for player i is given by the process

$$\pi_s^{i,*} = \bar{\delta}_i(\mu + \rho\beta p(s))Y_s^{\frac{1}{2}(1-\frac{1}{\ell})}.$$

If $\ell = 1$, the policy becomes deterministic, $\pi_s^{i,*} = \bar{\delta}_i(\mu + \rho\beta p(s))$, and the equilibrium wealth process solves

$$dX_s^{i,*} = \bar{\delta}_i(\mu + \rho\beta p(s))(\mu Y_s\, ds + \sqrt{Y_s}\, dW_s).$$

2.2 The common-noise MFG

We analyze the limit as $N \uparrow \infty$ of the N-player game studied in Section 2.1. We first give an intuitive and informal argument that leads to a candidate optimal strategy in the mean-field setting, and then propose a rigorous formulation for the MFG. The analysis follows closely the arguments developed in [12].

For the N-player game, we denote by $\eta_i = (x_i, \delta_i, c_i)$ the *type vector* for player i, where x_i is her initial wealth, and η_i and c_i are her risk tolerance coefficient and interaction parameter, respectively. Such type vectors induce an empirical measure

m_N, called the *type distribution*,

$$m_N(A) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\eta_i}(A), \text{ for Borel sets } A \subset \mathcal{Z},$$

which is a probability measure on the space $\mathcal{Z} := \mathbb{R} \times (0, \infty) \times (-\infty, 1]$.

We recall (*cf.* (17)) that the equilibrium strategies $(\pi_t^{i,*})_{t\in[0,T]}$, $i \in \mathcal{I}$, are given as the product of the common (type-independent) process $\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}$ and the modified risk tolerance parameter $\bar{\delta}_i = \delta_i + \frac{\varphi_N}{1-\psi_N} c_i$. Therefore, it is *only* the coefficient $\bar{\delta}_i$ that depends on the empirical distribution m_N through ψ_N and φ_N, as both these quantities can be obtained by averaging appropriate functions over m_N. Therefore, if we assume that m_N converges weakly to some limiting probability measure as $N \uparrow \infty$, we should intuitively expect that the corresponding equilibrium strategies also converge. This is possible, for instance, by letting the type vector $\eta = (x, \delta, c)$ be a random variable in the space $\mathcal{Z}$ with limiting distribution m, and take η_i as i.i.d. samples of η. The sample η_i is drawn and assigned to player i at initial time $t = 0$. We would then expect $(\pi^{i,*})_{t\in[0,T]}$ to converge to the process

$$\lim_{N\uparrow\infty} \pi_t^{i,*} = \left(\delta_i + \frac{\bar{\delta}}{1-\bar{c}} c_i\right)\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right), \tag{32}$$

where $\bar{c}$ and $\bar{\delta}$ represent the average interaction and risk tolerance coefficients.

Next, we introduce the mean-field game in the incomplete Itô-diffusion market herein, and we show that (32) indeed arises as its equilibrium strategy. We model a single representative player, whose type vector is a random variable with distribution m, and all players in the continuum act in this common incomplete market.

2.2.1 The Itô-diffusion common-noise MFG

To describe the heterogeneous population of players, we introduce the type vector

$$\eta = (x, \delta, c) \in \mathcal{Z}, \tag{33}$$

where $\delta > 0$ and $c \in (-\infty, 1]$ represent the risk tolerance coefficient and interaction parameter, and x is the initial wealth. This type vector is assumed to be independent of both W and W^Y, which drive the stock price process (1), and is assumed to have finite second moments.

To formulate the mean-field portfolio game, we now let the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ support W, W^Y as well as η. We assume that η has second moments under $\mathbb{P}$. We denote by $(\mathcal{F}_t^{MF})_{t\in[0,T]}$ the smallest filtration satisfying the usual assumptions for which η is $\mathcal{F}_0^{MF}$-measurable and both W, W^Y are adapted. As before, we denote by $(\mathcal{F}_t)_{t\in[0,T]}$ the natural filtration generated by W and W^Y, and by $(\mathcal{G}_t)_{t\in[0,T]}$ the one generated only by W^Y.

We also consider the wealth process $(X_t)_{t\in[0,T]}$ of the *representative player* solving

$$dX_t = \pi_t \left(\mu_t \, dt + \sigma_t \, dW_t\right), \tag{34}$$

with $X_0 = x \in \mathbb{R}$ and $\pi \in \mathcal{A}^{MF}$, where

$$\mathcal{A}^{MF} = \Bigg\{\pi : \text{self-financing, } \mathcal{F}_t^{MF}\text{-progressively measurable and } E_{\mathbb{P}}\left[\int_0^T \sigma_s^2 \pi_s^2 \, ds\right] < \infty\Bigg\}.$$

Similarly to the framework in [12], there exist two independent sources of randomness in the model: the first is due to the evolution of the stock price process, described by the Brownian motions W and W^Y. The second is given by η, which models the type of the player, *i.e.*, the triplet of initial wealth, risk tolerance, and interaction parameter in the population continuum. The first source of noise is *stochastic* and common to each player in the continuum while the second is *static*, being assigned at time zero and with the dynamic competition starting right afterwards.

In analogy to the N-player setting, the representative player optimizes the expected terminal utility, taking into account the performance of the average terminal wealth of the population, denoted by $\overline{X}$. As in [12], we introduce the following definition for the MFG considered herein.

Definition 2 For each $\pi \in \mathcal{A}^{MF}$, let $\overline{X} := E_{\mathbb{P}}[X_T|\mathcal{F}_T]$ with $(X_t)_{t\in[0,T]}$ solving (34), and consider the optimization problem

$$V(x) = \sup_{\pi\in\mathcal{A}^{MF}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta}\left(X_T - c\overline{X}\right)\right)\middle|\mathcal{F}_0^{MF}, X_0 = x\right]. \tag{35}$$

A strategy $\pi^* \in \mathcal{A}^{MF}$ is a mean-field equilibrium if π^* is the optimal strategy of the above problem when $\overline{X}^* := E_{\mathbb{P}}[X_T^*|\mathcal{F}_T]$ is used for $\overline{X}$, where $(X_t^*)_{t\in[0,T]}$ solves (34) with π^* being used.

Next, we state the main result.

Proposition 3 *If* $E_{\mathbb{P}}[c] < 1$, *there exists a unique wealth-independent MFG equilibrium* $(\pi_t^*)_{t\in[0,T]}$, *given by the* $\mathcal{F}_0^{MF} \vee \mathcal{G}_t$ *process*

$$\pi_t^* = \left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1 - E_{\mathbb{P}}[c]} c\right)\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right), \tag{36}$$

with ξ *as in* (10). *The corresponding optimal wealth is given by*

$$X_t^* = x + \left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1 - E_{\mathbb{P}}[c]} c\right)\int_0^t \left(\lambda_s + \frac{\rho}{1-\rho^2}\xi_s\right)(\lambda_s \, ds + dW_s), \tag{37}$$

and

$$V(x) = -\exp\left(-\frac{1}{\delta}(x-cm)\right) M_0^{\frac{1}{1-\rho^2}}$$
$$= -\exp\left(-\frac{1}{\delta}(x-cm)\right)\left(E_{\mathbb{Q}^{MM}}\left[e^{-\frac{1}{2}(1-\rho^2)\int_0^T \lambda_s^2\,ds}\right]\right)^{\frac{1}{1-\rho^2}},$$

where $m = E_{\mathbb{P}}[x]$. *If* $E_{\mathbb{P}}[c] = 1$, *there is no such Nash equilibrium.*

Proof We first observe that π^* in (36) is $\mathcal{F}_t^{MF}$-measurable since $\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right) \in \mathcal{G}_t$, and thus $\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right) \in \mathcal{F}_t$, while the factor $\left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}c\right) \in \mathcal{F}_0^{MF}$ (independent of $\mathcal{F}_t$). Furthermore, π^* is also square-integrable under standing assumptions, and thus admissible. To show that it is also indeed an equilibrium policy, we shall first define $\overline{X}$ using π^*, and then verify that the optimal strategy to the representative player's problem (35) coincides with π_t^* when this specific $\overline{X}$ is used in (35). To this end, we introduce the process $\overline{X}_t := E_{\mathbb{P}}[X_t^*|\mathcal{F}_t]$ with $(X_t^*)_{t\in[0,T]}$ as in (37). Then,

$$\begin{aligned}\overline{X}_t &= E_{\mathbb{P}}\left[x + \left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}c\right)\int_0^t\left(\lambda_s + \frac{\rho}{1-\rho^2}\xi_s\right)(\lambda_s\,ds + dW_s)\Big|\mathcal{F}_t\right]\\ &= m + \left(E_{\mathbb{P}}[\delta] + \frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}E_{\mathbb{P}}[c]\right)\int_0^t\left(\lambda_s + \frac{\rho}{1-\rho^2}\xi_s\right)(\lambda_s\,ds + dW_s)\\ &= m + \left(\frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}\right)\int_0^t\left(\lambda_s + \frac{\rho}{1-\rho^2}\xi_s\right)(\lambda_s\,ds + dW_s),\end{aligned}$$

where we have used that $\int_0^t\left(\lambda_s + \frac{\rho}{1-\rho^2}\xi_s\right)(\lambda_s\,ds + dW_s)$ is $\mathcal{G}_t$-measurable and thus $\mathcal{F}_t$-measurable, and that $\left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}c\right)$ is independent of $\mathcal{F}_t$.

Next, we introduce the auxiliary process $(\tilde{x}_t)_{t\in[0,T]}$, $\tilde{x}_t := X_t - c\overline{X}_t$, with $(X_t)_{t\in[0,T]}$ as in (34). Then,

$$d\tilde{x}_t = \widetilde{\pi}_t(\mu_t\,dt + \sigma_t\,dW_t) \quad \text{and} \quad \tilde{x}_0 = \tilde{x} := x - cm,$$

and $\widetilde{\pi}_t = \pi_t - c\left(\frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}\right)\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right)$. In turn, we consider the optimization problem

$$v(\tilde{x}) := \sup_{\widetilde{\pi}\in\mathcal{A}^{MF}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta}\tilde{x}_T\right)\Big|\mathcal{F}_0^{MF}, \tilde{x}_0 = \tilde{x}\right].$$

From Lemma 1, we deduce that the optimal strategy is given by

$$\widetilde{\pi}_t^* = \delta\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right),$$

and, thus,

$$\pi_t^* = \delta\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right) + c\left(\frac{E_{\mathbb{P}}[\delta]}{1-E_{\mathbb{P}}[c]}\right)\left(\frac{\lambda_t}{\sigma_t} + \frac{\rho}{1-\rho^2}\frac{\xi_t}{\sigma_t}\right).$$

The rest of the proof follows easily. □

If we view $\eta = (x, \delta, c)$ in the N-player game in Section 2.1 as i.i.d. samples on the space $\mathcal{Z}$ with distribution m, then $\lim_{N\uparrow\infty} \psi_N = E_{\mathbb{P}}[c]$ and $\lim_{N\uparrow\infty} \varphi_N = E_{\mathbb{P}}[\delta]$ a.s.. We then obtain the convergence of the corresponding optimal processes, namely, for $t \in [0, T]$,

$$\lim_{N\uparrow\infty} \pi_t^{i,*} = \pi_t^*, \quad \text{and} \quad \lim_{N\uparrow\infty} X_t^{i,*} = X_t^*.$$

2.2.2 The Markovian case

In analogy to the N-player case, we have the following result.

Proposition 4 *Assume that the stock price process follows the single factor model* (24)–(25). *Then, if* $E_{\mathbb{P}}[c] < 1$, *there exists a unique wealth-independent Markovian mean-field game equilibrium, given by the process* $(\pi_t^*)_{t\in[0,T]}$,

$$\pi_t^* = \pi^*(\eta, t, Y_t) = \left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1 - E_{\mathbb{P}}[c]} c\right) \left(\frac{\lambda(t, Y_t)}{\sigma(t, Y_t)} + \rho \frac{a(t, Y_t)}{\sigma(t, Y_t)} f_y(t, Y_t)\right),$$

with the $\mathcal{F}_0^{MF}$*-measurable random function* $\pi^*(\eta, t, y) : \mathcal{Z} \times [0, T] \times \mathbb{R}$,

$$\pi^*(\eta, t, y) := \left(\delta + \frac{E_{\mathbb{P}}[\delta]}{1 - E_{\mathbb{P}}[c]} c\right) \left(\frac{\lambda(t, y)}{\sigma(t, y)} + \rho \frac{a(t, y)}{\sigma(t, y)} f_y(t, y)\right).$$

If $E_{\mathbb{P}}[c] = 1$, *there is no such mean-field game stochastic equilibrium.*

3 Complete Itô-diffusion common market and CARA utilities with random risk tolerance coefficients

In this section, we focus on the complete common market case, but we extend the model by allowing random individual risk tolerance coefficients. We start with a background result for the single-player problem, which is new and interesting in its own right. Building on it, we analyze both the N-player and the MFG. The analysis shows that the randomness of the individual risk tolerance gives rise to virtual "personalized" markets, in that the original common risk premium process now differs across players, depending on their risk tolerance. This brings substantial complexity as the tractability coming from the original common market is now lost.

3.1 The Itô-diffusion market and random risk tolerance coefficients

We consider the complete analog of the Itô-diffusion market studied in Section 2. Specifically, we consider a market with a riskless bond (taken to be the numeraire and offering zero interest rate) and a stock whose price process $(S_t)_{t\in[0,T]}$ solves

$$dS_t = S_t\,(\mu_t\,dt + \sigma_t\,dW_t),$$

with $S_0 > 0$, and $(W_t)_{t\in[0,T]}$ being a Brownian motion in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The market coefficients $(\mu_t)_{t\in[0,T]}$ and $(\sigma_t)_{t\in[0,T]}$ are $\mathcal{F}_t$-adapted processes, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the natural filtration generated by W, and with $0 < c \le \sigma_t \le C$ and $|\mu_t| \le C$, $t \in [0,T]$, for some (possibly deterministic) constants c and C.

In this market, N players, indexed by $i \in \mathcal{I}$, $\mathcal{I} = \{1,2,\dots,N\}$, trade between the two accounts in $[0,T]$, with individual wealths $\left(X_t^i\right)_{t\in[0,T]}$ solving

$$dX_t^i = \pi_t^i\,(\mu_t\,dt + \sigma_t\,dW_t), \tag{38}$$

and $X_0^i = x_i \in \mathbb{R}$.

Each of the players, say player i, has *random risk tolerance*, δ_T^i, defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with the following properties:

Assumption 2 *For each $i \in \mathcal{I}$, the risk tolerance δ_T^i is an $\mathcal{F}_T$-measurable random variable with $\delta_T^i \ge \delta > 0$ and $E_{\mathbb{P}}\left(\delta_T^i\right)^2 < \infty$.*

The objective of each player is to optimize

$$V^i(x_1,\dots,x_i,\dots,x_N) = \sup_{\mathcal{A}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_T^i}\left(X_T^i - \frac{c_i}{N}\sum_{j=1}^N X_T^j\right)\right)\right.$$
$$\left.\Bigg| X_0^1 = x_1,\dots,X_0^i = x_i,\dots,X_0^N = x_N\right], \tag{39}$$

with $c_i \in (-\infty, 1]$, X^j, $j \in \mathcal{I}$, solving (38), and $\mathcal{A}$ defined similarly to (3).

As in Section 2.1, we are interested in a Nash equilibrium solution, which is defined as in Definition 1. Before we solve the underlying stochastic N-player game, we focus on the single-player case. This is a problem interesting in its own right and, to our knowledge, has not been studied before in such markets. A similar problem was considered in a single-period binomial model in [15] and in a special diffusion case in [16] in the context of indifference pricing of bonds. For generality, we present below the time-dependent case.

3.2 The single-player problem

We consider the optimization problem

$$v_t(x) = \sup_{\pi \in \mathcal{A}} E_{\mathbb{P}} \left[-e^{-\frac{1}{\delta_T} x_T} \middle| \mathcal{F}_t, x_t = x \right], \tag{40}$$

with $\delta_T \in \mathcal{F}_T$ satisfying Assumption 2 and $(x_s)_{s\in[t,T]}$ solving (38) with $x_t = x \in \mathbb{R}$.

We define $(Z_t)_{t\in[0,T]}$ by

$$Z_t = \exp\left(-\frac{1}{2}\int_0^t \lambda_s^2\, ds - \int_0^t \lambda_s\, dW_s\right),$$

and recall the associated (unique) risk neutral measure $\mathbb{Q}$, defined on $\mathcal{F}_T$ and given by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T. \tag{41}$$

We introduce the process $(\delta_t)_{t\in[0,T]}$,

$$\delta_t := E_{\mathbb{Q}}[\delta_T | \mathcal{F}_t], \tag{42}$$

which may be thought as the arbitrage-free price of the risk tolerance "claim" δ_T. We also introduce the measure $\hat{\mathbb{Q}}$, defined on $\mathcal{F}_T$, with

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{P}} = \frac{\delta_T}{E_{\mathbb{Q}}[\delta_T]} Z_T.$$

Direct calculations yield that under measure $\hat{\mathbb{Q}}$, the process $\left(\frac{S_t}{\delta_t}\right)_{t\in[0,T]}$ is an $\mathcal{F}_t$-martingale.

By the model assumptions and the martingale representation theorem, there exists an $\mathcal{F}_t$-adapted process $(\xi_t)_{t\in[0,T]}$ with $\xi \in \mathcal{L}^2(\mathbb{P})$ such that

$$d\delta_t = \xi_t \delta_t\, dW_t^{\mathbb{Q}}, \tag{43}$$

with $W_t^{\mathbb{Q}} = W_t + \int_0^t \lambda_s\, ds$. Next, we introduce the process

$$H_t := E_{\widetilde{\mathbb{Q}}}\left[\frac{1}{2}\int_t^T (\lambda_s - \xi_s)^2\, ds \middle| \mathcal{F}_t\right], \tag{44}$$

where $\widetilde{\mathbb{Q}}$ is defined on $\mathcal{F}_T$ by

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T (\lambda_s - \xi_s)^2\, ds - \int_0^T (\lambda_s - \xi_s)\, dW_s\right). \tag{45}$$

Under $\widetilde{\mathbb{Q}}$, the process $\left(W_t^{\widetilde{\mathbb{Q}}}\right)_{t\in[0,T]}$ with

$$W_t^{\widetilde{\mathbb{Q}}} := W_t + \int_0^t (\lambda_s - \xi_s)\, ds \tag{46}$$

is a standard Brownian motion, and $\left(\frac{1}{\delta_t} S_t\right)_{t\in[0,T]}$ is a martingale with dynamics

$$d\left(\frac{S_t}{\delta_t}\right) = (\sigma_t - \xi_t)\frac{S_t}{\delta_t}\, dW_t^{\widetilde{\mathbb{Q}}}.$$

Direct calculations yield

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{Q}} = \delta_T.$$

Alternatively, H_t may be also represented as

$$H_t = \frac{E_{\mathbb{Q}}[\delta_T \int_t^T \frac{1}{2}(\lambda_s - \xi_s)^2\, ds|\mathcal{F}_t]}{E_{\mathbb{Q}}[\delta_T|\mathcal{F}_t]} = E_{\mathbb{Q}}\left[\frac{\delta_T}{\delta_t}\int_t^T \frac{1}{2}(\lambda_s - \xi_s)^2\, ds\middle|\mathcal{F}_t\right], \tag{47}$$

which is obtained by using that

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{Q}} = \exp\left(-\frac{1}{2}\int_0^T \xi_s^2\, ds + \int_0^T \xi_s\, dW_s^{\mathbb{Q}}\right).$$

Finally, we introduce the processes $(M_t)_{t\in[0,T]}$ and $(\eta_t)_{t\in[0,T]}$ with

$$M_t = \mathbb{E}_{\widetilde{\mathbb{Q}}}\left[\frac{1}{2}\int_0^T (\lambda_s - \xi_s)^2\, ds\middle|\mathcal{F}_t\right] \quad \text{and} \quad dM_t = \eta_t\, dW_t^{\widetilde{\mathbb{Q}}}. \tag{48}$$

We are now ready to present the main result.

Proposition 5 *The following assertions hold:*

1. *The value function of* (40) *is given by*

$$v_t(x) = -\exp\left(-\frac{x}{\delta_t} - H_t\right),$$

   with δ and H as in (42) *and* (44).
2. *The optimal strategy $(\pi_s^*)_{s\in[t,T]}$ is given by*

$$\pi_s^* = \delta_s \frac{\lambda_s - \eta_s - \xi_s}{\sigma_s} + \frac{\xi_s}{\sigma_s} x_s^*, \tag{49}$$

   with ξ, η as in (43) *and* (48)*, and x^* solving* (38) *with π^* being used.*
3. *The optimal wealth $(x_s^*)_{s\in[t,T]}$ solves*

$$dx_s^* = \lambda_s\left(\delta_s(\lambda_s - \eta_s - \xi_s) + \xi_s x_s^*\right) ds + \left(\delta_s(\lambda_s - \eta_s - \xi_s) + \xi_s x_s^*\right) dW_s,$$

   with $x_t^* = x$, and is given by

$$x_s^* = x\Phi_{t,s} + \int_t^s \delta_u(\lambda_u - \xi_u)(\lambda_u - \eta_u - \xi_u)\Phi_{u,s}\,du + \int_t^s \delta_u(\lambda_u - \eta_u - \xi_u)\Phi_{u,s}\,dW_u, \quad (50)$$

where, for $0 \le u \le s \le T$,

$$\Phi_{u,s} := \exp\left(\int_u^s \left(\lambda_v - \frac{1}{2}\xi_v\right)\xi_v\,dv + \int_u^s \xi_v\,dW_v\right).$$

Using (50), (49) gives the explicit representation of the optimal policy,

$$\pi_s^* = \delta_s \frac{\lambda_s - \eta_s - \xi_s}{\sigma_s} + \frac{\xi_s}{\sigma_s}\left(x\Phi_{t,s} + \int_t^s \delta_u(\lambda_u - \xi_u)(\lambda_u - \eta_u - \xi_u)\Phi_{u,s}\,du + \int_t^s \delta_u(\lambda_u - \eta_u - \xi_u)\Phi_{u,s}\,dW_u\right).$$

3.2.1 The Markovian case

We assume that the stock price process $(S_t)_{t\in[0,T]}$ solves

$$dS_t = \mu(t,S_t)S_t\,dt + \sigma(t,S_t)S_t\,dW_t,$$

with the initial price $S_0 > 0$, and the functions $\mu(t,S_t)$ and $\sigma(t,S_t)$ satisfying appropriate conditions, similar to the ones in Subsection 2.1.1 and Assumption 1. The risk tolerance is assumed to have the functional representation

$$\delta_T = \delta(S_T),$$

for some function $\delta : \mathbb{R}^+ \to \mathbb{R}^+$ bounded from below and such that $E_{\mathbb{P}}\left[\delta^2(S_T)\right] < \infty$, (*cf.* Assumption 2).

The value function in (40) takes the form

$$V(t,x,S) = \sup_{\pi\in\mathcal{A}} E_{\mathbb{P}}\left[-e^{-\frac{1}{\delta(S_T)}x_T}\middle| x_t = x, S_t = S\right],$$

and, in turn, Proposition 5 yields

$$V(t,x,S) = -\exp\left(\frac{x}{\delta(t,S)} - H(t,S)\right),$$

with $\delta(t,S)$ and $H(t,S)$ solving

$$\delta_t + \frac{1}{2}\sigma^2(t,S)S^2\delta_{SS} = 0, \quad \delta(T,S) = \delta(S),$$

and

$$H_t + \frac{1}{2}\sigma^2(t,S)S^2 H_{SS} + \frac{1}{\delta(t,S)}\sigma^2(t,S)S^2\delta_S(t,S)H_S + \frac{1}{2}\left(\lambda(t,S) - \frac{1}{\delta(t,S)}\sigma(t,S)S\delta_S(t,S)\right)^2 = 0, \quad H(T,S) = 0.$$

Clearly,

$$\delta(t,S) = E_{\mathbb{Q}}\left[\delta(S_T)|\, S_t = S\right],$$

and

$$H(t,S) = E_{\widetilde{\mathbb{Q}}}\left[\int_t^T \frac{1}{2}\left(\lambda(u,S_u) - \sigma(u,S_u)S_u\frac{\delta_S(u,S_u)}{\delta(u,S_u)}\right)^2 du \middle| S_t = S\right],$$

and, furthermore,

$$\xi_t = \frac{\delta_S(t,S_t)}{\delta(t,S_t)}S_t\sigma(t,S_t) \quad \text{and} \quad \eta_t = H_S(t,S_t)S_t\sigma(t,S_t).$$

Using the above relations and (49), we derive the optimal investment process,

$$\pi_s^* = \delta(s,S_s)\left(\frac{\lambda(s,S_s)}{\sigma(s,S_s)} - S_s H_S(s,S_s)\right) + \delta_S(s,S_s)S_s\left(-1 + \frac{1}{\delta\,(s,S_s)}x_s^*\right).$$

For completeness, we note that if $\delta_T \equiv \delta > 0$, the above expression simplify to (see [18])

$$V(t,x,S) = -e^{-\frac{1}{\delta}x - H(t,S)},$$

with $H(t,S)$ solving

$$H_t + \frac{1}{2}\sigma^2(t,S)S^2 H_{SS} + \frac{1}{2}\lambda^2(t,S) = 0, \quad H(T,S) = 0.$$

The optimal strategy reduces to

$$\pi_s^* = \delta\left(\frac{\lambda(s,S_s)}{\sigma(s,S_s)} - S_s H_S(s,S_s)\right).$$

3.3 N-player game

We now study the N-player game. The concepts and various quantities are in direct analogy to those in Section 2.1 and, thus, we omit various intermediate steps and only focus on the new elements coming from the randomness of the risk tolerance coefficients.

Proposition 6 *For $i \in \mathcal{I}$, let*

$$\delta_t^i = E_{\mathbb{Q}}[\delta_T^i \big| \mathcal{F}_t],$$

with $\mathbb{Q}$ *as in* (41) *and* $\left(\xi_t^i\right)_{t\in[0,T]}$ *be such that*

$$d\delta_t^i = \xi_t^i \delta_t^i \, dW_t^{\mathbb{Q}}.$$

Define the measure $\widetilde{Q}^i$ *on* $\mathcal{F}_T$ *as*

$$\frac{d\widetilde{\mathbb{Q}}^i}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^T (\lambda_s - \xi_s^i)^2 \, ds - \int_0^T \left(\lambda_s - \xi_s^i\right) dW_s\right), \tag{51}$$

and the processes $(M_t^i)_{t\in[0,T]}$ *and* $(\eta_t)_{t\in[0,T]}$ *with*

$$M_t^i = \mathbb{E}_{\widetilde{\mathbb{Q}}^i}\left[\frac{1}{2}\int_0^T \left(\lambda_s - \xi_s^i\right)^2 ds \middle| \mathcal{F}_t\right] \quad \textit{and} \quad dM_t^i = \eta_t^i \, dW_t^{\widetilde{\mathbb{Q}}^i}. \tag{52}$$

Let also,

$$\psi_N = \frac{1}{N}\sum_{i=1}^N c_i,$$

and assume that $\psi_N < 1$. *Then*

1. *The player i's game value* (39) *is given by*

$$V^i(x_1, \ldots, x_i, \ldots, x_N) \\ = -\exp\left(-\frac{1}{E_{\mathbb{Q}}[\delta_T^i]}(x_i - \frac{c_i}{N}\Sigma_{j=1}^N x_j) - E_{\widetilde{\mathbb{Q}}^i}\left[\frac{1}{2}\int_0^T \left(\lambda_s - \xi_s^i\right)^2 ds\right]\right).$$

2. *The equilibrium strategies* $(\pi_t^{1,*}, \ldots, \pi_t^{N,*})_{t\in[0,T]}$ *are given by*

$$\pi_t^{i,*} = c_i \bar{\pi}_t^* + \frac{1}{\sigma_t}\left(\delta_t^i(\lambda_t - \xi_t^i - \eta_t^i) + (X_t^{i,*} - \frac{c_i}{N}\sum_{j=1}^N X_t^{j,*})\xi_t^i\right), \tag{53}$$

where $\bar{\pi}_t^* := \frac{1}{N}\Sigma_{j=1}^N \pi_t^{j,*}$ *is defined as*

$$\bar{\pi}_t^* = \frac{1}{1-\psi_N}\frac{1}{\sigma_t}\left(\lambda_t \varphi_N^1(t) - \varphi_N^2(t) + \varphi_N^3(t) - \varphi_N^4(t)\bar{X}_t^*\right), \tag{54}$$

with

$$\varphi_N^1(t) = \frac{1}{N}\Sigma_{j=1}^N \delta_t^j, \quad \varphi_N^2(t) = \frac{1}{N}\Sigma_{j=1}^N \delta_t^j(\xi_t^j + \eta_t^j),$$
$$\varphi_N^3(t) = \frac{1}{N}\Sigma_{j=1}^N X_t^{j,*}\xi_t^j, \quad \varphi_N^4(t) = \Sigma_{j=1}^N c_j \xi_t^j.$$

3. *The associated optimal wealth processes* $\left(X_t^{i,*}\right)_{t\in[0,T]}$ *are given by*

$$X_t^{i,*} = c_i \bar{X}_t^* + \left(\tilde{x}_i \Phi_{0,t}^i + \int_0^t (\lambda_s - \xi_s^i) \delta_s^i (\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, ds \right. \\ \left. + \int_0^t \delta_s^i (\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, dW_s \right), \quad (55)$$

with

$$\bar{X}_t^* := \frac{1}{1-\psi_N} \left(\frac{1}{N} \Sigma_{i=1}^N \left(\tilde{x}_i \Phi_{0,t}^i + \int_0^t \delta_s^i (\lambda_s - \xi_s^i)(\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, ds \right. \right. \\ \left. \left. + \int_0^t \delta_s^i (\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, dW_s \right) \right),$$

where $\tilde{x}_i = x_i - \frac{c_i}{N} \Sigma_{j=1}^N x_j$, *and*

$$\Phi_{s,t}^i := \exp\left(\int_s^t \left(\lambda_u - \frac{1}{2} \xi_u^i \right) \xi_u^i \, du + \int_s^t \xi_u^i \, dW_u \right). \quad (56)$$

Proof Using the dynamics of $X^1, \ldots, X^N$ in (38), problem (39) reduces to

$$v(\tilde{x}) = \sup_{\widetilde{\pi}^i \in \mathcal{A}} E_{\mathbb{P}} \left[-\exp\left(-\frac{1}{\delta_T^i} \widetilde{X}_T^i \right) \right],$$

where $\widetilde{X}_t^i = X_t^i - \frac{c_i}{N} \Sigma_{j=1}^N X_t^j$ satisfies $d\widetilde{X}_t^i = \widetilde{\pi}_t^i (\mu_t \, dt + \sigma_t \, dW_t)$ with $\widetilde{X}_0^i = \tilde{x}_i$. Taking $\pi^j \in \mathcal{A}$, $j \neq i$, as fixed and using Proposition 5, we deduce that $\pi^{i,*}$ satisfies

$$\widetilde{\pi}_t^{i,*} = \pi_t^{i,*} - \frac{c_i}{N} \left(\Sigma_{j \neq i} \pi_t^j + \pi_t^{i,*} \right) = \delta_t^i \frac{\lambda_t - \eta_t^i - \xi_t^i}{\sigma_t} + \frac{\xi_t^i}{\sigma_t} \widetilde{X}_t^{i,*}, \quad (57)$$

where $\widetilde{X}_t^{i,*}$ is the wealth process $\widetilde{X}_t^i$ associated with the strategy $\widetilde{\pi}_t^{i,*}$.

At equilibrium, π_t^j in (57) coincides with $\pi_t^{j,*}$. Therefore, averaging over $i \in \mathcal{I}$ gives

$$\bar{\pi}_t^* - \psi_N \bar{\pi}_t^* = \frac{1}{\sigma_t} \left(\lambda_t \varphi_N^1(t) - \varphi_N^2(t) + \varphi_N^3(t) - \varphi_N^4(t) \bar{X}_t^* \right).$$

Dividing both sides by $1 - \psi_N$ yields (54), and then (53) follows.

To obtain explicit expressions of $X_t^{i,*}$ and $\bar{X}_t^*$, we solve for $\widetilde{X}_t^{i,*}$ using the optimal strategy deduced in Section 3.2 (*cf.* (49)). We then obtain

$$\widetilde{X}_t^{i,*} = X_t^{i,*} - \frac{c_i}{N} \sum_{j=1}^N X_t^{j,*} = \tilde{x}_i \Phi_{0,t}^i + \int_0^t \delta_s^i (\lambda_s - \xi_s^i)(\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, ds \\ + \int_0^t \delta_s^i (\lambda_s - \eta_s^i - \xi_s^i) \Phi_{s,t}^i \, dW_s,$$

with $\Phi_{s,t}^i$ as in (56). We conclude by averaging over all $i \in \mathcal{I}$. □

3.4 The Itô-diffusion common-noise MFG

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space that supports the Brownian motion W as well as the *random* type vector

$$\theta = (x, \delta_T, c),$$

which is independent of W. As before, we denote by $(\mathcal{F}_t)_{t\in[0,T]}$ the natural filtration generated by W, and $(\mathcal{F}_t^{MF})_{t\in[0,T]}$ with $\mathcal{F}_t^{MF} = \mathcal{F}_t \vee \sigma(\theta)$. In the mean-field setting, we model the representative player. One may also think of a continuum of players whose initial wealth x and the interaction parameter c are random, chosen at initial time 0, similar to the MFG in Section 2.2 herein. However, now, their risk tolerance coefficients have *two* sources of randomness, related to their form and their terminal (at T) measurability, respectively. Specifically, at initial time 0, it is determined how these coefficients will depend on the final information, provided at T. For example, in the Markovian case, this amounts to (randomly) selecting at time 0 the functional form of $\delta(\cdot)$ and, in turn, the risk tolerance used for utility maximization is given by the random variable $\delta(S_T)$, which depends on the information $\mathcal{F}_T$ through S_T.

Similarly to (39), we are concerned with the optimization problem

$$V(x) = \sup_{\pi\in\mathcal{A}^{MF}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_T}\left(X_T^\pi - c\overline{X}\right)\right)\middle|\mathcal{F}_0^{MF},\ X_0 = x\right], \tag{58}$$

and the definition of the mean-field game is analogous to Definition 2.

Let the processes $(\delta_t)_{t\in[0,T]}$ and $(\xi_t)_{t\in[0,T]}$ be given by

$$\delta_t = E_{\mathbb{Q}}[\delta_T|\mathcal{F}_t^{MF}] \quad \text{and} \quad d\delta_t = \xi_t\delta_t\, dW_t^{\mathbb{Q}}, \tag{59}$$

with $\mathbb{Q}$ defined on $\mathcal{F}_T^{MF}$ by (41). The process $(\delta_t)_{t\in[0,T]}$ may be interpreted as the arbitrage-free price of the risk tolerance "claim" δ_T for this *representative* player. Let also $\widetilde{\mathbb{Q}}$ be defined on $\mathcal{F}_T^{MF}$ by

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{Q}} = \delta_T,$$

and consider the martingale $M_t = E_{\widetilde{\mathbb{Q}}}\left[\frac{1}{2}\int_0^T (\lambda_s - \xi_s)^2\, ds\middle|\mathcal{F}_t^{MF}\right]$ and $(\eta_t)_{t\in[0,T]}$ to be such that

$$dM_t = \eta_t\, dW_t^{\widetilde{\mathbb{Q}}}, \tag{60}$$

with $W_t^{\widetilde{\mathbb{Q}}} = W_t + \int_0^t (\lambda_s - \xi_s)\, ds$. The processes δ, ξ and η are all $\mathcal{F}_t^{MF}$-adapted.

We now state the main result of this section.

Proposition 7 *If* $E_{\mathbb{P}}[c] < 1$, *there exists a MFG equilibrium* $(\pi_t^*)_{t\in[0,T]}$, *given by*

$$\pi_t^* = \frac{c}{1-E_{\mathbb{P}}[c]}\frac{1}{\sigma_t}\left(\lambda_t E_{\mathbb{Q}}[\delta_T|\mathcal{F}_t] - E_{\mathbb{Q}}[\delta_T(\xi_t+\eta_t)|\mathcal{F}_t] + E_{\mathbb{P}}[X_t^*\xi_t|\mathcal{F}_t]\right.$$
$$\left. - E_{\mathbb{P}}[c\xi_t|\mathcal{F}_t]E_{\mathbb{P}}[X_t^*|\mathcal{F}_t]\right) + \frac{1}{\sigma_t}\left(\delta_t(\lambda_t - \xi_t - \eta_t) + (X_t^* - cE_{\mathbb{P}}[X_t^*|\mathcal{F}_t])\xi_t\right), \quad (61)$$

with δ, ξ *and* η *as in* (59) *and* (60)*, and* $(X_t^*)_{t\in[0,T]}$ *being the associated optimal wealth process, solving*

$$dX_t^* = \pi_t^*(\mu_t\,dt + \sigma_t\,dW_t). \quad (62)$$

The value of the MFG is given by

$$V(x) = -\exp\left(-\frac{1}{E_{\mathbb{Q}}[\delta_T|\mathcal{F}_0^{MF}]}(x-cm) - E_{\widetilde{\mathbb{Q}}}\left[\frac{1}{2}\int_0^T(\lambda_s-\xi_s)^2\,ds\middle|\mathcal{F}_0^{MF}\right]\right),$$

with $m = E_{\mathbb{P}}[x]$.

For the proof, we will need the following lemma.

Lemma 2 *If* X *is a* $\mathcal{F}_s^{MF}$*-measurable integrable random variable, then* $E_{\mathbb{P}}[X|\mathcal{F}_t] = E_{\mathbb{P}}[X|\mathcal{F}_s]$*, for* $s \in [0,t]$.

Proof Let $\mathcal{P} := \{A = C\cap D : C \in \mathcal{F}_s,\ D \in \sigma\{W_u - W_s, s \le u \le t\}\}$ and $\mathcal{L} = \{A \in \mathcal{F} : E_{\mathbb{P}}[X\mathbf{1}_A] = E_{\mathbb{P}}[E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_A]\}$. Then, the following assertions hold:

(1) $\mathcal{P}$ is a π-system since both $\mathcal{F}_s$ and $\sigma\{W_u - W_s, s\le u\le t\}$ are σ-algebras and closed under intersection. Also $\mathcal{F}_s \subseteq \mathcal{P}$ and $\sigma\{W_u - W_s, s \le u \le t\} \subseteq \mathcal{P}$ by taking $D = \Omega$ and $C = \Omega$.

(2) $\mathcal{P} \subseteq \mathcal{L}$. For any $A \in \mathcal{P}$, $A = C\cap D$ with $C \in \mathcal{F}_s,\ D \in \sigma\{W_u - W_s, s\le u\le t\}$, it holds that

$$E_{\mathbb{P}}[E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_A] = E_{\mathbb{P}}[E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_C\mathbf{1}_D] = E_{\mathbb{P}}[E_{\mathbb{P}}[X\mathbf{1}_C|\mathcal{F}_s]\mathbf{1}_D] = E_{\mathbb{P}}[X\mathbf{1}_C]E_{\mathbb{P}}[\mathbf{1}_D],$$

where we have consecutively used that $C \perp D$, the metastability of $\mathbf{1}_C$, and the independence between $\mathbf{1}_D$ and $\mathcal{F}_s$.

Furthermore, by the independence between $\mathbf{1}_D$ and $\mathcal{F}_s^{MF} = \mathcal{F}_t \vee \sigma(\theta)$, we deduce

$$E_{\mathbb{P}}[X\mathbf{1}_A] = E_{\mathbb{P}}[X\mathbf{1}_C\mathbf{1}_D] = E_{\mathbb{P}}[X\mathbf{1}_C]E_{\mathbb{P}}[\mathbf{1}_D],$$

and conclude that $A \in \mathcal{L}$. Therefore $\mathcal{P} \subseteq \mathcal{L}$.

(3) $\mathcal{L}$ is a λ-system. It is obvious that $\Omega \in \mathcal{L}$ and $A \in \mathcal{L}$ imply that $A^c \in \mathcal{L}$. For a sequence of disjoint sets $A_1, A_2, \ldots$ in $\mathcal{L}$, one has $\left|X\mathbf{1}_{\cup_{i=1}^\infty A_i}\right| \le |X|$ and, thus, by the dominated convergence theorem, we deduce that

$$E_{\mathbb{P}}[X\mathbf{1}_{\cup_{i=1}^\infty A_i}] = \sum_{i=1}^{\infty} E_{\mathbb{P}}[X\mathbf{1}_{A_i}]. \quad (63)$$

Similarly, by the inequalities $\|E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_{\cup_{i=1}^\infty A_i}\|_1 \le \|E_{\mathbb{P}}[X|\mathcal{F}_s]\|_1 \le \|X\|_1$, we have

$$E_{\mathbb{P}}[E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_{\cup_{i=1}^{\infty}A_i}] = \sum_{i=1}^{\infty} E_{\mathbb{P}}[E_{\mathbb{P}}[X|\mathcal{F}_s]\mathbf{1}_{A_i}]. \tag{64}$$

Since $A_i \in \mathcal{L}$, $\forall i$, the right-hand-sides of (63) and (64) are equal, which implies $\cup_{i=1}^{\infty} A_i \in \mathcal{L}$.

Therefore, by the π-λ theorem, we obtain that $\mathcal{F}_t = \sigma(\mathcal{F}_s \cup \sigma\{W_u - W_s, s \le u \le t\}) \subseteq \sigma(\mathcal{P}) \subseteq \mathcal{L}$. Noticing that $E_{\mathbb{P}}[X|\mathcal{F}_s]$ is $\mathcal{F}_t$-measurable by definition, we have that $E_{\mathbb{P}}[X|\mathcal{F}_t] = E_{\mathbb{P}}[X|\mathcal{F}_s]$. □

Proof (Proposition 7) Let $(X_t^{\alpha})_{t\in[0,T]}$ be given by $X_t^{\alpha} = x + \int_0^t \mu_s \alpha_s \, ds + \int_0^t \sigma_s \alpha_s \, dW_s$ for an admissible policy α_t ($\mathcal{F}_t^{MF}$-adapted) and define $\overline{X}_t := E_{\mathbb{P}}[X_t^{\alpha}|\mathcal{F}_t]$. Then,

$$\overline{X}_t = m + E_{\mathbb{P}}\left[\int_0^t \mu_s \alpha_s \, ds \Big| \mathcal{F}_s\right] + E_{\mathbb{P}}\left[\int_0^t \sigma_s \alpha_s \, dW_s \Big| \mathcal{F}_s\right].$$

Using Lemma 2, the adaptivity of μ_t, σ_t with respect to $\mathcal{F}_t$, and the definition of Itô integral, we rewrite the above as

$$\overline{X}_t = m + \int_0^t \mu_s E_{\mathbb{P}}[\alpha_s|\mathcal{F}_s] \, ds + \int_0^t \sigma_s E_{\mathbb{P}}[\alpha_s|\mathcal{F}_s] \, dW_s.$$

Direct arguments yield that the optimization problem (58) reduces to

$$V(\tilde{x}) = \sup_{\tilde{\pi}\in\mathcal{A}^{MF}} E_{\mathbb{P}}\left[-\exp\left(-\frac{1}{\delta_T}\widetilde{X}_T\right)\Big|\mathcal{F}_0^{MF}, \widetilde{X}_0 = \tilde{x}\right],$$

where $(\widetilde{X}_t)_{t\in[0,T]}$ solves

$$d\widetilde{X}_t \equiv d(X_t - c\overline{X}_t) = \widetilde{\pi}_t(\mu_t \, dt + \sigma_t \, dW_t), \tag{65}$$

with $\widetilde{X}_0 = \tilde{x} = x - cm$ and $\widetilde{\pi}_t := \pi_t - cE_{\mathbb{P}}[\alpha_t|\mathcal{F}_t]$. Then, (49) yields

$$\widetilde{\pi}_t^* = \delta_t \frac{\lambda_t - \eta_t - \xi_t}{\sigma_t} + \frac{\xi_t}{\sigma_t}\widetilde{X}_t^*, \tag{66}$$

with δ_t, ξ_t, η_t given in (59) and (60), and $(\widetilde{X}_t^*)_{t\in[0,T]}$ solving (65) with $\widetilde{\pi}^*$ being used. On the other hand, using that $\widetilde{\pi}_t^* = \pi_t^* - cE_{\mathbb{P}}[\alpha_t|\mathcal{F}_t]$, we obtain

$$\pi_t^* - cE_{\mathbb{P}}[\alpha_t|\mathcal{F}_t] = \delta_t \frac{\lambda_t - \eta_t - \xi_t}{\sigma_t} + \frac{\xi_t}{\sigma_t}\widetilde{X}_t^*.$$

In turn, using that, at equilibrium, $\alpha = \pi^*$, we get

$$(1 - E_{\mathbb{P}}[c])E_{\mathbb{P}}[\pi_t^*|\mathcal{F}_t] = \frac{1}{\sigma_t}\left(\lambda_t E_{\mathbb{P}}[\delta_t|\mathcal{F}_t] - E_{\mathbb{P}}[\delta_t(\xi_t + \eta_t)|\mathcal{F}_t] + E_{\mathbb{P}}[\widetilde{X}_t^*\xi_t|\mathcal{F}_t]\right).$$

Further calculations give

$$\pi_t^* = c\frac{1}{1-E_{\mathbb{P}}[c]}\frac{1}{\sigma_t}\left(\lambda_t E_{\mathbb{P}}[\delta_t|\mathcal{F}_t] - E_{\mathbb{P}}[\delta_t(\xi_t+\eta_t)|\mathcal{F}_t] + E_{\mathbb{P}}[X_t^*\xi_t|\mathcal{F}_t]\right.$$
$$\left.-E_{\mathbb{P}}[X_t^*|\mathcal{F}_t]E_{\mathbb{P}}[c\xi_t|\mathcal{F}_t]\right) + \frac{\delta_t(\lambda_t-\eta_t) - \delta_t\xi_t + X_t^*\xi_t - c\xi_t E_{\mathbb{P}}[X_t^*|\mathcal{F}_t]}{\sigma_t}. \quad (67)$$

Finally, we obtain

$$E_{\mathbb{P}}[\delta_t|\mathcal{F}_t] = E_{\mathbb{P}}[E_{\mathbb{Q}}[\delta_T|\mathcal{F}_t^{MF}]|\mathcal{F}_t] = E_{\mathbb{P}}\left[E_{\mathbb{P}}\left[\frac{\delta_T Z_T}{Z_t}\Big|\mathcal{F}_t^{MF}\right]\Big|\mathcal{F}_t\right]$$
$$= E_{\mathbb{P}}\left[\frac{\delta_T Z_T}{Z_t}\Big|\mathcal{F}_t\right] = E_{\mathbb{Q}}[\delta_T|\mathcal{F}_t],$$

and a similar derivation for $E_{\mathbb{P}}[\delta_t(\xi_t+\eta_t)|\mathcal{F}_t]$. We conclude by checking the admissibility of π^* which follows from model assumptions, the form of π^*, and equation (62). □

4 Conclusions and future research directions

In Itô-diffusion environments, we introduced and studied a family of N-player and common-noise mean-field games in the context of optimal portfolio choice in a common market. The players aim to maximize their expected terminal utility, which depends on their own wealth and the wealth of their peers.

We focused on two cases of exponential utilities, specifically, the classical CARA case and the extended CARA case with random risk tolerance. The former was considered for the incomplete market model while the latter for the complete one. We provided the equilibrium processes and the values of the games in explicit (incomplete market case) and in closed form (complete market case). We note that in the case of random risk tolerances, for which even the single-player case is interesting in its own right, the optimal strategy process depends on the state process, even if the preferences are of exponential type.

A natural extension is to consider power utilities (CRRA), which are also commonly used in models of portfolio choice. This extension, however, is by no means straightforward. Firstly, in the incomplete market case, the underlying measure depends on the individual risk tolerance, which is not the case for the CARA utilities considered herein (see (7) for the minimal martingale measure and (22)-(23) for the minimal entropy measure, respectively). Secondly, while it is formally clear how to formulate the random risk tolerance case for power utilities, its solution is far from obvious. The authors are working in both these directions.

Our results may be used to study such models when the dynamics of the common market and/or the individual preferences are not entirely known. This could extend the analysis to various problems in reinforcement learning (see, for example, the recent work [14] in a static setting). It is expected that results similar to the ones in

[19] could be derived and, in turn, used to build suitable algorithms (see, also, [7] for a Markovian case).

Acknowledgements RH was partially supported by the NSF grant DMS-1953035, and the Faculty Career Development Award and the Research Assistant Program Award at UCSB.

This work was presented at the SIAM Conference on Financial Mathematics and Engineering in 2021. The authors would like to thank the participants for fruitful comments and suggestions.

References

1. Suleyman Basak and Dmitry Makarov. Competition among portfolio managers and asset specialization. *Available at SSRN 1563567*, 2015.
2. René Carmona. *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*. SIAM, 2016.
3. George Chacko and Luis M Viceira. Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies*, 18(4):1369–1402, 2005.
4. Gilles-Edouard Espinosa and Nizar Touzi. Optimal investment under relative performance concerns. *Mathematical Finance*, 25(2):221–257, 2015.
5. Hans Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis*, 5(389-414):19–31, 1991.
6. Guanxing Fu, Xizhi Su, and Chao Zhou. Mean field exponential utility game: A probabilistic approach. *arXiv preprint arXiv:2006.07684*, 2020.
7. Xin Guo, Renyuan Xu, and Thaleia Zariphopoulou. Entropy regularization for mean field games with learning. *arXiv preprint arXiv:2010.00145*, 2020.
8. Minyi Huang, Roland P Malhamé, and Peter E Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems*, 6(3):221–252, 2006.
9. Minyi Huang and Son Luu Nguyen. Mean field games for stochastic growth with relative utility. *Applied Mathematics & Optimization*, 74(3):643–668, 2016.
10. Holger Kraft, André Meyer-Wehmann, and Frank Thomas Seifried. Dynamic asset allocation with relative wealth concerns in incomplete markets. *Journal of Economic Dynamics and Control*, 113:103857, 2020.
11. Daniel Lacker and Agathe Soret. Many-player games of optimal consumption and investment under relative performance criteria. *Mathematics and Financial Economics*, 14(2):263–281, 2020.
12. Daniel Lacker and Thaleia Zariphopoulou. Mean field and n-agent games for optimal investment under relative performance criteria. *Mathematical Finance*, 29(4):1003–1038, 2019.
13. Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
14. Yan Leng, Xiaowen Dong, Junfeng Wu, and Alex Pentland. Learning quadratic games on networks. In *International Conference on Machine Learning*, pages 5820–5830. PMLR, 2020.
15. Marek Musiela and Thaleia Zariphopoulou. A note on the term structure of risk aversion in utility-based pricing systems. *Technical report*, 2002.
16. Nathanael David Ringer. *Three essays on valuation and investment in incomplete markets*. PhD thesis, The University of Texas at Austin, 2011.
17. Richard Rouge and Nicole El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.
18. Ronnie Sircar and Thaleia Zariphopoulou. Bounds and asymptotic approximations for utility prices when volatility is random. *SIAM journal on control and optimization*, 43(4):1328–1353, 2005.

19. Haoran Wang, Thaleia Zariphopoulou, and Xun Yu Zhou. Reinforcement learning in continuous time and space: A stochastic control approach. *Journal of Machine Learning Research*, 21(198):1–34, 2020.
20. Mark Whitmeyer. Relative performance concerns among investment managers. *Annals of Finance*, 15(2):205–231, 2019.

A Variational Characterization of Langevin–Smoluchowski Diffusions

Ioannis Karatzas and Bertram Tschiderer

Abstract We show that Langevin–Smoluchowski measure on path space is invariant under time-reversal, followed by stochastic control of the drift with a novel entropic-type criterion. Repeated application of these forward-backward steps leads to a sequence of stochastic control problems, whose initial/terminal distributions converge to the Gibbs probability measure of the diffusion, and whose values decrease to zero along the relative entropy of the Langevin–Smoluchowski flow with respect to this Gibbs measure.

Key words: Langevin–Smoluchowski diffusion, relative entropy, Gibbs measure, time reversal, stochastic control, alternating forward-backward dynamics

MSC 2010 subject classifications: Primary 93E20, 94A17; secondary 35Q84, 60G44, 60J60

1 Introduction

Diffusions of Langevin–Smoluchowski type have some important properties. They possess invariant (Gibbs) probability measures described very directly in terms of their potentials and towards which, under appropriate conditions, their time-marginals converge as time increases to infinity and in a manner that conforms to the second law of thermodynamics: the relative entropy of the current distribution, with respect to the invariant one, decreases to zero. The seminal paper [24] revealed another remarkable, local aspect of this decrease towards equilibrium: the family

Ioannis Karatzas
Department of Mathematics, Columbia University, 2990 Broadway, New York, NY 10027, USA
e-mail: ik1@columbia.edu

Bertram Tschiderer
Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria
e-mail: bertram.tschiderer@univie.ac.at

The original version of this chapter was revised: The author affiliation has now been amended. The correction to this chapter is available at https://doi.org/10.1007/978-3-030-98519-6_19.

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_10

of time-marginals is, at (almost) every point in time, a curve of steepest descent among all probability density functions with finite second moment, when distances are measured according to the Wasserstein metric in configuration space.

We establish in this paper yet another variational property of such diffusions, this time a global one: their law is invariant under the combined effects of time-reversal, and of stochastic control of the drift under a novel, entropic-type criterion. Here, one minimizes over admissible controls the relative entropy of the "terminal" state with respect to the invariant measure, plus an additional term thought of as "entropic cost of time-reversal": the difference in relative entropy with respect to the Langevin–Smoluchowski measure on path space, computed based on the "terminal" state as opposed to on the entire path. Quite similar, but different, cost criteria have been considered in [12, 13, 16, 38, 39].

The setting under consideration bears similarities to the celebrated Schrödinger bridge problem, but also considerable differences. Both problems are posed on a fixed time horizon of finite length, and both involve the relative entropy with respect to the invariant measure. But here this entropy is modified by the addition of the above-mentioned entropic cost of time-reversal, and there is no fixed, target distribution on the terminal state. Yet the trajectory that emerges as the solution of the stochastic control problem has time-marginals that replicate exactly those of the original Langevin–Smoluchowski flow, whence the "invariance" property mentioned in the abstract.

We refer to [7, 8, 9, 30] for overviews on the classical Schrödinger bridge problem, to [44] for the related semimartingale transport problem, and to the recent paper [3] for a detailed study of the mean-field Schrödinger problem. A related controllability problem for a Fokker–Planck equation and its connection to Schrödinger systems and stochastic optimal control, is considered in [6]. More information about the Schrödinger equation, diffusion theory, and time reversal, can be found in the book [34].

1.1 Preview

In Section 2 we introduce the Langevin–Smoluchowski measure $\mathbb{P}$ on path space, under which the canonical process $(X(t))_{t \geqslant 0}$ has dynamics (3) with initial distribution $P(0)$. Then, in Section 3, this process is studied under time-reversal. That is, we fix a terminal time $T \in (0, \infty)$ and consider the time-reversed process $\overline{X}(s) = X(T-s)$, $0 \leqslant s \leqslant T$. Standard time-reversal theory shows that $\overline{X}$ is again a diffusion, and gives an explicit description of its dynamics.

Section 4 develops our main result, Theorem 1. An equivalent change of probability measure $\mathbb{P}^\gamma \sim \mathbb{P}$ adds to the drift of $\overline{X}$ a measurable, adapted process $\boldsymbol{\gamma}(T-s)$, $0 \leqslant s \leqslant T$. In broad brushes, this allows us to define, in terms of relative entropies, the quantities

$$H^\gamma := H(\mathbb{P}^\gamma \,|\, \mathbb{Q})\big|_{\sigma(\overline{X}(T))}, \qquad D^\gamma := H(\mathbb{P}^\gamma \,|\, \mathbb{P})\big|_{\sigma(\overline{X})} - H(\mathbb{P}^\gamma \,|\, \mathbb{P})\big|_{\sigma(\overline{X}(T))}. \quad (1)$$

Here, $\mathbb{Q}$ is the probability measure on path space, inherited from the Langevin–Smoluchowski dynamics (3) with initial distribution given by the invariant Gibbs probability measure Q. Theorem 1 establishes then the variational characterization

$$\inf_{\gamma} \left(H^{\gamma} + D^{\gamma} \right) = H\big(P(T) \,|\, \mathrm{Q}\big), \tag{2}$$

where $P(T)$ denotes the distribution of the random variable $X(T)$ under $\mathbb{P}$. The process γ_* that realizes the infimum in (2) gives rise to a probability measure $\mathbb{P}^{\gamma_*}$, under which the time-reversed diffusion $\overline{X}$ is of Langevin–Smoluchowski type in its own right, but now with initial distribution $P(T)$. In other words, with the constraint of minimizing the sum of the entropic quantities H^{γ} and D^{γ} of (1), Langevin–Smoluchowski measure on path space is invariant under time-reversal.

Sections 5 – 7 develop ramifications of the main result, including the following consistency property: starting with the time-reversal $\overline{X}$ of the Langevin–Smoluchowski diffusion, the solution of a related optimization problem, whose value is now $H(P(2T) \,|\, \mathrm{Q})$, leads to the original forward Langevin–Smoluchowski dynamics, but now with initial distribution $P(2T)$. Iterating these procedures we obtain an alternating sequence of forward-backward Langevin–Smoluchowski dynamics with initial distributions $(P(kT))_{k \in \mathbb{N}_0}$ converging to Q in total variation, along which the values of the corresponding optimization problems as in (2) are given by $(H(P(kT) \,|\, \mathrm{Q}))_{k \in \mathbb{N}}$ and decrease to zero.

2 The setting

Let us consider a Langevin–Smoluchowski diffusion process $(X(t))_{t \geqslant 0}$ of the form

$$\mathrm{d}X(t) = -\nabla\Psi\big(X(t)\big)\,\mathrm{d}t + \mathrm{d}W(t), \tag{3}$$

with values in $\mathbb{R}^n$. Here $(W(t))_{t \geqslant 0}$ is standard n-dimensional Brownian motion, and the "potential" $\Psi \colon \mathbb{R}^n \to [0, \infty)$ is a C^∞-function growing, along with its derivatives of all orders, at most exponentially as $|x| \to \infty$; we stress that no convexity assumptions are imposed on this potential. We posit also an "initial condition" $X(0) = \Xi$, a random variable independent of the driving Brownian motion and with given distribution $P(0)$. For concreteness, we shall assume that this initial distribution has a continuous probability density function $p_0(\,\cdot\,)$.

Under these conditions, the Langevin–Smoluchowski equation (3) admits a pathwise unique, strong solution, up until an "explosion time" $\mathfrak{e}$; such explosion never happens, i.e., $\mathbb{P}(\mathfrak{e} = \infty) = 1$, if in addition the second-moment condition (12) and the coercivity condition (11) below hold. The condition (11) propagates the finiteness of the second moment to the entire collection of time-marginal distributions $P(t) = \mathrm{Law}(X(t))$, $t \geqslant 0$, which are then determined uniquely. In fact, adapting the arguments in [42] to the present situation, we check that each time-marginal distribution $P(t)$ has probability density $p(t, \cdot)$ such that the resulting function

$(t,x) \mapsto p(t,x)$ is continuous and strictly positive on $(0,\infty)\times\mathbb{R}^n$; differentiable with respect to the temporal variable t for each $x \in \mathbb{R}^n$; smooth in the spatial variable x for each $t > 0$; and such that the logarithmic derivative $(t,x) \mapsto \nabla \log p(t,x)$ is continuous on $(0,\infty)\times\mathbb{R}^n$. These arguments also lead to the *Fokker–Planck* [20, 21, 41, 43], or *forward Kolmogorov* [28], equation

$$\partial p(t,x) = \tfrac{1}{2}\Delta p(t,x) + \operatorname{div}\big(\nabla\Psi(x)\,p(t,x)\big), \qquad (t,x) \in (0,\infty)\times\mathbb{R}^n \tag{4}$$

with initial condition $p(0,x) = p_0(x)$, for $x \in \mathbb{R}^n$.

Here and throughout this paper, ∂ denotes differentiation with respect to the temporal argument; whereas ∇, Δ and div stand, respectively, for gradient, Laplacian and divergence with respect to the spatial argument.

2.1 Invariant measure, likelihood ratio, and relative entropy

We introduce now the function

$$q(x) := \mathrm{e}^{-2\Psi(x)}, \qquad x \in \mathbb{R}^n \tag{5}$$

and note that it satisfies the stationary version

$$\tfrac{1}{2}\Delta q(x) + \operatorname{div}\big(\nabla\Psi(x)\,q(x)\big) = 0, \qquad x \in \mathbb{R}^n \tag{6}$$

of the forward Kolmogorov equation (4). We introduce also the σ-finite measure Q on the Borel subsets $\mathscr{B}(\mathbb{R}^n)$ of $\mathbb{R}^n$, which has density q as in (5) with respect to n-dimensional Lebesgue measure. This measure Q is invariant for the diffusion of (3); see Exercise 5.6.18 in [27]. When finite, Q can be normalized to an invariant probability measure for (3), to which the time-marginals $P(t)$ "converge" as $t \to \infty$; more about this convergence can be found in Section 7. We shall always assume tacitly that such a normalization has taken place when Q is finite, i.e., when

$$\mathrm{Q}(\mathbb{R}^n) = \int_{\mathbb{R}^n} q(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \mathrm{e}^{-2\Psi(x)}\,\mathrm{d}x < \infty. \tag{7}$$

One way to think of the above-mentioned convergence, is by considering the likelihood ratio

$$\ell(t,x) := \frac{p(t,x)}{q(x)}, \qquad (t,x) \in (0,\infty)\times\mathbb{R}^n. \tag{8}$$

It follows from (4), (6) that this function satisfies the *backward* Kolmogorov equation

$$\partial \ell(t,x) = \tfrac{1}{2}\Delta \ell(t,x) - \big\langle \nabla \ell(t,x), \nabla\Psi(x)\big\rangle, \qquad (t,x) \in (0,\infty)\times\mathbb{R}^n. \tag{9}$$

In terms of the likelihood ratio function (8), let us consider now for each $t \geqslant 0$ the relative entropy

$$H(P(t)\,|\,\mathrm{Q}) := \mathbb{E}_{\mathbb{P}}\big[\log \ell(t,X(t))\big] = \int_{\mathbb{R}^n} \log\Big(\frac{p(t,x)}{q(x)}\Big)\, p(t,x)\,\mathrm{d}x \tag{10}$$

of the probability distribution $P(t)$ with respect to the invariant measure Q. The expectation in (10) is well-defined in $[0,\infty]$, if Q is a probability measure. As we are not imposing this as a blanket assumption, we shall rely on [26, Appendix C], where it is shown that the relative entropy $H(P(t)\,|\,\mathrm{Q})$ is well-defined and takes values in $(-\infty,\infty]$, whenever $P(t)$ belongs to the space $\mathscr{P}_2(\mathbb{R}^n)$ of probability measures with finite second moment (see also [31] for a more general discussion). This, in turn, is the case whenever $P(0)$ has finite second moment, and the coercivity condition

$$\forall\, x \in \mathbb{R}^n, |x| \geqslant R: \qquad \langle x, \nabla\Psi(x)\rangle \geqslant -c\,|x|^2 \tag{11}$$

is satisfied by the potential Ψ in (3), for some real constants $c \geqslant 0$ and $R \geqslant 0$; see the first problem on page 125 of [20], or Appendix B in [26]. The prototypical such potential is $\Psi(x) = \frac{1}{2}|x|^2$, leading to Ornstein–Uhlenbeck dynamics in (3); but $\Psi \equiv 0$ and the "double well" $\Psi(x) = (x^2 - a^2)^2$ for $a > 0$, are also allowed. In particular, the coercivity condition (11) does not imply that the potential Ψ is convex.

We shall impose throughout Sections 2 – 6 *the coercivity condition* (11)*, as well as the finite second-moment condition*

$$\int_{\mathbb{R}^n} |x|^2\, p_0(x)\,\mathrm{d}x < \infty. \tag{12}$$

This amounts to $P(0) \in \mathscr{P}_2(\mathbb{R}^n)$, as has been already alluded to. In Section 7 we will see that these two conditions (11) and (12) are not needed when Q is a probability measure.

However, *we shall impose throughout the entire paper the crucial assumption that the initial relative entropy is finite*, i.e.,

$$H(P(0)\,|\,\mathrm{Q}) = \int_{\mathbb{R}^n} \log\Big(\frac{p_0(x)}{q(x)}\Big)\, p_0(x)\,\mathrm{d}x < \infty. \tag{13}$$

Under either the conditions "(11) + (12)", or the condition (7), the decrease of the relative entropy[1] function $[0,\infty) \ni t \mapsto H(P(t)\,|\,\mathrm{Q}) \in (-\infty,\infty]$ implies then that the quantity $H(P(t)\,|\,\mathrm{Q})$ in (10) is finite for all $t \geqslant 0$ whenever (13) holds.

In fact, under the conditions (11) – (13), the rate of decrease for the relative entropy, measured with respect to distances traveled in $\mathscr{P}_2(\mathbb{R}^n)$ in terms of the quadratic Wasserstein metric

$$W_2(\mu,\nu) = \Big(\inf_{Y\sim\mu,\,Z\sim\nu} \mathbb{E}|Y-Z|^2\Big)^{1/2}, \qquad \mu,\nu \in \mathscr{P}_2(\mathbb{R}^n)$$

[1] A classical aspect of thermodynamics; for a proof of this fact under the conditions "(11) + (12)" and without assuming finiteness of Q, see Theorem 3.1 in [25]; when Q is a probability measure, we refer to Appendix 1.

(cf. [1, 2, 45]) is, at Lebesgue-almost all times $t_0 \in [0,\infty)$, the *steepest possible* along the Langevin–Smoluchowski curve $(P(t))_{t \geqslant 0}$ of probability measures. Here, we are comparing the curve $(P(t))_{t \geqslant 0}$ against all such curves $(P^{\beta}(t))_{t \geqslant t_0}$ of probability measures generated as in (3) — but with an additional drift ∇B for suitable (smooth and compactly supported) perturbations B of the potential Ψ in (3). This local optimality property of Langevin–Smoluchowski diffusions is due to [24]; it was established by [25] in the form just described. We develop in this paper yet another, global this time, optimality property for such diffusions.

2.2 The probabilistic setting

In (10) and throughout this paper, we are denoting by $\mathbb{P}$ the unique probability measure on the space $\Omega = C([0,\infty);\mathbb{R}^n)$ of continuous, $\mathbb{R}^n$-valued functions, under which the canonical coordinate process $X(t,\omega) = \omega(t)$, $t \geqslant 0$ has the property that

$$W(t) := X(t) - X(0) + \int_0^t \nabla\Psi(X(\theta))\,\mathrm{d}\theta, \qquad t \geqslant 0 \tag{14}$$

is standard $\mathbb{R}^n$-valued Brownian motion, and independent of the random variable $X(0)$ with distribution

$$\mathbb{P}[X(0) \in A] = \int_A p_0(x)\,\mathrm{d}x, \qquad A \in \mathscr{B}(\mathbb{R}^n).$$

The $\mathbb{P}$-Brownian motion $(W(t))_{t \geqslant 0}$ of (14) is adapted to, in fact generates, the canonical filtration $\mathbb{F} = (\mathcal{F}(t))_{t \geqslant 0}$ with

$$\mathcal{F}(t) := \sigma\big(X(s) \colon\ 0 \leqslant s \leqslant t\big). \tag{15}$$

By analogy to the terminology "Wiener measure", we call $\mathbb{P}$ the "Langevin–Smoluchowski measure" associated with the potential Ψ.

3 Reversal of time

The densities $p(t,\cdot)$ and $q(\cdot)$ satisfy the forward Kolmogorov equations (4) and (6), respectively. Whereas, their likelihood ratio $\ell(t,\cdot)$ in (8) satisfies the *backward* Kolmogorov equation (9). This suggests that, in the study of relative entropy and of its temporal dissipation, it might make sense to look at the underlying Langevin–Smoluchowski diffusion under time-reversal. Such an approach proved very fruitful in [12], [38], [19] and [25]; it will be important in our context here as well.

Thus, we fix an arbitrary terminal time $T \in (0,\infty)$ and consider for $0 \leqslant s \leqslant T$ the time-reversed process

$$\overline{X}(s) := X(T-s), \qquad \overline{\mathcal{G}}(s) := \sigma\big(\overline{X}(u):\ 0 \leqslant u \leqslant s\big); \tag{16}$$

along with the filtration $\overline{\mathbb{G}} = (\overline{\mathcal{G}}(s))_{0 \leqslant s \leqslant T}$ this process generates. Then standard theory on time-reversal shows that the process $\overline{W} = (\overline{W}(s))_{0 \leqslant s \leqslant T}$, with

$$\overline{W}(s) := \overline{X}(s) - \overline{X}(0) + \int_0^s \Big(\nabla\Psi(\overline{X}(u)) - \nabla L(T-u, \overline{X}(u))\Big)\,\mathrm{d}u, \tag{17}$$

is a $(\overline{\mathbb{G}}, \mathbb{P})$-standard Brownian motion with values in $\mathbb{R}^n$ and independent of the random variable $\overline{X}(0) = X(T)$ (see, for instance, [17, 18, 22, 33, 35, 37] for the classical results; an extensive presentation of the relevant facts regarding the time reversal of diffusion processes can be found in Appendix G of [26]). Here

$$L(t,x) := \log \ell(t,x), \qquad (t,x) \in (0,\infty) \times \mathbb{R}^n$$

is the logarithm of the likelihood ratio function in (8); and on the strength of (9), this function solves the semilinear Schrödinger-type equation

$$\partial L(t,x) = \tfrac{1}{2}\Delta L(t,x) - \langle \nabla L(t,x), \nabla\Psi(x)\rangle + \tfrac{1}{2}|\nabla L(t,x)|^2. \tag{18}$$

Another way to express this, is by saying that the so-called *Hopf–Cole transform* $\ell = \mathrm{e}^L$ turns the semilinear equation (18), into the linear backward Kolmogorov equation (9). This observation is not new; it has been used in stochastic control to good effect by Fleming [14, 15], Holland [23], and in a context closer in spirit to this paper by Dai Pra and Pavon [13], Dai Pra [11].

4 A stochastic control problem

Yet another way to cast the equation (18), is in the *Hamilton–Jacobi–Bellman* form

$$\partial L(t,x) = \tfrac{1}{2}\Delta L(t,x) - \langle \nabla L(t,x), \nabla\Psi(x)\rangle - \min_{g \in \mathbb{R}^n} \Big(\langle \nabla L(t,x), g\rangle + \tfrac{1}{2}|g|^2\Big), \tag{19}$$

where the minimization is attained by the gradient $g_* = -\nabla L(t,x)$. This, in turn, suggests a *stochastic control problem* related to the backwards diffusive dynamics

$$\mathrm{d}\overline{X}(s) = \Big(\nabla L(T-s, \overline{X}(s)) - \nabla\Psi(\overline{X}(s))\Big)\,\mathrm{d}s + \mathrm{d}\overline{W}(s) \tag{20}$$

of (17), which we formulate now as follows.

For any measurable process $[0,T] \times \Omega \ni (t,\omega) \mapsto \boldsymbol{\gamma}(t,\omega) \in \mathbb{R}^n$ such that the time-reversed process $\boldsymbol{\gamma}(T-s)$, $0 \leqslant s \leqslant T$ is adapted to the backward filtration $\overline{\mathbb{G}}$ of (16), and which satisfies the condition

$$\mathbb{P}\Big[\int_0^T |\boldsymbol{\gamma}(T-s)|^2\,\mathrm{d}s < \infty\Big] = 1, \tag{21}$$

we consider the exponential $(\overline{\mathbb{G}}, \mathbb{P})$-local martingale

$$Z^{\gamma}(s) := \exp\Big(\int_0^s \langle \boldsymbol{\gamma}(T-u), \mathrm{d}\overline{W}(u) \rangle - \tfrac{1}{2} \int_0^s |\boldsymbol{\gamma}(T-u)|^2 \, \mathrm{d}u \Big) \tag{22}$$

for $0 \leqslant s \leqslant T$. We denote by $\boldsymbol{\Gamma}$ the collection of all processes $\boldsymbol{\gamma}$ as above, for which Z^{γ} is a true $(\overline{\mathbb{G}}, \mathbb{P})$-martingale. This collection is not empty: it contains all such uniformly bounded processes $\boldsymbol{\gamma}$, and quite a few more (e.g., conditions of Novikov [27, Corollary 3.5.13] and Kazamaki [40, Proposition VIII.1.14]).

Now, for every $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$, we introduce an equivalent probability measure $\mathbb{P}^{\gamma} \sim \mathbb{P}$ on path space, via

$$\frac{\mathrm{d}\mathbb{P}^{\gamma}}{\mathrm{d}\mathbb{P}}\bigg|_{\overline{\mathcal{G}}(s)} = Z^{\gamma}(s), \qquad 0 \leqslant s \leqslant T. \tag{23}$$

Then, by the Girsanov theorem [27, Theorem 3.5.1], the process

$$\overline{W}^{\gamma}(s) := \overline{W}(s) - \int_0^s \boldsymbol{\gamma}(T-u) \, \mathrm{d}u, \qquad 0 \leqslant s \leqslant T \tag{24}$$

is standard $\mathbb{R}^n$-valued $\mathbb{P}^{\gamma}$-Brownian motion of the filtration $\overline{\mathbb{G}}$, thus independent of the random variable $\overline{X}(0) = X(T)$. Under the probability measure $\mathbb{P}^{\gamma}$, the backwards dynamics of (20) take the form

$$\mathrm{d}\overline{X}(s) = \Big(\nabla L(T-s, \overline{X}(s)) + \boldsymbol{\gamma}(T-s) - \nabla \Psi\big(\overline{X}(s)\big) \Big) \, \mathrm{d}s + \mathrm{d}\overline{W}^{\gamma}(s); \tag{25}$$

and it follows readily from these dynamics and the semilinear parabolic equation (18), that the process

$$M^{\gamma}(s) := L(T-s, \overline{X}(s)) + \tfrac{1}{2} \int_0^s |\boldsymbol{\gamma}(T-u)|^2 \, \mathrm{d}u, \qquad 0 \leqslant s \leqslant T \tag{26}$$

is a local $\overline{\mathbb{G}}$-submartingale under $\mathbb{P}^{\gamma}$, with decomposition

$$\mathrm{d}M^{\gamma}(s) = \tfrac{1}{2} |\nabla L(T-s, \overline{X}(s)) + \boldsymbol{\gamma}(T-s)|^2 \, \mathrm{d}s + \Big\langle \nabla L(T-s, \overline{X}(s)), \mathrm{d}\overline{W}^{\gamma}(s) \Big\rangle. \tag{27}$$

In fact, introducing for $n \in \mathbb{N}_0$ the sequence

$$\sigma_n := \inf\Big\{ s \geqslant 0 \colon \int_0^s \Big(|\nabla L(T-u, \overline{X}(u))|^2 + |\boldsymbol{\gamma}(T-u)|^2 \Big) \, \mathrm{d}u \geqslant n \Big\} \wedge T \tag{28}$$

of $\overline{\mathbb{G}}$-stopping times with $\sigma_n \uparrow T$, we see that the stopped process $M^{\gamma}(\cdot \wedge \sigma_n)$ is a $\overline{\mathbb{G}}$-submartingale under $\mathbb{P}^{\gamma}$, for every $n \in \mathbb{N}_0$. In particular, we observe

$$\begin{aligned} H(P(T) \,|\, \mathrm{Q}) = \mathbb{E}_{\mathbb{P}}[L(T, X(T))] &= \mathbb{E}_{\mathbb{P}^{\gamma}}[L(T, \overline{X}(0))] \\ &\leqslant \mathbb{E}_{\mathbb{P}^{\gamma}}\Big[L(T-\sigma_n, \overline{X}(\sigma_n)) + \tfrac{1}{2} \int_0^{\sigma_n} |\boldsymbol{\gamma}(T-u)|^2 \, \mathrm{d}u \Big], \end{aligned} \tag{29}$$

since we have $\mathbb{P}^{\gamma} = \mathbb{P}$ on the σ-algebra $\overline{\mathcal{G}}(0) = \sigma(\overline{X}(0)) = \sigma(X(T))$. Now (29) holds for every $n \in \mathbb{N}_0$, thus

$$H(P(T)\,|\,\mathrm{Q}) \leqslant \liminf_{n\to\infty}\ \mathbb{E}_{\mathbb{P}^{\gamma}}\Big[L(T-\sigma_n,\overline{X}(\sigma_n)) + \tfrac{1}{2}\int_0^{\sigma_n} |\gamma(T-u)|^2\,\mathrm{d}u\Big]. \tag{30}$$

But as we remarked already, the minimum in (19) is attained by $g_* = -\nabla L(t,x)$; likewise, the drift term in (27) vanishes, if we select the process $\gamma_* \in \Gamma$ via

$$\gamma_*(t,\omega) := -\nabla L(t,\omega(t)), \qquad \text{thus} \qquad \gamma_*(T-s) = -\nabla L(T-s,\overline{X}(s)) \tag{31}$$

for $0 \leqslant s,t \leqslant T$. With this choice, the backwards dynamics of (25) take the form

$$\mathrm{d}\overline{X}(s) = -\nabla\Psi(\overline{X}(s))\,\mathrm{d}s + \mathrm{d}\overline{W}^{\gamma_*}(s); \tag{32}$$

that is, *precisely of the Langevin–Smoluchowski type* (3), but now with the "initial condition" $\overline{X}(0) = X(T)$ and independent driving $\overline{\mathbb{G}}$-Brownian motion $\overline{W}^{\gamma_*}$, under $\mathbb{P}^{\gamma_*}$. Since $\mathbb{P}^{\gamma_*} = \mathbb{P}$ holds on the σ-algebra $\overline{\mathcal{G}}(0) = \sigma(\overline{X}(0)) = \sigma(X(T))$, the initial distribution of $\overline{X}(0)$ under $\mathbb{P}^{\gamma_*}$ is equal to $P(T)$. Furthermore, with $\gamma = \gamma_*$, the process of (26), (27) becomes a $\mathbb{P}^{\gamma_*}$-local martingale, namely

$$M^{\gamma_*}(s) = L(T,X(T)) + \int_0^s \Big\langle \nabla L(T-u,\overline{X}(u)),\,\mathrm{d}\overline{W}^{\gamma_*}(u)\Big\rangle \tag{33}$$

for $0 \leqslant s \leqslant T$; and we have equality in (29), thus also

$$H(P(T)\,|\,\mathrm{Q}) = \lim_{n\to\infty}\ \mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[L(T-\sigma_n,\overline{X}(\sigma_n)) + \tfrac{1}{2}\int_0^{\sigma_n} |\gamma_*(T-u)|^2\,\mathrm{d}u\Big]. \tag{34}$$

We conclude that the infimum over $\gamma \in \Gamma$ of the right-hand side in (30) is attained by the process γ_* of (31), which gives rise to the Langevin–Smoluchowski dynamics (32) for the time-reversed process $\overline{X}(s) = X(T-s)$, $0 \leqslant s \leqslant T$, under $\mathbb{P}^{\gamma_*}$. We formalize this discussion as follows.

Theorem 1 *Consider the stochastic control problem of minimizing over the class* Γ *of measurable, adapted processes* γ *satisfying* (21) *and inducing an exponential martingale* Z^{γ} *in* (22)*, with the notation of* (28) *and with the backwards dynamics of* (25)*, the expected cost*

$$\mathcal{I}(\gamma) := \liminf_{n\to\infty}\ \mathbb{E}_{\mathbb{P}^{\gamma}}\Big[L(T-\sigma_n,\overline{X}(\sigma_n)) + \tfrac{1}{2}\int_0^{\sigma_n} |\gamma(T-u)|^2\,\mathrm{d}u\Big]. \tag{35}$$

Under the assumptions of Section 2*, the infimum* $\inf_{\gamma\in\Gamma}\mathcal{I}(\gamma)$ *is equal to the relative entropy* $H(P(T)\,|\,\mathrm{Q})$ *and is attained by the "score process"* γ_* *of* (31)*. This choice leads to the backwards Langevin–Smoluchowski dynamics* (32)*, and with* $\gamma = \gamma_*$ *the limit in* (35) *exists as in* (34)*.*

Proof It only remains to check that the minimizing process of (31) belongs indeed to the collection $\boldsymbol{\Gamma}$ of admissible processes. By its definition, this process $\boldsymbol{\gamma}_*$ is measurable, and its time-reversal is adapted to the backward filtration $\overline{\mathbb{G}}$ of (16). Theorem 4.1 in [25] gives

$$\mathbb{E}_{\mathbb{P}}\Big[\int_0^T |\nabla L(T-u, \overline{X}(u))|^2 \,\mathrm{d}u\Big] = \mathbb{E}_{\mathbb{P}}\Big[\int_0^T |\nabla L(\theta, X(\theta))|^2 \,\mathrm{d}\theta\Big] < \infty, \tag{36}$$

which implies *a fortiori* that the condition in (21) is satisfied for $\boldsymbol{\gamma} = \boldsymbol{\gamma}_*$.

We must also show that the process $Z^{\boldsymbol{\gamma}_*}$ defined in the manner of (22), is a true martingale. A very mild dose of stochastic calculus leads to

$$\mathrm{d}L(T-s, \overline{X}(s)) = \Big\langle \nabla L(T-s, \overline{X}(s)), \mathrm{d}\overline{W}(s)\Big\rangle + \tfrac{1}{2}|\nabla L(T-s, \overline{X}(s))|^2 \,\mathrm{d}s$$

on account of (18), (20). Therefore, we have

$$\begin{aligned}
&\int_0^s \langle \boldsymbol{\gamma}_*(T-u), \mathrm{d}\overline{W}(u)\rangle - \tfrac{1}{2}\int_0^s |\boldsymbol{\gamma}_*(T-u)|^2 \,\mathrm{d}u \\
&\quad = -\int_0^s \Big\langle \nabla L(T-u, \overline{X}(u)), \mathrm{d}\overline{W}(u)\Big\rangle - \tfrac{1}{2}\int_0^s |\nabla L(T-u, \overline{X}(u))|^2 \,\mathrm{d}u \\
&\quad = L(T, X(T)) - L(T-s, \overline{X}(s)) = \log\Big(\frac{\ell(T, X(T))}{\ell(T-s, \overline{X}(s))}\Big),
\end{aligned}$$

which expresses the exponential process of (22) with $\boldsymbol{\gamma} = \boldsymbol{\gamma}_*$ as

$$Z^{\boldsymbol{\gamma}_*}(s) = \frac{\ell(T, X(T))}{\ell(T-s, \overline{X}(s))}, \qquad 0 \leqslant s \leqslant T.$$

Now, let us argue that the process $Z^{\boldsymbol{\gamma}_*}$ is a true $(\overline{\mathbb{G}}, \mathbb{P})$-martingale. It is a positive local martingale, thus a supermartingale. It will be a martingale, if it has constant expectation. But $Z^{\boldsymbol{\gamma}_*}(0) \equiv 1$, so it is enough to show that $\mathbb{E}_{\mathbb{P}}[Z^{\boldsymbol{\gamma}_*}(T)] = 1$. Let us denote by $P(s, y; t, \xi)$ the transition kernel of the Langevin–Smoluchowski dynamics, so that $\mathbb{P}[X(s) \in \mathrm{d}y, X(t) \in \mathrm{d}\xi] = p(s, y)\, P(s, y; t, \xi)\, \mathrm{d}y\, \mathrm{d}\xi$ for $0 \leqslant s < t \leqslant T$ and $(y, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$. Then the invariance of Q gives

$$\int_{\mathbb{R}^n} q(y)\, P(0, y; T, \xi)\, \mathrm{d}y = q(\xi), \qquad \xi \in \mathbb{R}^n; \tag{37}$$

consequently

$$\mathbb{E}_{\mathbb{P}}\left[Z^{\gamma_*}(T)\right] = \mathbb{E}_{\mathbb{P}}\left[\frac{p(T,X(T))}{q(X(T))}\frac{q(X(0))}{p(0,X(0))}\right] \tag{38}$$

$$= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{p(T,\xi)}{q(\xi)}\frac{q(y)}{p(0,y)}\, p(0,y)\, P(0,y;T,\xi)\,\mathrm{d}y\,\mathrm{d}\xi \tag{39}$$

$$= \int_{\mathbb{R}^n} \frac{p(T,\xi)}{q(\xi)}\left(\int_{\mathbb{R}^n} q(y)\, P(0,y;T,\xi)\,\mathrm{d}y\right)\mathrm{d}\xi \tag{40}$$

$$= \int_{\mathbb{R}^n} p(T,\xi)\,\mathrm{d}\xi = 1, \tag{41}$$

implying that Z^{γ_*} is a true martingale and completing the proof of Theorem 1. □

Results related to Theorem 1 have been established in [12, 13, 16, 38, 39].

Remark 1 (Reincarnation of time-marginals) Let us denote by $\overline{P}_*(s)$ the distribution of the random variable $\overline{X}(s) = X(T-s)$ under the probability measure $\mathbb{P}^{\gamma_*}$, for $0 \leqslant s \leqslant T$. Since $(\overline{X}(s))_{0\leqslant s\leqslant T}$ is under $\mathbb{P}^{\gamma_*}$ a Langevin–Smoluchowski diffusion in its own right, we deduce

$$\overline{P}_*(s) = P(T+s), \qquad 0 \leqslant s \leqslant T \tag{42}$$

on the strength of uniqueness in distribution for the Langevin–Smoluchowski flow, and of its time-homogeneity. In other words, the branch $(P(T+s))_{0\leqslant s\leqslant T}$ of the original Langevin–Smoluchowski curve of time-marginals, gets "reincarnated" as $(\overline{P}_*(s))_{0\leqslant s\leqslant T}$, the curve of time-marginals arising from the solution of the stochastic control problem in Theorem 1. But now, under the probability measure $\mathbb{P}^{\gamma_*}$, the states of the Langevin–Smoluchowski diffusion $(\overline{X}(s))_{0\leqslant s\leqslant T}$ corresponding to the curve $(\overline{P}_*(s))_{0\leqslant s\leqslant T}$ traverse the time interval $[0,T]$ in the opposite temporal direction.

4.1 Entropic interpretation of the expected cost when $\mathbf{Q}(\mathbb{R}^n) < \infty$

Let us observe from (22) – (24) that

$$\log\left(\frac{\mathrm{d}\mathbb{P}^{\gamma}}{\mathrm{d}\mathbb{P}}\bigg|_{\mathcal{G}(\sigma_n)}\right) = \int_0^{\sigma_n} \langle \boldsymbol{\gamma}(T-u), \mathrm{d}\overline{W}^{\gamma}(u)\rangle + \frac{1}{2}\int_0^{\sigma_n} |\boldsymbol{\gamma}(T-u)|^2\,\mathrm{d}u \tag{43}$$

holds for every $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$ and $n \in \mathbb{N}_0$. Thus, as the $\mathbb{P}^{\gamma}$-expectation of the stochastic integral in (43) vanishes, the expected quadratic cost, or "energy", term in (35) is itself a relative entropy:

$$\mathbb{E}_{\mathbb{P}^{\gamma}}\left[\frac{1}{2}\int_0^{\sigma_n} |\boldsymbol{\gamma}(T-u)|^2\,\mathrm{d}u\right] = \mathbb{E}_{\mathbb{P}^{\gamma}}\left[\log\left(\frac{\mathrm{d}\mathbb{P}^{\gamma}}{\mathrm{d}\mathbb{P}}\bigg|_{\mathcal{G}(\sigma_n)}\right)\right].$$

By contrast, when Q is a probability measure on $\mathscr{B}(\mathbb{R}^n)$, and denoting by $\mathbb{Q}$ the probability measure induced on $\Omega = C([0,\infty);\mathbb{R}^n)$ by the canonical process driven by the dynamics (3) with Q as the distribution of $X(0)$, the first term in (35) can be cast as

$$\mathbb{E}_{\mathbb{P}^\gamma}\big[L(T-\sigma_n,\overline{X}(\sigma_n))\big] = \mathbb{E}_{\mathbb{P}^\gamma}\bigg[\log\Big(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big|_{\sigma(\overline{X}(\sigma_n))}\Big)\bigg] \tag{44}$$

$$= \mathbb{E}_{\mathbb{P}^\gamma}\bigg[\log\Big(\frac{\mathrm{d}\mathbb{P}^\gamma}{\mathrm{d}\mathbb{Q}}\Big|_{\sigma(\overline{X}(\sigma_n))}\Big) - \log\Big(\frac{\mathrm{d}\mathbb{P}^\gamma}{\mathrm{d}\mathbb{P}}\Big|_{\sigma(\overline{X}(\sigma_n))}\Big)\bigg]. \tag{45}$$

It follows that, in this case, the expected cost of (35) is equal to the sum $H_n^\gamma + D_n^\gamma$ of two non-negative quantities:

$$H_n^\gamma := \mathbb{E}_{\mathbb{P}^\gamma}\bigg[\log\Big(\frac{\mathrm{d}\mathbb{P}^\gamma}{\mathrm{d}\mathbb{Q}}\Big|_{\sigma(\overline{X}(\sigma_n))}\Big)\bigg],$$

the relative entropy of the probability measure $\mathbb{P}^\gamma$ with respect to the probability measure $\mathbb{Q}$, when both measures are restricted to the σ-algebra generated by the random variable $\overline{X}(\sigma_n)$; and

$$D_n^\gamma := \mathbb{E}_{\mathbb{P}^\gamma}\bigg[\log\Big(\frac{\mathrm{d}\mathbb{P}^\gamma}{\mathrm{d}\mathbb{P}}\Big|_{\overline{\mathcal{G}}(\sigma_n)}\Big)\bigg] - \mathbb{E}_{\mathbb{P}^\gamma}\bigg[\log\Big(\frac{\mathrm{d}\mathbb{P}^\gamma}{\mathrm{d}\mathbb{P}}\Big|_{\sigma(\overline{X}(\sigma_n))}\Big)\bigg], \tag{46}$$

the difference between the relative entropies of the probability measure $\mathbb{P}^\gamma$ with respect to the probability measure $\mathbb{P}$, when restricted to the σ-algebra generated by the collection of random variables $(\overline{X}(u\wedge\sigma_n))_{0\leqslant u\leqslant T}$ and by the random variable $\overline{X}(\sigma_n)$, respectively. The difference in (46) is non-negative, because conditioning on a smaller σ-algebra can only decrease the relative entropy; this difference can be thought of as an "entropic cost of time-reversal".

It develops from this discussion that the expected cost on the right-hand side of (35) is non-negative, when Q is a probability measure.

5 From local to square-integrable martingales

Whenever the process M^γ of (26), (27) happens to be a true submartingale under $\mathbb{P}^\gamma$ (as, for instance, with $\gamma \equiv 0$ on account of Theorem 4.1 in [25]), the expected cost (35) takes the form

$$\mathbb{E}_{\mathbb{P}^\gamma}\Big[L(0,X(0)) + \tfrac{1}{2}\int_0^T |\gamma(T-u)|^2\,\mathrm{d}u\Big].$$

Likewise, we derive from Theorem 1 the identity

$$H(P(T)\,|\,\mathrm{Q}) = \mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[L(0,X(0)) + \tfrac{1}{2}\int_0^T |\nabla L(T-u,\overline{X}(u))|^2\,\mathrm{d}u\Big],$$

whenever the process M^{γ_*} of (33) is a true $\mathbb{P}^{\gamma_*}$-martingale. This is the case, for instance, whenever

$$\mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\langle M^{\gamma_*}, M^{\gamma_*}\rangle(T)\Big] = \mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\int_0^T |\nabla L(T-u,\overline{X}(u))|^2\,\mathrm{d}u\Big] < \infty \tag{47}$$

holds; and then the stochastic integral in (33) is an $L^2(\mathbb{P}^{\gamma_*})$-bounded martingale (see, for instance, [27, Proposition 3.2.10] or [40, Corollary IV.1.25]). Using (42), we can express the expectation of (47) more explicitly as

$$\int_{\mathbb{R}^n}\int_0^T |\nabla \log \ell(t,x)|^2\, p(2T-t,x)\,\mathrm{d}t\,\mathrm{d}x. \tag{48}$$

The shift in the temporal variable makes it difficult to check whether the quantity in (48) is finite. At least, we have not been able to apply directly arguments similar to those in Theorem 4.1 of [25], where the expectation (47) is taken with respect to the probability measure $\mathbb{P}$, in the manner of (36) (and thus, the argument $2T-t$ in (48) is replaced by t). This problem is consonant with the fact that the expression in (44), (45) is not quite a relative entropy, but a linear combination of relative entropies.

The goal of this section is to find a square-integrable $\mathbb{P}^{\gamma_*}$-martingale $\overline{M}^{\gamma_*}$, which is closely related to the local martingale M^{γ_*} of (33). The idea is to correct the shift in the temporal variable appearing in (48), by reversing time once again.

First, we need to introduce some notation. We denote by $\overline{p}_*(s,\cdot)$ the probability density function of the random variable $\overline{X}(s) = X(T-s)$ under the probability measure $\mathbb{P}^{\gamma_*}$, for $0 \leqslant s \leqslant T$. From (42), we deduce the relation

$$\overline{p}_*(s,x) = p(T+s,x), \qquad (s,x) \in [0,T]\times\mathbb{R}^n.$$

For $(s,x) \in [0,T]\times\mathbb{R}^n$, the associated likelihood ratio function and its logarithm are defined respectively by

$$\lambda(s,x) := \frac{\overline{p}_*(s,x)}{q(x)}, \qquad \Lambda(s,x) := \log\lambda(s,x) = L(T+s,x). \tag{49}$$

From the definition (49), and the equations (18), (19), we see that the function $(s,x) \mapsto \Lambda(s,x)$ satisfies again the semilinear Schrödinger-type equation

$$-\partial\Lambda(s,x) + \tfrac{1}{2}\Delta\Lambda(s,x) - \langle\nabla\Lambda(s,x),\nabla\Psi(x)\rangle = \min_{b\in\mathbb{R}^n}\Big(\langle\nabla\Lambda(s,x),\,b\rangle + \tfrac{1}{2}|b|^2\Big). \tag{50}$$

In the setting introduced above, for each $0 \leqslant s \leqslant T$, the relative entropy with respect to Q of the distribution $\overline{P}_*(s)$ of $\overline{X}(s)$ under $\mathbb{P}^{\gamma_*}$ is

$$H(\overline{P}_*(s)\,|\,\mathrm{Q}) = \mathbb{E}_{\mathbb{P}^{\gamma_*}}\big[\Lambda(s,\overline{X}(s))\big] = \int_{\mathbb{R}^n} \log\Big(\frac{\overline{p}_*(s,x)}{q(x)}\Big)\overline{p}_*(s,x)\,\mathrm{d}x. \tag{51}$$

Again, the assumption that $H(P(0)\,|\,\mathrm{Q})$ is finite, and the decrease of the relative entropy function $[0,\infty) \ni t \mapsto H(P(t)\,|\,\mathrm{Q}) \in (-\infty,\infty]$, imply that the relative entropy in (51) is finite for all $0 \leqslant s \leqslant T$.

Finally, the relative Fisher information of $\overline{P}_*(s)$ with respect to Q is defined as

$$I(\overline{P}_*(s)\,|\,\mathrm{Q}) := \mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\big|\nabla\Lambda(s,\overline{X}(s))\big|^2\Big] = \int_{\mathbb{R}^n} |\nabla\Lambda(s,x)|^2\,\overline{p}_*(s,x)\,\mathrm{d}x. \tag{52}$$

5.1 Reversing time once again

Let us consider on the filtered probability space $(\Omega,\mathbb{F},\mathbb{P}^{\gamma_*})$ the canonical process $(X(t))_{0\leqslant t\leqslant T}$, whose time-reversal (16) satisfies the backwards Langevin–Smoluchowski dynamics (32). Reversing time once again, we find that the process $X(t) = \overline{X}(T-t)$, $0 \leqslant t \leqslant T$ satisfies the stochastic differential equation

$$\mathrm{d}X(t) = \Big(\nabla\Lambda(T-t,X(t)) - \nabla\Psi(X(t))\Big)\,\mathrm{d}t + \mathrm{d}W^{\gamma_*}(t), \tag{53}$$

where the process

$$W^{\gamma_*}(t) := \overline{W}^{\gamma_*}(T-t) - \overline{W}^{\gamma_*}(T) - \int_0^t \nabla\log\overline{p}_*(T-\theta,X(\theta))\,\mathrm{d}\theta, \qquad 0 \leqslant t \leqslant T$$

is Brownian motion on $(\Omega,\mathbb{F},\mathbb{P}^{\gamma_*})$. We recall here Proposition 4.1 from [26]. Comparing the equation (53) with (3), we see that the $\mathbb{P}^{\gamma_*}$-Brownian motion $(W^{\gamma_*}(t))_{0\leqslant t\leqslant T}$ and the $\mathbb{P}$-Brownian motion $(W(t))_{0\leqslant t\leqslant T}$ are related via

$$W(t) = W^{\gamma_*}(t) + \int_0^t \nabla\Lambda(T-\theta,X(\theta))\,\mathrm{d}\theta, \qquad 0 \leqslant t \leqslant T.$$

5.2 The dynamics of the relative entropy process

We look now at the *relative entropy process*

$$\Lambda(T-t,X(t)) = \log\Big(\frac{\overline{p}_*(T-t,X(t))}{q(X(t))}\Big), \qquad 0 \leqslant t \leqslant T \tag{54}$$

on $(\Omega,\mathbb{F},\mathbb{P}^{\gamma_*})$. Applying Itô's formula and using the equation (50), together with the forward dynamics (53), we obtain the following result.

Proposition 1 *On the filtered probability space* $(\Omega,\mathbb{F},\mathbb{P}^{\gamma_*})$*, the relative entropy process* (54) *is a submartingale with stochastic differential*

$$\mathrm{d}\Lambda(T-t,X(t)) = \tfrac{1}{2}|\nabla\Lambda(T-t,X(t))|^2\,\mathrm{d}t + \Big\langle\nabla\Lambda(T-t,X(t)),\,\mathrm{d}W^{\gamma_*}(t)\Big\rangle \tag{55}$$

for $0 \leqslant t \leqslant T$. *In particular, for* $0 \leqslant t \leqslant T$, *the process*

$$\overline{M}^{\gamma_*}(t) := \Lambda(T-t,X(t)) - \Lambda(T,X(0)) - \tfrac{1}{2}\int_0^t |\nabla\Lambda(T-\theta,X(\theta))|^2\,\mathrm{d}\theta$$

is an $L^2(\mathbb{P}^{\gamma_*})$*-bounded martingale, with stochastic integral representation*

$$\overline{M}^{\gamma_*}(t) = \int_0^t \Big\langle\nabla\Lambda(T-\theta,X(\theta)),\,\mathrm{d}W^{\gamma_*}(\theta)\Big\rangle, \qquad 0 \leqslant t \leqslant T. \tag{56}$$

Proof The last thing we need to verify for the proof of Proposition 1, is that

$$\mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\langle\overline{M}^{\gamma_*},\overline{M}^{\gamma_*}\rangle(T)\Big] = \mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\int_0^T |\nabla\Lambda(T-t,X(t))|^2\,\mathrm{d}t\Big] < \infty. \tag{57}$$

We observe that the expectation in (57) is equal to

$$\mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\int_0^T |\nabla\Lambda(s,\overline{X}(s))|^2\,\mathrm{d}s\Big] = \mathbb{E}_{\mathbb{P}}\Big[\int_T^{2T} |\nabla L(t,X(t))|^2\,\mathrm{d}t\Big]. \tag{58}$$

This is because (16) and (49) give the relation $\nabla\Lambda(s,\overline{X}(s)) = \nabla L(t,X(2T-t))$ with $t = T+s \in [T,2T]$; and because the $\mathbb{P}^{\gamma_*}$-distribution of $X(2T-t) = \overline{X}(s)$ is the same as the $\mathbb{P}$-distribution of $X(T+s) = X(t)$, on account of (42). But, as (36) holds for any finite time horizon $T > 0$, the quantity in (58) is finite as well. □

5.3 Relative entropy dissipation

Exploiting the trajectorial evolution of the relative entropy process (54), provided by Proposition 1, allows us to derive some immediate consequences on the decrease of the relative entropy function $[0,T] \ni s \longmapsto H(\overline{P}_*(s)\,|\,\mathrm{Q}) \in (-\infty,\infty)$ and its rate of dissipation. The submartingale-property of the relative entropy process (54) shows once more, that this function is non-decreasing. More precisely, we have the following rate of change for the relative entropy.

Corollary 1 *For all* $s, s_0 \geqslant 0$, *we have*

$$H(\overline{P}_*(s)\,|\,\mathrm{Q}) - H(\overline{P}_*(s_0)\,|\,\mathrm{Q}) = -\tfrac{1}{2}\int_{s_0}^{s} I(\overline{P}_*(u)\,|\,\mathrm{Q})\,\mathrm{d}u. \tag{59}$$

Proof Let $s, s_0 \geqslant 0$ and choose $T \geqslant \max\{s, s_0\}$. Taking expectations under $\mathbb{P}^{\gamma_*}$ in (55), and noting that the stochastic integral process in (56) is a martingale, leads to

$$\mathbb{E}_{\mathbb{P}^{\gamma_*}}\big[\Lambda(s,\overline{X}(s))\big]-\mathbb{E}_{\mathbb{P}^{\gamma_*}}\big[\Lambda(s_0,\overline{X}(s_0))\big]=-\tfrac{1}{2}\int_{s_0}^{s}\mathbb{E}_{\mathbb{P}^{\gamma_*}}\Big[\big|\nabla\Lambda(u,\overline{X}(u))\big|^2\Big]\,\mathrm{d}u.$$

Recalling the entropy (51) and the Fisher information (52), we obtain (59). □

Corollary 2 *For Lebesgue-almost every* $s \geqslant 0$*, the rate of relative entropy dissipation equals*

$$\frac{\mathrm{d}}{\mathrm{d}s}H(\overline{P}_*(s)\,|\,\mathrm{Q})=-\tfrac{1}{2}I(\overline{P}_*(s)\,|\,\mathrm{Q}).$$

6 From backwards dynamics "back" to forward dynamics

Starting with the forward Langevin–Smoluchowski dynamics (3), we have seen in Section 4 that the combined effects of time-reversal, and of stochastic control of the drift under an entropic-type criterion, lead to the backwards dynamics

$$\mathrm{d}\overline{X}(s)=-\nabla\Psi(\overline{X}(s))\,\mathrm{d}s+\mathrm{d}\overline{W}^{\gamma_*}(s),\qquad 0\leqslant s\leqslant T,\tag{60}$$

which are again of the Langevin–Smoluchowski type, but now viewed on the filtered probability space $(\Omega,\overline{\mathbb{G}},\mathbb{P}^{\gamma_*})$. We will see now that this universal property of Langevin–Smoluchowski measure is consistent in the following sense: starting with the backwards Langevin–Smoluchowski dynamics of (60), after another reversal of time, the solution of a related stochastic control problem leads to the original forward Langevin–Smoluchowski dynamics (3) we started with. This consistency property should come as no surprise, but its formal proof requires the results of Section 5, which perhaps appeared artificial at first sight.

Let us recall from Subsection 5.1 that reversing time in (60) leads to the forward dynamics

$$\mathrm{d}X(t)=\Big(\nabla\Lambda(T-t,X(t))-\nabla\Psi(X(t))\Big)\,\mathrm{d}t+\mathrm{d}W^{\gamma_*}(t),\qquad 0\leqslant t\leqslant T\tag{61}$$

on the filtered probability space $(\Omega,\mathbb{F},\mathbb{P}^{\gamma_*})$. By analogy with Section 4, we define an equivalent probability measure $\Pi^{\boldsymbol{\beta}}\sim\mathbb{P}^{\gamma_*}$ as follows.

For any measurable process $[0,T]\times\Omega\ni(t,\omega)\mapsto\boldsymbol{\beta}(t,\omega)\in\mathbb{R}^n$, adapted to the forward filtration $\mathbb{F}$ of (15) and satisfying the condition

$$\mathbb{P}^{\gamma_*}\bigg[\int_0^T|\boldsymbol{\beta}(t)|^2\,\mathrm{d}t<\infty\bigg]=1,\tag{62}$$

we consider the exponential $(\mathbb{F},\mathbb{P}^{\gamma_*})$-local martingale

$$Z^{\boldsymbol{\beta}}(t):=\exp\bigg(\int_0^t\langle\boldsymbol{\beta}(\theta),\mathrm{d}W^{\gamma_*}(\theta)\rangle-\tfrac{1}{2}\int_0^t|\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta\bigg)\tag{63}$$

for $0 \leqslant t \leqslant T$. We denote by $\mathcal{B}$ the collection of all processes $\boldsymbol{\beta}$ as above, for which $Z^{\boldsymbol{\beta}}$ is a true $(\mathbb{F}, \mathbb{P}^{\boldsymbol{\gamma}_*})$-martingale.

Now, for every $\boldsymbol{\beta} \in \mathcal{B}$, we introduce an equivalent probability measure $\Pi^{\boldsymbol{\beta}} \sim \mathbb{P}^{\boldsymbol{\gamma}_*}$ on path space, via

$$\frac{\mathrm{d}\Pi^{\boldsymbol{\beta}}}{\mathrm{d}\mathbb{P}^{\boldsymbol{\gamma}_*}}\bigg|_{\mathcal{F}(t)} = Z^{\boldsymbol{\beta}}(t), \qquad 0 \leqslant t \leqslant T. \tag{64}$$

Then we deduce from the Girsanov theorem that, under the probability measure $\Pi^{\boldsymbol{\beta}}$, the process

$$W^{\boldsymbol{\beta}}(t) := W^{\boldsymbol{\gamma}_*}(t) - \int_0^t \boldsymbol{\beta}(\theta)\,\mathrm{d}\theta, \qquad 0 \leqslant t \leqslant T \tag{65}$$

is $\mathbb{R}^n$-valued $\mathbb{F}$-Brownian motion, and the dynamics (61) become

$$\mathrm{d}X(t) = \Big(\nabla\Lambda(T-t, X(t)) + \boldsymbol{\beta}(t) - \nabla\Psi(X(t))\Big)\,\mathrm{d}t + \mathrm{d}W^{\boldsymbol{\beta}}(t). \tag{66}$$

We couple these dynamics with the stochastic differential (55) and deduce that the process

$$N^{\boldsymbol{\beta}}(t) := \Lambda(T-t, X(t)) + \tfrac{1}{2}\int_0^t |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta, \qquad 0 \leqslant t \leqslant T \tag{67}$$

is a local $\Pi^{\boldsymbol{\beta}}$-submartingale with decomposition

$$\mathrm{d}N^{\boldsymbol{\beta}}(t) = \tfrac{1}{2}\big|\nabla\Lambda(T-t, X(t)) + \boldsymbol{\beta}(t)\big|^2\,\mathrm{d}t + \Big\langle \nabla\Lambda(T-t, X(t)), \,\mathrm{d}\overline{W}^{\boldsymbol{\beta}}(t)\Big\rangle. \tag{68}$$

In fact, introducing for $n \in \mathbb{N}_0$ the sequence

$$\tau_n := \inf\left\{t \geqslant 0 \colon \int_0^t \Big(|\nabla\Lambda(T-\theta, X(\theta))|^2 + |\boldsymbol{\beta}(\theta)|^2\Big)\,\mathrm{d}\theta \geqslant n\right\} \wedge T \tag{69}$$

of $\mathbb{F}$-stopping times with $\tau_n \uparrow T$, we see that the stopped process $N^{\boldsymbol{\beta}}(\cdot \wedge \tau_n)$ is an $\mathbb{F}$-submartingale under $\Pi^{\boldsymbol{\beta}}$, for every $n \in \mathbb{N}_0$. In particular, we observe

$$\begin{aligned} H(P(2T)\,|\,\mathrm{Q}) = H(\overline{P}_*(T)\,|\,\mathrm{Q}) &= \mathbb{E}_{\mathbb{P}^{\boldsymbol{\gamma}_*}}\big[\Lambda(T, \overline{X}(T))\big] = \mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\big[\Lambda(T, X(0))\big] \\ &\leqslant \mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\Big[\Lambda(T-\tau_n, X(\tau_n)) + \tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta\Big], \end{aligned} \tag{70}$$

since we have $\Pi^{\boldsymbol{\beta}} = \mathbb{P}^{\boldsymbol{\gamma}_*}$ on the $\boldsymbol{\sigma}$-algebra $\mathcal{F}(0) = \boldsymbol{\sigma}(X(0)) = \boldsymbol{\sigma}(\overline{X}(T))$. Now (70) holds for every $n \in \mathbb{N}_0$, thus

$$H(P(2T)\,|\,\mathrm{Q}) \leqslant \liminf_{n\to\infty}\ \mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\Big[\Lambda(T-\tau_n, X(\tau_n)) + \tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta\Big]. \tag{71}$$

But the minimum in (50) is attained by $b_* = -\nabla\Lambda(s, x)$; likewise, the drift term in (68) vanishes, if we select the process $\boldsymbol{\beta}_* \in \mathcal{B}$ via

$$\boldsymbol{\beta}_*(t,\omega) := -\nabla\Lambda(T-t,\omega(t)), \qquad \text{thus} \qquad \boldsymbol{\beta}_*(t) = -\nabla\Lambda(T-t,X(t)) \tag{72}$$

for $0 \leqslant t \leqslant T$. With this choice, the forward dynamics of (66) take the form

$$\mathrm{d}X(t) = -\nabla\Psi(X(t))\,\mathrm{d}t + \mathrm{d}W^{\boldsymbol{\beta}_*}(t); \tag{73}$$

that is, *precisely the forward Langevin–Smoluchowski dynamics* (3) we started with, but now with the "initial condition" $X(0) = \overline{X}(T)$ and independent driving $\mathbb{F}$-Brownian motion $W^{\boldsymbol{\beta}_*}$, under the probability measure $\Pi^{\boldsymbol{\beta}_*}$. Since $\Pi^{\boldsymbol{\beta}_*} = \mathbb{P}^{\boldsymbol{\gamma}_*}$ holds on the $\boldsymbol{\sigma}$-algebra $\mathcal{F}(0) = \boldsymbol{\sigma}(X(0)) = \boldsymbol{\sigma}(\overline{X}(T))$, the initial distribution of $X(0)$ under $\Pi^{\boldsymbol{\beta}_*}$ is equal to $P(2T)$. Furthermore, with $\boldsymbol{\beta} = \boldsymbol{\beta}_*$, the process of (67), (68) becomes a $\Pi^{\boldsymbol{\beta}_*}$-local martingale, namely

$$N^{\boldsymbol{\beta}_*}(t) = \Lambda(T,X(0)) + \int_0^t \Big\langle \nabla\Lambda(T-\theta,X(\theta))\,,\,\mathrm{d}\overline{W}^{\boldsymbol{\beta}_*}(\theta)\Big\rangle$$

for $0 \leqslant t \leqslant T$; and we have equality in (70), thus also

$$H(P(2T)\,|\,\mathrm{Q}) = \lim_{n\to\infty} \mathbb{E}_{\Pi^{\boldsymbol{\beta}_*}}\Big[\Lambda(T-\tau_n,X(\tau_n)) + \tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}_*(\theta)|^2\,\mathrm{d}\theta\Big]. \tag{74}$$

We conclude that the infimum over $\boldsymbol{\beta} \in \mathcal{B}$ of the right-hand side in (71) is attained by the process $\boldsymbol{\beta}_*$ of (72), which gives rise to the Langevin–Smoluchowski dynamics (73) for the process $(X(t))_{0\leqslant t\leqslant T}$, under $\Pi^{\boldsymbol{\beta}_*}$. We formalize this result as follows.

Theorem 2 *Consider the stochastic control problem of minimizing over the class $\mathcal{B}$ of measurable, adapted processes $\boldsymbol{\beta}$ satisfying* (62) *and inducing an exponential martingale $Z^{\boldsymbol{\beta}}$ in* (63)*, with the notation of* (69) *and with the forward dynamics of* (66)*, the expected cost*

$$\mathcal{J}(\boldsymbol{\beta}) := \liminf_{n\to\infty} \mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\Big[\Lambda(T-\tau_n,X(\tau_n)) + \tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta\Big]. \tag{75}$$

Under the assumptions of Section 2*, the infimum $\inf_{\boldsymbol{\beta}\in\mathcal{B}}\mathcal{J}(\boldsymbol{\beta})$ is equal to the relative entropy $H(P(2T)\,|\,\mathrm{Q})$ and is attained by the "score process" $\boldsymbol{\beta}_*$ of* (72)*. This choice leads to the forward Langevin–Smoluchowski dynamics* (73)*, and with $\boldsymbol{\beta} = \boldsymbol{\beta}_*$ the limit in* (75) *exists as in* (74)*.*

Proof We have to show that the minimizing process $\boldsymbol{\beta}_*$ belongs to the collection $\mathcal{B}$ of admissible processes. By its definition in (72), the process $\boldsymbol{\beta}_*$ is measurable, and adapted to the forward filtration $\mathbb{F}$ of (15). Thanks to (57) in Proposition 1, we have

$$\mathbb{E}_{\mathbb{P}^{\boldsymbol{\gamma}_*}}\Big[\int_0^T |\boldsymbol{\beta}_*(t)|^2\,\mathrm{d}t\Big] = \mathbb{E}_{\mathbb{P}^{\boldsymbol{\gamma}_*}}\Big[\int_0^T \big|\nabla\Lambda(T-t,X(t))\big|^2\,\mathrm{d}t\Big] < \infty,$$

which implies a fortiori that the condition in (62) is satisfied for $\boldsymbol{\beta} = \boldsymbol{\beta}_*$.

It remains to check that the process $Z^{\boldsymbol{\beta}_*}$ defined in the manner of (63), is a true martingale. From Proposition 1 we have the stochastic differential

$$\mathrm{d}\Lambda(T-t, X(t)) = \tfrac{1}{2}|\nabla\Lambda(T-t, X(t))|^2\,\mathrm{d}t + \Big\langle \nabla\Lambda(T-t, X(t)), \mathrm{d}W^{\boldsymbol{\gamma}_*}(t)\Big\rangle,$$

and therefore

$$\begin{aligned}
&\int_0^t \langle \boldsymbol{\beta}_*(\theta), \mathrm{d}W^{\boldsymbol{\gamma}_*}(\theta)\rangle - \tfrac{1}{2}\int_0^t |\boldsymbol{\beta}_*(\theta)|^2\,\mathrm{d}\theta \\
&\qquad = -\int_0^t \Big\langle \nabla\Lambda(T-\theta, X(\theta)), \mathrm{d}W^{\boldsymbol{\gamma}_*}(\theta)\Big\rangle - \tfrac{1}{2}\int_0^t |\nabla\Lambda(T-\theta, X(\theta))|^2\,\mathrm{d}\theta \\
&\qquad = \Lambda(T, X(0)) - \Lambda(T-t, X(t)) = \log\Big(\frac{\lambda(T, X(0))}{\lambda(T-t, X(t))}\Big),
\end{aligned}$$

which expresses the exponential process of (63) with $\boldsymbol{\beta} = \boldsymbol{\beta}_*$ as

$$Z^{\boldsymbol{\beta}_*}(t) = \frac{\lambda(T, X(0))}{\lambda(T-t, X(t))}, \qquad 0 \leqslant t \leqslant T.$$

The process $Z^{\boldsymbol{\beta}_*}$ is a positive local martingale, thus a supermartingale. To see that it is a true $(\mathbb{F}, \mathbb{P}^{\boldsymbol{\gamma}_*})$-martingale, it suffices to argue that it has constant expectation. But $Z^{\boldsymbol{\beta}_*}(0) \equiv 1$, so we have to show $\mathbb{E}_{\mathbb{P}^{\boldsymbol{\gamma}_*}}[Z^{\boldsymbol{\beta}_*}(T)] = 1$. We denote again by $P(s, y; t, \xi)$ the transition kernel of the Langevin–Smoluchowski dynamics, note that

$$\mathbb{P}^{\boldsymbol{\gamma}_*}\big[\overline{X}(s) \in \mathrm{d}y, \overline{X}(t) \in \mathrm{d}\xi\big] = \overline{p}_*(s, y)\, P(s, y; t, \xi)\,\mathrm{d}y\,\mathrm{d}\xi$$

for $0 \leqslant s < t \leqslant T$ and $(y, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$, and recall the invariance property (37) of Q, to deduce $\mathbb{E}_{\mathbb{P}^{\boldsymbol{\gamma}_*}}[Z^{\boldsymbol{\beta}_*}(T)] = 1$ in the manner of (38) – (41). This implies that $Z^{\boldsymbol{\beta}_*}$ is a true martingale and completes the proof of Theorem 2. □

6.1 Entropic interpretation of the expected cost when $\mathrm{Q}(\mathbb{R}^n) < \infty$

By analogy with Subsection 4.1, we interpret now the expected cost on the right-hand side of (75) in terms of relative entropies. From (63) – (65), we deduce that

$$\log\Big(\frac{\mathrm{d}\Pi^{\boldsymbol{\beta}}}{\mathrm{d}\mathbb{P}^{\boldsymbol{\gamma}_*}}\Big|_{\mathcal{F}(\tau_n)}\Big) = \int_0^{\tau_n} \langle \boldsymbol{\beta}(\theta), \mathrm{d}W^{\boldsymbol{\beta}}(\theta)\rangle + \tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta \tag{76}$$

holds for every $\boldsymbol{\beta} \in \mathcal{B}$ and $n \in \mathbb{N}_0$. Therefore, as the $\Pi^{\boldsymbol{\beta}}$-expectation of the stochastic integral in (76) vanishes, the expected quadratic cost, or "energy", term in (75) is equal to the relative entropy

$$\mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\Big[\tfrac{1}{2}\int_0^{\tau_n} |\boldsymbol{\beta}(\theta)|^2\,\mathrm{d}\theta\Big] = \mathbb{E}_{\Pi^{\boldsymbol{\beta}}}\Big[\log\Big(\frac{\mathrm{d}\Pi^{\boldsymbol{\beta}}}{\mathrm{d}\mathbb{P}^{\boldsymbol{\gamma}_*}}\Big|_{\mathcal{F}(\tau_n)}\Big)\Big].$$

In order to interpret the first term in (75), let us assume that Q, and thus also the induced measure $\mathbb{Q}$ on path space, are probability measures. Then we have

$$\begin{aligned}\mathbb{E}_{\Pi^\beta}\big[\Lambda(T-\tau_n, X(\tau_n))\big] &= \mathbb{E}_{\Pi^\beta}\bigg[\log\bigg(\frac{\mathrm{d}\mathbb{P}^{\gamma_*}}{\mathrm{d}\mathbb{Q}}\bigg|_{\sigma(X(\tau_n))}\bigg)\bigg] \\ &= \mathbb{E}_{\Pi^\beta}\bigg[\log\bigg(\frac{\mathrm{d}\Pi^\beta}{\mathrm{d}\mathbb{Q}}\bigg|_{\sigma(X(\tau_n))}\bigg) - \log\bigg(\frac{\mathrm{d}\Pi^\beta}{\mathrm{d}\mathbb{P}^{\gamma_*}}\bigg|_{\sigma(X(\tau_n))}\bigg)\bigg].\end{aligned}$$

We conclude that, in this case, the expected cost of (75) is equal to the sum $H_n^\beta + D_n^\beta$ of two non-negative quantities:

$$H_n^\beta := \mathbb{E}_{\Pi^\beta}\bigg[\log\bigg(\frac{\mathrm{d}\Pi^\beta}{\mathrm{d}\mathbb{Q}}\bigg|_{\sigma(X(\tau_n))}\bigg)\bigg],$$

the relative entropy of the probability measure Π^β with respect to the probability measure $\mathbb{Q}$ when both are restricted to the σ-algebra generated by the random variable $X(\tau_n)$; and

$$D_n^\beta := \mathbb{E}_{\Pi^\beta}\bigg[\log\bigg(\frac{\mathrm{d}\Pi^\beta}{\mathrm{d}\mathbb{P}^{\gamma_*}}\bigg|_{\mathcal{F}(\tau_n)}\bigg)\bigg] - \mathbb{E}_{\Pi^\beta}\bigg[\log\bigg(\frac{\mathrm{d}\Pi^\beta}{\mathrm{d}\mathbb{P}^{\gamma_*}}\bigg|_{\sigma(X(\tau_n))}\bigg)\bigg],$$

the difference between the relative entropies of the probability measure Π^β with respect to the probability measure $\mathbb{P}^{\gamma_*}$, when restricted to the σ-algebra generated by the collection of random variables $(X(\theta \wedge \tau_n))_{0 \leqslant \theta \leqslant T}$ and by the random variable $X(\tau_n)$, respectively.

7 The case of finite invariant measure, and an iterative procedure

Let us suppose now that the diffusion process $(X(t))_{t \geqslant 0}$ as in (3) is well-defined, along with the curve $P(t) = \mathrm{Law}(X(t))$, $t \geqslant 0$ of its time-marginals; and that the invariant measure Q of Subsection 2.1 is finite, i.e., (7) holds, and is thus normalized to a probability measure.

Then, neither the coercivity condition (11), nor the finite second-moment condition (12), are needed for the results of Sections 4 – 6. The reason is that the relative entropy $H(P(t)\,|\,\mathrm{Q})$ is now well-defined and non-negative, as both $P(t)$ and Q are probability measures. Since the function $t \mapsto H(P(t)\,|\,\mathrm{Q})$ is decreasing and the initial relative entropy $H(P(0)\,|\,\mathrm{Q})$ is finite on account of (13), it follows that this function takes values in $[0,\infty)$. It can also be shown in this case that

$$\lim_{t\to\infty} \downarrow H(P(t)\,|\,\mathrm{Q}) = 0, \tag{77}$$

i.e., the relative entropy decreases down to zero; see [19, Proposition 1.9] for a quite general version of this result. This, in turn, implies that the time-marginals

$(P(t))_{t\geqslant 0}$ converge to Q in total variation as $t\to\infty$, on account of the *Pinsker–Csiszár inequality*

$$2\,\|P(t)-\mathrm{Q}\|_{\mathrm{TV}}^2 \leqslant H(P(t)\,|\,\mathrm{Q}).$$

The entropic decrease to zero is actually exponentially fast, whenever the Hessian of the potential Ψ dominates a positive multiple of the identity matrix; see, e.g., [4], [32, Section 5], [36, Proposition 1'], [45, Formal Corollary 9.3], or [26, Remark 3.23]. As another consequence of (77), the initial relative entropy $H(P(0)\,|\,\mathrm{Q})$ can be expressed as

$$H(P(0)\,|\,\mathrm{Q}) = \tfrac{1}{2}\,\mathbb{E}_{\mathbb{P}}\Big[\int_0^\infty |\nabla L(t,X(t))|^2\,\mathrm{d}t\Big]. \tag{78}$$

We prove the claims (77) and (78) in Appendix 1.

In this context, i.e., with (7) replacing (11) and (12), and always under the standing assumption (13), Theorem 4.1 in [25] continues to hold, as do the results in Sections 4 – 6. By combining time-reversal with stochastic control of the drift, these results lead to an alternating sequence of forward and backward Langevin–Smoluchowski dynamics, with time-marginals starting at $P(0)$ and converging along $(P(kT))_{k\in\mathbb{N}_0}$ in total variation to the invariant probability measure Q. Along the way, the values of the corresponding stochastic control problems decrease along $(H(P(kT)\,|\,\mathrm{Q}))_{k\in\mathbb{N}}$ to zero.

Appendix 1: The decrease of the relative entropy without convexity assumption

We present a probabilistic proof of the claims (77) and (78), which complements the proof of the more general Proposition 1.9 in [19]. We stress that no convexity assumptions are imposed on the potential Ψ.

Proof of (77) ***and*** (78)***:*** Since Q is assumed to be a probability measure in Section 7, the relative entropy $H(P(t)\,|\,\mathrm{Q})$ is non-negative for every $t\geqslant 0$. Thus, [25, Corollary 4.3] gives the inequality

$$\begin{aligned} H(P(0)\,|\,\mathrm{Q}) &= H(P(T)\,|\,\mathrm{Q}) + \tfrac{1}{2}\,\mathbb{E}_{\mathbb{P}}\Big[\int_0^T |\nabla L(t,X(t))|^2\,\mathrm{d}t\Big] \\ &\geqslant \tfrac{1}{2}\,\mathbb{E}_{\mathbb{P}}\Big[\int_0^T |\nabla L(t,X(t))|^2\,\mathrm{d}t\Big] \end{aligned} \tag{79}$$

for every $T\in(0,\infty)$. Letting $T\uparrow\infty$ in (79), we deduce from the monotone convergence theorem that

$$\tfrac{1}{2}\,\mathbb{E}_{\mathbb{P}}\Big[\int_0^\infty |\nabla L(t,X(t))|^2\,\mathrm{d}t\Big] \leqslant H(P(0)\,|\,\mathrm{Q}). \tag{80}$$

By analogy with Subsection 2.2, we denote by $\mathbb{Q}$ the Langevin–Smoluchowski measure associated with the potential Ψ, but now with distribution

$$\mathbb{Q}[X(0) \in A] = \mathrm{Q}(A) = \int_A q(x)\,\mathrm{d}x, \qquad A \in \mathscr{B}(\mathbb{R}^n)$$

for the random variable $X(0)$. Since Q is a probability measure, the Langevin–Smoluchowski measure $\mathbb{Q}$ is a well-defined probability measure on the path space $\Omega = C([0,\infty);\mathbb{R}^n)$.

Let us recall now the likelihood ratio of (8). We denote the corresponding likelihood ratio process by $\vartheta(t) := \ell(t, X(t))$, $t \geqslant 0$. The following remarkable insight comes from Pavon [38] and Fontbona–Jourdain [19]: For any given $T \in (0,\infty)$, the time-reversed likelihood ratio process

$$\overline{\vartheta}(s) := \ell(T-s, \overline{X}(s)) = \frac{p(T-s, X(T-s))}{q(\overline{X}(s))}, \qquad 0 \leqslant s \leqslant T \tag{81}$$

is a $\mathbb{Q}$-martingale of the backwards filtration $\overline{\mathbb{G}} = (\overline{\mathcal{G}}(s))_{0\leqslant s\leqslant T}$ in (16). For a simple proof of this result in the setting of this paper we refer to [26, Appendix E].

Let us pick arbitrary times $0 \leqslant t_1 < t_2 < \infty$. For any given $T \in (t_2,\infty)$, the martingale property of the process (81) amounts, with $s_1 = T - t_1$ and $s_2 = T - t_2$, to

$$\mathbb{E}_{\mathbb{Q}}\Big[\overline{\vartheta}(s_1)\,|\,\overline{\mathcal{G}}(s_2)\Big] = \overline{\vartheta}(s_2) \quad \Longleftrightarrow \quad \mathbb{E}_{\mathbb{Q}}\Big[\vartheta(t_1)\,|\,\boldsymbol{\sigma}(X(\theta) \colon t_2 \leqslant \theta \leqslant T)\Big] = \vartheta(t_2).$$

Because $T \in (t_2,\infty)$ is arbitrary, this gives

$$\mathbb{E}_{\mathbb{Q}}\big[\vartheta(t_1)\,|\,\mathcal{H}(t_2)\big] = \vartheta(t_2), \qquad \mathcal{H}(t) := \boldsymbol{\sigma}(X(\theta) \colon t \leqslant \theta < \infty). \tag{82}$$

In other words, the likelihood ratio process $(\vartheta(t))_{t\geqslant 0}$ is a backwards $\mathbb{Q}$-martingale of the filtration $\mathbb{H} = (\mathcal{H}(t))_{t\geqslant 0}$. We denote by $\mathcal{H}(\infty) := \bigcap_{t\geqslant 0}\mathcal{H}(t)$ the tail $\boldsymbol{\sigma}$-algebra of the Langevin–Smoluchowski diffusion $(X(t))_{t\geqslant 0}$. The ergodicity of this process under the probability measure $\mathbb{Q}$ implies that the tail $\boldsymbol{\sigma}$-algebra $\mathcal{H}(\infty)$ is $\mathbb{Q}$-trivial, i.e., $\mathcal{H}(\infty) = \{\varnothing, \Omega\}$ modulo $\mathbb{Q}$; see Appendix 2 for a proof of this claim.

We recall now the martingale version of Theorem 9.4.7 (backwards submartingale convergence) in [10]. This says that $(\vartheta(t))_{t\geqslant 0}$ is a $\mathbb{Q}$-uniformly integrable family, that the limit

$$\vartheta(\infty) := \lim_{t\to\infty} \vartheta(t) \tag{83}$$

exists $\mathbb{Q}$-a.e., that the convergence in (83) holds also in $L^1(\mathbb{Q})$, and that for every $t \geqslant 0$ we have

$$\mathbb{E}_{\mathbb{Q}}\big[\vartheta(t)\,|\,\mathcal{H}(\infty)\big] = \vartheta(\infty), \qquad \mathbb{Q}\text{-a.e.} \tag{84}$$

But since the tail $\boldsymbol{\sigma}$-algebra $\mathcal{H}(\infty)$ is $\mathbb{Q}$-trivial, the random variable $\vartheta(\infty)$ is $\mathbb{Q}$-a.e. constant, and (84) identifies this constant as $\vartheta(\infty) \equiv 1$.

In terms of the function $f(x) := x\log x$ for $x > 0$ (and with $f(0) := 0$), we can express the relative entropy $H(P(t)\,|\,\mathrm{Q})$ as

$$H(P(t)\,|\,\mathrm{Q}) = \mathbb{E}_{\mathbb{P}}\big[\log \vartheta(t)\big] = \mathbb{E}_{\mathbb{Q}}\big[f(\vartheta(t))\big] \geqslant 0, \qquad t \geqslant 0. \tag{85}$$

The convexity of f, in conjunction with (82), shows that the process $(f(\vartheta(t)))_{t\geqslant 0}$ is a backwards $\mathbb{Q}$-submartingale of the filtration $\mathbb{H}$, with decreasing expectation as in (85). By appealing to the backwards submartingale convergence theorem [10, Theorem 9.4.7] once again, we deduce that $(f(\vartheta(t)))_{t\geqslant 0}$ is a $\mathbb{Q}$-uniformly integrable family, which converges, a.e. and in L^1 under $\mathbb{Q}$, to

$$\lim_{t\to\infty} f(\vartheta(t)) = f(\vartheta(\infty)) = f(1) = 0.$$

In particular,

$$\lim_{t\to\infty} \downarrow H(P(t)\,|\,\mathrm{Q}) = \lim_{t\to\infty} \mathbb{E}_{\mathbb{Q}}[f(\vartheta(t))] = \mathbb{E}_{\mathbb{Q}}[\lim_{t\to\infty} f(\vartheta(t))] = 0, \tag{86}$$

proving (77). From (86) and (79) it follows now that (80) holds as equality, proving (78). □

Appendix 2: The triviality of the tail σ-algebra $\mathcal{H}(\infty)$

We recall the filtered probability space $(\Omega, \mathcal{F}(\infty), \mathbb{F}, \mathbb{Q})$. Here, $\Omega = C([0,\infty);\mathbb{R}^n)$ is the path space of continuous functions, $\mathcal{F}(\infty) = \sigma(\bigcup_{t\geqslant 0}\mathcal{F}(t))$, the canonical filtration $\mathbb{F} = (\mathcal{F}(t))_{t\geqslant 0}$ is as in (15), and the Langevin–Smoluchowski measure $\mathbb{Q}$ is represented by

$$\mathbb{Q}(B) = \int_{\mathbb{R}^n} \mathbb{P}^x(B)\,\mathrm{dQ}(x), \qquad B \in \mathscr{B}(\Omega), \tag{87}$$

where $\mathbb{P}^x$ denotes the Langevin–Smoluchowski measure with initial distribution δ_x, for every $x \in \mathbb{R}^n$, and $\mathscr{B}(\Omega)$ is the Borel σ-field[2] on Ω.

For every $s \geqslant 0$, we define a measurable map $\theta_s : \Omega \to \Omega$, called *shift transformation*, by requiring that $\theta_s(\omega)(t) = \omega(s+t)$ hold for all $\omega \in \Omega$ and $t \geqslant 0$. A Borel set $B \in \mathscr{B}(\Omega)$ is called *shift-invariant* if $\theta_s^{-1}(B) = B$ holds for any $s \geqslant 0$. Since the Gibbs probability measure Q is the unique invariant measure for the Langevin–Smoluchowski diffusion $(X(t))_{t\geqslant 0}$, [5, Theorem 3.8] implies that the probability measure $\mathbb{Q}$ of (87) is *ergodic*, meaning that $\mathbb{Q}(B) \in \{0,1\}$ holds for every shift-invariant set B. As a consequence of the ergodicity of $\mathbb{Q}$, the *Birkhoff Ergodic Theorem* [5, Theorem 3.4] implies that, for every $A \in \mathscr{B}(\mathbb{R}^n)$, the limit

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{1}_A(X(s))\,\mathrm{d}s = \mathrm{Q}(A) \tag{88}$$

exists $\mathbb{Q}$-a.e.

Lemma 1 *The tail σ-algebra $\mathcal{H}(\infty)$ is $\mathbb{Q}$-trivial, i.e., $\mathcal{H}(\infty) = \{\varnothing, \Omega\}$ modulo $\mathbb{Q}$.*

---

[2] There are several equivalent ways to define the Borel σ-field on Ω. Two possible constructions appear in Problems 2.4.1 and 2.4.2 in [27].

Proof We follow a reasoning similar to that in [19, Remark 1.10]. According to [29, Theorem 1.3.9], it suffices to show that the Langevin–Smoluchowski diffusion $(X(t))_{t \geqslant 0}$ is *recurrent in the sense of Harris*, i.e.,

$$\mathbb{P}^x \Big[\int_0^\infty \mathbb{1}_A(X(s))\, \mathrm{d}s = \infty \Big] = 1 \tag{89}$$

is satisfied for every $x \in \mathbb{R}^n$ and all $A \in \mathscr{B}(\mathbb{R}^n)$ with $\mathrm{Q}(A) > 0$.

For the proof of (89), we fix $x \in \mathbb{R}^n$ and $A \in \mathscr{B}(\mathbb{R}^n)$ with $\mathrm{Q}(A) > 0$. By its definition, the event

$$B := \Big\{ \int_0^\infty \mathbb{1}_A(X(s))\, \mathrm{d}s = \infty \Big\}$$

is shift-invariant. Thus, by the ergodicity of $\mathbb{Q}$, the probability $\mathbb{Q}(B)$ is equal to either zero or one. Outside the set B, we have

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \mathbb{1}_A(X(s))\, \mathrm{d}s = 0.$$

But since $\mathrm{Q}(A) > 0$, the $\mathbb{Q}$-a.e. limit (88) implies that $\mathbb{Q}(B^{\mathrm{c}}) = 0$ and hence $\mathbb{Q}(B) = 1$. From the definition (87) of the probability measure $\mathbb{Q}$ it follows that $\mathbb{P}^x(B) = 1$ for Q-a.e. $x \in \mathbb{R}^n$. Since Q is equivalent to Lebesgue measure, we also have that $\mathbb{P}^x(B) = 1$ for Lebesgue-a.e. $x \in \mathbb{R}^n$.

Furthermore, the shift-invariance of B and the Markov property of the Langevin–Smoluchowski diffusion give

$$\mathbb{P}^x(B) = \mathbb{P}^x(\theta_t^{-1}(B)) = \mathbb{E}_{\mathbb{P}^x}[\mathbb{P}^{X(t)}(B)] = T_t(\mathbb{P}^x(B)) = \int_{\mathbb{R}^n} P(0,x;t,\mathrm{d}y)\, \mathbb{P}^y(B) \tag{90}$$

for every $t \geqslant 0$. Here, $P(0,x;t,y)$ denotes the transition kernel of the Langevin–Smoluchowski dynamics, so that $\mathbb{P}^x[X(t) \in \mathrm{d}y] = P(0,x;t,y)\,\mathrm{d}y$; and T_t denotes the operator

$$T_t f(x) := \int_{\mathbb{R}^n} P(0,x;t,\mathrm{d}y)\, f(y), \qquad (t,x) \in [0,\infty) \times \mathbb{R}^n$$

acting on bounded measurable functions $f \colon \mathbb{R}^n \to \mathbb{R}$. Since $(T_t)_{t \geqslant 0}$ is a strong Feller semigroup under the assumptions of this paper, the function $\mathbb{R}^n \ni x \mapsto T_t f(x)$ is continuous. Now (90) implies the continuity of the function $\mathbb{R}^n \ni x \mapsto \mathbb{P}^x(B)$. On the other hand, we have already seen that the function $\mathbb{R}^n \ni x \mapsto \mathbb{P}^x(B) \in [0,1]$ is Lebesgue-a.e. equal to one. But such a function is constant everywhere, i.e., $\mathbb{P}^x(B) = 1$ for every $x \in \mathbb{R}^n$, proving (89). □

Acknowledgements We are indebted to Walter Schachermayer for joint work that this paper complements, and we are grateful to Promit Ghosal for his interest in this work and for encouraging us to consider its iterative aspects. We would also like to thank Michele Pavon for pointing out relevant literature to us.

I. Karatzas acknowledges support from the National Science Foundation (NSF) under grant NSF-DMS-20-04997. B. Tschiderer acknowledges support by the Austrian Science Fund (FWF) under grant P28661 and by the Vienna Science and Technology Fund (WWTF) through project MA16-021.

References

1. L. Ambrosio and N. Gigli. A User's Guide to Optimal Transport. In *Modelling and Optimisation of Flows on Networks*, volume **2062** of *Lecture Notes in Math.*, pages 1–155. Springer, Berlin, Heidelberg, 2013.
2. L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lect. Math. ETH Zürich. Birkhäuser, Basel, second edition, 2008.
3. J. Backhoff, G. Conforti, I. Gentil, and C. Léonard. The mean field SchrÃűdinger problem: ergodic behavior, entropy estimates and functional inequalities. *Probab. Theory Relat. Fields*, **178**(1–2):475–530, 2020.
4. D. Bakry and M. Émery. Diffusions hypercontractives. In J. Azéma and M. Yor, editors, *Sémin. Probab. XIX 1983/84*, volume **1123** of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, Heidelberg, 1985.
5. L. Bellet. Ergodic Properties of Markov Processes. In S. Attal, A. Joye, and C.-A. Pillet, editors, *Open Quantum Systems II — The Markovian Approach*, volume **1881** of *Lecture Notes in Math.*, pages 1–39. Springer, Berlin, Heidelberg, 2006.
6. A. Blaquière. Controllability of a Fokker–Planck equation, the Schrödinger system, and a related stochastic optimal control (revised version). *Dyn. Control*, **2**(3):235–253, 1992.
7. Y. Chen, T. Georgiou, and M. Pavon. On the Relation Between Optimal Transport and Schrödinger Bridges: A Stochastic Control Viewpoint. *J. Optim. Theory Appl.*, **169**(2):671–691, 2016.
8. Y. Chen, T. Georgiou, and M. Pavon. Stochastic control, entropic interpolation and gradient flows on Wasserstein product spaces. In *Proceedings of the 22nd International Symposium on Mathematical Theory of Networks and Systems*, Minneapolis, MN, USA, July 11–15, 2016.
9. Y. Chen, T. Georgiou, and M. Pavon. Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge. arXiv e-print, arXiv:2005.10963, 2020.
10. K. Chung. *A Course in Probability Theory*. Academic Press, San Diego, third edition, 2001.
11. P. Dai Pra. A stochastic control approach to reciprocal diffusion processes. *Appl. Math. Optim.*, **23**(1):313–329, 1991.
12. P. Dai Pra and M. Pavon. Variational path-integral representations for the density of a diffusion process. *Stochastics*, **26**(4):205–226, 1989.
13. P. Dai Pra and M. Pavon. On the Markov processes of Schrödinger, the Feynman–Kac formula and stochastic control. In M. Kaashoek, J. van Schuppen, and A. Ran, editors, *Realization and Modelling in System Theory*, volume **3** of *Prog. Syst. Control Theory*, pages 497–504. Birkhäuser, Boston, 1990.
14. W. Fleming. Exit probabilities and optimal stochastic control. *Appl. Math. Optim.*, **4**(1):329–346, 1977.
15. W. Fleming. Logarithmic transformations and stochastic control. In W. Fleming and L. Gorostiza, editors, *Advances in Filtering and Optimal Stochastic Control*, volume **42** of *Lect. Notes Control Inf. Sci.*, pages 131–141. Springer, Berlin, Heidelberg, 1982.
16. W. Fleming and S.-J. Sheu. Stochastic variational formula for fundamental solutions of parabolic PDE. *Appl. Math. Optim.*, **13**(1):193–204, 1985.
17. H. Föllmer. An entropy approach to the time reversal of diffusion processes. In M. Métivier and É. Pardoux, editors, *Stochastic Differential Systems — Filtering and Control*, volume **69** of *Lect. Notes Control Inf. Sci.*, pages 156–163. Springer, Berlin, Heidelberg, 1985.

18. H. Föllmer. Time reversal on Wiener space. In S. Albeverio, P. Blanchard, and L. Streit, editors, *Stochastic Processes — Mathematics and Physics*, volume **1158** of *Lecture Notes in Math.*, pages 119–129. Springer, Berlin, Heidelberg, 1986.
19. J. Fontbona and B. Jourdain. A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. *Ann. Probab.*, **44**(1):131–170, 2016.
20. A. Friedman. *Stochastic Differential Equations and Applications — Volume 1*, volume **28** of *Probab. Math. Statist. Ser. Monogr. Textb.* Academic Press, New York, 1975.
21. C. Gardiner. *Stochastic Methods — A Handbook for the Natural and Social Sciences*, volume **13** of *Springer Ser. Synergetics*. Springer, Berlin, Heidelberg, fourth edition, 2009.
22. U. Haussmann and É. Pardoux. Time Reversal of Diffusions. *Ann. Probab.*, **14**(4):1188–1205, 1986.
23. C. Holland. A new energy characterization of the smallest eigenvalue of the Schrödinger equation. *Comm. Pure Appl. Math.*, **30**(6):755–765, 1977.
24. R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker–Planck Equation. *SIAM J. Math. Anal.*, **29**(1):1–17, 1998.
25. I. Karatzas, W. Schachermayer, and B. Tschiderer. A trajectorial approach to the gradient flow properties of Langevin–Smoluchowski diffusions. arXiv e-print, arXiv:2008.09220, 2020.
26. I. Karatzas, W. Schachermayer, and B. Tschiderer. Trajectorial Otto calculus. arXiv e-print, arXiv:1811.08686, 2020.
27. I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*, volume **113** of *Grad. Texts in Math.* Springer, New York, 1988.
28. A. Kolmogorov. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.*, **104**(1):415–458, 1931.
29. H. Kunita. *Stochastic flows and stochastic differential equations*, volume **24** of *Cambridge Stud. Adv. Math.* Cambridge Univ. Press, Cambridge, 1997.
30. C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, **34**(4):1533–1574, 2014.
31. C. Léonard. Some Properties of Path Measures. In C. Donati-Martin, A. Lejay, and A. Rouault, editors, *Sémin. Probab. XLVI*, volume **2123** of *Lecture Notes in Math.*, pages 207–230. Springer International Publishing, Switzerland, 2014.
32. P. Markowich and C. Villani. On the trend to equilibrium for the Fokker–Planck equation: an interplay between physics and functional analysis. In *Sixth Workshop on Partial Differential Equations — Part II*, volume **19** of *Mat. Contemp.*, pages 1–29. Soc. Brasil. Mat., Rio de Janeiro, 2000.
33. P. Meyer. Sur une transformation du mouvement brownien due à Jeulin et Yor. In J. Azéma, M. Yor, and P. Meyer, editors, *Sémin. Probab. XXVIII*, volume **1583** of *Lecture Notes in Math.*, pages 98–101. Springer, Berlin, Heidelberg, 1994.
34. M. Nagasawa. *Schrödinger Equations and Diffusion Theory*, volume **86** of *Monogr. Math.* Birkhäuser, Basel, 1993.
35. E. Nelson. *Dynamical Theories of Brownian motion*. Math. Notes. Princeton Univ. Press, Princeton, NJ, second edition, 2001.
36. F. Otto and C. Villani. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *J. Funct. Anal.*, **173**(2):361–400, 2000.
37. É. Pardoux. Grossissement d'une filtration et retournement du temps d'une diffusion. In J. Azéma and M. Yor, editors, *Sémin. Probab. XX 1984/85*, volume **1204** of *Lecture Notes in Math.*, pages 48–55. Springer, Berlin, Heidelberg, 1986.
38. M. Pavon. Stochastic control and nonequilibrium thermodynamical systems. *Appl. Math. Optim.*, **19**(1):187–202, 1989.
39. M. Pavon and A. Wakolbinger. On Free Energy, Stochastic Control, and SchrÃűdinger Processes. In G. Di Masi, A. Gombani, and A. Kurzhansky, editors, *Modeling, Estimation and Control of Systems with Uncertainty*, volume **10** of *Prog. Syst. Control Theory*, pages 334–348. Birkhäuser, Boston, MA, 1991.
40. D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume **293** of *Grundlehren Math. Wiss.* Springer, Berlin, Heidelberg, third edition, 1999.

41. H. Risken. *The Fokker–Planck Equation — Methods of Solution and Applications*, volume **18** of *Springer Ser. Synergetics*. Springer, Berlin, Heidelberg, second edition, 1996.
42. L. Rogers. Smooth Transition Densities for OneâĂŘDimensional Diffusion. *Bull. Lond. Math. Soc.*, **17**(2):157–161, 1985.
43. Z. Schuss. Singular Perturbation Methods in Stochastic Differential Equations of Mathematical Physics. *SIAM Rev.*, **22**(2):119–155, 1980.
44. X. Tan and N. Touzi. Optimal transportation under controlled stochastic dynamics. *Ann. Probab.*, **41**(5):3201–3240, 2013.
45. C. Villani. *Topics in Optimal Transportation*, volume **58** of *Grad. Stud. Math.* Amer. Math. Soc., Providence, RI, 2003.

Incomplete Stochastic Equilibria with Exponential Utilities Close to Pareto Optimality

Constantinos Kardaras, Hao Xing, and Gordan Žitković

Abstract We study existence and uniqueness of continuous-time stochastic Radner equilibria in an incomplete markets model. An assumption of "smallness" type—imposed through the new notion of "closeness to Pareto optimality"—is shown to be sufficient for existence and uniqueness. Central role in our analysis is played by a fully-coupled nonlinear system of quadratic BSDEs.

Introduction

The equilibrium problem

The focus of the present paper is the problem of existence and uniqueness of a competitive (Radner) equilibrium in an incomplete continuous-time stochastic model of a financial market. A discrete version of our model was introduced by Radner in [26] as an extension of the classical Arrow-Debreu framework, with the goal of understanding how asset prices in financial (or any other) markets are formed, under minimal assumption on the ingredients or the underlying market structure. One of those assumptions is often market completeness; more precisely, it is usually postulated that the range of various types of transactions the markets allow is such that the wealth distribution among agents, after all the trading is done, is Pareto optimal, i.e., that no further redistribution of wealth can make one agent better off

Constantinos Kardaras
Department of Statistics, London School of Economics and Political Science, e-mail: k.kardaras@lse.ac.uk

Hao Xing
Department of Finance, Questrom School of Business, Boston University, e-mail: haoxing@bu.edu

Gordan Žitković
Department of Mathematics, University of Texas at Austin, e-mail: gordanz@math.utexas.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_11

without hurting somebody else. Real markets are not complete; in fact, as it turns out, the precise way in which completeness fails matters greatly for the output and should be understood as an a-priori constraint. Indeed, it is instructive to ask the following questions: Why are markets incomplete in the first place? Would rational economic agents not continue introducing new assets into the market, as long as it is still useful? The answer is that they, indeed, would, were it not for exogenously-imposed constraints out there, no markets exist for most contingencies; those markets that do exist are heavily regulated, transactions costs are imposed, short selling is sometimes prohibited, liquidity effects render replication impossible, etc. Instead of delving into the modeling issues regarding various types of *completeness constraints*, we point the reader to [31] where a longer discussion of such issues can be found.

The "fast-and-slow" model

The particular setting we subscribe to here is one of the simplest from the financial point of view. It, nevertheless, exhibits many of the interesting features found in more general incomplete structures and admits a straightforward continuous-time formulation. It corresponds essentially to the so-called "fast-and-slow" completeness constraint, introduced in [31].

One of the ways in which the "fast-and-slow" completeness constraint can be envisioned is by allowing for different speeds at which information of two different kinds is incorporated and processed. The discrete-time version of the model is described in detail in [25, p. 213], where it goes under the heading of "short-lived" asset models. Therein, at each node in the event tree, the agents have access to a number of short-lived assets, i.e., assets whose life-span ends in one unit of time, at which time all the dividends are distributed. The prices of such assets are determined in the equilibrium, but their number is typically not sufficient to guarantee local (and therefore global) completeness of the market. In our, continuous time model, the underlying filtration is generated by two independent Brownian motions (B and W). Positioned the "node" (ω, t), we think of dB_t and dW_t as two independent symmetric random variables, realized at time $t + dt$, with values $\pm\sqrt{dt}$. Allowing the agents to insure each other only with respect to the risks contained in dB, we denote the (equilibrium) price of such an "asset" by $-\lambda_t\, dt$. As already hinted to above, one possible economic rationale behind this type of constraint is obtained by thinking of dB as the readily-available (fast) information, while dW models slower information which will be incorporated into the process λ_t indirectly, and only at later dates. For simplicity, we also fix the spot interest rate to 0, allowing agents to transfer wealth from t to $t + dt$ costlessly and profitlessly. While, strictly speaking, this feature puts us in the partial-equilibrium framework, this fact will not play a role in our analysis, chiefly because our agents draw their utility only from the terminal wealth (which is converted to the consumption good at that point).

For mathematical convenience, and to be able to access the available continuous-time results, we concatenate all short-lived assets with payoffs dB_t and prices $-\lambda_t\, dt$

into a single asset $B_t^\lambda = B_t + \int_0^t \lambda_u \, du$. It should not be thought of as an asset that carries a dividend at time T, but only as a single-object representation of the family of all infinitesimal, short-lived assets.

As a context for the ”fast-and-slow” constraint, we consider a finite number I of agents; we assume that all of their utility functions are of exponential type, but allow for idiosyncratic risk-aversion parameters and non-traded random endowments. The exponential nature of the agents' utilities is absolutely crucial for all of our results as it induces a “backward” structure to our problem, which, while still very difficult to analyze, allows us to make a significant step forward.

The representative-agent approach, and its failure in incomplete markets

The classical and nearly ubiquitous approach to existence of equilibria in complete markets is using the so-called representative-agent approach. Here, the agents' endowments are first aggregated and then split in a Pareto-optimal way. Along the way, a pricing measure is produced, and then, a-posteriori, a market is constructed whose unique martingale measure is precisely that particular pricing measure. As long as no completeness constraints are imposed, this approach works extremely well, pretty much independently of the shape of the agents' utility functions (see, e.g., [14, 13, 18, 19, 20, 9, 1, 30] for a sample of continuous-time literature). A convenient exposition of some of these and many other results, together with a thorough classical literature overview can be found in the Notes section of Chapter 4. of [21]).

The incomplete case requires a completely different approach and what were once minute details, now become salient features. The failure of representative-agent methods under incompleteness are directly related to the inability of the market to achieve Pareto optimality by wealth redistribution. Indeed, when not every transaction can be implemented through the market, one cannot reduce the search for the equilibrium to a finite-dimensional “manifold” of Pareto-optimal allocations. Even more dramatically, the whole nature of what is considered a solution to the equilibrium problem changes. In the complete case, one simply needs to identify a market-clearing valuation measure. In the present “fast-and-slow” formulation, the very family of all replicable claims (in addition to the valuation measure) has to be determined. This significantly impacts the “dimensionality” of the problem and calls for a different toolbox.

Our probabilistic-analytic approach

The direction of the present paper is partially similar to that of [31], where a much simpler model of the “fast-and-slow” type is introduced and considered. Here, however, the setting is different and somewhat closer to [29] and [8]. The fast component is modeled by an independent Brownian motion, instead of the one-jump process. Also, unlike in any of the above papers, pure PDE techniques are largely replaced or supplemented by probabilistic ones, and much stronger results are obtained.

Doing away with the Markovian assumption, we allow for a collection of unbounded random variables, satisfying suitable integrability assumptions, to act as random endowments and characterize the equilibrium as a (functional of a) solution to a nonlinear system of quadratic Backward Stochastic Differential Equations (BSDE). Unlike single quadratic BSDE, whose theory is by now quite complete (see e.g., [23, 5, 6, 12, 15, 3] for a sample), the systems of quadratic BSDEs are much less understood. The main difficulty is that the comparison theorem may fail to hold for BSDE systems (see [17]). Moreover, Frei and dos Reis (see [16]) constructed a quadratic BSDE system which has bounded terminal condition but admits no solution. The strongest general-purpose result seems to be the one of Tevzadze (see [28]), which guarantees existence under an "$\mathbb{L}^\infty$-smallness" condition placed on the terminal conditions.

Like in [28], but unlike in [31] or [8], our general result imposes no regularity conditions on the agents' random endowments. Unlike [28], however, our smallness conditions come in several different forms. First, we show existence and uniqueness when the random-endowment allocation among agents is close to a Pareto optimal one. In contrast to [28], we allow here for unbounded terminal conditions (random endowments), and measure their size using an "entropic" BMO-type norm strictly weaker than the $\mathbb{L}^\infty$-norm. In addition, the equilibrium established is unique in a global sense (as in [24], where a different quadratic BSDE system is studied).

Another interesting feature of our general result is that it is largely independent of the number of agents. This leads to the following observation: the equilibrium exists as soon as "sufficiently many sufficiently homogeneous" (under an appropriate notion of homogeneity) agents share a given total endowment, which is not assumed to be small. This is precisely the natural context of a number of competitive equilibrium models with a large number of small agents, none of whom has a dominating sway over the price.

Another parameter our general result is independent of is the time horizon T. Indirectly, this leads to our second existence and uniqueness result which holds when the time horizon is sufficiently small, but the random endowments are not limited in size. Under the additional assumption of Malliavin differentiabilty, a lower bound on how small the horizon has to be to guarantee existence and uniqueness turns out to be inversely proportional to the size of the (Malliavin) derivatives of random endowments. This extends [8, Theorem 3.1] to a non-Markovian setting. Interestingly, both the $\mathbb{L}^\infty$-smallness of the random endowments and the smallness of the time-horizon are implied by the small-entropic-BMO-norm condition mentioned above, and the existence theorems under these conditions can be seen as special cases of our general result.

Some notational conventions

As we will be dealing with various classes of vector-valued random variables and stochastic processes, we try to introduce sufficiently compact notation to make reading more palatable.

A time horizon $T > 0$ is fixed throughout. An equality sign between random variables signals almost-sure equality, while one between two processes signifies Lebesgue-almost everywhere, almost sure equality; any two processes that are equal in this sense will be identified; this, in particular, applied to indistinguishable càdlàg processes. Given a filtered probability space $(\Omega, \mathcal{F}_T, \mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ satisfying the usual conditions, $\mathcal{T}$ denotes the set of all $[0,T]$-valued $\mathbb{F}$-stopping times, and $\mathcal{P}^2$ denotes the set of all predictable processes $\{\mu_t\}_{t\in[0,T]}$ such that $\int_0^T \mu_t^2\, dt < \infty$, a.s. The integral $\int_0^\cdot \mu_u\, dB_u$ of $\mu \in \mathcal{P}^2$ with respect to an $\mathbb{F}$-Brownian motion B is alternatively denoted by $\mu \cdot B$, while the stochastic (Doléans-Dade) exponential retains the standard notation $\mathcal{E}(\cdot)$. The $\mathbb{L}^p$-spaces, $p \in [1,\infty]$ are all defined with respect to $(\Omega, \mathcal{F}_T, \mathbb{P})$ and $\mathbb{L}^0$ denotes the set of ($\mathbb{P}$-equivalence classes) of finite-valued random variables on this space. For a continuous adapted process $\{Y_t\}_{t\in[0,T]}$, we set

$$||Y||_{\mathcal{S}^\infty} = || \sup_{t\in[0,T]} |Y_t| \,||_{\mathbb{L}^\infty},$$

and denote the space of all such Y with $||Y||_{\mathcal{S}^\infty} < \infty$ by $\mathcal{S}^\infty$. For $p \geq 1$, the space of all $\mu \in \mathcal{P}^2$ with $||\mu||_{H^p}^p = \mathbb{E}\left[\int_0^T |\mu_u|^p\, du\right] < \infty$ is denoted by H^p, an alias for the Lebesgue space $\mathbb{L}^p$ on the product $[0,T] \times \Omega$.

Given a probability measure $\hat{\mathbb{P}}$ and a $\hat{\mathbb{P}}$-martingale M, we define its BMO-norm by

$$||M||^2_{\mathrm{BMO}(\hat{\mathbb{P}})} = \sup_{\tau\in\mathcal{T}} \left\| \mathbb{E}_\tau^{\hat{\mathbb{P}}}[\langle M\rangle_T - \langle M\rangle_\tau] \right\|_{\mathbb{L}^\infty},$$

where $\mathbb{E}_\tau^{\hat{\mathbb{P}}}[\cdot]$ denotes the conditional expectation $\mathbb{E}^{\hat{\mathbb{P}}}[\cdot|\mathcal{F}_\tau]$ with respect to $\mathcal{F}_\tau$, computed under $\hat{\mathbb{P}}$. The set of all $\hat{\mathbb{P}}$-martingales M with finite $||M||_{\mathrm{BMO}(\hat{\mathbb{P}})}$ is denoted by $\mathrm{BMO}(\hat{\mathbb{P}})$, or, simply, BMO, when $\hat{\mathbb{P}} = \mathbb{P}$. When applied to random variables, $X \in \mathrm{BMO}(\hat{\mathbb{P}})$ means that $X = M_T$, for some $M \in \mathrm{BMO}(\hat{\mathbb{P}})$. In the same vein, we define (for some, and then any, $(\hat{\mathbb{P}}, \mathbb{F})$-Brownian motion B)

$$\mathrm{bmo}(\hat{\mathbb{P}}) = \{\mu \in \mathcal{P}^2 \,:\, \mu \cdot B \in \mathrm{BMO}(\hat{\mathbb{P}})\},$$

with the norm $||\mu||_{\mathrm{bmo}(\hat{\mathbb{P}})} = ||\mu \cdot B||_{\mathrm{BMO}(\hat{\mathbb{P}})}$. The same convention as above is used: the dependence on $\hat{\mathbb{P}}$ is suppressed when $\hat{\mathbb{P}} = \mathbb{P}$.

Many of our objects will take values in $\mathbb{R}^I$, for some fixed $I \in \mathbb{N}$. Those are typically denoted by bold letters such as $\boldsymbol{E}, \boldsymbol{G}, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\alpha}$, etc. If specific components are needed, they will be given a superscript - e.g., $\boldsymbol{G} = (G^i)_i$. Unquantified variables i, j always range over $\{1, 2, \ldots, I\}$. The topology of $\mathbb{R}^k$ is induced by the Euclidean norm $|\cdot|_2$, defined by $|\boldsymbol{x}|_2 = \sqrt{\sum_k |x^k|^2}$ for $\boldsymbol{x} \in \mathbb{R}^k$. All standard operations and relations (including the absolute value $|\cdot|$ and order $\leq$) between $\mathbb{R}^k$-valued variables are considered componentwise.

1 The Equilibrium Problem and its BSDE Reformulation

We work on a filtered probability space $(\Omega, \mathcal{F}_T, \mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$, where $\mathbb{F}$ is the standard augmentation of the filtration generated by a two-dimensional standard Brownian motion $\{(B_t, W_t)\}_{t\in[0,T]}$. The augmented natural filtrations $\mathbb{F}^B$ and $\mathbb{F}^W$ of the two Brownian motions B and W will also be considered below.

1.1 The financial market, its agents, and equilibria

Our model of a financial market features one liquidly traded **risky asset**, whose value, denoted in terms of a prespecified numéraire which we normalize to 1, is given by

$$dB^\lambda_t = \lambda_t\, dt + dB_t, \quad t \in [0,T], \tag{1}$$

for some $\lambda \in \mathcal{P}^2$. Given that it will play a role of a “free parameter” in our analysis, the volatility in (1) is normalized to 1; this way, λ can simultaneously be interpreted as the **market price of risk**. The reader should consult the section ‘The “fast-and-slow” model’ in the introduction for the proper economic interpretation of this asset as a concatenation of a continuum of infinitesimally-short-lived securities.

We assume there is a finite number $I \in \mathbb{N}$ of **economic agents**, all of whom trade the risky asset as well as the aforementioned riskless, numéraire, asset of constant value 1. The preference structure of each agent is modeled in the von Neumann-Morgenstern framework via the following two elements:

i) an exponential **utility function** with **risk tolerance coefficient** $\delta^i > 0$:

$$U^i(x) = -\exp(-x/\delta^i), \quad x \in \mathbb{R}, \text{ and}$$

ii) a **random endowment** $E^i \in \mathbb{L}^0(\mathcal{F}_T)$.

The pair $(\boldsymbol{E}, \boldsymbol{\delta})$, where $\boldsymbol{E} = (E^i)_i$, $\boldsymbol{\delta} = (\delta^i)_i$, of endowments and risk-tolerance coefficients fully characterizes the behavior of the agents in the model; we call it the **population characteristics—$\boldsymbol{E}$** is the **initial allocation** and $\boldsymbol{\delta}$ the **risk profile**. In general, any $\mathbb{R}^I$-valued random vector will be refereed to as an **allocation**.

Each agent maximizes the expected utility of trading and random endowment:

$$\mathbb{E}\left[U^i(\pi \cdot B^\lambda_T + E^i)\right] \to \max. \tag{2}$$

Here $\{\pi_t\}_{t\in[0,T]}$ is a one-dimensional process which represents the number of shares of the asset kept by the agent at time t. As usual, this strategy is financed by investing in or borrowing from the interestless numéraire asset, as needed. To describe the admissible strategies of the agent, we follow the convention in [11]:

For $\lambda \in \mathcal{P}^2$, we denote by $\mathcal{M}_a^\lambda$ the set of absolutely continuous local martingale measures for B^λ, i.e., all probability measures $\mathbb{Q} \ll \mathbb{P}$ such that $\mathbb{E}^{\mathbb{Q}}[h(B^\lambda_\tau - B^\lambda_\sigma)] = 0$ for all pairs of stopping times $\sigma \le \tau \le T$ and for all bounded $\mathcal{F}_\sigma$-measurable random variables h. For a probability measure $\mathbb{Q} \ll \mathbb{P}$, let $H(\mathbb{Q}|\mathbb{P})$ be the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$, i.e., $H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\right] \ge 0$. For $\lambda \in \mathcal{P}^2$ such that $\mathcal{M}^\lambda \neq \emptyset$, where

$$\mathcal{M}^\lambda = \{\mathbb{Q} \in \mathcal{M}_a^\lambda \,|\, H(\mathbb{Q}|\mathbb{P}) < \infty\},$$

a strategy π is said to be **λ-admissible** if $\pi \in \mathcal{A}^\lambda$, where

$$\mathcal{A}^\lambda = \left\{\pi \in \mathcal{P}^2 \,|\, \pi \cdot B^\lambda \text{ is a } \mathbb{Q}\text{-martingale for all } \mathbb{Q} \in \mathcal{M}^\lambda\right\}.$$

We note that the set $\mathcal{A}^\lambda$ corresponds - up to finiteness of the utility - exactly to the set Θ_2 in [11]. This admissible class contains, in particular, all $\pi \in \mathcal{P}^2$ such that $\pi \cdot B^\lambda$ is bounded (uniformly in t and ω).

Definition 1 (Equilibrium)

Given a population with characteristics $(\boldsymbol{E}, \boldsymbol{\delta})$, a process $\lambda \in \mathcal{P}^2$ with $\mathcal{M}^\lambda \neq \emptyset$ is called an **equilibrium (market price of risk)** if there exists an I-tuple $(\pi^i)_i$ such that

i) each π^i is an *optimal strategy* for the agent i under λ, i.e.

$$\pi^i \in \operatorname{argmax}_{\pi \in \mathcal{A}^\lambda} \mathbb{E}\left[U^i(\pi \cdot B^\lambda_T + E^i)\right],$$

ii) the *market clears*, i.e., $\sum_i \pi^i = 0$.

The set of all equilibria is denoted by $\Lambda_{\boldsymbol{\delta}}(\boldsymbol{E}, \mathbb{P})$, or simply, $\Lambda_{\boldsymbol{\delta}}(\boldsymbol{E})$, when the the probability $\mathbb{P}$ is clear from the context.

Remark 1 The assumptions on the agents' random endowments that we introduce below and the proof techniques we employ make it clear that bmo is a natural space to search for equilibria in. There is, however, no compelling economic argument to include bmo into the *definition* of an equilibrium, so we do not. It turns out, nevertheless, that whenever an equilibrium λ is mentioned in the rest of the paper it will be in the bmo context, and we will assume automatically that any equilibrium market price of risk belongs to bmo. In particular, all uniqueness statements we make will be with respect to bmo as the ambient space.

1.2 A simple risk-aware reparametrization

It turns out that a simple reparametrization in our "ingredient space" leads to substantial notational simplification. It also sheds some light on the economic meaning of various objects. The main idea is to think of the risk- tolerance coefficients as numéraires, as they naturally carry the same currency units as wealth. When expressed in risk-tolerance units, the random endowments and strategies become unitless and we introduce the following notation

$$\boldsymbol{G} = \tfrac{1}{\boldsymbol{\delta}}\boldsymbol{E}, \text{ i.e., } G^i = \tfrac{1}{\delta^i}E^i, \quad \text{and} \quad \boldsymbol{\rho} = \tfrac{1}{\boldsymbol{\delta}}\boldsymbol{\pi}, \text{ i.e., } \rho^i = \tfrac{1}{\delta^i}\pi^i. \tag{3}$$

Since $\mathcal{A}^\lambda$ is invariant under this reparametrization, the equilibrium conditions become

$$\rho^i \in \operatorname{argmax}_{\rho\in\mathcal{A}_i^\lambda} \mathbb{E}\left[U(\rho\cdot B_T^\lambda + G^i)\right] \quad \text{and} \quad \textstyle\sum_i \alpha^i\rho^i = 0, \tag{4}$$

where $U(x) = -\exp(-x)$, and $\alpha^i = \delta^i/(\sum_j \delta^j) \in (0,1)$ - with $\sum_i \alpha^i = 1$ - are the **(relative) weights** of the agents. The set of all equilibria with risk-denominated random endowments $\boldsymbol{G} = (G^i)_i$ and relative weights $\boldsymbol{\alpha} = (\alpha^i)_i$ is denoted by $\Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G},\mathbb{P})$ (this notation overload should not cause any confusion in the sequel).

Since the market-clearing condition in (4) now involves the relative weights α^i as "conversion rates", it is useful to introduce the **aggregation operator** $A : \mathbb{R}^I \to \mathbb{R}$ by

$$A[\boldsymbol{x}] = \textstyle\sum_i \alpha^i x^i, \quad \text{for } \boldsymbol{x} \in \mathbb{R}^I, \tag{5}$$

so that the market-clearing condition now simply reads $A[\boldsymbol{\rho}] = 0$, pointwise.

1.3 A solution of the single-agent utility-maximization problem

Before we focus on the questions of existence and uniqueness of an equilibrium, we start with the single agent's optimization problem. Here we suppress the index i and first introduce an assumptions on the risk-denominated random endowment:

$$G \text{ is bounded from above and } G \in \text{EBMO}, \tag{6}$$

where EBMO denotes the set of all $G \in \mathbb{L}^0$ for which there exists (necessarily unique) processes m^G and n^G in bmo, as well a constant X_0^G, such that $G = X_T^G$, where

$$X_t^G = X_0^G + \int_0^t m_u^G\, dB_u + \int_0^t n_u^G\, dW_u + \tfrac{1}{2}\int_0^t \left((m_u^G)^2 + (n_u^G)^2\right) du. \tag{7}$$

The supermartingale X^G admits the following representation

$$X_t^G = -\log\mathbb{E}_t[\exp(-G)], \text{ so that } U(X_t^G) = \mathbb{E}_t[U(G)] \text{ for } t \in [0,T], \tag{8}$$

and can be interpreted as the certainty-equivalent process (without access to the market) of G, expressed in the units of risk tolerance.

Remark 2

1. When G is bounded from above, as we require it to be in (6), a sufficient condition for $G \in \mathrm{EBMO}$ is $e^{-G} \in \mathrm{BMO}$. This follows directly from the boundedness of the (exponential) martingale $e^{-X^G_t}$ away from zero.
2. The condition (6) amounts to the membership $M^G \in \mathrm{BMO}$, where $M^G = m^G \cdot B + n^G \cdot W$. Then $-M^G \in \mathrm{BMO}$ and, by Theorem 3.1, p. 54 in [22], $\mathcal{E}(-M^G)$ satisfies the reverse Hölder inequality (R_p) with some $p > 1$. Therefore, for $\varepsilon < p-1$, we have

$$\mathbb{E}[e^{-(1+\varepsilon)G}] = \mathbb{E}[e^{-(1+\varepsilon)(X^G_0 + M^G_T + \frac{1}{2}\langle M^G\rangle_T)}]$$
$$= e^{-(1+\varepsilon)X^G_0}\,\mathbb{E}\left[\left(\mathcal{E}(-M^G)_T\right)^{1+\varepsilon}\right] < \infty.$$

On the other hand, by (1) above, we clearly have $\mathbb{L}^\infty \subseteq \mathrm{EBMO}$, so

$$G \in \mathbb{L}^\infty \Rightarrow G \in \mathrm{EBMO} \Rightarrow \mathbb{E}[e^{-(1+\varepsilon)G}] < \infty \text{ for some } \varepsilon > 0.$$

In particular our condition (6), while implied by the boundedness of G, itself implies the conditions $G^+ = \max\{G, 0\} \in \mathbb{L}^\infty$, $e^{-G} \in \cup_{p>1}\mathbb{L}^p$, imposed in [11].

We recall in Proposition 1 some results about the nature of the optimal solution to the utility-maximization problem (2) from [11]; the proof if given in Section 3 below.

Proposition 1 (Single agent's optimization problem: existence and duality)
Suppose that $\lambda \in bmo$ and that G satisfies (6). *Then both primal and dual problems have finite values and the following statements hold:*

1. *There exists a unique $\rho^{\lambda,G} \in \mathcal{A}^\lambda$ such that*

$$\rho^{\lambda,G} \in \operatorname*{argmax}_{\rho \in \mathcal{A}^\lambda} \mathbb{E}\left[U(\rho \cdot B^\lambda_T + G)\right].$$

2. *There exists a unique $\mathbb{Q}^{\lambda,G} \in \mathcal{M}^\lambda$ such that*

$$\mathbb{Q}^{\lambda,G} \in \operatorname*{argmin}_{\mathbb{Q} \in \mathcal{M}^\lambda}(H(\mathbb{Q}|\mathbb{P}) + \mathbb{E}^{\mathbb{Q}}[G]).$$

3. *There exists a constant $c^{\lambda,G}$ such that*

$$c^{\lambda,G} + \rho^{\lambda,G} \cdot B^\lambda_T + G = -\log(Z^{\lambda,G}_T), \text{ where } Z^{\lambda,G}_T = \frac{d\mathbb{Q}^{\lambda,G}}{d\mathbb{P}}. \tag{9}$$

The process $\rho^{\lambda,G}$ and the probability measure $\mathbb{Q}^{\lambda,G}$ are called the **primal** and the **dual optimizers**, respectively. While they were first obtained by convex-duality methods, they also admit a BSDE representation (see, e.g., [27]), where a major role is played by (the risk-denominated version) of the so-called **certainty-equivalent process**:

$$Y^{\lambda,G}_t = U^{-1}\left(\mathbb{E}_t\left[U(\rho^{\lambda,G} \cdot B^\lambda_T - \rho^{\lambda,G} \cdot B^\lambda_t + G)\right]\right), \quad t \in [0,T]. \tag{10}$$

The optimality of $\rho^{\lambda,G}$ implies that

$$U(Y_t^{\lambda,G}) = \operatorname*{esssup}_{\rho\in\mathcal{A}^\lambda} \mathbb{E}_t\left[U(\rho\cdot B_T^\lambda - \rho\cdot B_t^\lambda + G)\right],\quad t\in[0,T]. \tag{11}$$

Hence $Y_t^{\lambda,G}$ can be interpreted as the risk-denominated certainty equivalent of the agent i, when he/she trades optimally from t onwards, starting from no wealth. Finally, with

$$Z_t^{\lambda,G} = \mathbb{E}_t\left[\frac{d\mathbb{Q}^{\lambda,G}}{d\mathbb{P}}\right] = \mathcal{E}(-\lambda\cdot B - \nu^{\lambda,G}\cdot W)_t,\ t\in[0,T] \text{ for some } \nu^{\lambda,G}\in\mathcal{P}^2, \tag{12}$$

we have the following BSDE characterization for single agent's optimization problem.

Lemma 1 (Single agent's optimization problem: a BSDE characterization)
For $\lambda\in bmo$ and G satisfying (6)*, let $Y^{\lambda,G}$ be as in* (10)*, let $\mu^{\lambda,G}=\lambda-\rho^{\lambda,G}$ and let $\nu^{\lambda,G}$ be defined by* (12)*. Then the triplet $(Y^{\lambda,G},\mu^{\lambda,G},\nu^{\lambda,G})$ is the unique solution to the BSDE*

$$dY_t = \mu_t\,dB_t + \nu_t\,dW_t + \left(\tfrac12\nu_t^2 - \tfrac12\lambda_t^2 + \lambda_t\mu_t\right)dt,\quad Y_T = G, \tag{13}$$

in the class where $(\mu,\nu)\in bmo$. Such a unique solution also satisfies $Y^{\lambda,G}-X^G\in\mathcal{S}^\infty$.

Given the results of Propositions 1 and 1 above, we fix the notation $Y^{\lambda,G}$, $\mu^{\lambda,G}$, $\nu^{\lambda,G}$, $\mathbb{Q}^{\lambda,G}$, $Z^{\lambda,G}$ and $\rho^{\lambda,G}$ for λ and G. We also introduce the vectorized versions $\boldsymbol{Y}^{\lambda,\boldsymbol{G}}$, $\boldsymbol{\mu}^{\lambda,\boldsymbol{G}}$, $\boldsymbol{\nu}^{\lambda,\boldsymbol{G}}$, $\mathbb{Q}^{\lambda,\boldsymbol{G}}$, and $\boldsymbol{Z}^{\lambda,\boldsymbol{G}}$, so that, e.g., $\boldsymbol{\mu}^{\lambda,\boldsymbol{G}} = (\mu^{\lambda,G^i})_i$ and $\boldsymbol{G}=(G^i)_i$.

1.4 A BSDE characterization of equilibria

The BSDE-based description in Lemma 1 of the solution of a single agent's optimization problem is the main ingredient in the following characterization, whose proof is given in Subsection 3.3 below. We use the risk-aware parametrization introduced in Subsection 1.2, and remind the reader that $\Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G})$ denotes the set of all equilibria in bmo when $\boldsymbol{G}=(G^i)_i$ are the agents' risk-denominated random endowments and $\boldsymbol{\alpha}=(\alpha^i)_i$ are the relative weights.

Theorem 1 (BSDE characterization of equilibria)
For a process $\lambda\in bmo$, and an allocation $\boldsymbol{G}$ which satisfies (6) *componentwise, the following are equivalent:*

1. *$\lambda\in\Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G})$, i.e., λ is an equilibrium for the population $(\boldsymbol{G},\boldsymbol{\alpha})$.*
2. *$\lambda = A[\boldsymbol{\mu}]$ for some solution $(\boldsymbol{Y},\boldsymbol{\mu},\boldsymbol{\nu})$ of the BSDE system:*

$$d\boldsymbol{Y}_t = \boldsymbol{\mu}_t\,dB_t + \boldsymbol{\nu}_t\,dW_t + \left(\tfrac12\boldsymbol{\nu}_t^2 - \tfrac12 A[\boldsymbol{\mu}_t]^2 + A[\boldsymbol{\mu}_t]\boldsymbol{\mu}_t\right)dt,\quad \boldsymbol{Y}_T=\boldsymbol{G}, \tag{14}$$

with $(\boldsymbol{\mu},\boldsymbol{\nu})\in bmo^{\boldsymbol{I}}$.

Remark 3

1. Spelled out "in coordinates", the system (14) becomes

$$\begin{cases} dY_t^i = \mu_t^i\, dB_t + \nu_t^i\, dW_t + \left(\tfrac{1}{2}(\nu_t^i)^2 - \tfrac{1}{2}(\sum_j \alpha^j \mu_t^j)^2 + (\sum_j \alpha^j \mu_t^j)\mu_t^i\right) dt, \\ Y_T^i = G^i,\ i \in \{1, 2, \ldots, I\}, \end{cases} \tag{14}$$

   and the market-clearing condition $\lambda = A[\boldsymbol{\mu}_t]$ reads $\lambda = \sum_j \alpha^j \mu^j$.
2. While quite meaningless from the competitive point of view, in the case $I = 1$ of the above characterization still admits a meaningful interpretation. The notion of an equilibrium here corresponds to the choice of λ under which an agent, with risk-denominated random endowment $G \in$ EBMO would choose not to invest in the market at all. The system (14) reduces to a single equation

$$dY_t = \mu_t\, dB_t + \nu_t\, dW_t + (\tfrac{1}{2}\mu_t^2 + \tfrac{1}{2}\nu_t^2)\, dt,\ Y_T = G,$$

   which admits a unique solution, namely $Y = X^G$, so that $\lambda = m^G$ is the unique equilibrium. This case also singles out the space EBMO as the natural environment for the random endowments G^i in this context.

2 Main Results

We first present our main result, then discuss its implications on models with short time horizons or a large population of agents. All proofs are postponed until Section 3.

2.1 Equilibria close to Pareto optimality

Whenever equilibrium is discussed, Pareto optimality is a key concept. Passing to the more-convenient risk-aware notation, we remind the reader the following definition, where, as usual, $A[\boldsymbol{x}] = \sum_i \alpha^i x^i$:

Definition 2 For $\xi \in \mathbb{L}^0(\mathcal{F}_T)$, an allocation $\boldsymbol{\xi}$ is called ξ-**feasible** if $A[\boldsymbol{\xi}] \le \xi$. An allocation $\boldsymbol{\xi}$ is said to be **Pareto optimal** if there is no $A[\boldsymbol{\xi}]$-feasible allocation $\tilde{\boldsymbol{\xi}}$, such that $\mathbb{E}[U(\tilde{\xi}^i)] \ge \mathbb{E}[U(\xi^i)]$ for all i, and $\mathbb{E}[U(\tilde{\xi}^i)] > \mathbb{E}[U(\xi^i)]$ for some i.

In our setting, Pareto optimal allocations admit a very simple characterization; this is a direct consequence of the classical result [4] of Borch so we omit the proof.

Lemma 2 *A (sufficiently integrable) allocation $\boldsymbol{\xi}$ is Pareto optimal if and only if its components agree up to a constant, i.e., if there exist $\xi^c \in \mathbb{L}^0(\mathcal{F}_T)$ and constants $(c^i)_i$ such that $\xi^i = \xi^c + c^i$ for all i.*

Next, we introduce a concept which plays a central role in our main result. Given a population with the (risk-denominated) initial allocation $\boldsymbol{G}$ whose components satisfy (6), let $(m^i, n^i) \in \text{bmo}$ be an alias for the pair (m^{G^i}, n^{G^i}) defined in (7). We define **distance to Pareto optimality** $H(\boldsymbol{G})$ of $\boldsymbol{G}$ by

$$H(\boldsymbol{G}) = \inf_{\xi^c} \max_i ||(m^i - m^c, n^i - n^c)||_{\text{bmo}(\mathbb{P}^c)},$$

where the infimum is taken over the set of $\xi^c \in \text{EBMO}$, with $(m^c, n^c) = (m^{\xi^c}, n^{\xi^c})$ as in (7), and the probability measure $\mathbb{P}^c$ is given by

$$d\mathbb{P}^c/d\mathbb{P} = \mathcal{E}(-m^c \cdot B - n^c \cdot W)_T = \exp(-\xi^c)/\mathbb{E}[\exp(-\xi^c)]. \tag{15}$$

Remark 4

1. Suppose that $H(\boldsymbol{G}) = 0$ and that the infimum is attained. Then $(m^i, n^i) = (m^c, n^c)$, for all i, implying that all components of $\boldsymbol{G}$ coincide with ξ^c up to some additive constants, making $\boldsymbol{G}$ Pareto optimal. On the other hand, since each agent has exponential utility, shifting all components of $\boldsymbol{G}$ by the same amount ξ^c is equivalent to a measure change from $\mathbb{P}$ to $\mathbb{P}^c$. Therefore, $\lambda \in \Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G}, \mathbb{P})$ if and only if $\lambda - m^c \in \Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G} - \xi^c, \mathbb{P}^c)$, i.e., translation in endowments does not affect the wellposedness of the equilibrium. As a consequence, to show $\Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G}, \mathbb{P}) \neq \emptyset$, it suffices to prove $\Lambda_{\boldsymbol{\alpha}}(\boldsymbol{G} - \xi^c, \mathbb{P}^c) \neq \emptyset$ for some ξ^c, which is the strategy we follow below.
2. Our "distance to Pareto optimality" is conceptually similar to the "coefficient of resource utilization" of Debreu (see [10]), well known in economics. There, however, seems to be no simple and direct mathematical connection between the two.

In our first main result below, we assume that $\boldsymbol{G}$ is sufficiently close to *some* Pareto optimal allocation, i.e., that $H(\boldsymbol{G}) \leq \epsilon^*$, for some sufficiently small ϵ^*:

Theorem 2 (Existence and uniqueness close to Pareto optimality)

Let (6) *hold for all components in* $\boldsymbol{G}$. *There exists a sufficiently small constant* ϵ^*, *independent of the number of agents I, such that if*

$$H(\boldsymbol{G}) \leq \epsilon^*, \tag{16}$$

Then there exists a unique equilibrium $\lambda \in bmo$. *Moreover, the triplet* $(\boldsymbol{Y}^{\lambda,\boldsymbol{G}}, \boldsymbol{\mu}^{\lambda,\boldsymbol{G}}, \boldsymbol{\nu}^{\lambda,\boldsymbol{G}})$, *defined in Lemma 1, is the unique solution to* (14) *with* $(\boldsymbol{\mu}^{\lambda,\boldsymbol{G}}, \boldsymbol{\nu}^{\lambda,\boldsymbol{G}}) \in bmo^{\boldsymbol{I}}$.

Remark 5 A similar global uniqueness has been obtained in [24, Theorem 4.1] for a different quadratic BSDE system arising from a price impact model.

The proof of Theorem 2 will be presented in Section 2.1. For the time being, let us discuss two important cases in which (16) holds:

- First, given $\xi^c \in \text{EBMO}$ and $1 \leq i \leq I$, let X^{G^i} and X^{ξ^c} be defined by (7) with terminal conditions G^i and ξ^c, respectively. A simple calculation shows that

$$d(X_t^{G^i} - X_t^{\xi^c}) = (m_t^i - m_t^c)\,dB_t^c + (n_t^i - n_t^c)\,dW_t^c + \tfrac{1}{2}\Big((m_t^i - m_t^c)^2 + (n_t^i - n_t^c)^2\Big)\,dt,$$

with the terminal condition $G^i - \xi^c$, for a two-dimensional $\mathbb{P}^c$-Brownian motion (B^c, W^c), where $\mathbb{P}^c$ is given by (15). If, furthermore, $G^i - \xi^c \in \mathbb{L}^\infty$, it follows that

$$\begin{aligned}||(m^i - m^c, n^i - n^c)||^2_{\mathrm{bmo}(\mathbb{P}^c)} &= 2\sup_\tau ||\mathbb{E}_\tau^{\mathbb{P}^c}[X_T^{G^i} - \xi^c] - (X_\tau^{G^i} - \xi_\tau^c)||_{\mathbb{L}^\infty} \\ &\leq 4||G^i - \xi^c||_{\mathbb{L}^\infty}.\end{aligned}$$

Therefore, assumption (16) holds, if

$$\inf_{\xi^c} \max_i ||G^i - \xi^c||_{\mathbb{L}^\infty} \leq \frac{(\epsilon^*)^2}{4}. \tag{17}$$

- The second case in which (16) can be verified is in the case of a "large" number of agents. Indeed, an interesting feature of (17) is its lack of dependence on I, leading to the existence of equilibria in an economically meaningful asymptotic regime. Given a **total endowment** $E_\Sigma \in \mathbb{L}^\infty$ to be shared among I agents, i.e., $\sum_i E^i = E_\Sigma$, one can ask the following question: how many and what kind of agents need to share this total endowment so that they can form a financial market in which an equilibrium exists? The answer turns out to be “sufficiently many sufficiently homogeneous agents”. In order show that, we first make precise what we mean by sufficiently homogeneous. For the population characteristics $\boldsymbol{E} = (E^i)_i$ and $\boldsymbol{\delta} = (\delta^i)_i$, with $\boldsymbol{E} \in (\mathbb{L}^\infty)^I$, we define the **endowment heterogeneity index** $\chi^E(\boldsymbol{E}) \in [0,1]$ by

$$\chi^E(\boldsymbol{E}) = \max_{i,j} \frac{||E^i - E^j||_{\mathbb{L}^\infty}}{||E^i||_{\mathbb{L}^\infty} + ||E^j||_{\mathbb{L}^\infty}}.$$

We think of a population of agents as “sufficiently homogeneous” if $\chi^E(\boldsymbol{E}) \leq \chi_0^E$ for some, given, critical index χ_0^E. With this in mind, we have the following corollary of Theorem 2:

Corollary 1 (Existence of equilibria for sufficiently many sufficiently homogeneous agents)

Given a critical endowment homogeneity index $\chi_0^E \in [0, \frac{1}{2})$*, a critical risk tolerance* $\delta_0 > 0$*, as well as the total endowment* $E_\Sigma \in \mathbb{L}^\infty$*, there exists* $I_0 = I_0(||E_\Sigma||_{\mathbb{L}^\infty}, \chi_0^E, \delta_0) \in \mathbb{N}$*, so that any population* $(\boldsymbol{E}, \boldsymbol{\delta}) = (E^i, \delta^i)_i$ *satisfying*

$$I \geq I_0, \quad \textstyle\sum_i E^i = E_\Sigma, \quad \chi^E(E^i) \leq \chi_0^E, \quad \textit{and} \quad \min_i \delta^i \geq \delta_0,$$

admits an equilibrium.

Condition (17) can be thought of as a smallness-in-size assumption placed on the random endowments, possibly after translation. It turns out that it can be “traded” for a smallness-in-time condition which we now describe. We start by briefly recalling the notion of Malliavin differentiation on the Wiener space. Let Φ be the set of random variables ζ of the form $\zeta = \varphi(\mathcal{I}(h^1), \dots, \mathcal{I}(h^k))$, where $\varphi \in C_b^\infty(\mathbb{R}^k, \mathbb{R})$ (smooth functions with bounded derivatives of all orders) for some k, $h^j = (h^{j,b}, h^{j,w}) \in \mathbb{L}^2([0,T];\mathbb{R}^2)$ and $\mathcal{I}(h^j) = h^{j,b} \cdot B_T + h^{j,w} \cdot W_T$, for each $j = 1,\dots,k$. If $\zeta \in \Phi$, we define its **Malliavin derivative** as the 2-dimensional process

$$D_\theta \zeta = \sum_{j=1}^{k} \frac{\partial \varphi}{\partial x_j}(\mathcal{I}(h^1),\dots,\mathcal{I}(h^k))h^j_\theta, \quad \theta \in [0,T].$$

We denote by $D^b_\theta \zeta$ and $D^w_\theta \zeta$ the two components of $D_\theta \zeta$ and for $\zeta \in \Phi$, $p \geq 1$, define the norm

$$||\zeta||_{1,p} = \left[\mathbb{E}\left[|\zeta|^p + \left(\int_0^T |D_\theta \zeta|^2 d\theta\right)^{p/2}\right]\right]^{1/p}.$$

For $p \in [1,\infty)$, the Banach space $\mathbb{D}^{1,p}$ is the closure of Φ under $||\cdot||_{1,p}$. For $p = \infty$, we define $\mathbb{D}^{1,\infty}$ as the set of all those $G \in \mathbb{D}^{1,1}$ with $D^b G, D^w G \in \mathcal{S}^\infty$.

Corollary 2 (Existence of equilibria on sufficiently small time horizons)

Suppose that (6) *holds for all components of* $\boldsymbol{G}$ *and that there exists* $\xi^c \in EBMO$ *such that* $G^i - \xi^c \in \mathbb{D}^{1,\infty}$ *for all i. Then a unique equilibrium exists as soon as*

$$T < T^* = \frac{(\epsilon^*)^2}{\max_i \left(||D^b(G^i - \xi^c)||^2_{\mathcal{S}^\infty} + ||D^w(G^i - \xi^c)||^2_{\mathcal{S}^\infty}\right)}. \tag{18}$$

Remark 6 In a Markovian setting where $\boldsymbol{G} = \boldsymbol{g}(B_T, W_T)$, for some functions $\boldsymbol{g} = (g^i)_i$, we only need to assume there exists some $g^c \in \mathbb{L}^\infty$ such that $\partial_b(g^i - g^c), \partial_w(g^i - g^c) \in \mathbb{L}^\infty$, for any i, where $\partial_b(g^i - g^c)$ and $\partial_w(g^i - g^c)$ are weak derivatives of $g^i - g^c$. A similar "smallness in time" result has been proven in [8, Theorem 3.1] (and in [30] in a simpler model) in a Markovian setting. Corollary 2 extends the result of [8] to a non-Markovian setting.

3 Proofs

3.1 Proof of Proposition 1

For $\lambda \in$ bmo, we record that $\mathcal{M}^\lambda \neq \emptyset$. Indeed, thanks to the bmo property of λ, the process $Z^\lambda = \mathcal{E}(-\lambda \cdot B)$ is a martingale and satisfies the reverse Hölder inequality R_p for some $p > 1$ (see [22, Theorem 3.1]). That, in turn, implies the reverse Hölder inequality $R \log R$, and, so, the probability $\mathbb{Q}^\lambda$ defined via $d\mathbb{Q}^\lambda/d\mathbb{P} = Z^\lambda_T$ satisfies $H(\mathbb{Q}^\lambda|\mathbb{P}) < \infty$, and, consequently $\mathbb{Q}^\lambda \in \mathcal{M}^\lambda$.

The statements of Proposition 1 will follow from [11, Theorem 2.2], once we verify that Z^λ satisfies the reverse Hölder inequality $R \log R$ under $\overline{\mathbb{P}}$ as well, where $d\overline{\mathbb{P}}/d\mathbb{P} = e^{-G}/\mathbb{E}[e^{-G}]$. For that, we note that $e^{-G}/\mathbb{E}[e^{-G}] = \mathcal{E}(-m^G \cdot B - n^G \cdot W)_T$, where (m^G, n^G) is as in (7). Given $\lambda \in$ bmo, the bmo property of (m^G, n^G) and [22, Theorem 3.6] imply that $\lambda - m^G \in \mathrm{bmo}(\overline{\mathbb{P}})$, and, so, $Z^\lambda = \mathcal{E}(-(\lambda - m)\cdot \overline{B})_T$, where $\overline{B} = \int_0^\cdot m_u du + B$ is a $\overline{\mathbb{P}}$-martingale. It remains to use the same argument as in the previous paragraph to show that Z^λ indeed satisfies the reverse Hölder inequality $R\log R$ under $\overline{\mathbb{P}}$.

3.2 Proof of Lemma 1

Let $(m,n) = (m^G, n^G)$ from (7); more generally, we suppress the superscripts λ and G throughout to increase legibility. A combination of (9) and (10) yields that

$$Y = -c - \rho\cdot B^\lambda - \log Z,$$

and a simple calculation confirms that (Y,μ,ν) satisfies (13). Next, we show $Y - X \in \mathcal{S}^\infty$. We start by defining the probability measure $\overline{\mathbb{P}}$ via $d\overline{\mathbb{P}}/d\mathbb{P} = \mathcal{E}(-m\cdot B - n\cdot W)_T$ so that under $\overline{\mathbb{P}}$, $D = Y - X$ is the certainty-equivalent process corresponding to the zero endowment. By (11), we have $D \geq 0$ as well as

$$\begin{aligned} dD_t &= (\mu_t - m_t)\, d\overline{B} + (\nu_t - n_t)\, d\overline{W} \\ &\quad + \left(\tfrac{1}{2}(\nu_t - n_t)^2 - \tfrac{1}{2}(\lambda_t - m_t)^2 + (\lambda_t - m_t)(\mu_t - m_t)\right) dt, \text{ with } D_T = 0, \end{aligned} \tag{19}$$

where $\overline{B} = B + \int_0^\cdot m_u du$ and $\overline{W} = W + \int_0^\cdot n_u du$ are $\overline{\mathbb{P}}$-Brownian motions. Using the notation $\mathbb{Q}^\lambda$, as well as the argument of Proof of Proposition 1 above, we can deduce that $\mathbb{Q}^\lambda \in \mathcal{M}^{\lambda - m}$ (where $\mathbb{P}$ in the definition of $\mathcal{M}^{\lambda-m}$ is replaced by $\overline{\mathbb{P}}$). We claim that

$$D_\tau \leq H_\tau(\mathbb{Q}^\lambda|\overline{\mathbb{P}}), \qquad \text{for any } \tau \in \mathcal{T}. \tag{20}$$

Proposition 1, applied under $\overline{\mathbb{P}}$ and with zero random endowment produces the dual optimizer $\mathbb{Q}^{\lambda,G}$, with $\overline{\mathbb{P}}$-density $Z^{\lambda-m,G}$. If we project both sides of the equality $\overline{c}^{\lambda,G} + \rho^{\lambda,G}\cdot B^\lambda_T = -\log(Z_T^{\lambda-m,G})$ under $\mathbb{Q}^{\lambda,G}$ onto $\mathcal{F}_\tau$ we obtain

$$D_\tau = H_\tau(\mathbb{Q}^{\lambda,G}|\overline{\mathbb{P}}).$$

No integrability issues arise here since $H(\mathbb{Q}^{\lambda,G}|\overline{\mathbb{P}}) < \infty$ and $\rho^{\lambda,G}\cdot B^\lambda$ is a $\mathbb{Q}^{\lambda,G}$-martingale (by part (iii) of Proposition 1). The required inequality (20) follows from the optimality of $\mathbb{Q}^{\lambda,G}$ in part (ii) of Proposition 1.

The right-hand side of (20) can be written as

$$H_\tau(\mathbb{Q}^\lambda|\overline{\mathbb{P}}) = \mathbb{E}_\tau^{\mathbb{Q}^\lambda}\left[\tfrac{1}{2}\int_\tau^T (\lambda_t - m_t)^2 dt - \int_\tau^T (\lambda_t - m_t) dB_t^\lambda\right] \leq \tfrac{1}{2}||\lambda - m||^2_{\mathrm{bmo}(\mathbb{Q}^\lambda)}.$$

Given that both λ and m belong to bmo we have $\lambda - m \in \mathrm{bmo}(\mathbb{Q}^\lambda)$ by [22, Theorem 3.6]. Therefore, we can combine (20) and the fact that $D \geq 0$ to conclude that $D \in \mathcal{S}^\infty$. Consequently, it suffices to apply the standard bmo-estimate for quadratic BSDEs (see Lemma 9) to (19), to obtain $(\mu - m, \nu - n) \in \mathrm{bmo}(\overline{\mathbb{P}})$. Since $(m,n) \in \mathrm{bmo}$, another application of [22, Theorem 3.6] confirms that $(\mu, \nu) \in \mathrm{bmo}$.

Lastly, we show that there can be at most one solution to (13) with $(\mu, \nu) \in \mathrm{bmo}$. Let (Y, μ, ν) and $(\tilde{Y}, \tilde{\mu}, \tilde{\nu})$ be two solutions with $(\mu, \nu), (\tilde{\mu}, \tilde{\nu}) \in \mathrm{bmo}$. For $\delta Y = \tilde{Y} - Y$, we have

$$d(\delta Y)_t = \delta\mu_t dB_t^\lambda + \delta\nu_t dW_t^{\overline{\nu}}, \quad \delta Y_T = 0.$$

Here $\delta\mu = \tilde{\mu} - \mu$, $\delta\nu = \tilde{\nu} - \nu$, $\overline{\nu} = \frac{1}{2}(\nu + \tilde{\nu})$, and $W^{\overline{\nu}} = W + \int_0^\cdot \overline{\nu}_t dt$ is a $\mathbb{Q}^{\lambda,\overline{\nu}}$-Brownian motion, where $\mathbb{Q}^{\lambda,\overline{\nu}}$ is defined via $d\mathbb{Q}^{\lambda,\overline{\nu}}/d\mathbb{P} = \mathcal{E}(-\lambda \cdot B - \overline{\nu} \cdot W)_T$. By [22, Theorem 3.6], both $\delta\mu \cdot B^\lambda$ and $\delta\nu \cdot W^{\overline{\nu}}$ are BMO($\mathbb{Q}^{\lambda,\overline{\nu}}$)-martingales. Hence $\delta Y_T = 0$ implies that $\delta Y = 0$ and, consequently, $\delta\mu = \delta\nu = 0$.

3.3 Proof of Theorem 1

(1) $\Rightarrow$ (2). Given an equilibrium $\lambda \in \Lambda_\alpha(\boldsymbol{G})$ and $i \in \{1, 2, \ldots, I\}$, let ρ^{λ, G^i} be the primal optimizer of agent i, and let (Y^i, μ^i, ν^i) be defined as in Lemma 1 where (13) has the terminal condition $Y_T^i = G^i$. Since λ is an equilibrium, $\sum_i \alpha^i \rho^{\lambda, G^i} = 0$, and so $\lambda = \lambda - \sum_i \alpha^i \rho^{\lambda, G^i} = \sum_i \alpha^i \mu^i$, for $\mu^i = \lambda - \rho^{\lambda, G^i}$, implying that $(\boldsymbol{Y}, \boldsymbol{\mu}, \boldsymbol{\nu}) = (Y^i, \mu^i, \nu^i)_i$ solves the system (14). The property $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \mathrm{bmo}^I$ follows from Lemma 1.

(2) $\Rightarrow$ (1). Given a solution $(\boldsymbol{Y}, \boldsymbol{\mu}, \boldsymbol{\nu})$ of (14), we set $\lambda = \sum_i \alpha^i \mu^i$. This way, individual equations in (14) turn into BSDEs of the form (13). If we set $\rho^{\lambda, i} = \lambda - \mu^i$ the market clearing condition $\sum_i \alpha_i \rho^{\lambda, i} = 0$ holds. Since $(\mu^i, \nu^i) \in \mathrm{bmo}$ the uniqueness part of Lemma 1 implies that λ, ρ^i maximizes single-agents' utilities.

3.4 Proof of Theorem 2

In order to prove Theorem 2, we start with a refinement of the classical result on uniform equivalence of bmo spaces (see Theorem 3.6, p. 62 in [22]), based on a result of Chinkvinidze and Mania (see [7]).

Lemma 3 *Let $\sigma \in bmo$ be such that $||\sigma||_{bmo} =: \sqrt{2}R$ for some $R < 1$. If $\hat{\mathbb{P}} \sim \mathbb{P}$ is such that $d\hat{\mathbb{P}} = \mathcal{E}(\sigma \cdot \tilde{B})_T \, d\mathbb{P}$, for some $\mathbb{F}$-Brownian motion $\tilde{B}$, then, for all $\zeta \in bmo$, we have*

$$(1 + R)^{-1} ||\zeta||_{bmo} \le ||\zeta||_{bmo(\hat{\mathbb{P}})} \le (1 - R)^{-1} ||\zeta||_{bmo}. \tag{21}$$

Proof Since $M = \sigma \cdot \tilde{B}$ is a BMO-martingale, Theorem 3.6. in [22] states that the spaces bmo and bmo($\hat{\mathbb{P}}$) coincide and that the norms $||\cdot||_{\mathrm{bmo}}$ and $||\cdot||_{\mathrm{bmo}(\hat{\mathbb{P}})}$ are uniformly equivalent. This norm equivalence is refined in [7]; Theorem 2 there implies that

$$(1 + R)^{-1} ||\zeta||_{\mathrm{bmo}} \le ||\zeta||_{\mathrm{bmo}(\hat{\mathbb{P}})} \le (1 + \hat{R}) ||\zeta||_{\mathrm{bmo}}, \text{ where } \hat{R} = \sqrt{\tfrac{1}{2} ||\sigma||^2_{\mathrm{bmo}(\hat{\mathbb{P}})}}. \tag{22}$$

Clearly, only the second inequality in (21) needs to be discussed; it is obtained by substituting $\zeta = \sigma$ into the second inequality in (22):

$$\sqrt{2}\hat{R} = ||\sigma||_{\mathrm{bmo}(\hat{\mathbb{P}})} = (1+\hat{R})||\sigma||_{\mathrm{bmo}} \leq \sqrt{2}(1+\hat{R})R, \text{ so that } (1+\hat{R}) \leq (1-R)^{-1}.$$

Coming back to Theorem 2, suppose that (16) is satisfied. Then there exists $\xi^c \in \mathrm{EBMO}$ such that

$$\max_i ||(m^i - m^c, n^i - n^c)||_{\mathrm{bmo}(\mathbb{P}^c)} \leq \epsilon^*. \tag{23}$$

To simplify notation, we introduce $\boldsymbol{m} = (m^i)_i$ and $\boldsymbol{n} = (n^i)_i$. A calculation shows that (component-by-component)

$$\begin{aligned} d(\boldsymbol{Y}_t - \xi^c_t) =& (\boldsymbol{\mu}_t - m^c_t)\, dB^c_t + (\boldsymbol{\nu}_t - n^c_t)\, dW^c_t \\ &+ \left(\tfrac{1}{2}(\boldsymbol{\nu}_t - n^c_t)^2 - \tfrac{1}{2}(\lambda_t - m^c_t)^2 + (\lambda_t - m^c_t)(\boldsymbol{\mu}_t - m^c_t)\right) dt, \\ \boldsymbol{Y}^T - \xi^c_T =& \boldsymbol{G} - \xi^c, \end{aligned}$$

where $\lambda = A[\boldsymbol{\mu}]$, $\xi^c_t = -\log(\mathbb{E}_t[\exp(-\xi^c)])$, and B^c, W^c are $\mathbb{P}^c$-Brownian motions. This is exactly the type of system covered in (14). Therefore, to ease notation, we treat, throughout this section, $\mathbb{P}$ as $\mathbb{P}^c$, B as B^c, W as W^c, and $\boldsymbol{G}, \lambda, \boldsymbol{\mu}, \boldsymbol{\nu}$ as their shifted versions, i.e., eg. $\boldsymbol{G}$ as $\boldsymbol{G} - \xi^c$, λ as $\lambda - m^c$, etc. As a result, (23) translates to

$$\max_i ||(m^i, n^i)||_{\mathrm{bmo}} \leq \epsilon^*. \tag{24}$$

We proceed by setting up a framework for the Banach fixed-point theorem. First observe that since $(m^i, n^i) \in \mathrm{bmo}$ for all i, then bmo is a natural space in which the fixed-point theorem can be applied. Given $\lambda \in \mathrm{bmo}$ and $\boldsymbol{G} = (G^i)_i$, let $\boldsymbol{Y}^\lambda = (Y^{\lambda, G^i})_i$ and $\boldsymbol{X} = (X^{G^i})_i$, denote the agents' certainty-equivalent processes with and without assess the market, respectively; we also set $(\boldsymbol{\mu}^{\lambda,\boldsymbol{G}}, \boldsymbol{\nu}^{\lambda,\boldsymbol{G}}) = (\mu^{\lambda,G^i}, \nu^{\lambda,G^i})_i$, where $(\mu^{\lambda,G^i}, \nu^{\lambda,G^i})_i$ is defined in Lemma 1. This allows us to define (a simple transformation of) the **excess-demand map**

$$F : \lambda \mapsto A[\boldsymbol{\mu}^{\lambda,\boldsymbol{G}}],$$

where the aggregation operator $A[\cdot]$ is defined in (5). The significance of this map lies in the simple fact that λ is an equilibrium if and only if $F(\lambda) = \lambda$, i.e., if λ is a fixed point of F.

Before proceeding to studying properties of F, we first record the following a-priori estimate on λ in equilibrium.

Lemma 4 *If $\lambda \in bmo$ is an equilibrium, then*

$$||\lambda||_{bmo} \leq \max_i ||(m^i, n^i)||_{bmo}.$$

Proof Aggregating all single equations in (14) and (7), we obtain

$$dA[\boldsymbol{Y}_t^\lambda - \boldsymbol{X}_t] = (\lambda_t - A[\boldsymbol{m}_t])dB_t + A[\boldsymbol{\nu}_t^\lambda - \boldsymbol{n}_t]dW_t + \tfrac{1}{2}(\lambda_t^2 + A[(\boldsymbol{\nu}_t^\lambda)^2])dt - \tfrac{1}{2}A[\boldsymbol{m}_t^2 + \boldsymbol{n}_t^2]dt.$$

Let $(\sigma_n)_n$ be a reducing sequence for local martingale part above. For any $\tau \in \mathcal{T}$, integrating the previous dynamics from $\tau \wedge \sigma_n$ to σ_n and projecting onto $\mathcal{F}_\tau$ yields

$$\begin{aligned}\mathbb{E}_\tau\left[A[\boldsymbol{Y}_{\sigma_n}^\lambda - \boldsymbol{X}_{\sigma_n}]\right] - A[\boldsymbol{Y}_{\tau\wedge\sigma_n}^\lambda - \boldsymbol{X}_{\tau\wedge\sigma_n}] =& \\ = \tfrac{1}{2}\mathbb{E}_\tau\left[\int_{\tau\wedge\sigma_n}^{\sigma_n} (\lambda_t^2 + A[(\boldsymbol{\nu}_t^\lambda)^2])dt\right] - \tfrac{1}{2}\mathbb{E}_\tau\left[\int_{\tau\wedge\sigma_n}^{\sigma_n} A[\boldsymbol{m}_t^2 + \boldsymbol{n}_t^2]dt\right].& \quad (25)\end{aligned}$$

Sending $n \to \infty$, since $\boldsymbol{Y}^\lambda - \boldsymbol{X} \geq 0$ (component-by-component) and is also bounded (see Lemma 1) and $A[\boldsymbol{X}_T] = A[\boldsymbol{G}] = A[\boldsymbol{Y}_T^\lambda]$, we obtain

$$\begin{aligned}||\lambda||_{\text{bmo}}^2 &\leq ||\lambda^2 + A[(\boldsymbol{\nu}^\lambda)^2]||_{\text{bmo}} \leq ||A[\boldsymbol{m}^2 + \boldsymbol{n}^2]||_{\text{bmo}} \\ &\leq A[||(\boldsymbol{m}, \boldsymbol{n})||_{\text{bmo}}^2] \leq \max_i ||(m^i, n^i)||_{\text{bmo}}^2.\end{aligned}$$

For the third inequality, note that $\mathbb{E}_\tau[\int_\tau^T A[\boldsymbol{m}_t^2 + \boldsymbol{n}_t^2]dt] \leq A[||(\boldsymbol{m}, \boldsymbol{n})||_{\text{bmo}}^2]$ holds for all stopping times τ. □

For arbitrary $\lambda \in \text{bmo}$, the following estimate gives an explicit upper bound on the (nonnegative) difference $D^{\lambda,i} = Y^{\lambda,i} - X^i$.

Lemma 5 *Suppose that* $||\lambda||_{bmo} < \sqrt{2}$. *Then,*

$$0 \leq \sqrt{D^{\lambda,i}} \leq \frac{||\lambda||_{bmo} + ||(m^i, n^i)||_{bmo}}{\sqrt{2} - ||\lambda||_{bmo}}, \quad \textit{for all } i.$$

Proof Let $\mathbb{Q}^\lambda$ be the probability such that $d\mathbb{Q}^\lambda = Z_T^\lambda d\mathbb{P}$, where $Z^\lambda = \mathcal{E}(-\lambda \cdot B)$. Since $\mathbb{Q}^\lambda \in \mathcal{M}^\lambda$, then the argument that leads to (20) also implies that

$$Y_\tau^{\lambda,i} \leq H_\tau(\mathbb{Q}^\lambda|\mathbb{P}) + \mathbb{E}_\tau^{\mathbb{Q}^\lambda}[G^i], \quad \text{for any } \tau \in \mathcal{T}. \quad (26)$$

On the right-hand side of (26),

$$H_\tau(\mathbb{Q}^\lambda|\mathbb{P}) = \mathbb{E}_\tau^{\mathbb{Q}^\lambda}\left[\tfrac{1}{2}\int_\tau^T \lambda_u^2 du - \int_\tau^T \lambda_u dB_u^\lambda\right] \leq \tfrac{1}{2}||\lambda||_{\text{bmo}(\mathbb{Q}^\lambda)}^2.$$

Since $||\lambda||_{\text{bmo}(\mathbb{Q}^\lambda)} \leq \sqrt{2}||\lambda||_{\text{bmo}}/\left(\sqrt{2} - ||\lambda||_{\text{bmo}}\right)$, as follows from Lemma 3, we obtain

$$H_\tau(\mathbb{Q}^\lambda|\mathbb{P}) \leq \frac{||\lambda||_{\text{bmo}}^2}{\left(\sqrt{2} - ||\lambda||_{\text{bmo}}\right)^2}.$$

Furthermore, recalling that $X_T^i = G^i$ and $dX_t^i = m_t^i dB_t + n_t^i dW_t + \tfrac{1}{2}((m_t^i)^2 + (n_t^i)^2)dt$, we note that

$$\mathbb{E}_\tau^{\mathbb{Q}^\lambda}[G^i] = \mathbb{E}_\tau[(Z_T^\lambda/Z_\tau^\lambda)G^i] = \mathbb{E}_\tau[(Z_T^\lambda/Z_\tau^\lambda)X_T^i].$$

Given that Z^λ is a BMO-martingale and $||(m^i, n^i)||_{\text{bmo}} < \infty$, the integration-by-parts formula implies that

$$\begin{aligned}
&\mathbb{E}_\tau[(Z_T^\lambda/Z_\tau^\lambda)X_T^i] = \\
&\quad = X_\tau^i - \mathbb{E}_\tau\left[\int_\tau^T (Z_u^\lambda/Z_\tau^\lambda)\lambda_u m_u^i du\right] + \tfrac{1}{2}\mathbb{E}_\tau\left[\int_\tau^T (Z_u^\lambda/Z_\tau^\lambda)\left((m_u^i)^2 + (n_u^i)^2\right) du\right] \\
&\quad = X_\tau^i - \mathbb{E}_\tau^{\mathbb{Q}^\lambda}\left[\int_\tau^T \lambda_u m_u^i du\right] + \tfrac{1}{2}\mathbb{E}_\tau^{\mathbb{Q}^\lambda}\left[\int_\tau^T \left((m_u^i)^2 + (n_u^i)^2\right) du\right].
\end{aligned}$$

A use of Holder's inequality then gives

$$\begin{aligned}
\mathbb{E}_\tau^{\mathbb{Q}^\lambda}[G^i] - X_\tau^i &\le ||\lambda||_{\text{bmo}(\mathbb{Q}^\lambda)}||m^i||_{\text{bmo}(\mathbb{Q}^\lambda)} + \tfrac{1}{2}||(m^i, n^i)||^2_{\text{bmo}(\mathbb{Q}^\lambda)} \\
&\le \frac{2||\lambda||_{\text{bmo}}||(m^i, n^i)||_{\text{bmo}} + ||(m^i, n^i)||^2_{\text{bmo}}}{\left(\sqrt{2} - ||\lambda||_{\text{bmo}}\right)^2},
\end{aligned}$$

where, again, the last inequality follows from Lemma 3. A Combination of the above estimates shows that

$$D_\tau^{\lambda,i} = Y_\tau^{\lambda,i} - X_\tau^i \le \left(\frac{||\lambda||_{\text{bmo}} + ||(m^i, n^i)||_{\text{bmo}}}{\sqrt{2} - ||\lambda||_{\text{bmo}}}\right)^2,$$

which completes the proof. □

Lemma 6 *Suppose that $\lambda \in bmo$ satisfies*

$$||\lambda||_{bmo} < \frac{\sqrt{2} - ||(m^i, n^i)||_{bmo}}{2}.$$

Then, it holds that

$$\begin{aligned}
&||(\mu^{\lambda,i}, \nu^{\lambda,i})||_{bmo} \le \\
&\quad \le \frac{(\sqrt{2} + ||(m^i, n^i)||_{bmo})||(m^i, n^i)||_{bmo} + ||\lambda||_{bmo}(||\lambda||_{bmo} + ||(m^i, n^i)||_{bmo})}{\sqrt{2} - 2||\lambda||_{bmo} - ||(m^i, n^i)||_{bmo}}.
\end{aligned}$$

In particular, the previous is also a bound for both $||\mu^{\lambda,i}||_{bmo}$ and $||\nu^{\lambda,i}||_{bmo}$.

Proof Set $\boldsymbol{Y} = \boldsymbol{Y}^\lambda$, $\boldsymbol{\mu} = \boldsymbol{\mu}^\lambda$ and $\boldsymbol{\nu} = \boldsymbol{\nu}^\lambda$ to increase legibility, and define

$$f^i = \frac{||\lambda||_{\text{bmo}} + ||(m^i, n^i)||_{\text{bmo}}}{\sqrt{2} - ||\lambda||_{\text{bmo}}},$$

and $\boldsymbol{D} = \boldsymbol{Y} - \boldsymbol{X}$. Note that $D_T^i = 0$ and $0 \le D^i \le (f^i)^2$ from Lemma 5. Since

$$dD_t^i = (\mu_t^i - m_t^i)dB_t + (\nu_t^i - n_t^i)dW_t + \tfrac{1}{2}\left((\nu_t^i)^2 - \lambda_t^2 + 2\mu_t^i\lambda_t - (m_t^i)^2 - (n_t^i)^2\right)dt,$$

an application of Itô's lemma gives

$$\begin{aligned} d(D^i_t)^2 =& 2D^i_t(\mu^i_t - m^i_t)dB_t + 2D^i_t(\nu^i_t - n^i_t)dW_t \\ &+ D^i_t\left((\nu^i_t)^2 - \lambda^2_t + 2\mu^i_t\lambda_t - (m^i_t)^2 - (n^i_t)^2\right)dt \\ &+ \left((\mu^i_t - m^i_t)^2 + (\nu^i_t - n^i_t)^2\right)dt. \end{aligned}$$

Next, we take a reducing sequence $(\sigma_n)_n$ for the local martingales on the right-hand side above, as well as and an arbitrary $\tau \in \mathcal{T}$. If we integrate the above dynamics between $\sigma_n \wedge \tau$ and σ_n, and use the facts that $(\nu^i)^2 \geq 0$, $\lambda^2 - 2\mu^i\lambda \leq (\mu^i - \lambda)^2$, and $D^i \geq 0$, we obtain

$$\begin{aligned} (D^i_{\sigma_n})^2 \geq (D^i_{\sigma_n}) - (D^i_{\tau\wedge\sigma_n})^2 \geq & 2\int_{\tau\wedge\sigma_n}^{\sigma_n} D^i_t(\mu^i_t - m^i_t)dB_t + 2\int_{\tau\wedge\sigma_n}^{\sigma_n} D^i_t(\nu^i_t - n^i_t)dW_t \\ &- \int_{\tau\wedge\sigma_n}^{\sigma_n} D^i_t\left((\mu^i_t - \lambda_t)^2 + (m^i_t)^2 + (n^i_t)^2\right)dt \\ &+ \int_{\tau\wedge\sigma_n}^{\sigma_n} \left((\mu^i_t - m^i_t)^2 + (\nu^i_t - n^i_t)^2\right)dt. \end{aligned}$$

Given that $D^i \leq (f^i)^2$, a projection of both sides above on $\mathcal{F}_\tau$ yields

$$\begin{aligned} &\mathbb{E}_\tau\left[\int_{\tau\wedge\sigma_n}^{\sigma_n} \left((\mu^i_t - m^i_t)^2 + (\nu^i_t - n^i_t)^2\right)dt\right] \\ &\leq \mathbb{E}_\tau[D^i_{\sigma_n}] + (f^i)^2\mathbb{E}_\tau\left[\int_{\tau\wedge\sigma_n}^{\sigma_n} \left((\mu^i - \lambda)^2 + (m^i_t)^2 + (n^i_t)^2\right)dt\right]. \end{aligned}$$

Sending $n \to \infty$ first on the right-hand side then the left, helped by the facts that D^i is bounded and $D^i_T = 0$, implies that

$$||(\mu^i, \nu^i) - (m^i, n^i)||^2_{\text{bmo}} \leq (f^i)^2\left(||\mu^i - \lambda||^2_{\text{bmo}} + ||(m^i, n^i)||^2_{\text{bmo}}\right).$$

Taking square roots on both sides, and using the elementary inequality $\sqrt{x^2 + y^2} \leq |x| + |y|$ for any x, y, and the fact that $||\mu^i - \lambda||_{\text{bmo}} \leq ||(\mu^i, \nu^i)||_{\text{bmo}} + ||\lambda||_{\text{bmo}}$, we obtain

$$||(\mu^i, \nu^i) - (m^i, n^i)||_{\text{bmo}} \leq f^i\left(||\lambda||_{\text{bmo}} + ||(\mu^i, \nu^i)||_{\text{bmo}} + ||(m^i, n^i)||_{\text{bmo}}\right).$$

Finally, since $||(\mu^i, \nu^i)||_{\text{bmo}} \leq ||(\mu^i, \nu^i) - (m^i, n^i)||_{\text{bmo}} + ||(m^i, n^i)||_{\text{bmo}}$, it follows that

$$(1 - f^i)||(\mu^i, \nu^i)||_{\text{bmo}} \leq ||(m^i, n^i)||_{\text{bmo}} + f^i\left(||\lambda||_{\text{bmo}} + ||(m^i, n^i)||_{\text{bmo}}\right),$$

from which the result follows after simple algebra. □

Define $\mathcal{B}(r) = \{\lambda \in \text{bmo} : ||\lambda||_{\text{bmo}} \leq r\}$. The following result shows that the excess-demand map F maps $\mathcal{B}(r)$ into itself for an appropriate choice of r.

Lemma 7 *There exists a sufficiently small ϵ^* independent of the number of the agents I, such that whenever $\max_i ||(m^i, n^i)||_{bmo} \le \sqrt{2}\epsilon$ for $\epsilon \le \epsilon^*$, F maps $\mathcal{B}(2\epsilon)$ into itself.*

Proof Suppose that $\max_i ||(m^i, n^i)||_{\text{bmo}} \le \sqrt{2}\epsilon$ for some $\epsilon \in (0, 1)$ determined later. Let us consider $\lambda \in \mathcal{B}(\sqrt{2}\epsilon a)$, where $a \in [1, 1/\epsilon)$ will also be determined later. Our goal is to choose a sufficiently small ϵ such that $A[\boldsymbol{\mu}^\lambda] \in \mathcal{B}(\sqrt{2}\epsilon a)$for some $a \in [1, 1/\epsilon)$, whenever λ is chosen from the same ball. If this task is successful, given $a \ge 1$, Lemma 4 implies that all possible equilibria are already in the same ball. Hence the local uniqueness immediately implies global uniqueness in bmo.

For $\lambda \in \mathcal{B}(\sqrt{2}\epsilon a)$, Lemma 5 gives

$$0 \le \sqrt{D^{\lambda,i}} \le \frac{\epsilon(1+a)}{1-a\epsilon} =: \phi(\epsilon, a).$$

Note that ϕ is an increasing function of both arguments. For Lemma 6 we need $\phi < 1$. Therefore, only $\epsilon \in (0, 1)$ such that $\phi(\epsilon, 1) < 1$ can be used, i.e., $\epsilon \in (0, 1/3)$. Taking $\epsilon \in (0, 1/3)$ and $a \in [1, 1/\epsilon)$, in order to have $\phi(\epsilon, a) < 1$, it is necessary and sufficient that

$$a < \frac{1-\epsilon}{2\epsilon} =: \overline{a}(\epsilon).$$

Note that $\overline{a}$ is decreasing in ϵ with $\overline{a}(0+) = \infty$ and $\overline{a}(1/3) = 1$, and that $\overline{a}(\epsilon) < 1/\epsilon$ holds for all $\epsilon \in (0, 1/3)$.

Now, in order to have $||\mu^{\lambda,i}||_{\text{bmo}} \le \sqrt{2}\epsilon a$, by Lemma 6 we need to ensure that

$$\frac{2(1+\epsilon)\epsilon + 2a\epsilon^2(1+a)}{\sqrt{2}(1-2a\epsilon-\epsilon)} \le a\sqrt{2}\epsilon,$$

or, equivalently, that

$$q(a, \epsilon) := 3\epsilon a^2 - (1-2\epsilon)a + (1+\epsilon) \le 0.$$

Fix $a > 1$, say $a = \sqrt{2}$, there exists a sufficiently small ϵ^* such that $q(\sqrt{2}, \epsilon) \le 0$ for any $\epsilon \le \epsilon^*$. Note that the choice of ϵ^* is independent of the number of the agent I. For such choice of ϵ, we have $||\mu^{\lambda,i}||_{\text{bmo}} \le 2\epsilon$ for all i. As a weighted sum of individual component, $||F[\lambda]||_{\text{bmo}} \le A[||\boldsymbol{\mu}^\lambda||_{\text{bmo}}]$, hence $F[\lambda] \in \mathcal{B}(2\epsilon)$ as well. □

Finally we check that F is a contraction on $\mathcal{B}(2\epsilon)$ for sufficiently small ϵ.

Lemma 8 *There exists a sufficiently small ϵ^* independent of the number of the agents I, such that whenever $\max_i ||(m^i, n^i)||_{bmo} \le \sqrt{2}\epsilon$ for $\epsilon \le \epsilon^*$, F is a contraction on $\mathcal{B}(2\epsilon)$.*

Proof We drop the superscript i to increase legibility. Set $\delta Y = Y^\lambda - Y^{\tilde{\lambda}}$, and note that $||\delta Y||_{\mathcal{S}^\infty} < \infty$ from Lemma 5 and $\delta Y_T = 0$. Set $(\mu, \nu) = (\mu^\lambda, \nu^\lambda)$ and $(\tilde{\mu}, \tilde{\nu}) = (\mu^{\tilde{\lambda}}, \nu^{\tilde{\lambda}})$. Denote $\bar{\lambda} = (\lambda + \tilde{\lambda})/2$, $\bar{\mu} = (\mu + \tilde{\mu})/2$, and $\bar{\nu} = (\nu + \tilde{\nu})/2$. Calculation using (13) gives

$$\begin{aligned} d\delta Y_t &= (\mu_t - \tilde{\mu}_t)dB_t + (\nu_t - \tilde{\nu}_t)dW_t + \tfrac{1}{2}\left(\nu_t^2 - \tilde{\nu}_t^2 + \tilde{\lambda}_t^2 - \lambda_t^2 + 2\mu_t\lambda_t - 2\tilde{\mu}_t\tilde{\lambda}_t\right)dt \\ &= (\mu_t - \tilde{\mu}_t)dB_t^{\bar{\lambda}} + (\nu_t - \tilde{\nu}_t)dW_t^{\bar{\nu}} - (\lambda_t - \tilde{\lambda}_t)(\bar{\lambda}_t - \bar{\mu}_t)dt, \end{aligned}$$

where $B^{\bar\lambda} = B + \int_0^\cdot \lambda_t dt$, $W^{\bar\nu} = W + \int_0^\cdot \bar\nu_t dt$ are Brownian motions under $\mathbb{Q}^{\bar\lambda,\bar\nu}$, and $\mathbb{Q}^{\bar\lambda,\bar\nu}$ is defined via $d\mathbb{Q}^{\bar\lambda,\bar\nu}/d\mathbb{P} = \mathcal{E}(-\bar\lambda\cdot B - \bar\nu\cdot W)_T$. For an arbitrary $\tau \in \mathcal{T}$, integrating the previous dynamics on $[\tau, T]$, taking conditional expectation $\mathbb{E}_\tau^{\mathbb{Q}^{\bar\lambda,\bar\nu}}$ on both sides, (both local martingales are BMO($\mathbb{Q}^{\bar\lambda,\bar\nu}$)-martingales, due to $\mu, \tilde\mu, \nu, \tilde\nu \in$ bmo from Lemma 6 and [22, Theorem 3.6]), and finally using $\delta Y_T = 0$, we obtain

$$|\delta Y_\tau| \leq \mathbb{E}_\tau^{\mathbb{Q}^{\bar\lambda,\bar\nu}}\left[\int_\tau^T |\lambda_t - \tilde\lambda_t||\bar\lambda_t - \bar\mu_t|dt\right] \leq ||\bar\lambda - \bar\mu||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}||\lambda - \tilde\lambda||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}.$$

This implies that

$$||\delta Y||_{\mathcal{S}^\infty} \leq ||\bar\lambda - \bar\mu||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}||\lambda - \tilde\lambda||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}. \tag{27}$$

To establish the Lipschitz continuity of F, we use Itô's formula to get

$$\begin{aligned} d(\delta Y_t)^2 =& 2\delta Y_t(\mu_t - \tilde\mu_t)dB_t^{\bar\lambda} + 2\delta Y_t(\nu_t - \tilde\nu_t)dW_t^{\bar\nu} - 2\delta Y_t(\lambda_t - \tilde\lambda_t)(\bar\lambda_t - \bar\mu_t)dt \\ &+ \left((\mu_t - \tilde\mu_t)^2 + (\nu_t - \tilde\nu_t)^2\right)dt. \end{aligned}$$

For an arbitrary $\tau \in \mathcal{T}$, an integration of the above dynamics between τ and T, and using (27) and $\delta Y_T = 0$, yields that

$$\begin{aligned} &\mathbb{E}_\tau^{\mathbb{Q}^{\bar\lambda,\bar\nu}}\left[\int_\tau^T \left((\mu_t - \tilde\mu_t)^2 + (\nu_t - \tilde\nu_t)^2\right)dt\right] \leq \\ &\qquad \leq 2||\delta Y||_{\mathcal{S}^\infty}\mathbb{E}_\tau^{\mathbb{Q}^{\bar\lambda,\bar\nu}}\left[\int_\tau^T (\lambda_t - \tilde\lambda_t)(\bar\lambda_t - \bar\mu_t)dt\right] \\ &\qquad \leq 2||\bar\lambda - \bar\mu||^2_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}||\lambda - \tilde\lambda||^2_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}, \end{aligned}$$

which, in turn, implies that

$$||(\tilde\mu,\tilde\nu) - (\mu,\nu)||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})} \leq \sqrt{2}||\bar\lambda - \bar\mu||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}||\lambda - \tilde\lambda||_{\text{bmo}(\mathbb{Q}^{\bar\lambda,\bar\nu})}.$$

Note that Lemma 6 and the estimates in Lemma 7 also imply that $||\bar\nu||_{\text{bmo}} \leq 2\epsilon$, where 2ϵ is taken from Lemma 7. Therefore, $||(\bar\lambda,\bar\nu)||_{\text{bmo}} \leq 4\epsilon$ and, similarly, $||\bar\lambda - \bar\mu||_{\text{bmo}} \leq 4\epsilon$. Therefore, it follows from Lemma 3 that

$$\begin{aligned} ||(\tilde\mu,\tilde\nu) - (\mu,\nu)||_{\text{bmo}} &\leq \sqrt{2}\frac{1+2\sqrt{2}\epsilon}{(1-2\sqrt{2}\epsilon)^2}||\bar\lambda - \bar\mu||_{\text{bmo}}||\lambda - \tilde\lambda||_{\text{bmo}} \\ &\leq \frac{1+2\sqrt{2}\epsilon}{(1-2\sqrt{2}\epsilon)^2}8\epsilon||\lambda - \tilde\lambda||_{\text{bmo}}. \end{aligned}$$

Choosing sufficiently small ϵ so that $\frac{1+2\sqrt{2}\epsilon}{(1-2\sqrt{2}\epsilon)^2}8\epsilon < 1$, the proof is complete after aggregating all components. □

Proof (of Theorem 2) We have shown in the sequence of lemmas above that, when (24) holds, the excess-demand map F is a contraction on $\mathcal{B}(2\epsilon)$ and that $(\boldsymbol{\mu}^\lambda, \boldsymbol{\nu}^\lambda) \in \text{bmo}^{\boldsymbol{I}}$. The Banach fixed point theorem implies that F has a unique fixed point λ with $||\lambda||_{\text{bmo}} \leq 2\epsilon$. Therefore the system (14) admits a solution $(\boldsymbol{Y}, \boldsymbol{\mu}, \boldsymbol{\nu})$ with $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \text{bmo}^{\boldsymbol{I}}$. Hence λ is an equilibrium by Theorem 1. For the uniqueness of equilibrium, Lemma 4 implies that any equilibrium λ satisfies $||\lambda||_{\text{bmo}} \leq \max_i ||(m^i, n^i)||_{\text{bmo}} \leq \sqrt{2}\epsilon$. However, we have already shown that there can be only one equilibrium λ in $\mathcal{B}(2\epsilon)$. Therefore we immediately have global uniqueness of equilibrium. Given the unique λ, by Lemma 1, $(\boldsymbol{Y}, \boldsymbol{\mu}, \boldsymbol{\nu})$ is the unique solution to (14) with $(\boldsymbol{\mu}, \boldsymbol{\nu}) \in \text{bmo}^{\boldsymbol{I}}$.□

3.5 Proof of Corollary 1

Summing both sides of $||E^i - E^j||_{\mathbb{L}^\infty} \leq \chi_0^E(||E^i||_{\mathbb{L}^\infty} + ||E^j||_{\mathbb{L}^\infty})$ over j, we obtain

$$\begin{aligned} I||E^i||_{\mathbb{L}^\infty} - ||E_\Sigma||_{\mathbb{L}^\infty} \leq \|IE^i - \textstyle\sum_j E^j\|_{\mathbb{L}^\infty} &\leq \textstyle\sum_j ||E^i - E^j||_{\mathbb{L}^\infty} \\ &\leq \chi_0^E I||E^i||_{\mathbb{L}^\infty} + \chi_0^E \textstyle\sum_j ||E^j||_{\mathbb{L}^\infty}, \end{aligned}$$

which implies

$$(1 - \chi_0^E)||E^i||_{\mathbb{L}^\infty} \leq \tfrac{1}{I}||E_\Sigma||_{\mathbb{L}^\infty} + \chi_0^E \tfrac{1}{I} \textstyle\sum_j ||E^j||_{\mathbb{L}^\infty}.$$

Summing both sides of the previous inequality over i yields

$$\textstyle\sum_i ||E^i||_{\mathbb{L}^\infty} \leq \frac{1}{1-2\chi_0^E}||E_\Sigma||.$$

The previous two inequalities combined then imply

$$||E^i||_{\mathbb{L}^\infty} \leq \tfrac{1}{1-2\chi_0^E}\tfrac{1}{I}||E_\Sigma||_{\mathbb{L}^\infty}, \quad \text{for all } i.$$

Therefore

$$\max_i \frac{||E^i||_{\mathbb{L}^\infty}}{\delta^i} \leq \tfrac{1}{1-2\chi_0^E}\tfrac{1}{I\delta_0}||E_\Sigma||_{\mathbb{L}^\infty},$$

where the right-hand side is strictly less than $(\epsilon^*)^2/4$ for sufficiently large I. Hence (17) is satisfied when $I \geq I_0$, for some I_0, and the existence of equilibrium follows from Theorem 2.

3.6 Proof of Corollary 2

Throughout the proof, we treat $\boldsymbol{G}$ as $\boldsymbol{G} - \xi^c$ and suppress the superscript i when we work with each component.

Recalling (6) and Remark 2, we have $\mathbb{E}[G^2] < \infty$, which combined with the assumption $D^b G, D^w G \in \mathcal{S}^\infty$ implies $G \in \mathbb{D}^{1,2}$. Let $G = \mathbb{E}[G] + M_T$, where $M_T = \overline{m} \cdot B_T + \overline{n} \cdot W_T$ for some $(\overline{m}, \overline{n})$. Clark-Ocone formula implies that $\mathbb{E}_\theta[D_\theta G] = (\overline{m}_\theta, \overline{n}_\theta)$, for any $\theta \le T$, hence $(\overline{m}, \overline{n}) \in \mathcal{S}^\infty$ as well. As a result, there exists a constant C such that $\langle M \rangle_T \le CT$, implying that G has at most Gaussian tail by Bernstein inequality (see Equation (4.i) in [2]), hence $\mathbb{E}[\exp(-2G)] < \infty$. Now combining the previous inequality with $D^b G, D^w G \in \mathcal{S}^\infty$, we obtain $\exp(-G) \in \mathbb{D}^{1,2}$, consequently, $V_t = \mathbb{E}_t[\exp(-G)] \in \mathbb{D}^{1,2}$ and

$$D^k_\theta V_t = -\mathbb{E}_t[e^{-G} D^k_\theta G] \quad \text{for all } \theta \le t \le T \text{ and } k = b \text{ or } w.$$

Applying Clark-Ocone formula to V_t yields

$$V_t = \mathbb{E}[V_t] + \int_0^t \mathbb{E}_\theta[D^b_\theta V_t] dB_\theta + \int_0^t \mathbb{E}_\theta[D^w_\theta V_t] dW_\theta.$$

On the other hand, $dV_\theta = -V_\theta m_\theta dB_\theta - V_\theta n_\theta dW_\theta$. Therefore $\mathbb{E}_\theta[D^b_\theta V_t] = -V_\theta m_\theta$ and $\mathbb{E}_\theta[D^w_\theta V_t] = -V_\theta m_\theta$, for $\theta \le t$. Hence,

$$m_\theta = -\frac{\mathbb{E}_\theta[D^b_\theta V_t]}{V_\theta} = \frac{\mathbb{E}_\theta[e^{-G} D^b_\theta G]}{\mathbb{E}_\theta[e^{-G}]} \le ||D^b G||_{\mathcal{S}^\infty},$$

which implies $||m||_{\mathcal{S}^\infty} \le ||D^w G||_{\mathcal{S}^\infty}$, and similarly, $||n||_{\mathcal{S}^\infty} \le ||D^w G||_{\mathcal{S}^\infty}$.

The statement now follows from Theorem 2 since, for $T < T^*$, where T^* is given in Corollary 2, we have

$$\begin{aligned} \max_i ||(m^i, n^i)||^2_{\text{bmo}} &< T^* \max_i (||m^i||^2_{\mathcal{S}^\infty} + ||n^i||^2_{\mathcal{S}^\infty}) \\ &\le T^* \max_i (||D^b G^i||^2_{\mathcal{S}^\infty} + ||D^w G^i||^2_{\mathcal{S}^\infty}) \le (\epsilon^*)^2. \end{aligned}$$

3.7 An a-priori bmo-estimate

Lemma 9 (An a-priori bmo-estimate for a single BSDE)
Given $\lambda \in \mathcal{P}^2$, let (Y, μ, ν) be a solution of the BSDE

$$dY_t = \mu_t\, dB_t + \nu_t\, dW_t + (\tfrac{1}{2}\nu_t^2 - \tfrac{1}{2}\lambda_t^2 + \mu_t \lambda_t)\, dt, \; Y_T = \xi.$$

If $Y \in \mathcal{S}^\infty$, then $(\mu, \nu) \in bmo$.

Proof For $\beta > 1$ and two stopping times $\tau \le \sigma \in \mathcal{T}$, Itô's formula yields

$$
\begin{aligned}
e^{-\beta Y_\sigma} \geq & e^{-\beta Y_\sigma} - e^{-\beta Y_\tau} = -\beta \int_\tau^\sigma e^{-\beta Y_u} (\mu_u dB_u + \nu_u dW_u) \\
& - \beta \int_\tau^\sigma e^{-\beta Y_u} \Big(\tfrac{1}{2}\nu_u^2 - \tfrac{1}{2}\lambda_u^2 + \lambda_u \mu_u \Big) du + \tfrac{1}{2}\beta^2 \int_\tau^\sigma e^{-\beta Y_u} (\mu_u^2 + \nu_u^2) du \\
\geq & -\beta \int_\tau^\sigma e^{-\beta Y_u} (\mu_u dB_u + \nu_u dW_u) + \tfrac{1}{2}(\beta^2 - \beta) \int_\tau^\sigma e^{-\beta Y_u} (\mu_u^2 + \nu_u^2)\, du,
\end{aligned}
$$

where we used the elementary fact that $a^2 - b^2 + 2bc \leq a^2 + c^2$, for all a, b, c. We pick a reducing sequence $\{\sigma_n\}_{n\in\mathbb{N}}$ for the stochastic integral above, project onto $\mathcal{F}_\tau$, and then let $n \to \infty$ to get

$$
\begin{aligned}
e^{\beta ||Y||_{\mathcal{S}^\infty}} & \geq \tfrac{1}{2}(\beta^2 - \beta)\mathbb{E}_\tau[\textstyle\int_\tau^T e^{\beta Y_u} (\mu_u^2 + \nu_u^2)\, du] \\
& \geq \tfrac{1}{2}(\beta^2 - \beta) e^{-\beta ||Y||_{\mathcal{S}^\infty}} \mathbb{E}_\tau[\int_\tau^T (\mu_u^2 + \nu_u^2)\, dt].
\end{aligned}
$$

This implies

$$
\mathbb{E}_\tau \left[\int_\tau^T (\mu_u^2 + \nu_u^2)\, du \right] \leq \frac{2}{\beta^2 - \beta} e^{2\beta ||Y||_{\mathcal{S}^\infty}}.
$$

Since the above inequality holds for arbitrary $\tau \in \mathcal{T}$, the statement follows. □

References

1. Anderson, R.M., Raimondo, R.C.: Equilibrium in continuous-time financial markets: Endogenously dynamically complete markets. Econometrica **76**(4), 841–907 (2008)
2. Barlow, M., Jacka, S., Yor, M.: Inequality for a pair of processes stopped at a random time. Proc. London Math. Soc. **52**(3), 142–172 (1986)
3. Barrieu, P., El Karoui, N.: Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs. Ann. Probab. **41**, 1831–2853 (2013)
4. Borch, K.: The safety loading of reinsurance premiums. Skand. Aktuarietidskr. **1960**, 163–184 (1961) (1960)
5. Briand, P., Hu, Y.: BSDE with quadratic growth and unbounded terminal value. Probab. Theory Relat. Fields **136**(4), 604–618 (2006)
6. Briand, P., Hu, Y.: Quadratic BSDEs with convex generators and unbounded terminal conditions. Probab. Theory Relat. Fields **141**(3-4), 543–567 (2008)
7. Chikvinidze, B., Mania, M.: New proofs of some results on bounded mean oscillation martingales using Backward stochastic differential equations. J. Theor. Probab. **27**, 1213–1228 (2014)
8. Choi, J.H., Larsen, K.: Taylor approximation of incomplete Radner equilibrium models (2014). To appear in Finance Stoch.
9. Dana, R.A., Pontier, M.: On existence of an Arrow-Radner equilibrium in the case of complete markets. A remark. Math. Oper. Res. **17**(1), 148–163 (1992)
10. Debreu, G.: The coefficient of resource utilization. Econometrica **19**(3), 273–292 (1951)
11. Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M., Stricker, C.: Exponential hedging and entropic penalties. Math. Finance **12**(2), 99–123 (2002)
12. Delbaen, F., Hu, Y., Bao, X.: Backward SDEs with superquadratic growth. Probab. Theory Relat. Fields **150**, 145–192 (2011)

13. Duffie, D.: Stochastic equilibria: existence, spanning number, and the "no expected financial gain from trade" hypothesis. Econometrica **54**(5), 1161–1183 (1986)
14. Duffie, D., Huang, C.F.: Implementing Arrow-Debreu equilibria by continuous trading of few long-lived securities. Econometrica **53**(6), 1337–1356 (1985)
15. Elie, R., Briand, P.: A simple constructive approach to quadratic BSDEs with or without delay. Stoch. Process. Appl. **123**(8), 2921–2939 (2013)
16. Frei, C., dos Reis, G.: A financial market with interacting investors: does an equilibrium exist? Math. Financ. Econ. **4**(3), 161–182 (2011)
17. Hu, Y., Peng, S.: On the comparison theorem for multidimensional BSDEs. C. R. Math. Acad. Sci. Paris **343**(2), 135–140 (2006)
18. Karatzas, I., Lakner, P., Lehoczky, J.P., Shreve, S.E.: Equilibrium in a simplified dynamic, stochastic economy with heterogeneous agents. In: Stochastic analysis, pp. 245–272. Academic Press, Boston, MA (1991)
19. Karatzas, I., Lehoczky, J.P., Shreve, S.E.: Existence and uniqueness of multi-agent equilibrium in a stochastic, dynamic consumption/investment model. Math. Oper. Res. **15**(1), 80–128 (1990)
20. Karatzas, I., Lehoczky, J.P., Shreve, S.E.: Equilibrium models with singular asset prices. Math. Finance **1**, 11–29 (1991)
21. Karatzas, I., Shreve, S.E.: Methods of mathematical finance, *Applications of Mathematics (New York)*, vol. 39. Springer-Verlag, New York (1998)
22. Kazamaki, N.: Continuous exponential martingales and BMO, *Lecture Notes in Mathematics*, vol. 1579. Springer-Verlag, Berlin (1994)
23. Kobylanski, M.: Backward stochastic differential equations and partial differential equations with quadratic growth. Ann. Probab. **28**(2), 558–602 (2000)
24. Kramkov, D., Pulido, S.: A system of quadratic BSDEs arising in a price impact model (2014). To appear in Ann. Appl. Probab.
25. Magill, M., Quinzii, M.: Theory of Incomplete Markets, Volume 1. MIT Press, Cambridge and London (1996)
26. Radner, R.: Equilibrium under uncertainty. Econometrica **36**(1), 31–58 (1982)
27. Rouge, R., El Karoui, N.: Pricing via utility maximization and entropy. Math. Finance **10**(2), 259–276 (2000). INFORMS Applied Probability Conference (Ulm, 1999)
28. Tevzadze, R.: Solvability of backward stochastic differential equations with quadratic growth. Stochastic Process. Appl. **118**(3), 503–515 (2008)
29. Zhao, Y.: Stochastic equilibria in a general class of incomplete brownian market environments. Ph.D. thesis, The University of Texas at Austin (2012)
30. Žitković, G.: Financial equilibria in the semimartingale setting: complete markets and markets with withdrawal constraints. Finance Stoch. **10**(1), 99–119 (2006)
31. Žitković, G.: An example of a stochastic equilibrium with incomplete markets. Finance Stoch. **16**(2), 177–206 (2012)

Finite Markov Chains Coupled to General Markov Processes and An Application to Metastability I

Thomas G. Kurtz and Jason Swanson

Abstract We consider a diffusion given by a small noise perturbation of a dynamical system driven by a potential function with a finite number of local minima. The classical results of Freidlin and Wentzell show that the time this diffusion spends in the domain of attraction of one of these local minima is approximately exponentially distributed and hence the diffusion should behave approximately like a Markov chain on the local minima. By the work of Bovier and collaborators, the local minima can be associated with the small eigenvalues of the diffusion generator. Applying a Markov mapping theorem, we use the eigenfunctions of the generator to couple this diffusion to a Markov chain whose generator has eigenvalues equal to the eigenvalues of the diffusion generator that are associated with the local minima and establish explicit formulas for conditional probabilities associated with this coupling. The fundamental question then becomes to relate the coupled Markov chain to the approximate Markov chain suggested by the results of Freidlin and Wentzel.

1 Introduction

Fix $\varepsilon > 0$ and consider the stochastic process,

$$X_\varepsilon(t) = X_\varepsilon(0) - \int_0^t \nabla F(X_\varepsilon(s))\,ds + \sqrt{2\varepsilon}\,W(t), \tag{1}$$

where $F \in C^3(\mathbb{R}^d)$ and W is a standard d-dimensional Brownian motion. For the precise assumptions on F, see Section 3.1. Let φ be the solution to the differential

Thomas G. Kurtz
University of Wisconsin-Madison, e-mail: kurtz@math.wisc.edu

Jason Swanson
University of Central Florida, e-mail: Jason.Swanson@ucf.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_12

equation $\varphi' = -\nabla F(\varphi)$. We will use φ_x to denote the solution with $\varphi_x(0) = x$. The process X_ε is a small-noise perturbation of the deterministic process φ.

Suppose that $\mathcal{M} = \{x_0, \dots, x_m\}$ is the set of local minima of the potential function F. The points x_j are stable points for the process φ. For X_ε, however, they are not stable. The process X_ε will initially gravitate toward one of the x_j and move about randomly in a small neighborhood of this point. But after an exponential amount of time, a large fluctuation of the noise term will move the process X_ε out of the domain of attraction of x_j and into the domain of attraction of one of the other minima. We say that each point x_j is a point of *metastability* for the process X_ε.

If X is a cadlag process in a complete, separable metric space S adapted to a right continuous filtration (assumptions that are immediately satisfied for all processes considered here) and H is either open or closed, then $\tau_H^X = \inf\{t > 0 : X(t) \text{ or } X(t-) \in H\}$ is a stopping time (see, for example, [8, Proposition 1.5]). If $x \in S$, let $\tau_x^X = \tau_{\{x\}}^X$. We may sometimes also write $\tau^X(H)$, and if the process is understood, we may omit the superscript.

Let

$$D_j = \{x \in \mathbb{R}^d : \lim_{t\to\infty} \varphi_x(t) = x_j\} \tag{2}$$

be the domains of attraction of the local minima. It is well-known (see, for example, [9], [4, Theorem 3.2], [5, Theorems 1.2 and 1.4], and [7]) that as $\varepsilon \to 0$, $\tau^{X_\varepsilon}(D_j^c)$ is asymptotically exponentially distributed under P^{x_j}. It is therefore common to approximate the process X_ε by a continuous time Markov chain on the set $\mathcal{M}$ (or equivalently on $\{0, \dots, m\}$). In fact, metastability can be defined in terms of convergence, in an appropriate sense, to a continuous time Markov chain. (See the survey article [15] for details.) Beltrán and Landim [2, 3] introduced a general method for proving the metastability of a Markov chain. Along similar lines, Rezakhanlou and Seo [19] developed such a method for diffusions. For an alternative approach using intertwining relations, see [1].

In this project, for each $\varepsilon > 0$, we wish to capture this approximate Markov chain behavior by coupling X_ε to a continuous time Markov chain, Y_ε, on $\{0, \dots, m\}$. We will refer to the indexed collection of coupled processes, $\{(X_\varepsilon, Y_\varepsilon) : \varepsilon > 0\}$, as a *coupling sequence*. Our objective is to investigate the possibility of constructing a coupling sequence which satisfies

$$P(X_\varepsilon(t) \in D_j \mid Y_\varepsilon(t) = j) \to 1 \tag{3}$$

as $\varepsilon \to 0$, for all j. We also want the transition rate for Y_ε to go from i to j to be asymptotically equivalent as $\varepsilon \to 0$ to the transition rate for X_ε to go from a neighborhood of x_i to a neighborhood of x_j. That is, we would like

$$E^i[\tau_j^{Y_\varepsilon}] \sim E^{x_i}[\tau_{B_\rho(x_0)}^{X_\varepsilon}] \tag{4}$$

as $\varepsilon \to 0$, for all i and j, where $B_\rho(x)$ is the ball of radius ρ centered at x.

In this paper (Part I), we develop our general coupling construction. The construction goes beyond the specific case of interest here. It is a construction that builds a coupling between a Markov process on a complete and separable metric space and a continuous-time Markov chain where the generators of the two processes have common eigenvalues. The coupling is done in such a way that observations of the chain yield quantifiable conditional probabilities about the process. This coupling construction is built in Section 2 and uses the Markov mapping theorem (Theorem 19). In Section 3, we apply this construction method to reversible diffusions in $\mathbb{R}^d$ driven by a potential function with a finite number of local minima.

With this coupling construction in hand, we can build the coupling sequences described above. In our follow-up work (Part II), we take up the question of the existence and uniqueness of a coupling sequence that satisfies requirements (3) and (4).

2 The general coupling

2.1 Assumptions and definitions

Given a Markov process X with generator A satisfying Assumption 1, we will use the Markov mapping theorem to construct a coupled pair, (X,Y), in such a way that for a specified class of initial distributions, Y is a continuous-time Markov chain on a finite state space. The construction then allows us to explicitly compute the conditional distribution of X given observations of Y.

For explicit definitions of the notation used here and throughout, see the Appendix.

Assumption Let E be a complete and separable metric space.

(i) $A \subset \overline{C}(E) \times \overline{C}(E)$.
(ii) A has a stationary distribution $\varpi \in \mathcal{P}(E)$, which implies $\int_E Af\, d\varpi = 0$ for all $f \in \mathcal{D}(A)$.
(iii) For some m, there exist signed measures $\varpi_1, \ldots, \varpi_m$ on E and positive real numbers $\lambda_1, \ldots, \lambda_m$ such that, for each $k \in \{1, \ldots, m\}$ and $f \in \mathcal{D}(A)$,

$$\int_E Af\, d\varpi_k = -\lambda_k \int_E f\, d\varpi_k, \tag{5}$$

$$\varpi_k(dx) = \eta_k(x)\varpi(dx), \text{ where } \eta_k \in \overline{C}(E), \tag{6}$$

$$\varpi_k(E) = 0. \tag{7}$$

We define $\varpi_0 = \varpi$ and $\eta_0 = 1$. □

Remark If $(1,0) \in A$, then (5) implies (7). □

Remark In what follows, we will make use of the assumption that the functions η_k are continuous. However, this assumption can be relaxed by appealing to the methods in Kurtz and Stockbridge [14]. □

Assumption Let E be a complete and separable metric space. Let $A \subset \overline{C}(E)\times\overline{C}(E)$, $m \in \mathbb{N}$, $Q \in \mathbb{R}^{(m+1)\times(m+1)}$, and $\xi^{(1)},\dots,\xi^{(m)} \in \mathbb{R}^{m+1}$.

(i) A and m satisfy Assumption 1.
(ii) Q is the generator of a continuous-time Markov chain with state space $E_0 = \{0,1,\dots,m\}$ and eigenvalues $\{0,-\lambda_1,\dots,-\lambda_m\}$.
(iii) The vectors $\xi^{(1)},\dots,\xi^{(m)}$ are right eigenvectors of Q, corresponding to the eigenvalues $-\lambda_1,\dots,-\lambda_m$.
(iv) For each $i \in \{0,1,\dots,m\}$, the function

$$\alpha_i(x) = 1 + \sum_{k=1}^{m} \xi_i^{(k)} \eta_k(x) \tag{8}$$

satisfies $\alpha_i(x) > 0$ for all $x \in E$.

We define $\xi^{(0)} = (1,\dots,1)^T$, so that the function $\alpha : E \to \mathbb{R}^{m+1}$ is given by $\alpha = \sum_{k=0}^{m} \xi^{(k)}\eta_k$. □

Remark Given (A,m,Q) satisfying (i) and (ii) of Assumption 4, it is always possible to choose vectors $\xi^{(1)},\dots,\xi^{(m)}$ satisfying (iii) and (iv). This follows from the fact that each η_k is a bounded function. □

Definition Suppose $(A,m,Q,\xi^{(0)},\dots,\xi^{(m)})$ satisfies Assumption 4. For $0 \le j \ne i \le m$, define

$$q_{ij}(x) = Q_{ij}\frac{\alpha_j(x)}{\alpha_i(x)}. \tag{9}$$

Note that $q_{ij} \in C(E)$. Let $S = E \times E_0$. Define $B \subset \overline{C}(S)\times C(S)$ by

$$Bf(x,i) = Af(x,i) + \sum_{j\ne i} q_{ij}(x)(f(x,j) - f(x,i)), \tag{10}$$

where we take

$$\mathcal{D}(B) = \{f(x,i) = f_1(x)f_2(i) : f_1 \in \mathcal{D}(A), f_2 \in B(E_0)\} \tag{11}$$

In particular, $Af(x,i) = f_2(i)Af_1(x)$.

For each $i \in E_0$, define the measure $\alpha(i,\cdot)$ on E by

$$\alpha(i,\Gamma) = \int_\Gamma \alpha_i(x)\varpi(dx), \tag{12}$$

for all $\Gamma \in \mathcal{B}(E)$. Note that by (8), (7), and (6), these are probability measures. □

2.2 Construction of the coupling

We are now ready to construct our coupled pair, (X,Y), which will have generator B, to prove, for appropriate initial conditions, that the marginal process Y is a Markov chain with generator Q, and to establish our conditional probability formulas. We first require two lemmas.

Lemma 1 *In the setting of Definition 6, let X be a cadlag solution of the martingale problem for A. Then there exists a cadlag process Y such that (X,Y) solves the (local) martingale problem for B. If X is Markov, then (X,Y) is Markov. If the martingale problem for A is well-posed, then the martingale problem for B is well-posed.*

Remark We are not requiring the q_{ij} to be bounded, so for the process we construct,

$$f(X(t),Y(t)) - f(X(0),Y(0)) - \int_0^t Bf(X(s),Y(s))\,ds$$

may only be a local martingale. □

Proof (Proof of Lemma 1) Let $X(t)$ be a cadlag solution to the martingale problem for A. Let $\{N_{ij} : i,j \in E_0, i \neq j\}$ be a family of independent unit rate Poisson processes, which is independent of X. Then the equation

$$Y(t) = k + \sum_{i\neq j}(j-i)N_{ij}\left(\int_0^t 1_{\{i\}}(Y(s))q_{ij}(X(s))\,ds\right) \tag{13}$$

has a unique solution, and as in [12], the process $Z = (X,Y)$ is a solution of the (local) martingale problem for B. If X is Markov, the uniqueness of the solution of (13) ensures that (X,Y) is Markov. Similarly, A well-posed implies B is well posed. □

Lemma 2 *Let A satisfy Assumption 1. Taking $\psi(x,i) = 1 + \sum_{j\neq i} q_{ij}(x) \geq 1$, if A satisfies Condition 17, then B satisfies Condition 17 with E replaced by $S = E \times E_0$.*

Proof Since $\mathcal{D}(A)$ is closed under multiplication, $\mathcal{D}(B)$ defined in (11) is closed under multiplication.

Since we are assuming that $\mathcal{R}(A) \subset \bar{C}(E)$, for each $f \in \mathcal{D}(B)$, there exists $c_f > 0$ such that $|Bf(x,i)| \leq c_f\psi(x)$.

Condition 17(iii) for A and the separability of $B(E_0)$ implies Condition 17(iii) for B_0.

Since A is a pre-generator and B is a perturbation of A by a jump operator, B_0 is a pre-generator. □

Theorem Suppose A satisfies Condition 17 and $(A,m,Q,\xi^{(1)},\ldots,\xi^{(m)})$ satisfies Assumption 4. Let B be given by (10) and for $p_i \geq 0$, $\sum_{i=0}^m p_i = 1$, define

$$\nu(\Gamma\times\{i\}) = p_i\alpha(i,\Gamma), \quad \Gamma \in \mathcal{B}(E), i \in E_0.$$

If $\widetilde{Y}$ is a cadlag E_0-valued Markov chain with generator Q and initial distribution $\{p_i\}$, then there exists a solution (X, Y) of the martingale problem for (B, ν) such that Y and $\widetilde{Y}$ have the same distribution on $D_{E_0}[0, \infty)$, and

$$P(X(t) \in \Gamma \mid \mathcal{F}_t^Y) = \alpha(Y(t), \Gamma), \tag{14}$$

for all $t \geq 0$ and all $\Gamma \in \mathcal{B}(E)$. □

Proof We apply Theorem 19 to the operator $B \subset \overline{C}(S) \times C(S)$.

Let $\gamma : S \to E_0$ be the coordinate projection. Let $\widetilde{\alpha}$ be the transition function from E_0 into S given by the product measure $\widetilde{\alpha}(i, \cdot) = \alpha(i, \cdot) \otimes \delta_i^{E_0}$, where $\alpha(i, \cdot)$ is given by (12). Then $\widetilde{\alpha}(i, \gamma^{-1}(i)) = 1$ and

$$\widetilde{\psi}(i) \equiv \int_S \psi(z)\widetilde{\alpha}(i, dz) = \int_E \psi(x, i)\alpha_i(x)\varpi(dx) = 1 + \sum_{j \neq i} Q_{ij} < \infty,$$

for each $i \in E_0$. Define

$$C = \left\{ \left(\int_S f(z)\widetilde{\alpha}(\cdot, dz), \int_S Bf(z)\widetilde{\alpha}(\cdot, dz) \right) : f \in \mathcal{D}(B) \right\} \subset \mathbb{R}^{m+1} \times \mathbb{R}^{m+1}.$$

The result follow by Theorem 19, if we can show that $Cv = Qv$ for every vector $v \in \mathcal{D}(C)$. Given $f \in \mathcal{D}(B)$, let

$$\overline{f}(i) = \int_S f(z)\widetilde{\alpha}(i, dz) = \int_E f(x, i)\alpha(i, dx) = \int_E f(x, i)\alpha_i(x)\varpi(dx).$$

Note that

$$C\overline{f}(i) = \int_E Bf(x, i)\alpha_i(x)\varpi(dx).$$

Since $\lambda_0 = 0$, by (5) and the definition of $q_{ij}(x)$,

$$C\overline{f}(i) = -\sum_{k=0}^{m} \xi_i^{(k)} \lambda_k \int_E f(x, i)\eta_k(dx) + \sum_{j \neq i} Q_{ij} \int_E \alpha_j(x)(f(x, j) - f(x, i))\varpi(dx).$$

By assumption $Q\xi^{(k)} = -\lambda_k \xi^{(k)}$, so $-\xi_i^{(k)} \lambda_k = \sum_{j=0}^m Q_{ij}\xi_j^{(k)}$ and

$$\begin{aligned}
-\sum_{k=0}^{m} \xi_i^{(k)} \lambda_k \int_E f(x, i)\eta_k(dx) &= \sum_{k=0}^{m} \sum_{j=0}^{m} Q_{ij}\xi_j^{(k)} \int_E f(x, i)\eta_k(dx) \\
&= \sum_{j=0}^{m} Q_{ij} \sum_{k=0}^{m} \xi_j^{(k)} \int_E f(x, i)\eta_k(dx) \\
&= \sum_{j=0}^{m} Q_{ij} \int_E f(x, i)\alpha_j(x)\varpi(dx).
\end{aligned}$$

This gives

$$C\overline{f}(i) = Q_{ii} \int_E f(x,i)\alpha_i(x)\varpi(dx) + \sum_{j \neq i} Q_{ij} \int_S f(x,j)\alpha_j(x)\varpi(dx)$$
$$= \sum_{j=0}^{m} Q_{ij}\overline{f}(j) = Q\overline{f}(i).$$

It follows that $\widetilde{Y}$ is a solution to the martingale problem for (C, p).

By Theorem 19(a), there exists a solution $Z = (X, Y)$ of the martingale problem for (B, ν) such that $Y = \gamma(Z)$ and $\widetilde{Y}$ have the same distribution on $D_{E_0}[0, \infty)$. Theorem 19(b) implies (14). □

Remark In what follows, we may still write expectations with the notation E^x or E^i, even when we have a coupled process, (X, Y). The meaning will be determined by context, depending on whether the integrand of the expectation involves only X or only Y. □

3 Reversible diffusions

3.1 Assumptions on the potential function

We now consider the special case of our coupling when X is a reversible diffusion on $\mathbb{R}^d$ driven by a potential function F and a small white noise perturbation. We will need to use several results from the literature about the eigenvalues and eigenfunctions of the generator of X. We assume the following on F.

Assumption (i) $F \in C^3(\mathbb{R}^d)$ and $\lim_{|x|\to\infty} F(x) = \infty$.
(ii) F has $m + 1 \geq 2$ local minima $\mathcal{M} = \{x_0, \ldots, x_m\}$.
(iii) There exist constants $a_i > 0$ and $c_i > 0$ such that $a_2 < 2a_1 - 2$, and

$$c_1|x|^{a_1} - c_2 \leq |\nabla F(x)|^2 \leq c_3|x|^{a_2} + c_4, \tag{15}$$
$$c_1|x|^{a_1} - c_2 \leq (|\nabla F(x)| - 2\Delta F(x))^2 \leq c_3|x|^{a_2} + c_4. \tag{16}$$

Remark Note that $2 < a_1 \leq a_2$. To see this, observe that (15) implies $a_1 \leq a_2$. Thus, $a_1 \leq a_2 < 2a_1 - 2$, which implies $a_1 > 2$. □

Lemma 3 *Under Assumption 10, there exist constants $\widetilde{c}_i > 0$ such that*

$$\widetilde{c}_1|x|^{\widetilde{a}_1} - \widetilde{c}_2 \leq |F(x)| \leq \widetilde{c}_3|x|^{\widetilde{a}_2} + \widetilde{c}_4, \tag{17}$$

where $\widetilde{a}_i = a_i/2 + 1$.

Proof Since

$$F(x) = F(0) + \int_0^1 \nabla F(sx) \cdot x\, ds,$$

it follows from (15) that

$$|F(x)| \leq |F(0)| + |x|(c_3|x|^{a_2} + c_4)^{1/2},$$

and the upper bound in (17) follows immediately.

Since $F \to \infty$, there exists $C > 0$ such that $F(x) > -C$ for all $x \in \mathbb{R}^d$, and since $|\nabla F| \to \infty$, there exists $R > 0$ such that $|\nabla F(x)| \geq 1$ whenever $|x| \geq R$.

Recall that φ_x satisfies $\varphi_x' = -\nabla F(\varphi_x)$ and $\varphi_x(0) = x$, and define

$$T_x = \inf\{t \geq 0 : |\varphi_x(t)| < R\}.$$

Suppose there exists x such that $T_x = \infty$. Then, for all $t > 0$,

$$\begin{aligned}
-C < F(\varphi_x(t)) &= F(x) + \int_0^t \nabla F(\varphi_x(s)) \cdot \varphi_x'(s)\, ds \\
&= F(x) - \int_0^t |\nabla F(\varphi_x(s))|^2\, ds \\
&\leq F(x) - t.
\end{aligned}$$

Therefore, $F(x) \geq t - C$ for all t, a contradiction, and we must have $T_x < \infty$ for all $x \in \mathbb{R}^d$.

Let $L = \sup_{|x| \leq R} F(x)$. By (15) and the fact that $F \to \infty$, we may choose $R' \geq R$ and $C' > 0$ such that $F(x) > L$ and $|\nabla F(x)| \geq C'|x|^{a_1/2}$ whenever $|x| > R'$.

Fix $x \in \mathbb{R}^d$ with $|x| > 2R'$, so that $F(x) > L$. Since $|\varphi_x(T_x)| = R$, it follows that $F(\varphi_x(T_x)) \leq L$. By the continuity of φ_x, we may choose $T' \in (0, T_x]$ such that $F(\varphi_x(T')) = L$. We then have

$$\begin{aligned}
L &= F(x) + \int_0^{T'} \nabla F(\varphi_x(t)) \cdot \varphi_x'(t)\, dt \\
&= F(x) - \int_0^{T'} |\nabla F(\varphi_x(t))||\varphi_x'(t)|\, dt.
\end{aligned}$$

Let $T'' = \inf\{t \geq 0 : |\varphi_x(t)| < |x|/2\}$. Note that $F(\varphi_x(T')) = L$ implies $|\varphi_x(T')| \leq R' < |x|/2$, and therefore $T'' \leq T'$. Moreover, for all $t < T''$, we have $|\varphi_x(t)| \geq |x|/2 > R'$, which implies

$$|\nabla F(\varphi_x(t))| \geq C'|\varphi_x(t)|^{a_1/2} \geq C' \left(\frac{|x|}{2}\right)^{a_1/2}.$$

Thus,

$$L \leq F(x) - C' \left(\frac{|x|}{2}\right)^{a_1/2} \int_0^{T''} |\varphi_x'(t)|\, dt.$$

But $\int_0^{T''} |\varphi_x'(t)|\,dt$ is the length of φ_x from $t = 0$ to $t = T''$, which is bounded below by

$$|\varphi_x(T'') - \varphi_x(0)| \geq |\varphi_x(0)| - |\varphi_x(T'')| = |x| - \frac{|x|}{2} = \frac{|x|}{2}.$$

Therefore, for all $|x| > 2R'$, we have $F(x) \geq C''|x|^{a_1/2+1} - |L|$, where $C'' = 2^{-a_1/2-1}C'$, and this proves the lower bound in (17). □

3.2 Spectral properties of the generator

Having established our assumptions on F, we now turn our attention to the diffusion process, X_ε, given by (1). To simplify notation, we may sometimes omit the ε. The process X has generator $A = \varepsilon\Delta - \nabla F \cdot \nabla$. To show that A meets the requirements of our coupling from Section 2, we must prove certain results about its eigenvalues and eigenfunctions. For this, we begin with some notation, a lemma, and two results from the literature.

Define $\pi(x) = \pi_\varepsilon(x) = e^{-F(x)/2\varepsilon}$. Let

$$V = V_\varepsilon := \frac{\Delta\pi}{\pi} = \frac{1}{4\varepsilon^2}|\nabla F|^2 - \frac{1}{2\varepsilon}\Delta F. \tag{18}$$

Lemma 4 *Let V_ε be given by* (18)*, where F satisfies Assumption 10. Recall the constants a_i from* (15)-(16)*. For all $\varepsilon \in (0,1)$, there exist constants $c_{i,\varepsilon} > 0$ such that*

$$c_{1,\varepsilon}|x|^{a_1} - c_{2,\varepsilon} \leq V_\varepsilon(x) \leq c_{3,\varepsilon}|x|^{a_2} + c_{4,\varepsilon}.$$

In particular, $V_\varepsilon \to \infty$ for all $\varepsilon \in (0,1)$.

Proof Fix $\varepsilon \in (0,1)$. By (15) and (16), for x sufficiently large,

$$c|x|^{a_1} \leq (|\nabla F(x)| - 2\Delta F)^2 \leq C|x|^{a_2},$$

and

$$c|x|^{a_1} \leq |\nabla F(x)|^2 \leq C|x|^{a_2},$$

for some $0 < c \leq C < \infty$. Note that

$$4V_1 = |\nabla F|^2 - 2\Delta F = (|\nabla F| - 2\Delta F) + (|\nabla F|^2 - |\nabla F|).$$

Hence, for x sufficiently large, $V_1(x) \leq C_1|x|^{a_2}$. Also,

$$V_1(x) \geq \frac{1}{4}(c|x|^{a_1} - C_2|x|^{a_2/2}).$$

Since $a_1 > a_2/2$, it follows that for x sufficiently large, $V_1(x) \geq \widetilde{c}|x|^{a_1}$. Therefore, there exist constants $\widetilde{c_i} > 0$ such that

$$\widetilde{c}_1|x|^{a_1} - \widetilde{c}_2 \le V_1(x) \le \widetilde{c}_3|x|^{a_2} + \widetilde{c}_4,$$

and

$$\widetilde{c}_1|x|^{a_1} - \widetilde{c}_2 \le |\nabla F(x)|^2 \le \widetilde{c}_3|x|^{a_2} + \widetilde{c}_4,$$

for all $x \in \mathbb{R}^d$. Note that

$$V_\varepsilon = \frac{1}{\varepsilon}\left(V_1 + \left(\frac{1-\varepsilon}{4\varepsilon}\right)|\nabla F|^2\right),$$

so that

$$\frac{1}{\varepsilon}V_1 \le V_\varepsilon \le \frac{1}{\varepsilon}V_1 + \frac{1}{\varepsilon^2}|\nabla F|^2.$$

From here, the lemma follows easily. □

The following two theorems are from [6]. Theorem 12 is a consequence of [6, Theorem 4.5.4] and [6, Lemma 4.2.2]. Theorem 13 is part of [6, Theorem 2.1.4].

Theorem Let $H = -\Delta + W$, where W is continuous with $W \to \infty$. Let λ denote the smallest eigenvalue of H, and ψ the corresponding eigenfunction, normalized so that $\|\psi\|_{L^2(\mathbb{R}^d)} = 1$. Define $Uf = \psi f$ and $\widetilde{H} = U^{-1}(H - \lambda)U$. If

$$\widehat{c}_1|x|^{\widehat{a}_1} - \widehat{c}_2 \le |W(x)| \le \widehat{c}_3|x|^{\widehat{a}_2} + \widehat{c}_4,$$

where $\widehat{a}_i > 0$, $\widehat{c}_i > 0$, and $\widehat{a}_2 < 2\widehat{a}_1 - 2$, then $e^{-\widetilde{H}t}$ is an ultracontractive symmetric Markov semigroup on $L^2(\mathbb{R}^d, \psi(x)^2\,dx)$. That is, for each $t \ge 0$, the operator $e^{-\widetilde{H}t}$ is a bounded operator mapping $L^2(\mathbb{R}^d, \psi(x)^2\,dx)$ to $L^\infty(\mathbb{R}^d, \psi(x)^2\,dx)$. □

Theorem Let e^{-Ht} be an ultracontractive symmetric Markov semigroup on $L^2(\Omega, \mu)$, where Ω is a locally compact, second countable Hausdorff space and μ is a Borel measure on Ω. If $\mu(\Omega) < \infty$, then each eigenfunction of H belongs to $L^\infty(\Omega, \mu)$. □

This next proposition establishes the spectral properties of A that are needed to carry out the construction of our coupling.

Proposition Fix $\varepsilon > 0$. The operator $H = -\Delta + V_\varepsilon$ is a self-adjoint operator on $L^2(\mathbb{R}^d)$ with discrete, nonnegative spectrum $\widehat{\lambda}_k \uparrow \infty$ and corresponding orthonormal eigenfunctions ψ_k. Each ψ_k is locally Hölder continuous. Moreover, $\widehat{\lambda}_0 = 0$ is simple and ψ_0 is proportional to π. We define μ by $\mu(dx) = \pi(x)^2\,dx$ and $\varpi = Z^{-1}\mu$, where $Z = \mu(\mathbb{R}^d)$. The operator $\widetilde{H}$ given by $\widetilde{H}f = \pi^{-1}H(\pi f)$ is a self-adjoint operator on $L^2(\varpi)$ with eigenvalues $\widehat{\lambda}_k$ and orthogonal eigenfunctions $\widehat{\eta}_k = \psi_k/\pi$. The functions $\widehat{\eta}_k$ have norm one in $L^2(\mu)$, whereas the functions $\eta_k = Z^{1/2}\widehat{\eta}_k$ have norm one in $L^2(\varpi)$.

For $f \in C_c^\infty(\mathbb{R}^d)$, we have $-\varepsilon\widetilde{H}f = \varepsilon\Delta f - \nabla F \cdot \nabla f$. Hence, if we define A by

$$A = \{(f, -\varepsilon\widetilde{H}f) : f \in C_c^\infty(\mathbb{R}^d)\},$$

then A is the generator for the diffusion process given by (1). For each $x \in \mathbb{R}^d$, (1) has a unique, global solution for all time, so that the process X with $X(0) = x$ is a solution to the martingale problem for (A, δ_x). The operator A is graph separable, and $\mathcal{D}(A)$ is separating and closed under multiplication. The measure ϖ is a stationary distribution for A. Moreover,

$$\int Af\,d\varpi_k = -\lambda_k \int f\,d\varpi_k,$$

where $\varpi_k(dx) = \eta_k(x)\varpi(dx)$ and $\lambda_k = \varepsilon\widehat{\lambda}_k$. The signed measures ϖ_k satisfy $\varpi_k(\mathbb{R}^d) = 0$, and each η_k belongs to $\overline{C}(\mathbb{R}^d)$, the space of bounded, continuous functions on $\mathbb{R}^d$. □

Proof Note that $V \to \infty$ by Lemma 4. Therefore, by [18, Theorem XIII.67], we have that H is a self-adjoint operator on $L^2(\mathbb{R}^d)$ with compact resolvent. It follows (see [6, pp. 108–109, 119–120, and Proposition 1.4.3]) that H has a purely discrete spectrum and there exists a complete, orthonormal set of eigenfunctions $\{\psi_k\}_{k=0}^{\infty}$ with corresponding eigenvalues $\widehat{\lambda}_k \uparrow \infty$. Moreover, $\widehat{\lambda}_0$ is simple and ψ_0 is strictly positive.

Since V is locally bounded, and $(-\Delta + V - \widehat{\lambda}_k)\psi_k = 0$, [10, Theorem 8.22] implies that, for each compact $K \subset \mathbb{R}^d$, ψ_k is Hölder continuous on K with exponent $\gamma(K)$.

Define $U : L^2(\mu) \to L^2(\mathbb{R}^d)$ by $Uf = \pi f$, so that $\widetilde{H} = U^{-1}HU$. Since U is an isometry, $\widetilde{H}$ is self-adjoint on $L^2(\mu)$ and has the same eigenvalues as H. Note that, for any $f \in \mathcal{D}(\widetilde{H})$, it follows from Green's identity that

$$\begin{aligned}\langle f, \widetilde{H}f\rangle_{L^2(\mu)} &= \langle \pi f, H(\pi f)\rangle_{L^2(\mathbb{R}^d)} = \int |\nabla(\pi f)|^2 + \int V(\pi f)^2 \\ &= \int |\nabla(\pi f)|^2 + \int (\Delta\pi)\pi f^2 = \int |\nabla(\pi f)|^2 - \int \nabla\pi \cdot \nabla(\pi f^2).\end{aligned}$$

Using the product rule, $\nabla(gh) = g\nabla h + h\nabla g$, this simplifies to

$$\begin{aligned}\langle f, \widetilde{H}f\rangle_{L^2(\mu)} &= \int (|\nabla\pi|^2 f^2 + 2f\pi(\nabla f\cdot\nabla\pi) + |\nabla f|^2\pi^2 - |\nabla\pi|^2 f^2 - \pi(\nabla(f^2)\cdot\nabla\pi)) \\ &= \int (2f\pi(\nabla f\cdot\nabla\pi) + |\nabla f|^2\pi^2 - \pi(\nabla(f^2)\cdot\nabla\pi)) = \int |\nabla f|^2\pi^2,\end{aligned}$$

showing that $\widetilde{H}$ cannot have a negative eigenvalue. Hence, $\widehat{\lambda}_0 \geq 0$.

By (17), we have $\pi \in L^2(\mathbb{R}^d)$, so that $\pi \in \mathcal{D}(H)$ with $H\pi = 0$. Hence, since $\widehat{\lambda}_0$ is nonnegative and has multiplicity one, it follows that $\widehat{\lambda}_0 = 0$ and ψ_0 is proportional to π.

Observe that, if $f \in C_c^\infty$, then, using the product rule for the Laplacian and the identity $V = \Delta\pi/\pi$, we have

$$-\widetilde{H}f = -\frac{1}{\pi}H(\pi f) = \frac{1}{\pi}(\Delta(\pi f) - V\pi f) = \frac{1}{\pi}(f\Delta\pi + 2\nabla\pi\cdot\nabla f + \pi\Delta f - f\Delta\pi).$$

Since $2\varepsilon\nabla\pi/\pi = -\nabla F$, we have $-\varepsilon\widetilde{H}f = \varepsilon\Delta f - \nabla F\cdot\nabla f$.

Since ∇F is locally Lipschitz, (1) has a unique solution up to an explosion time (see [17, Theorem V.38]). Since $\lim_{|x|\to\infty} F = \infty$ by assumption and $\lim_{|x|\to\infty} AF(x) = \infty$ by Lemma 4.2, it follows that F is a Liapunov function for X_ε proving that X_ε does not explode.

By [13, Remark 2.5], A is graph separable. Clearly $\mathcal{D}(A)$ is closed under multiplication. Since $\mathcal{D}(A)$ separates points and $\mathbb{R}^d$ is complete and separable, $\mathcal{D}(A)$ is separating (see [8, Theorem 3.4.5]).

If $f \in C_c^\infty$, then

$$\int Af\, d\varpi = -\varepsilon\langle 1, \widetilde{H}f\rangle_{L^2(\varpi)} = -\varepsilon\langle \widetilde{H}1, f\rangle_{L^2(\varpi)} = 0,$$

so that ϖ is a stationary distribution for A. For $k \geq 1$, since $\varpi_k(dx) = \eta_k(x)\varpi(dx)$, we have

$$\int Af\, d\varpi_k = -\varepsilon\langle \eta_k, \widetilde{H}f\rangle_{L^2(\varpi)} = -\varepsilon\langle \widetilde{H}\eta_k, f\rangle_{L^2(\varpi)} = -\lambda_k \int f\, d\varpi_k.$$

Also, $\varpi_k(\mathbb{R}^d) = \langle \eta_k, 1\rangle_{L^2(\varpi)} = 0$, since η_k and $\eta_0 = 1$ are orthogonal.

Finally, since $\eta_k = Z^{1/2}\psi_k/\pi$ and ψ_k is locally Hölder continuous, it follows that each η_k belongs to $C(\mathbb{R}^d)$, and the fact that they are bounded follows from Theorems 12 and 13. □

3.3 The coupled process

By Proposition 14, the pair (A, m) satisfies Assumption 1 with $E = \mathbb{R}^d$, so we have the following.

Theorem Let A be the generator for (1) where F satisfies Assumption 10, and let $(-\lambda_0, \eta_0), \ldots, (-\lambda_m, \eta_m)$ be the first $m + 1$ eigenvalues and eigenvectors of A. Let $Q \in \mathbb{R}^{(m+1)\times(m+1)}$ be the generator of a continuous-time Markov chain with state space $E_0 = \{0, 1, \ldots, m\}$ and eigenvalues $\{0, -\lambda_1, \ldots, -\lambda_m\}$ and eigenvectors $\xi^{(1)}, \ldots, \xi^{(m)}$ such that α_i defined by (8) is strictly positive. Let B be defined as in Definition 6.

Let $\widetilde{Y}$ be a continuous time Markov chain with generator Q and initial distribution $p = (p_0, \ldots, p_m) \in \mathcal{P}(E_0)$. Then there exists a cadlag Markov process (X, Y) with generator B and initial distribution ν given by

$$\nu(\Gamma \times \{i\}) = p_i \alpha(i, \Gamma), \quad \Gamma \in \mathcal{B}(\mathbb{R}^d), \tag{19}$$

such that Y and $\widetilde{Y}$ have the same distribution on $D_{E_0}[0, \infty)$, and

$$P(X(t) \in \Gamma \mid Y(t) = j) = \int_\Gamma \alpha_j(x)\, \varpi(dx), \tag{20}$$

for all $t \geq 0$, all $0 \leq j \leq m$, and all $\Gamma \in \mathcal{B}(E)$. □

Remark That Q with these properties exists can be seen from [16, Theorem 1]. Remark 5 ensures the existence of the eigenvectors. □

Proof Note that under the assumptions of the theorem, $(A, m, Q, \xi^{(1)}, \dots, \xi^{(m)})$ satisfies Assumption 4. By Proposition 14, the rest of the hypotheses of Theorem 8 are also satisfied. Consequently, the process (X, Y) exists, and by uniqueness of the martingale problem for B, (X, Y) is Markov. □

We can now construct the coupling sequences described in the introduction. For each $\varepsilon > 0$, choose a matrix Q_ε and eigenvectors $\xi_\varepsilon^{(1)}, \dots, \xi_\varepsilon^{(m)}$ that satisfy the assumptions of Theorem 15. If $(X_\varepsilon, Y_\varepsilon)$ is the Markov process described in Theorem 15, then the family, $\{(X_\varepsilon, Y_\varepsilon) : \varepsilon > 0\}$, forms a coupling sequence.

The coupling sequence is determined by the collection, $\{Q_\varepsilon, \xi_\varepsilon^{(1)}, \dots, \xi_\varepsilon^{(m)} : \varepsilon > 0\}$. By making different choices for the matrices and eigenvectors, we can obtain different coupling sequences. In our follow-up paper, we will consider the question of existence and uniqueness of a coupling sequence that satisfies conditions (3) and (4).

Acknowledgements This paper was completed while the first author was visiting the University of California, San Diego with the support of the Charles Lee Powell Foundation. The hospitality of that institution, particularly that of Professor Ruth Williams, was greatly appreciated.

Appendix

Let E be a complete and separable metric space, $\mathcal{B}(E)$ the σ-algebra of Borel subsets of E, and $\mathcal{P}(E)$ the family of Borel probability measures on E. Let $M(E)$ be the collection of all real-valued, Borel measurable functions on E, and $B(E) \subset M(E)$ the Banach space of bounded functions with $\|f\|_\infty = \sup_{x \in E} |f(x)|$. Let $\overline{C}(E) \subset B(E)$ be the subspace of bounded continuous functions, while $C(E)$ denotes the collection of continuous, real-valued functions on E. A collection of functions $D \subset \overline{C}(E)$ is *separating* if $\mu, \nu \in \mathcal{P}(E)$ and $\int f\, d\mu = \int f\, d\nu$ for all $f \in D$ implies $\mu = \nu$.

Condition (i) $B \subset \overline{C}(E) \times C(E)$ and $\mathcal{D}(B)$ is closed under multiplication and separating.

(ii) There exists $\psi \in C(E)$, $\psi \geq 1$, such that for each $f \in \mathcal{D}(B)$, there exists a constant c_f such that

$$|Bf(x)| \leq c_f \psi(x), \quad x \in E.$$

(We write Bf even though we do not exclude the possibility that B is multivalued. In the multivalued case, each element of Bf must satisfy the inequality.)

(iii) There exists a countable subset $B_c \subset B$ such that every solution of the (local) martingale problem for B_c is a solution of the (local) martingale problem for B.

(iv) $B_0 f \equiv \psi^{-1} Bf$ is a pre-generator, that is, B_0 is dissipative and there are sequences of functions $\mu_n : E \to \mathcal{P}(E)$ and $\lambda_n : E \to [0, \infty)$ such that for each

$(f,g) \in B$,

$$g(x) = \lim_{n\to\infty} \lambda_n(x) \int_E (f(y) - f(x))\mu_n(x,dy) \tag{21}$$

for each $x \in E$. □

Remark Condition 17(iii) holds if B_0 is graph-separable, that is, there is a countable subset $B_{0,c}$ of B_0 such that B_0 is a subset of the bounded, pointwise closure of $B_{0,c}$.

An operator is a pre-generator if for each $x \in E$, there exists a solution of the martingale problem for (B, δ_x). □

For a measurable E_0-valued process Y, where E_0 is a complete and separable metric space, let

$$\widehat{\mathcal{F}}_t^Y = \text{completion of } \sigma\left(\int_0^r g(Y(s))\,ds : r \le t, g \in B(E_0)\right) \vee \sigma(Y(0)).$$

Theorem Let (S,d) and (E_0,d_0) be complete, separable metric spaces. Let B satisfy Condition 17. Let $\gamma : S \to E_0$ be measurable, and let $\widetilde{\alpha}$ be a transition function from E_0 into S (that is, $\widetilde{\alpha} : E_0 \times \mathcal{B}(S) \to \mathbb{R}$ satisfies $\widetilde{\alpha}(y,\cdot) \in \mathcal{P}(S)$ for all $y \in E_0$ and $\widetilde{\alpha}(\cdot,\Gamma) \in B(E_0)$ for all $\Gamma \in \mathcal{B}(S)$) satisfying $\int h \circ \gamma(z)\,\widetilde{\alpha}(y,dz) = h(y)$, $y \in E_0$, $h \in B(E_0)$, that is, $\widetilde{\alpha}(y, \gamma^{-1}(y)) = 1$. Assume that $\widetilde{\psi}(y) \equiv \int_S \psi(z)\widetilde{\alpha}(y,dz) < \infty$ for each $y \in E_0$ and define

$$C = \left\{\left(\int_S f(z)\widetilde{\alpha}(\cdot,dz), \int_S Bf(z)\widetilde{\alpha}(\cdot,dz)\right) : f \in \mathcal{D}(B)\right\}.$$

Let $\mu \in \mathcal{P}(E_0)$ and define $\nu = \int \widetilde{\alpha}(y,\cdot)\,\mu(dy)$.

a) If $\widetilde{Y}$ satisfies $\int_0^t E[\widetilde{\psi}(\widetilde{Y}(s))]\,ds < \infty$ a.s. for all $t > 0$ and $\widetilde{Y}$ is a solution of the martingale problem for (C,μ), then there exists a solution Z of the martingale problem for (B,ν) such that $\widetilde{Y}$ has the same distribution on $M_{E_0}[0,\infty)$ as $Y = \gamma \circ Z$. If Y and $\widetilde{Y}$ are cadlag, then Y and $\widetilde{Y}$ have the same distribution on $D_{E_0}[0,\infty)$.
b) Let $\mathbf{T}^Y = \{t : Y(t) \text{ is } \widehat{\mathcal{F}}_t^Y \text{ measurable}\}$ (which holds for Lebesgue-almost every t). Then for $t \in \mathbf{T}^Y$,

$$P(Z(t) \in \Gamma \mid \widehat{\mathcal{F}}_t^Y) = \widetilde{\alpha}(Y(t),\Gamma), \quad \Gamma \in \mathcal{B}(S).$$

c) If, in addition, uniqueness holds for the martingale problem for (B,ν), then uniqueness holds for the $M_{E_0}[0,\infty)$-martingale problem for (C,μ). If $\widetilde{Y}$ has sample paths in $D_{E_0}[0,\infty)$, then uniqueness holds for the $D_{E_0}[0,\infty)$-martingale problem for (C,μ).
d) If uniqueness holds for the martingale problem for (B,ν), then Y restricted to $\mathbf{T}^Y$ is a Markov process. □

Remark If Y is cadlag with no fixed points of discontinuity (that is $Y(t) = Y(t-)$ a.s. for all t), then $\widehat{\mathcal{F}}_t^Y = \mathcal{F}_t^Y$ for all t. □

Remark The main precursor of this Markov mapping theorem is [13, Corollary 3.5]. The result stated here is a special case of Corollary 3.3 of [11]. □

References

1. Luca Avena, Fabienne Castell, Alexandre Gaudillière, and Clothilde Melot. Approximate and exact solutions of intertwining equations through random spanning forests, 2017.
2. J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.*, 140(6):1065–1114, 2010.
3. J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains II, the nonreversible case. *J. Stat. Phys.*, 149(4):598–618, 2012.
4. Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.
5. Anton Bovier, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.
6. E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.
7. Michael Eckhoff. Precise asymptotics of small eigenvalues of reversible diffusions in the metastable regime. *Ann. Probab.*, 33(1):244–299, 2005.
8. Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
9. M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.
10. David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983.
11. T. G. Kurtz and G. Nappo. The filtered martingale problem. In *The Oxford handbook of nonlinear filtering*, pages 129–165. Oxford Univ. Press, Oxford, 2011.
12. Thomas G. Kurtz. *Approximation of population processes*, volume 36 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1981.
13. Thomas G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.*, 3:no. 9, 29 pp. (electronic), 1998.
14. Thomas G. Kurtz and Richard H. Stockbridge. Stationary solutions and forward equations for controlled and singular martingale problems. *Electron. J. Probab.*, 6:no. 17, 52 pp. (electronic), 2001.
15. Claudio Landim. Metastable Markov chains. *Probab. Surv.*, 16:143–227, 2019.
16. Hazel Perfect. Methods of constructing certain stochastic matrices. *Duke Math. J.*, 20:395–404, 1953.
17. Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
18. Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.
19. Fraydoun Rezakhanlou and Insuk Seo. Scaling limit of small random perturbation of dynamical systems, 2018.

Finite Markov Chains Coupled to General Markov Processes and An Application to Metastability II

Thomas G. Kurtz and Jason Swanson

Abstract We consider a diffusion given by a small noise perturbation of a dynamical system driven by a potential function with a finite number of local minima. The classical results of Freidlin and Wentzell show that the time this diffusion spends in the domain of attraction of one of these local minima is approximately exponentially distributed and hence the diffusion should behave approximately like a Markov chain on the local minima. By the work of Bovier and collaborators, the local minima can be associated with the small eigenvalues of the diffusion generator. In Part I of this work [10], by applying a Markov mapping theorem, we used the eigenfunctions of the generator to couple this diffusion to a Markov chain whose generator has eigenvalues equal to the eigenvalues of the diffusion generator that are associated with the local minima and established explicit formulas for conditional probabilities associated with this coupling. The fundamental question now becomes to relate the coupled Markov chain to the approximate Markov chain suggested by the results of Freidlin and Wentzel. In this paper, we take up this question and provide a complete analysis of this relationship in the special case of a double-well potential in one dimension.

1 Introduction

In the interest of self-containment, we will first recap the essential definitions from Part I of this work [10]. Fix $\varepsilon > 0$ and consider the stochastic process,

$$X_\varepsilon(t) = X_\varepsilon(0) - \int_0^t \nabla F(X_\varepsilon(s))\, ds + \sqrt{2\varepsilon}\, W(t), \tag{1}$$

Thomas G. Kurtz
University of Wisconsin-Madison, e-mail: kurtz@math.wisc.edu

Jason Swanson
University of Central Florida, e-mail: Jason.Swanson@ucf.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_13

where $F \in C^3(\mathbb{R}^d)$ and W is a standard d-dimensional Brownian motion. Let φ be the solution to the differential equation $\varphi' = -\nabla F(\varphi)$. We will use φ_x to denote the solution with $\varphi_x(0) = x$. The process X_ε is a small-noise perturbation of the deterministic process φ.

Suppose $F \in C^3(\mathbb{R}^d)$ and $\lim_{|x|\to\infty} F(x) = \infty$, and that $\mathcal{M} = \{x_0, \dots, x_m\}$ is the set of local minima of the F, with $m \geq 1$. The points x_j are stable points for the process φ. For X_ε, however, they are not stable. The process X_ε will initially gravitate toward one of the x_j and move about randomly in a small neighborhood of this point. But after an exponential amount of time, a large fluctuation of the noise term will move the process X_ε out of the domain of attraction of x_j and into the domain of attraction of one of the other minima. We say that each point x_j is a point of *metastability* for the process X_ε.

If X is a cadlag process in a complete, separable metric space S adapted to a right continuous filtration (assumptions that are immediately satisfied for all processes considered here) and H is either open or closed, then $\tau_H^X = \inf\{t > 0 : X(t) \text{ or } X(t-) \in H\}$ is a stopping time (see, for example, [6, Proposition 1.5]). If $x \in S$, let $\tau_x^X = \tau_{\{x\}}^X$. We may sometimes also write $\tau^X(H)$, and if the process is understood, we may omit the superscript.

Let

$$D_j = \{x \in \mathbb{R}^d : \lim_{t\to\infty} \varphi_x(t) = x_j\} \tag{2}$$

be the domains of attraction of the local minima. It is well-known (see, for example, [7], [1, Theorem 3.2], [2, Theorems 1.2 and 1.4], and [4]) that as $\varepsilon \to 0$, $\tau^{X_\varepsilon}(D_j^c)$ is asymptotically exponentially distributed under P^{x_j}. It is therefore common to approximate the process X_ε by a continuous time Markov chain on the set $\mathcal{M}$ (or equivalently on $\{0, \dots, m\}$).

In this project, for each $\varepsilon > 0$, we wish to capture this approximate Markov chain behavior by coupling X_ε to a continuous time Markov chain, Y_ε, on $\{0, \dots, m\}$. We refer to the indexed collection of coupled processes, $\{(X_\varepsilon, Y_\varepsilon) : \varepsilon > 0\}$ as a *coupling sequence*.

In [10], we developed a general coupling procedure that goes beyond the specific case of interest here. It is a construction that builds a coupling between a Markov process on a complete and separable metric space and a continuous-time Markov chain where the generators of the two processes have common eigenvalues. The coupling is done in such a way that observations of the chain yield quantifiable conditional probabilities about the process.

We then applied this construction to the special case of a reversible diffusion on $\mathbb{R}^d$ driven by a potential function and a small white noise perturbation. We summarize here the results in this special case. Assume there exist constants $a_i > 0$ and $c_i > 0$ such that $a_2 < 2a_1 - 2$, and

$$c_1|x|^{a_1} - c_2 \leq |\nabla F(x)|^2 \leq c_3|x|^{a_2} + c_4, \tag{3}$$

$$c_1|x|^{a_1} - c_2 \leq (|\nabla F(x)| - 2\Delta F(x))^2 \leq c_3|x|^{a_2} + c_4. \tag{4}$$

Let

$$A = \{(f, -\varepsilon \widetilde{H} f) : f \in C_c^\infty(\mathbb{R}^d)\}$$

be the generator for (1), and let $(-\lambda_0, \eta_0), \dots, (-\lambda_m, \eta_m)$ be the first $m+1$ eigenvalues and eigenfunctions of A. By [10, Proposition 3.7], the functions η_k are continuous and bounded. We may therefore choose a matrix, $Q \in \mathbb{R}^{(m+1)\times(m+1)}$, and vectors, $\xi^{(1)}, \dots, \xi^{(m)}$, such that

(i) Q is the generator of a continuous-time Markov chain with state space $E_0 = \{0, 1, \dots, m\}$,
(ii) $\xi^{(k)}$ is a right eigenvector of Q with eigenvalue $-\lambda_k$, and
(iii) for $0 \le i \le m$, the functions,

$$\alpha_i(x) = 1 + \sum_{k=1}^m \xi_i^{(k)} \eta_k(x),$$

are strictly positive.

We then choose a probability measure, $p = (p_0, \dots, p_m)$, on E_0, define the measure ν on $\mathbb{R}^d \times E_0$ by

$$\nu(\Gamma \times \{i\}) = p_i \alpha(i, \Gamma), \quad \Gamma \in \mathcal{B}(\mathbb{R}^d), \tag{5}$$

and let $(X_\varepsilon, Y_\varepsilon)$ be the cadlag Markov process on $\mathbb{R}^d \times E_0$ with initial distribution ν and generator,

$$Bf(x,i) = Af(x,i) + \sum_{j \neq i} Q_{ij} \frac{\alpha_j(x)}{\alpha_i(x)}(x)(f(x,j) - f(x,i)). \tag{6}$$

Note that all of these objects (A, λ_k, η_k, Q, $\xi^{(k)}$, p, and so on) depend on ε, though this dependence is suppressed in the notation for readability.

By [10, Theorem 3.8], the process X_ε solves (1), the process Y_ε has generator Q, and

$$P(X(t) \in \Gamma \mid Y(t) = j) = \int_\Gamma \alpha_j(x)\, \varpi(dx), \tag{7}$$

for all $t \ge 0$, all $0 \le j \le m$, and all $\Gamma \in \mathcal{B}(E)$.

In this way, for each $\varepsilon > 0$, we create a coupling, $(X_\varepsilon, Y_\varepsilon)$. We referred to the indexed collection of coupled processes, $\{(X_\varepsilon, Y_\varepsilon) : \varepsilon > 0\}$, as a *coupling sequence*. Our objective is to investigate the possibility of constructing a coupling sequence which satisfies both

$$P(X_\varepsilon(t) \in D_j \mid Y_\varepsilon(t) = j) \to 1 \tag{8}$$

and

$$E^i[\tau_j^{Y_\varepsilon}] \sim E^{x_i}[\tau_{B_\rho(x_0)}^{X_\varepsilon}] \tag{9}$$

as $\varepsilon \to 0$, for all i and j, where $B_\rho(x)$ is the ball of radius ρ centered at x.

In the current paper, we consider this question in the case of a double-well potential in one dimension. That is, suppose $d = 1$ and $\mathcal{M} = \{x_0, x_1\}$, where $x_0 < 0 < x_1$. Let F be decreasing on $(-\infty, x_0)$ and $(0, x_1)$, and increasing on $(x_0, 0)$ and (x_1, ∞), and satisfy $F(x_0) < F(x_1)$. Then the domains of attraction are $D_0 = (-\infty, 0)$ and

$D_1 = (0, \infty)$. There are many possible coupling sequences, so for each such sequence, we can ask if it satisfies any of the following:

$$P(X_\varepsilon(t) < 0 \mid Y_\varepsilon(t) = 0) \to 1, \tag{10}$$
$$P(X_\varepsilon(t) > 0 \mid Y_\varepsilon(t) = 1) \to 1, \tag{11}$$
$$E^1[\tau_0^{Y_\varepsilon}] \sim E^{x_1}[\tau_{B_\rho(x_0)}^{X_\varepsilon}], \tag{12}$$
$$E^0[\tau_1^{Y_\varepsilon}] \sim E^{x_0}[\tau_{B_\rho(x_1)}^{X_\varepsilon}], \tag{13}$$

as $\varepsilon \to 0$, where $0 < \rho < |x_0| \wedge x_1$.

Let $-\lambda_\varepsilon$ be the second eigenvalue of the generator of X_ε. It is known (see, for example, [12, 13] or [1, 2]), that in (18) and (19), we have

$$E^{x_1}[\tau_{B_\rho(x_0)}^{X_\varepsilon}] \sim \frac{2\pi}{|F''(0)F''(x_1)|^{1/2}} e^{(F(0)-F(x_1))/\varepsilon} \sim \frac{1}{\lambda_\varepsilon},$$
$$E^{x_0}[\tau_{B_\rho(x_1)}^{X_\varepsilon}] \sim \frac{2\pi}{|F''(0)F''(x_0)|^{1/2}} e^{(F(0)-F(x_0))/\varepsilon}.$$

Thus, (12) and (13) are equivalent to (36) and (37), respectively. Moreover, Theorem 9 shows that, in our coupling construction, (10) is equivalent to the assertion that, given $Y(t) = 0$, the distribution of $X(t)$ is asymptotically equivalent to the stationary distribution, conditioned to be on $(-\infty, 0)$. Theorem 10 gives the analogous equivalency for (11).

In Section 4, we will show that, in our coupling construction, (11) implies (12), which implies (10), and (13) implies (12). We also show by example that there are no other implications among these conditions. For example, we can couple X_ε and Y_ε so that (10), (12), and (13) are satisfied, but (11) is not. In other words, it is possible to build the Markov chain with asymptotically the same transition rates as the process, but the two do not remain synchronized, in the sense that (11) fails. Or, as another example, we can couple the processes so that (10)-(12) are satisfied, but (13) is not. In other words, we can have a coupling where the Markov chain accurately tracks the diffusion, but the transition rates of the two processes are not the same.

In the case of the double-well potential, for fixed $\varepsilon > 0$, the dynamics of the coupling $(X_\varepsilon, Y_\varepsilon)$ are uniquely determined by two parameters, $\xi_{1,\varepsilon}$ and $\xi_{2,\varepsilon}$ (see Lemma 11). If we identify coupling sequences whose parameters are asymptotically equivalent as $\varepsilon \to 0$, then there is a unique coupling sequence satisfying (10)-(13). Heuristically, we build this sequence by choosing the ξ's so that $\alpha_j \approx c_{j,\varepsilon} 1_{D_j}$. More specifically, we choose them so that $\alpha_0 = -\eta_1/\eta_1(\infty) + 1$ and $\alpha_1 = \eta_1/|\eta_1(-\infty)| + 1$. We then prove sharp enough bounds on the behavior of η_1 to show that the approximation $\alpha_j \approx c_{j,\varepsilon} 1_{D_j}$ is sufficiently accurate.

The outline of the paper is as follows. In Section 2, we address the issue of how the minima should be ordered so that they correspond to the eigenvalues of the generator of the diffusion. This is a necessary prerequisite for attaining the asymptotic behavior in (8) and (9). In Section 3, we specialize to the case of the double-well potential in $d = 1$. We begin there with the study the structure of the second eigenfunction. In particular, we narrow down the location of the nodal point, show that the eigenfunction is asymptotically flat near the minima, and establish key estimates on the behavior of the eigenfunction near the saddle point. Then, in Section 4, we use these results to give a complete analysis of our coupling sequences for the double-well potential.

2 Ordering the local minima

Heretofore, no mention has been made of the order in which the local minima, $\mathcal{M} = \{x_0, \ldots, x_m\}$, are listed. No particular order is necessary in order to construct a coupling sequence. But if that sequence is to exhibit the behavior in (8) and (9), then the minima should be ordered so that they correspond with the eigenvalues of A.

To describe this ordering, we first establish some notation and terminology. For any two sets $A, B \subset \mathbb{R}^d$, define the set of paths from A to B as

$$\mathcal{P}^*(A,B) = \{\omega \in C([0,1];\mathbb{R}^d) : \omega(0) \in A, \omega(1) \in B\}.$$

Given $F : \mathbb{R}^d \to \mathbb{R}$, the height of the saddle, or communication height, between A and B is defined as

$$\widehat{F}(A,B) = \inf_{\omega \in \mathcal{P}^*(A,B)} \sup_{t \in [0,1]} F(\omega(t)).$$

The set of minimal paths from A to B is

$$\mathcal{P}(A,B) = \{\omega \in \mathcal{P}^*(A,B) : \sup_{t \in [0,1]} F(\omega(t)) = \widehat{F}(A,B)\}.$$

A gate, $G(A,B)$, is a minimal subset of $\{z \in \mathbb{R}^d : F(z) = \widehat{F}(A,B)\}$ such that all minimal paths intersect $G(A,B)$. In general, $G(A,B)$ is not unique. The set of saddle points, $\mathcal{S}(A,B)$, is the union of all gates.

Assumption (i) For $x, y \in \mathcal{M}$, $G(x,y)$ is unique and consists of a finite set of isolated points $\{z_i^*(x,y)\}$.
(ii) The Hessian matrix of F is non-degenerate at each $x \in \mathcal{M}$ and at each saddle point $z_i^*(x,y)$.
(iii) The minima $\mathcal{M} = \{x_0, \ldots, x_m\}$ can be labeled in such a way that, with $\mathcal{M}_k = \{x_0, \ldots, x_k\}$, each saddle point $z^*(x_k, \mathcal{M}_{k-1})$ is unique, the Hessian matrix of F at $z^*(x_k, \mathcal{M}_{k-1})$ is non-degenerate, and

$$\widehat{F}(x_k, \mathcal{M}_k \setminus x_k) - F(x_k) < \widehat{F}(x_i, \mathcal{M}_k \setminus x_i) - F(x_i), \tag{14}$$

for all $0 \le i < k \le m$. □

We shall assume our potential function F satisfies Assumption 1, and that the minima are ordered as in (iii).

3 Structure of the second eigenfunction

3.1 Tools and preliminary results

From this point forward, we take $d = 1$. Note that $\varepsilon\eta_k'' - F'\eta_k' = -\lambda_k\eta_k$ for all integers $k \geq 0$. We will make use of the fact that the eigenfunctions satisfy the integral equations in the following lemma.

Lemma 1 *For any $k \in \mathbb{N}$,*

$$\eta_k(x) = \eta_k(\infty) - \frac{\lambda_k}{\varepsilon}\int_x^\infty \int_x^u \exp\left(\frac{F(v)-F(u)}{\varepsilon}\right)\eta_k(u)\,dv\,du \tag{15}$$

$$= \eta_k(-\infty) - \frac{\lambda_k}{\varepsilon}\int_{-\infty}^x \int_u^x \exp\left(\frac{F(v)-F(u)}{\varepsilon}\right)\eta_k(u)\,dv\,du. \tag{16}$$

Proof Fix $k \in \mathbb{N}$. Since η_k is bounded by Proposition [10, Proposition 3.7], we may choose $C_1 > 0$ such that $|\eta_k(x)| \leq C_1$ for all $x \in \mathbb{R}$. Now fix $x \in \mathbb{R}$. Since $a_1 > 2$, we may choose $\alpha \in (1, a_1/2)$. By [10, Lemma 3.3], assumptions (3) and (4) imply that

$$\widetilde{c}_1|x|^{\widetilde{a}_1} - \widetilde{c}_2 \leq |F(x)| \leq \widetilde{c}_3|x|^{\widetilde{a}_2} + \widetilde{c}_4, \tag{17}$$

where $\widetilde{a}_i = a_i/2 + 1$. It follows that $\lim_{u\to\infty} u^{-\alpha}e^{F(u)/\varepsilon} = \infty$. Also by (3), for u sufficiently large, $|u^{-\alpha}F'(u)| \geq C|u|^{a_1/2-\alpha}$ for some $C > 0$. Hence, by L'Hôpital's rule,

$$\lim_{u\to\infty}\frac{\int_x^u e^{F(v)/\varepsilon}\,dv}{u^{-\alpha}e^{F(u)/\varepsilon}} = \lim_{u\to\infty}\frac{1}{-\alpha u^{-(\alpha+1)} + u^{-\alpha}F'(u)} = 0,$$

and so we may choose $C_2 > 0$ such that $\int_x^u e^{F(v)/\varepsilon}\,dv \leq C_2 u^{-\alpha}e^{F(u)/\varepsilon}$ for all $u \geq x$. Therefore,

$$\int_x^\infty \int_x^u \left|\exp\left(\frac{F(v)-F(u)}{\varepsilon}\right)\eta_k(u)\right| dv\,du \leq C_1C_2\int_x^\infty u^{-\alpha}\,du < \infty,$$

and so the right-hand side of (15) is well-defined.

Let

$$y(x) = \eta_k(\infty) - \frac{\lambda_k}{\varepsilon}\int_x^\infty \int_x^u \exp\left(\frac{F(v)-F(u)}{\varepsilon}\right)\eta_k(u)\,dv\,du.$$

Then

$$y'(x) = \frac{\lambda_k}{\varepsilon}\int_x^\infty \exp\left(\frac{F(x)-F(u)}{\varepsilon}\right)\eta_k(u)\,du,$$

and

$$y''(x) = -\frac{\lambda_k}{\varepsilon}\eta_k(x) + F'(x)\frac{\lambda_k}{\varepsilon^2}\int_x^\infty \exp\left(\frac{F(x)-F(u)}{\varepsilon}\right)\eta_k(u)\,du.$$

Thus, $\varepsilon y'' - F'y' = -\lambda_k \eta_k = \varepsilon \eta_k'' - F'\eta_k'$, so that $y - \eta_k$ is an eigenfunction corresponding to λ_0. That is, y and η_k differ by a constant. But $y(\infty) = \eta_k(\infty)$, so $y = \eta_k$ and this proves (15).

By replacing F with $x \mapsto F(-x)$, equation (15) gives

$$\eta_k(-x) = \eta_k(-\infty) - \frac{\lambda_k}{\varepsilon} \int_x^\infty \int_x^u \exp\left(\frac{F(-v) - F(-u)}{\varepsilon}\right) \eta_k(-u)\, dv\, du,$$

which gives

$$\begin{aligned}\eta_k(x) &= \eta_k(-\infty) - \frac{\lambda_k}{\varepsilon} \int_{-x}^\infty \int_{-u}^x \exp\left(\frac{F(v') - F(-u)}{\varepsilon}\right) \eta_k(-u)\, dv'\, du \\ &= \eta_k(-\infty) - \frac{\lambda_k}{\varepsilon} \int_{-\infty}^x \int_{u'}^x \exp\left(\frac{F(v') - F(u')}{\varepsilon}\right) \eta_k(u')\, dv'\, du',\end{aligned}$$

proving (16). □

We now assume that for some fixed $\widetilde{x}_0 < 0 < \widetilde{x}_1$:

(i) F is strictly decreasing on $(-\infty, \widetilde{x}_0)$ and $(0, \widetilde{x}_1)$, and strictly increasing on $(\widetilde{x}_0, 0)$ and $(\widetilde{x}_1, \infty)$.
(ii) $F''(\widetilde{x}_0) > 0$, $F''(0) < 0$, $F''(\widetilde{x}_1) > 0$.
(iii) $F(\widetilde{x}_0) \neq F(\widetilde{x}_1)$.

Then $\mathcal{M} = \{\widetilde{x}_0, \widetilde{x}_1\}$ and $m = 1$. If $F(\widetilde{x}_0) < F(\widetilde{x}_1)$, then

$$\widehat{F}(\widetilde{x}_1, \{\widetilde{x}_0\}) - F(\widetilde{x}_1) = F(0) - F(\widetilde{x}_1) < F(0) - F(\widetilde{x}_0) = \widehat{F}(\widetilde{x}_0, \{\widetilde{x}_1\}) - F(\widetilde{x}_0),$$

which would imply $x_0 = \widetilde{x}_0$, and $x_1 = \widetilde{x}_1$. On the other hand, if $F(\widetilde{x}_1) < F(\widetilde{x}_0)$, then $x_0 = \widetilde{x}_1$ and $x_1 = \widetilde{x}_0$. For now, we will not assume either ordering of the local minima, so that our assumptions are symmetric under the reflection $x \mapsto -x$. Because of this, results that are stated in terms of $\widetilde{x}_0$ can be applied to $\widetilde{x}_1$ by replacing $F(x)$ with $F(-x)$.

Let $\eta = \eta_1$ and $\lambda = \lambda_1$. By Courant's nodal domain theorem [3, Section VI.6, p.454], replacing η by $-\eta$ if necessary, there exists $r = r_\varepsilon \in \mathbb{R}$ such that

$$\eta(x) \begin{cases} < 0 & \text{if } x < r_\varepsilon, \\ = 0 & \text{if } x = r_\varepsilon, \\ > 0 & \text{if } x > r_\varepsilon. \end{cases}$$

It therefore follows from Lemma 1 that η is strictly increasing.

By [2, Theorem 1.2],

$$\lambda = \frac{|F''(0)F''(x_1)|^{1/2}}{2\pi} e^{-(F(0)-F(x_1))/\varepsilon} (1 + O(\varepsilon^{1/2} |\log \varepsilon|)). \tag{18}$$

By [1, (3.3)], we have

$$E^{x_j}[\tau^X_{B_\rho(x_{1-j})}] \sim \frac{2\pi}{|F''(0)F''(x_j)|^{1/2}}\, e^{(F(0)-F(x_j))/\varepsilon}, \tag{19}$$

for $0 < \rho < |\widetilde{x}_0| \wedge |\widetilde{x}_1|$. And the following special case of [2, Proposition 3.3] gives us a way to estimate the shape of the eigenfunction.

Theorem Let $h(y) = P^y(\tau^X_{(x_0-\varepsilon,x_0+\varepsilon)} < \tau^X_{r_\varepsilon})$ and $\phi(y) = |\eta(y)|/|\eta(x_0+\varepsilon)|$. Then there exists $C,\alpha,\varepsilon_0 > 0$ such that

$$h(y) \le \phi(y) \le h(y)(1+C\varepsilon^{\alpha/2}),$$

for all $y < r_\varepsilon$ and all $\varepsilon \in (0,\varepsilon_0)$. □

To apply this result, we will use the following two lemmas, which formulate the Freidlin and Wentzell results in our specific case.

Lemma 2 *Let $a < \widetilde{a} < \widetilde{x}_0 < \widetilde{b} < b < 0$ and fix $\delta > 0$. Then there exists $\varepsilon_0 > 0$ such that*

$$\exp\left(\frac{1-\delta}{\varepsilon}(F(a)\wedge F(b) - F(\widetilde{x}_0))\right) \le E^x[\tau^X_{(a,b)^c}]$$
$$\le \exp\left(\frac{1+\delta}{\varepsilon}(F(a)\wedge F(b) - F(\widetilde{x}_0))\right),$$

for all $\widetilde{a} \le x \le \widetilde{b}$ and all $\varepsilon \in (0,\varepsilon_0)$. The analogous result also holds when $0 < a < \widetilde{a} < \widetilde{x}_1 < \widetilde{b} < b$.

Proof By Theorem 18, $\varepsilon \log E^x[\tau^X_{(a,b)^c}] \to L := F(a)\wedge F(b) - F(\widetilde{x}_0)$ as $\varepsilon \to 0$, uniformly in x on $[\widetilde{a},\widetilde{b}]$. Thus, there exists ε_0 such that $\varepsilon \in (0,\varepsilon_0)$ implies $\varepsilon \log E^x[\tau_{(a,b)^c}] \le (1+\delta)L$, which gives the upper bound. The lower bound is deduced similarly. □

Lemma 3 *Let $a < \widetilde{x}_0 < b < 0$ or $0 < a < \widetilde{x}_1 < b$ and define $G = (a,b)$. Assume $F(a) \ne F(b)$ and choose $y \in \{a,b\}$ such that $F(y) = F(a)\vee F(b)$. Then, for all compact $K \subset G$ and all $\gamma > 0$, there exists $\varepsilon_0 > 0$ such that*

$$\exp\left(-\frac{|F(a)-F(b)|+\gamma}{\varepsilon}\right) \le P^x(X(\tau^X_{G^c}) = y) \le \exp\left(-\frac{|F(a)-F(b)|-\gamma}{\varepsilon}\right),$$

for all $x \in K$ and all $\varepsilon \in (0,\varepsilon_0)$.

Proof We prove only the case where $a < \widetilde{x}_0 < b$ and $F(a) > F(b)$, so that $y = a$. The proofs of the other cases are similar. We use Theorem 20, Proposition 21, and Lemma 12. Note that, according to the discussion preceding Theorem 20, we have $V_G(x,y) = V(x,y)$ for all $x,y \in [a,b]$.

Fix $x \in K$. In this case,

$$M_G = V_G(\{\widetilde{x}_0\},\{a,b\}) = V_G(\widetilde{x}_0,a)\wedge V_G(\widetilde{x}_0,b) = 2(F(b)-F(\widetilde{x}_0)),$$

and

$$\begin{aligned}M_G(x,a) &= \min\{V_G(\widetilde{x}_0,x)+V_G(x,a), V_G(\widetilde{x}_0,\{a,b\})+V_G(x,a),\\ &\qquad V_G(x,\widetilde{x}_0)+V_G(\widetilde{x}_0,a)\}\\ &= \min\{2(F(x)-F(\widetilde{x}_0))+V_G(x,a), 2(F(b)-F(\widetilde{x}_0))+V_G(x,a),\\ &\qquad 2(F(a)-F(\widetilde{x}_0))\}\end{aligned}$$

If $a < x < \widetilde{x}_0$, then $V_G(x,a) = 2(F(a)-F(x))$, so that

$$\begin{aligned}M_G(x,a) &= 2\min\{F(a)-F(\widetilde{x}_0), F(b)-F(\widetilde{x}_0)+F(a)-F(\widetilde{x}_0), F(a)-F(\widetilde{x}_0)\}\\ &= 2(F(a)-F(\widetilde{x}_0)).\end{aligned}$$

If $\widetilde{x}_0 \le x < b$, then $V_G(x,a) = 2(F(a)-F(\widetilde{x}_0))$, so that

$$\begin{aligned}M_G(x,a) &= 2\min\{F(x)+F(a)-2F(\widetilde{x}_0), F(b)+F(a)-2F(\widetilde{x}_0), F(a)-F(\widetilde{x}_0)\}\\ &= 2(F(a)-F(\widetilde{x}_0)).\end{aligned}$$

Thus, $M_G(x,a) - M_G = 2(F(a)-F(b))$, and the result follows from Theorem 20.□

3.2 Location of the nodal point

Our first order of business is to identify an interval in which the nodal point (that is, the zero of the second eigenfunction) is asymptotically located. The essential feature of the interval is that it is bounded away from the minima as $\varepsilon \to 0$.

The statement of this result is Corollary 4. To prove this result, we need four lemmas, all concerning stopping times of X.

Lemma 4 *There exists* $R > 0$ *such that* $\sup\{E^x[\tau_K^X] : x \in \mathbb{R}^d, \varepsilon \in (0,1)\} < \infty$, *where* $K = \overline{B_R(0)}$.

Proof In this proof, for $r > 0$, let $\sigma_r = \tau_{(-\infty,r]}^{F(X)} = \inf\{t \ge 0 : F(X(t)) \le r\}$.

Choose $C_1, C_2, L > 0$ such that

(i) $V(x) \ge C_1|x|^{a_1}$,
(ii) $C_1|x|^{a_1} \le |\nabla F(x)|^2 \le C_2|x|^{a_2}$, and
(iii) $C_1|x|^{\widetilde{a}_1} \le F(x) \le C_2|x|^{\widetilde{a}_2}$,

for all $|x| > L$, where $\widetilde{a}_j$ are as in (17). Choose $R > L$ such that

$$I := (1 \vee \sup_{|x|\le L} F(x), C_1 R^{\widetilde{a}_1}] \cap \mathbb{N} \ne \emptyset,$$

and choose $b \in I$.

Suppose $\omega \in \{\tau_K > t\}$. Then, for all $s \le t$, we have that $|X(s)| > R > L$, and so it follows that $F(X(s)) \ge C_1|X(s)|^{\widetilde{a}_1} > C_1 R^{\widetilde{a}_1} \ge b$. Thus, $\omega \in \{\sigma_b > t\}$, and we have shown that $\tau_K \le \sigma_b$ a.s. It therefore suffices to show that $E^x[\sigma_b]$ is bounded above by a constant that does not depend on x or ε.

Fix $\varepsilon \in (0,1)$. Let $r = a_1/\widetilde{a}_2$ and $C_3 = C_1 C_2^{-r}$. We will first prove that if $x \in \mathbb{R}^d$, $n \in \mathbb{N}$, and $b \le n < F(x) \le n+1$, then

$$E^x[\sigma_n] \le 2C_3^{-1} n^{-r}. \tag{20}$$

Let x and n satisfy the assumptions. Using Itô's rule, we can write

$$F(X(t)) = F(x) + \sqrt{2\varepsilon}\, M(t) - 2\varepsilon \int_0^t \psi(X(s))\, ds, P^x\text{-a.s.}$$

where $M(t) = \int_0^t \nabla F(X(s))\, dW(s)$ and $\psi = \varepsilon V + |\nabla F|^2/(4\varepsilon)$. Let $\widetilde{W}(s) = M(T(s))$, where the stopping time $T(s)$ is defined by $T(s) = \inf\{t \ge 0 : [M]_t > s\}$. By [9, Theorem 3.4.6], $\widetilde{W}$ is a standard Brownian motion, and $M(t) = \widetilde{W}([M]_t)$. Moreover, by [9, Problem 3.4.5], $s < [M]_t$ if and only if $T(s) < t$, and $[M]_{T(s)} = s$ for all $s \ge 0$.

Let

$$\widehat{W}(t) = \widetilde{W}(t) - \frac{1}{2\sqrt{2\varepsilon}}\, t,$$

and define $\widetilde{\sigma}_n = \tau^{\sqrt{2\varepsilon}\widehat{W}}_{(-\infty, n-F(x)]} = \inf\{t \ge 0 : \widehat{W}(t) \le (n - F(x))/\sqrt{2\varepsilon}\}$. We will prove that $[M]_{\sigma_n} \le \widetilde{\sigma}_n$ a.s. Note that

$$\begin{aligned}
\{\widetilde{\sigma}_n < [M]_{\sigma_n}\} &= \bigcup_{s \in \mathbb{Q}} \left(\{s < [M]_{\sigma_n}\} \cap \left\{ \widehat{W}(s) \le \frac{n - F(x)}{\sqrt{2\varepsilon}} \right\} \right) \\
&= \bigcup_{s \in \mathbb{Q}} \left(\{T(s) < \sigma_n\} \cap \left\{ \widehat{W}([M]_{T(s)}) \le \frac{n - F(x)}{\sqrt{2\varepsilon}} \right\} \right).
\end{aligned}$$

On the event $\{T(s) < \sigma_n\}$, we have, for all $u \le T(s)$,

$$F(X(u)) > n \ge b > \sup_{|x| \le L} F(x), \tag{21}$$

where the first inequality comes from the definition of σ_n. It follows that $|X(u)| > L$. Thus, by (i), we have $V(X(u)) > 0$, and so $\psi(X(u)) > |\nabla F(X(u))|^2/(4\varepsilon)$. Hence,

$$\begin{aligned}
n < F(X(T(s))) &\le F(x) + \sqrt{2\varepsilon} M(T(s)) - \frac{1}{2} \int_0^{T(s)} |\nabla F(X(u))|^2\, du \\
&= F(x) + \sqrt{2\varepsilon}\, \widetilde{W} B([M]_{T(s)}) - \frac{1}{2} [M]_{T(s)} \\
&= F(x) + \sqrt{2\varepsilon}\, \widehat{W}([M]_{T(s)}).
\end{aligned}$$

Therefore, $\widehat{W}([M]_{T(s)}) > (n - F(x))/\sqrt{2\varepsilon}$ a.s. on the event $\{T(s) < \sigma_n\}$, which shows that $P(\widetilde{\sigma}_n < [M]_{\sigma_n}) = 0$.

Note that for all $|x| > L$, we have

$$|\nabla F(x)|^2 \geq C_1|x|^{a_1} = C_1(|x|^{\widetilde{a}_2})^{a_1/\widetilde{a}_2} \geq C_1(C_2^{-1}F(x))^{a_1/\widetilde{a}_2} = C_3F(x)^r.$$

Thus, as in (21), we obtain

$$\widetilde{\sigma}_n \geq [M]_{\sigma_n} = \int_0^{\sigma_n} |\nabla F(X(u))|^2\,du \geq C_3 \int_0^{\sigma_n} F(X(u))^r\,du \geq C_3 n^r \sigma_n.$$

Hence, using [9, Exercise 3.5.10], which gives the Laplace transform of $\widetilde{\sigma}_n$, we have

$$E^x[\sigma_n] \leq C_3^{-1}n^{-r}E^x[\widetilde{\sigma}_n] = 2C_3^{-1}n^{-r}(F(x)-n) \leq 2C_3^{-1}n^{-r},$$

which proves (20). It now follows by induction and the Markov property that

$$E^x[\sigma_b] \leq 2C_3^{-1}\sum_{j=b}^{n} j^{-r},$$

whenever $b \leq n < F(x) \leq n+1$. Since

$$\widetilde{a}_2 = \frac{a_2}{2} + 1 < \frac{2a_1 - 2}{2} + 1 = a_1,$$

it follows that $r > 1$. Hence, $C_4 := \sum_{j=b}^{\infty} j^{-r} < \infty$. Since $\sigma_b = 0$, P^x-a.s., whenever $F(x) \leq b$, we have that $E^x[\sigma_b] \leq 2C_3^{-1}C_4$ for all $x \in \mathbb{R}^d$. □

Lemma 5 *Let $x < \widetilde{x}_0$. Then there exists $\varepsilon_0 > 0$ such that*

$$\sup\{E^y[\tau_x^X] : y < x, \varepsilon \in (0, \varepsilon_0)\} < \infty.$$

Proof Choose $R > |x|$ as in Lemma 4, so that there exists $C_1 > 0$ such that $E^y[\tau_{-R}^X] \leq C_1$ for all $y < -R$ and all $\varepsilon \in (0,1)$.

Suppose $-R < x < \widetilde{x}_0$ and $\varepsilon \in (0,1)$. Let $J = (-R-1, x)$. Since $\tau_{J^c}^X \leq \tau_x^X$ P^{-R}-a.s., the strong Markov property gives

$$E^{-R}[\tau_x^X] = E^{-R}[\tau_{J^c}^X] + E^{-R}[E^{X(\tau_{J^c}^X)}[\tau_x^X]] = E^{-R}[\tau_{J^c}^X] + p_\varepsilon E^{-R-1}[\tau_x^X],$$

where $p_\varepsilon = P^{-R}(X(\tau_{J^c}^X) = -R-1)$. Also by the strong Markov property and Lemma 4,

$$E^{-R-1}[\tau_x^X] = E^{-R-1}[\tau_{-R}^X] + E^{-R}[\tau_x^X] \leq C_1 + E^{-R}[\tau_x^X].$$

Thus,

$$E^{-R}[\tau_x^X] \leq \frac{E^{-R}[\tau_{J^c}^X] + p_\varepsilon C_1}{1 - p_\varepsilon}.$$

By Theorem 19, there exists $C_2 > 0$, $T > 0$, and $\varepsilon_0 \in (0,1)$ such that for all $\varepsilon \in (0, \varepsilon_0)$,

$$E^{-R}[\tau_{J^c}^X] = \int_0^\infty P(\tau_{J^c}^X > t)\,dt \leq T + \int_T^\infty e^{-\varepsilon^{-2}C_2(t-T)}\,dt \leq T + \frac{\varepsilon_0^2}{C_2} =: C_3.$$

Choose $0 < r < |\widetilde{x}_0|$ such that $F(\widetilde{x}_0 + r) < F(-R-1)$, and choose $\gamma < F(-R-1) - F(\widetilde{x}_0 + r)$. By Lemma 3, making ε_0 smaller, if necessary, we have

$$p_\varepsilon \le P^{-R}(X(\tau^X_{(-R-1,\widetilde{x}_0+r)^c}) = -R-1) \le \exp\left(-\frac{F(-R-1) - F(\widetilde{x}_0 + r) - \gamma}{\varepsilon}\right),$$

for all $\varepsilon \in (0, \varepsilon_0)$. By making ε_0 even smaller, if necessary, we have $p_\varepsilon < 1/2$ for all $\varepsilon \in (0, \varepsilon_0)$. Thus,

$$E^{-R}[\tau^X_x] \le 2C_3 + C_1 =: C_4,$$

for all $\varepsilon \in (0, \varepsilon_0)$.

Now, if $y < -R < x < x_0$, then

$$E^y[\tau^X_x] = E^y[\tau^X_{-R}] + E^{-R}[\tau^X_x] \le C_1 + C_4,$$

for all $\varepsilon \in (0, \varepsilon_0)$, and if $-R \le y < x < x_0$, then

$$C_4 \ge E^{-R}[\tau^X_x] = E^{-R}[\tau^X_y] + E^y[\tau^X_x] \ge E^y[\tau^X_x],$$

for all $\varepsilon \in (0, \varepsilon_0)$. □

Lemma 6 *For all $\widetilde{x}_0 < x < 0$ and all $\delta > 0$, there exists $C > 0$ and $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ and all $y < x$, we have*

$$E^y[\tau^X_x] \le C \exp\left(\frac{1+\delta}{\varepsilon}(F(x) - F(\widetilde{x}_0))\right).$$

Proof Suppose $\widetilde{x}_0 < x < 0$ and fix $\delta > 0$. Choose $R > |\widetilde{x}_0|$ as in Lemma 4, so that there exists $C_1 > 0$ such that $E^y[\tau^X_{-R}] \le C_1$ for all $y < -R$ and all $\varepsilon \in (0,1)$. By making R larger, if necessary, we may assume $F(x) < F(-R-1)$. Let $J := (-R-1, x)$. As in the proof of Lemma 5,

$$E^{-R}[\tau^X_x] \le \frac{E^{-R}[\tau^X_{J^c}] + p_\varepsilon C_1}{1 - p_\varepsilon},$$

where $p_\varepsilon = P^{-R}(X(\tau^X_{J^c}) = -R-1)$. Using Lemma 3, we may choose $\varepsilon_0 > 0$ such that $p_\varepsilon \le 1/2$ for all $\varepsilon \in (0, \varepsilon_0)$, giving

$$E^{-R}[\tau^X_x] \le 2E^{-R}[\tau^X_{J^c}] + C_1.$$

As in the proof of Lemma 5, if $y < -R$, then

$$E^y[\tau^X_x] = E^y[\tau^X_{-R}] + E^{-R}[\tau^X_x] \le E^{-R}[\tau^X_x] + C_1,$$

and if $-R \le y$, then

$$E^y[\tau^X_x] \le E^{-R}[\tau^X_y] + E^y[\tau^X_x] = E^{-R}[\tau^X_x] \le E^{-R}[\tau^X_x] + C_1.$$

Thus,

$$E^y[\tau_x^X] \le 2E^{-R}[\tau_{J^c}^X] + 2C_1,$$

for all $y < x$ and all $\varepsilon \in (0, \varepsilon_0)$.

By Lemma 2, making ε_0 smaller if necessary, we have

$$E^{-R}[\tau_{J^c}^X] \le \exp\left(\frac{1+\delta}{2\varepsilon}(F(x) - F(\widetilde{x}_0))\right),$$

for all $\varepsilon \in (0, \varepsilon_0)$, which proves the lemma with $C = 2 + 2C_1$. □

Lemma 7 *Let* $\varpi_\eta(dx) = |\eta(x)| 1_{(-\infty, r_\varepsilon)}(x) \varpi(dx)$ *and* $\widehat{\varpi} = \varpi_\eta((-\infty, r_\varepsilon))^{-1} \varpi_\eta$. *It then follows that* $P^{\widehat{\varpi}}(\tau_{r_\varepsilon}^X > t) = e^{-\lambda t}$ *for all* $t \ge 0$.

Proof Let $I = (-\infty, r_\varepsilon)$. Let X^I denote X killed upon leaving I. Note that X^I with $X^I(0) = x$ solves the martingale problem for (A^I, δ_x), where $A^I = \{(f, Af) : f \in C_c^\infty(\mathbb{R}), f(r) = 0\}$. Choose $\varphi_n \in C_c^\infty(\mathbb{R})$ such that $0 \le \varphi_n \le 1$, $\varphi_n(r) = 0$, and $\varphi_n \to 1_I$ pointwise. Then

$$P^{\widehat{\varpi}}(\tau_r > t) = P^{\widehat{\varpi}}(X^I(t) \in I) = E^{\widehat{\varpi}}[1_I(X^I(t))] = \lim_{n\to\infty} h_n(t),$$

where $h_n(t) = E^{\widehat{\varpi}}[\varphi_n(X^I(t))]$. Let $P_t^I f(x) = E^x[f(X^I(t))]$. Fix $t \ge 0$ and let $\psi_n = P_t^I \varphi_n$. Then

$$h_n(t) = \int_I \psi_n \, d\widehat{\varpi} = -\frac{1}{\varpi_\eta(I)} \int_I \psi_n \eta \, d\varpi,$$

so that

$$\begin{aligned} h_n'(t) &= -\frac{1}{\varpi_\eta(I)} \int_I (A^I \psi_n) \eta \, d\varpi = -\frac{1}{\varpi_\eta(I)} \int_I (\varepsilon \psi_n'' - F' \psi_n') \eta \, d\varpi \\ &= -\frac{1}{\varpi_\eta(I)} \int_I \psi_n (\varepsilon \eta'' - F' \eta') \, d\varpi = \frac{\lambda}{\varpi_\eta(I)} \int_I \psi_n \eta \, d\varpi = -\lambda h_n(t). \end{aligned}$$

Thus, $h_n(t) = h_n(0) e^{-\lambda t}$. Note that $h_n(0) = \int_I \varphi_n \, d\widehat{\varpi} \to \widehat{\varpi}(I) = 1$ as $n \to \infty$. It therefore follows that $P^{\widehat{\varpi}}(\tau_r > t) = e^{-\lambda t}$. □

Theorem Let $x \in (\widetilde{x}_0, 0)$ satisfy $F(x) - F(\widetilde{x}_0) < F(0) - F(x_1)$. Then there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, we have $x < r_\varepsilon$. □

Proof Choose $\delta > 0$ such that $(1+\delta)(F(x) - F(\widetilde{x}_0)) < F(0) - F(x_1)$. By Lemma 6, there exists $\varepsilon_0 > 0$ and $C_1 > 0$ such that

$$E^y[\tau_x^X] \le C_1 \exp\left(\frac{1+\delta}{\varepsilon}(F(x) - F(\widetilde{x}_0))\right),$$

for all $\varepsilon \in (0, \varepsilon_0)$ and all $y < x$. By (18), there exists a constant $C_2 > 0$, not depending on ε, such that $\lambda \le C_2 e^{-(F(0) - F(x_1))/\varepsilon}$. By making ε_0 smaller if necessary, we may assume

$$\varepsilon \log(C_1 C_2) < F(0) - F(x_1) - (1+\delta)(F(x) - F(\widetilde{x}_0)),$$

for all $\varepsilon \in (0, \varepsilon_0)$.

Fix $\varepsilon < \varepsilon_0$. Suppose $r_\varepsilon \le x$. By Lemma 7,

$$C_2^{-1} \exp\left(\frac{1}{\varepsilon}(F(0) - F(x_1))\right) \le \lambda^{-1} = E^{\widehat{\varpi}}[\tau^X_{r_\varepsilon}] = \int_{-\infty}^{r_\varepsilon} E^y[\tau^X_{r_\varepsilon}]\widehat{\varpi}(dy)$$
$$\le \int_{-\infty}^{r_\varepsilon} E^y[\tau^X_x]\widehat{\varpi}(dy) \le \sup_{y<r_\varepsilon} E^y[\tau^X_x] \le \sup_{y<x} E^y[\tau^X_x] \le C_1 \exp\left(\frac{1+\delta}{\varepsilon}(F(x) - F(\widetilde{x}_0))\right),$$

which implies

$$\exp\left(\frac{F(0) - F(x_1) - (1+\delta)(F(x) - F(\widetilde{x}_0))}{\varepsilon}\right) \le C_1 C_2,$$

a contradiction. □

Corollary Suppose $F(\widetilde{x}_0) < F(\widetilde{x}_1)$, so that $x_0 = \widetilde{x}_0$ and $x_1 = \widetilde{x}_1$. Choose $\xi \in (x_0, 0)$ such that $F(\xi) - F(x_0) = F(0) - F(x_1)$. Then for all $\delta > 0$, there exists $\varepsilon_0 > 0$ such that $r_\varepsilon \in (\xi - \delta, \delta)$ for all $0 < \varepsilon < \varepsilon_0$. □

Proof Without loss of generality, we may assume $\xi - \delta > x_0$ and $\delta < x_1$. Taking $x = \xi - \delta$ in Theorem 3, we may choose ε_1 such that $\xi - \delta < r_\varepsilon$ for all $\varepsilon < \varepsilon_1$. For the upper bound on r_ε, we apply Theorem 3 to $x \mapsto F(-x)$. In this case, the theorem says that if $x \in (-x_1, 0)$ satisfies $F(-x) - F(x_1) < F(0) - F(x_0)$, then there exists $\varepsilon_2 > 0$ such that $x < \widetilde{r}_\varepsilon$ for all $\varepsilon < \varepsilon_2$, where $\widetilde{r}_\varepsilon$ is the nodal point of $x \mapsto -\eta(-x)$, that is, $\widetilde{r}_\varepsilon = -r_\varepsilon$. Taking $x = -\delta$ and $\varepsilon_0 = \varepsilon_1 \wedge \varepsilon_2$ finishes the proof. □

3.3 Behavior near the minima

Corollary 4 divides the domain of the second eigenfunction, η, into three intervals: two infinite half-lines that each contain one of the two minima, and a bounded interval separating the half-lines that contains the nodal point. Our next order of business is to show that η is asymptotically flat on the infinite half-lines. Theorem 5 gives this result for the half-line containing $\widetilde{x}_0$. Applying Theorem 5 to $x \mapsto F(-x)$ gives the result for the half-line containing $\widetilde{x}_1$.

We begin with a lemma. Recall $a_j, \widetilde{a}_j$ and $c_j, \widetilde{c}_j$ from (3), (4), and (17). In applying this lemma, note that

$$\frac{\widetilde{a}_2}{\widetilde{a}_1} = \frac{a_2+2}{a_1+2} < \frac{2a_1}{a_1+2} < \frac{a_1}{2},$$

where the first inequality comes from $a_2 < 2a_1 - 2$ and the second from $a_1 > 2$.

Lemma 8 *Let $x \in (\widetilde{x}_0, 0)$. Suppose p satisfies*

$$\frac{2}{a_1} < p < \frac{\widetilde{a}_1}{\widetilde{a}_2} \le 1.$$

Then there exists $u_0 < -1$ *and* $C > 0$ *such that*

$$e^{-F(u)/\varepsilon} \int_u^x e^{F(v)/\varepsilon}\, dv \le C\varepsilon |u|^{-pa_1/2},$$

for all $u < u_0$ *and all* $\varepsilon > 0$.

Proof Choose $t < \widetilde{x}_0$ such that $F(t) = F(x)$. Using (3), we may choose $u_0 < -1$ and $C' > 0$ such that

(i) $-|u_0|^p < t$,
(ii) $F(\theta) > 0$ and $|F'(\theta)| \ge C'|\theta|^{a_1/2}$, for all $\theta < -|u_0|^p$, and
(iii) $\widetilde{c}_3|u|^{p\widetilde{a}_2-\widetilde{a}_1} < \dfrac{\widetilde{c}_1}{2}$ and $\widetilde{c}_4 - \widetilde{c}_2 \le \dfrac{\widetilde{c}_1}{4}|u|^{\widetilde{a}_1}$, for all $u < u_0$.

Let $G(u) = \int_u^x e^{F(v)/\varepsilon}\, dv$ and $H(u) = e^{F(u)/\varepsilon}$. Fix $u < u_0$ and let $v = -|u|^p < -|u_0|^p$. Note that $u < v$.

By Cauchy's generalized law of the mean,

$$\frac{G(u) - G(v)}{H(u) - H(v)} = \frac{G'(\theta)}{H'(\theta)},$$

for some $u < \theta < v$. From this, we get

$$\begin{aligned} \frac{G(u)}{H(u)} &= \frac{G(v)}{H(u)} + \frac{G'(\theta)}{H'(\theta)}\left(1 - \frac{H(v)}{H(u)}\right) \\ &= \frac{G(v)}{H(u)} + \frac{\varepsilon}{|F'(\theta)|}\left(1 - \frac{H(v)}{H(u)}\right) \\ &\le \frac{G(v)}{H(u)} + \frac{\varepsilon}{|F'(\theta)|}. \end{aligned}$$

By (ii),

$$\frac{\varepsilon}{|F'(\theta)|} \le \frac{\varepsilon}{C'|\theta|^{a_1/2}} \le \frac{\varepsilon}{C'|v|^{a_1/2}} = \frac{\varepsilon}{C'|u|^{pa_1/2}}.$$

It therefore suffices to show that

$$\frac{G(v)}{H(u)} \le C''\varepsilon|u|^{-pa_1/2}, \tag{22}$$

for some constant C'' that does not depend on u or ε.

By (17),

$$\begin{aligned} F(v) - F(u) &\le \widetilde{c}_3|v|^{\widetilde{a}_2} + \widetilde{c}_4 - \widetilde{c}_1|u|^{\widetilde{a}_1} - \widetilde{c}_2 \\ &= (\widetilde{c}_3|u|^{p\widetilde{a}_2-\widetilde{a}_1} - \widetilde{c}_1)|u|^{\widetilde{a}_1} + \widetilde{c}_4 - \widetilde{c}_2 \le -\frac{\widetilde{c}_1}{4}|u|^{\widetilde{a}_1}, \end{aligned}$$

where the last inequality comes from (iii). By (i), we have $v < t$, so that $F(v) > F(w)$ for all $w \in (v, x)$. Hence,

$$\begin{aligned}\frac{G(v)}{H(u)} = \int_v^x e^{(F(w)-F(u))/\varepsilon}\,dw \le |v|e^{(F(v)-F(u))/\varepsilon} \\ \le |u|^p \exp\left(-\frac{\widetilde{c}_1}{4\varepsilon}|u|^{\widetilde{a}_1}\right) \\ = \left(\varepsilon|u|^{-pa_1/2}\right)\frac{1}{\varepsilon}|u|^{p\widetilde{a}_1}\exp\left(-\frac{\widetilde{c}_1}{4\varepsilon}|u|^{\widetilde{a}_1}\right).\end{aligned}$$

Since $x \mapsto x^p e^{-\widetilde{c}_1 x/4}$ is bounded on $[0,\infty)$, this proves (22). □

Theorem Let $x \in (\widetilde{x}_0, 0)$ satisfy $F(x) - F(\widetilde{x}_0) < F(0) - F(x_1)$. Then there exists $C > 0$ and $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$,

$$\left|1 - \frac{\eta(x)}{\eta(-\infty)}\right| \le \frac{C}{\varepsilon}\exp\left(-\frac{1}{\varepsilon}(F(0) - F(x_1) - F(x) + F(\widetilde{x}_0))\right). \tag{23}$$

Proof Again by (18), there exists a constant $C_1 > 0$, not depending on ε, such that $\lambda \le C_1 e^{-(F(0)-F(x_1))/\varepsilon}$.

Let ε_0 be as in Theorem 3, and let $\varepsilon \in (0, \varepsilon_0)$. Choose $t < \widetilde{x}_0$ such that $F(t) = F(x)$. By Theorem 3, $x < r_\varepsilon$. Since η is increasing, $\eta(u) < 0$ for all $u \le x$. Therefore, by (16),

$$\begin{aligned}0 < \eta(x) - \eta(-\infty) &= \frac{\lambda}{\varepsilon}\int_{-\infty}^x \int_u^x e^{(F(v)-F(u))/\varepsilon}|\eta(u)|\,dv\,du \\ &\le \frac{\lambda}{\varepsilon}|\eta(-\infty)|\int_{-\infty}^x \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du.\end{aligned}$$

Thus,

$$\begin{aligned}\left|1 - \frac{\eta(x)}{\eta(-\infty)}\right| &\le \frac{\lambda}{\varepsilon}\int_{-\infty}^x \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du \\ &\le \frac{C_1}{\varepsilon}e^{-(F(0)-F(x_1))/\varepsilon}\int_{-\infty}^x \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du.\end{aligned} \tag{24}$$

Choose p as in Lemma 8. Then there exist $u_0 < 0$ and $C_2 > 0$ such that

$$\int_{-\infty}^{u_0} \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du \le C_3\varepsilon,$$

where $C_3 = C_2|u_0|^{1-pa_1/2}/(pa_1/2 - 1)$. By the proof of Lemma 8, we have $u_0 < t$, and so

$$\int_{u_0}^t \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du \le \int_{u_0}^t (x-u)\,du \le |u_0|^2.$$

Lastly,

$$\int_t^x \int_u^x e^{(F(v)-F(u))/\varepsilon}\,dv\,du \le \int_t^x \int_u^x e^{(F(x)-F(\widetilde{x}_0))/\varepsilon}\,dv\,du \le |u_0|^2 e^{(F(x)-F(\widetilde{x}_0))/\varepsilon}.$$

Thus,

$$\int_{-\infty}^{x}\int_{u}^{x} e^{(F(v)-F(u))/\varepsilon}\,dv\,du \le C_3\varepsilon + |u_0|^2 + |u_0|^2 e^{(F(x)-F(\widetilde{x}_0))/\varepsilon}$$
$$\le C_4 e^{(F(x)-F(\widetilde{x}_0))/\varepsilon},$$

where $C_4 = (C_3\varepsilon_0 + |u_0|^2)e^{-(F(x)-F(\widetilde{x}_0))/\varepsilon_0} + |u_0|^2$. Finally, combining this with (24), we obtain (23), where $C = C_1C_4$. □

3.4 Behavior near the nodal point

From this point forward, for definiteness, we assume $F(\widetilde{x}_0) < F(\widetilde{x}_1)$, so that $x_0 = \widetilde{x}_0$ and $x_1 = \widetilde{x}_1$.

Having shown that η is asymptotically flat near the minima, we would now like to show that it behaves, weakly, like a simple function that is constant on the domains of attraction defined in (2). That is, we want to show that $\int_{D_0} \eta\, d\varpi \sim \eta(x_0)\varpi(D_0)$ and $\int_{D_1} \eta\, d\varpi \sim \eta(x_1)\varpi(D_1)$. (Note that we cannot use Theorem 22 since η depends on ε.) Combined with $\int \eta\, d\varpi = 0$, this would give us the relative magnitudes of $\eta(x_0)$ and $\eta(x_1)$. By Theorem 5, this is equivalent to understanding the relative magnitudes of $\eta(-\infty)$ and $\eta(\infty)$, respectively.

Lemma 9 *Choose $\delta \in (0, x_1)$ such that $\xi - \delta \in (x_0, 0)$. Let k be a positive integer and let $g : \mathbb{R} \to \mathbb{R}$ be bounded. If g is continuous at x_0 and x_1, then*

$$\int_{-\infty}^{\xi-\delta} g(x)|\eta(x)|^k e^{-F(x)/\varepsilon}\,dx \sim g(x_0)|\eta(-\infty)|^k \sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}\, e^{-F(x_0)/\varepsilon}, \tag{25}$$

and

$$\int_{\delta}^{\infty} g(x)|\eta(x)|^k e^{-F(x)/\varepsilon}\,dx \sim g(x_1)|\eta(\infty)|^k \sqrt{\frac{2\pi\varepsilon}{F''(x_1)}}\, e^{-F(x_1)/\varepsilon}, \tag{26}$$

as $\varepsilon \to 0$.

Proof By writing $g = g^+ - g^-$, g^+ and g^- nonnegative, we may assume without loss of generality that g is nonnegative. By Corollary 4 and the fact that η is increasing, we have that, for ε sufficiently small, $|\eta(x)| \le |\eta(-\infty)|$ for all $x \in (-\infty, \xi - \delta)$. Thus,

$$\int_{-\infty}^{\xi-\delta} g(x)|\eta(x)|^k e^{-F(x)/\varepsilon}\,dx \le |\eta(-\infty)|^k \int_{-\infty}^{\xi-\delta} g(x)e^{-F(x)/\varepsilon}\,dx.$$

Similarly,

$$\int_{-\infty}^{\xi-\delta} g(x)|\eta(x)|^k e^{-F(x)/\varepsilon}\,dx \ge |\eta(\xi-\delta)|^k \int_{-\infty}^{\xi-\delta} g(x)e^{-F(x)/\varepsilon}\,dx.$$

Hence, by Theorem 5,

$$\int_{-\infty}^{\xi-\delta} g(x)|\eta(x)|^k e^{-F(x)/\varepsilon}\,dx \sim |\eta(-\infty)|^k \int_{-\infty}^{\xi-\delta} g(x)e^{-F(x)/\varepsilon}\,dx.$$

By Theorem 22, this proves (25). Replacing F with $x \mapsto F(-x)$, Theorem 5 shows that $\eta(\delta) \sim \eta(\infty)$. Thus, the same argument can be used to obtain (26). □

Lemma 10 *There exists* $\delta_0 > 0$ *such that for all* $\delta \in (0,\delta_0)$*,*

$$\int_{\xi-\delta}^{\delta} \eta(x)e^{-F(x)/\varepsilon}\,dx = o\left(\int_{-\infty}^{\xi-\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx + \int_{\delta}^{\infty} |\eta(x)|e^{-F(x)/\varepsilon}\,dx\right),$$

as $\varepsilon \to 0$.

Proof Without loss of generality, we may assume $F(x_0) = 0$. Let $\gamma = (F(0) - F(x_1))/4 > 0$. By the continuity of F, we may choose $\delta_0 > 0$ such that $F(-\delta_0) > F(x_1)$ and

$$F(-\delta/2) - F(x_0 - \delta) - F(x_1) > 2\gamma, \tag{27}$$

for all $\delta \in (0,\delta_0)$.

Let $\delta \in (0,\delta_0)$ be arbitrary. By Theorem 2 applied to $x \mapsto F(-x)$, there exists $\delta' > 0$ and $0 < \varepsilon_0 < x_1$ such that

$$|\eta(x)| \le (1+\delta')|\eta(x_1-\varepsilon)|P^x(\tau^X_{x_1-\varepsilon} < \tau^X_{r_\varepsilon}),$$

for all $x \in (\xi-\delta,-\delta)\cap(r_\varepsilon,\infty)$ and all $\varepsilon \in (0,\varepsilon_0)$. For any such x and ε, since X is continuous and

$$x_0 - \delta < r_\varepsilon < x < -\delta/2 < x_1 - \varepsilon,$$

it follows that on $\{\tau^X_{x_1-\varepsilon} < \tau^X_{r_\varepsilon}\}$, we have $\tau^X_{-\delta/2} < \tau^X_{x_0-\delta}$, P^x-a.s. Hence,

$$|\eta(x)| \le (1+\delta')|\eta(x_1-\varepsilon)|P^x(\tau^X_{-\delta/2} < \tau^X_{x_0-\delta}).$$

By making ε_0 smaller, if necessary, and using Theorem 5 applied to $x \mapsto F(-x)$, this gives

$$|\eta(x)| \le (1+\delta')^2|\eta(\infty)|P^x(\tau^X_{-\delta/2} < \tau^X_{x_0-\delta}),$$

for all $x \in (\xi-\delta,-\delta)\cap(r_\varepsilon,\infty)$ and all $\varepsilon \in (0,\varepsilon_0)$. By (27), we may apply Lemma 3, so that by making ε_0 smaller, if necessary, we obtain

$$|\eta(x)| \le (1+\delta')^2|\eta(\infty)|\exp\left(-\frac{1}{\varepsilon}(F(-\delta/2) - F(x_0-\delta) - 2\gamma),\right) \tag{28}$$

for all $x \in (\xi-\delta,-\delta)\cap(r_\varepsilon,\infty)$ and all $\varepsilon \in (0,\varepsilon_0)$. By (27), for fixed $x \in (\xi-\delta,-\delta)\cap(r_\varepsilon,\infty)$ and $\varepsilon \in (0,\varepsilon_0)$, we may write

$$|\eta(x)| \le (1+\delta')^2|\eta(\infty)|e^{-F(x_1)/\varepsilon}.$$

For fixed $x \in (-\infty, r_\varepsilon]$, by the monotonicity of η, we have $|\eta(x)| \le |\eta(-\infty)|$. Therefore, for all $x \in (\xi - \delta, -\delta)$ and all $\varepsilon \in (0, \varepsilon_0)$, we have

$$|\eta(x)| \le (1+\delta')^2(|\eta(-\infty)| + |\eta(\infty)|e^{-F(x_1)/\varepsilon}).$$

By Proposition 23, after making ε_0 smaller, if necessary, we have

$$\begin{aligned}\int_{\xi-\delta}^{-\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx &\le (1+\delta')^2(|\eta(-\infty)| + |\eta(\infty)|e^{-F(x_1)/\varepsilon}) \int_{\xi-\delta}^{-\delta} e^{-F(x)/\varepsilon}\,dx \\ &\le (1+\delta')^3(|\eta(-\infty)| + |\eta(\infty)|e^{-F(x_1)/\varepsilon}) \frac{\varepsilon}{F'(\xi-\delta)}. \qquad (29)\end{aligned}$$

Let $m = \min\{F'(\xi-\delta), F'(-\delta), |F'(\delta)|\}$. Choose $c \in \{-\delta, \delta\}$ such that $F(c) = F(-\delta) \wedge F(\delta)$. By Proposition 23, by making ε_0 smaller, if necessary, we also have

$$\begin{aligned}\int_{-\delta}^{\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx &\le (|\eta(-\infty)| + |\eta(\infty)|) \left(\int_{-\delta}^{0} e^{-F(x)/\varepsilon}\,dx + \int_{0}^{\delta} e^{-F(x)/\varepsilon}\,dx\right) \\ &\le (1+\delta')(|\eta(-\infty)| + |\eta(\infty)|) \frac{2\varepsilon}{m} e^{-F(c)/\varepsilon} \\ &\le (1+\delta')(|\eta(-\infty)| + |\eta(\infty)|e^{-F(x_1)/\varepsilon}) \frac{2\varepsilon}{m} \qquad (30)\end{aligned}$$

Combining (29) and (30) gives

$$\int_{\xi-\delta}^{\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx \le (1+\delta')^3(|\eta(-\infty)| + |\eta(\infty)|e^{-F(x_1)/\varepsilon}) \frac{3\varepsilon}{m}. \qquad (31)$$

Using Lemma 9, again making ε_0 smaller, if necessary, we have

$$\begin{aligned}&\int_{\xi-\delta}^{\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx \\ &\le (1+\delta')^4 \left(\sqrt{\frac{F''(x_0)}{2\pi\varepsilon}} \int_{-\infty}^{\xi-\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx \right. \\ &\qquad \left. + \sqrt{\frac{F''(x_1)}{2\pi\varepsilon}} \int_{\delta}^{\infty} |\eta(x)|e^{-F(x)/\varepsilon}\,dx\right) \frac{3\varepsilon}{m} \\ &\le \frac{3\varepsilon^{1/2}(1+\delta')^4\sqrt{F''(x_0) \vee F''(x_1)}}{m} \\ &\qquad \times \left(\int_{-\infty}^{\xi-\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx + \int_{\delta}^{\infty} |\eta(x)|e^{-F(x)/\varepsilon}\,dx\right),\end{aligned}$$

which completes the proof. □

Remark Although we have narrowed down the location of the nodal point, r_ε, to the interval $(\xi - \delta, \delta)$, the work in [8] suggests that the nodal point actually converges to

ξ. Moreover, the caption to [8, Fig. 3], states that a step function with discontinuity at ξ is a candidate limit for η as $\varepsilon \to 0$. However, (28) shows that $\eta(x) = o(\eta(\infty))$ for all $x < 0$. In fact, together with Theorem 5 applied to $x \mapsto F(-x)$, it follows that $\eta/\eta(\infty) \to 1_{(0,\infty)}$, pointwise on $\mathbb{R} \setminus \{0\}$. □

Proposition We have

$$\frac{\eta(\infty)}{|\eta(-\infty)|} \sim \sqrt{\frac{F''(x_1)}{F''(x_0)}}\, e^{(F(x_1)-F(x_0))/\varepsilon},$$

as $\varepsilon \to 0$. □

Proof Choose δ such that Lemma 9 and Lemma 10 hold. Let

$$\begin{aligned}
\kappa_{1,\varepsilon} &= \int_{-\infty}^{\xi-\delta} \eta(x)e^{-F(x)/\varepsilon}\,dx = -\int_{-\infty}^{\xi-\delta} |\eta(x)|e^{-F(x)/\varepsilon}\,dx,\\
\kappa_{2,\varepsilon} &= \int_{\delta}^{\infty} \eta(x)e^{-F(x)/\varepsilon}\,dx = \int_{\delta}^{\infty} |\eta(x)|e^{-F(x)/\varepsilon}\,dx,\\
\kappa_{3,\varepsilon} &= \int_{\xi-\delta}^{\delta} \eta(x)e^{-F(x)/\varepsilon}\,dx.
\end{aligned}$$

Since $\int_{\mathbb{R}} \eta(x)e^{-F(x)/\varepsilon}\,dx = 0$, we have that $|\kappa_{1,\varepsilon}| = |\kappa_{2,\varepsilon}| + \kappa_{3,\varepsilon}$. By Lemma 10, we also have that $\kappa_{3,\varepsilon} = o(|\kappa_{1,\varepsilon}| + |\kappa_{2,\varepsilon}|)$.

Since $\kappa_{3,\varepsilon} = o(|\kappa_{1,\varepsilon}| + |\kappa_{2,\varepsilon}|)$, there exists $\varepsilon_0 > 0$ such that $|\kappa_{1,\varepsilon}| + |\kappa_{2,\varepsilon}| > 0$ and

$$\frac{|\kappa_{3,\varepsilon}|}{|\kappa_{1,\varepsilon}| + |\kappa_{2,\varepsilon}|} < 1,$$

for all $\varepsilon \in (0, \varepsilon_0)$. Hence, for any such ε, we may write

$$\frac{2\left(\frac{\kappa_{3,\varepsilon}}{|\kappa_{1,\varepsilon}|+|\kappa_{2,\varepsilon}|}\right)}{1-\left(\frac{\kappa_{3,\varepsilon}}{|\kappa_{1,\varepsilon}|+|\kappa_{2,\varepsilon}|}\right)} = \frac{2\kappa_{3,\varepsilon}}{|\kappa_{1,\varepsilon}| + |\kappa_{2,\varepsilon}| - \kappa_{3,\varepsilon}} = \frac{\kappa_{3,\varepsilon}}{|\kappa_{2,\varepsilon}|},$$

which implies $|\kappa_{2,\varepsilon}| > 0$ for all such ε, and also shows that $\kappa_{3,\varepsilon}/|\kappa_{2,\varepsilon}| \to 0$ as $\varepsilon \to 0$. Therefore, $|\kappa_{1,\varepsilon}|/|\kappa_{2,\varepsilon}| = 1 + \kappa_{3,\varepsilon}/|\kappa_{2,\varepsilon}| \to 1$ as $\varepsilon \to 0$. That is, $|\kappa_{1,\varepsilon}| \sim |\kappa_{2,\varepsilon}|$. Applying Lemma 9 finishes the proof. □

In the following theorem, we improve the results of Lemma 9 in the case $k = 1$, to extend the intervals of integration to include the entire domains of attraction.

Theorem If $g \in L^\infty(\mathbb{R})$ is continuous at x_0 and x_1, then

$$\int_{-\infty}^{0} g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx \sim g(x_0)\eta(-\infty)\sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}\, e^{-F(x_0)/\varepsilon}, \tag{32}$$

and

$$\int_0^\infty g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx \sim g(x_1)\eta(\infty)\sqrt{\frac{2\pi\varepsilon}{F''(x_1)}}\,e^{-F(x_1)/\varepsilon}, \tag{33}$$

as $\varepsilon \to 0$, provided the integrals exist for sufficiently small ε. Consequently,

$$\int g\eta\,d\varpi \sim (g(x_0) - g(x_1))\eta(-\infty), \tag{34}$$

as $\varepsilon \to 0$. □

Proof Without loss of generality, we may assume $F(x_0) = 0$. Choose δ so that Lemma 10 applies. By (31) and Proposition 7,

$$\left|\int_{\xi-\delta}^0 g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx\right| \le \|g\|_\infty(1+\delta')^4\left(1+\sqrt{\frac{F''(x_1)}{F''(x_0)}}\right)|\eta(-\infty)|\frac{6\varepsilon}{m},$$

for ε sufficiently small, where $m = \min\{F'(\xi-\delta), F'(-\delta), |F'(\delta)|\}$. Thus, to prove (32), it suffices to show that

$$\int_{-\infty}^{\xi-\delta} g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx \sim g(x_0)\eta(-\infty)\sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}.$$

But this follows from (25) with $k = 1$ and the fact that $\eta < 0$ on $(-\infty, \xi-\delta)$.

Using Proposition 7, to prove (33), it suffices to show that

$$\int_0^\infty g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx \sim -g(x_1)\eta(-\infty)\sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}.$$

As above, by (31) and Proposition 7, it suffices to show that

$$\int_\delta^\infty g(x)\eta(x)e^{-F(x)/\varepsilon}\,dx \sim -g(x_1)\eta(-\infty)\sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}.$$

But this follows from (26), Proposition 7, and the fact that $\eta > 0$ on (δ, ∞). Finally, combining these results with Proposition 7 and Theorem 22, we obtain

$$\eta(-\infty)^{-1}\int g\eta\,d\varpi \to g(x_0) - g(x_1),$$

as $\varepsilon \to 0$. □

4 Asymptotic behavior of the coupled process

Recall that we are assuming F is a double-well potential in one dimension, with $x_0 < 0 < x_1$ and $F(x_0) < F(x_1)$. Here, the x_j's are the local minima and 0 is the local maximum.

Our construction of the coupling is dependent on our choice of $Q \in \mathbb{R}^{2\times 2}$ and $\xi = \xi^{(1)}$ in the coupling construction outlined in the introduction (see [10, Theorem 3.8] for more details). We begin with a lemma that characterizes all the admissible choices for Q and ξ.

Lemma 11 *Let $\xi_0, \xi_1 \in \mathbb{R}$. Define $a_j = \lambda\xi_j/(\xi_j - \xi_{1-j})$. Then*

$$Q = \begin{pmatrix} -a_0 & a_0 \\ a_1 & -a_1 \end{pmatrix}$$

is the generator of a continuous-time Markov chain with state space $E_0 = \{0,1\}$, eigenvalues $\{0,-\lambda\}$, and corresponding eigenvectors $(1,1)^T$ and $\xi = (\xi_1,\xi_2)^T$ satisfying $\alpha_j = 1 + \xi_j\eta > 0$ if and only if the following conditions hold:

(i) $-\dfrac{1}{\eta(\infty)} \le \xi_j \le \dfrac{1}{|\eta(-\infty)|}$, *for $j = 0,1$, and*
(ii) $\xi_0\xi_1 < 0$.

Proof Note that the a_j are defined precisely so that Q has the given eigenvalues and eigenvectors. Also, $\alpha_j = 1 + \xi_j\eta > 0$ if and only if (i). And the a_j are both positive if and only if (ii). □

For any such choice of ξ as in Lemma 11, we obtain a coupled process (X,Y) with generator B given by (6) and initial distribution ν given by (5). This process is cadlag, X satisfies the SDE given by (1), Y is a continuous-time Markov chain with generator Q, and, by (7),

$$P(X(t) \in \Gamma \mid Y(t) = j) = \int_\Gamma \alpha_j(x)\,\varpi(dx) = \varpi(\Gamma) + \xi_j \int_\Gamma \eta(x)\,\varpi(dx), \qquad (35)$$

for $j = 0,1$ and all Borel sets $\Gamma \subset \mathbb{R}$. Recall that $\varpi = \mu(\mathbb{R})^{-1}\mu$ and $\mu(dx) = e^{-F(x)/\varepsilon}\,dx$.

For each fixed $\varepsilon > 0$, we may choose a different ξ. Hence, all of these objects, in fact, depend on ε. We will, however, suppress that dependence in the notation.

Theorem The following are equivalent to (10):

(a) $\xi_0 = o(|\eta(-\infty)|^{-1})$ as $\varepsilon \to 0$.
(b) $E[g(X(t)) \mid Y(t) = 0] - E^{\varpi}[g(X(0)) \mid X(0) < 0] \to 0$ as $\varepsilon \to 0$, for each $t \ge 0$ and each bounded, measurable $g : \mathbb{R} \to \mathbb{R}$ that is continuous at x_0 and x_1. □

Proof Note that

$$E[g(X(t)) \mid Y(t)=0] - E^{\varpi}[g(X(0)) \mid X(0)<0]$$
$$= \int g(x)(1+\xi_0\eta(x))\,\varpi(dx) - \varpi((-\infty,0))^{-1}\int_{-\infty}^{0} g(x)\,\varpi(dx).$$

Since $\varpi((-\infty,0))^{-1} \to 1$ and $\left|\int_0^\infty g\,d\varpi\right| \le \|g\|_\infty \varpi((0,\infty)) \to 0$, in order to prove that (a) and (b) are equivalent, it suffices to show that $\xi_0 = o(|\eta(-\infty)|^{-1})$ if and only if $\xi_0 \int g\eta\,d\varpi \to 0$ for all g satisfying the hypotheses. But this follows from (34).

That (b) implies (10) is trivial. Assume (10). Since

$$P(X(t)<0 \mid Y(t)=0) = \varpi((-\infty,0)) + \xi_0 \int_{-\infty}^{0} \eta(x)\,\varpi(dx),$$

and $\varpi((-\infty,0)) \to 1$, it follow that $\xi_0 \int_{-\infty}^0 \eta(x)\,\varpi(dx) \to 0$. By (34) with $g = 1_{(-\infty,0)}$, we have $\int_{-\infty}^0 \eta(x)\,\varpi(dx) \sim \eta(-\infty)$, and (a) follows. □

Theorem The following are equivalent to (11):

(a) $\xi_1 \sim |\eta(-\infty)|^{-1}$.
(b) $E[g(X(t)) \mid Y(t)=1] - E^{\varpi}[g(X(0)) \mid X(0)>0] \to 0$ as $\varepsilon \to 0$, for each $t \ge 0$ and each bounded, measurable $g : \mathbb{R} \to \mathbb{R}$ that is continuous at x_0 and x_1.

Moreover, (11) implies (10). □

Proof Note that

$$E[g(X(t)) \mid Y(t)=1] - E^{\varpi}[g(X(0)) \mid X(0)>0]$$
$$= \int g(x)(1+\xi_1\eta(x))\,\varpi(dx) - \varpi((0,\infty))^{-1}\int_0^\infty g(x)\,\varpi(dx)$$
$$= \int g\,d\varpi - \varpi((0,\infty))^{-1}\int_{(0,\infty)} g\,d\varpi + \xi_1 \int g\eta\,d\varpi$$

To prove that (a) and (b) are equivalent, by (39), it suffices to show that $\xi_1 \sim |\eta(-\infty)|^{-1}$ if and only if $\xi_1 \int g\eta\,d\varpi \to -(g(x_0)-g(x_1))$ for all g satisfying the hypotheses. But this follows from (34).

That (b) implies (11) is trivial. Assume (11). Since

$$P(X(t)>0 \mid Y(t)=1) = \varpi((0,\infty)) + \xi_1 \int_0^\infty \eta(x)\,\varpi(dx),$$

and $\varpi((0,\infty)) \to 0$, it follows that $\xi_1 \int_0^\infty \eta(x)\,\varpi(dx) \to 1$. By (34) with $g = 1_{(0,\infty)}$, we have $\int_0^\infty \eta(x)\,\varpi(dx) \sim |\eta(-\infty)|$, and (a) follows.

Finally, assume (11). Then (a) holds. By Lemma 11, we have $-\eta(\infty)^{-1} \le \xi_0 < 0$ for sufficiently small ε. In particular, $|\xi_0| \le \eta(\infty)^{-1}$, so Theorem 9(a) follows from Proposition 7. □

Theorem Let $0 < \rho < |x_0| \wedge x_1$. Then $\xi_0 = o(\xi_1)$ if and only if

$$E^1[\tau_0^Y] \sim E^{x_1}[\tau^X_{B_\rho(x_0)}] \sim \lambda^{-1} \sim \frac{2\pi}{|F''(0)F''(x_1)|^{1/2}} e^{(F(0)-F(x_1))/\varepsilon}, \tag{36}$$

as $\varepsilon \to 0$. And $\xi_1/\xi_0 \sim \eta(\infty)/\eta(-\infty)$ if and only if

$$E^0[\tau_1^Y] \sim E^{x_0}[\tau^X_{B_\rho(x_1)}] \sim \frac{2\pi}{|F''(0)F''(x_0)|^{1/2}} e^{(F(0)-F(x_0))/\varepsilon}, \tag{37}$$

as $\varepsilon \to 0$. Moreover, (37) implies (36), which implies (10). Also, (11) implies (36).□

Proof By (19) and (18), we need only determine the asymptotics of a_0 and a_1. Recall that $a_j = \lambda\xi_j/(\xi_j - \xi_{1-j})$. Thus,

$$E^j[\tau^Y_{1-j}] = a_j^{-1} = \lambda^{-1}\left(1 - \frac{\xi_{1-j}}{\xi_j}\right) \sim \left(1 - \frac{\xi_{1-j}}{\xi_j}\right) \frac{2\pi}{|F''(0)F''(x_1)|^{1/2}} e^{(F(0)-F(x_1))/\varepsilon}, \tag{38}$$

so the first biconditional follows immediately. The second biconditional then follows from Proposition 7.

By Proposition 7, we have (37) implies (36). By Lemma 11, we have $|\xi_0/\xi_1| \geq |\xi_0\eta(-\infty)|$, so that $\xi_0 = o(\xi_1)$ implies $\xi_0 = o(|\eta(-\infty)|^{-1})$. Hence, (36) implies that Theorem 9(a) holds, which is equivalent to (10).

Finally, suppose (11) holds. By Theorems 9 and 10, we have that $\xi_1 \sim |\eta(-\infty)|^{-1}$ and $\xi_0 = o(|\eta(-\infty)|^{-1})$, so that $\xi_0 = o(\xi_1)$, which is equivalent to (36). □

Theorem The Markov chain fully tracks the diffusion, in the sense that (10)-(13) all hold, if and only if $\xi_0 \sim -\eta(\infty)^{-1}$ and $\xi_1 \sim |\eta(-\infty)|^{-1}$. □

Proof Suppose (10)-(13) hold. Then, by Theorem 10, we have $\xi_1 \sim |\eta(-\infty)|^{-1}$. Since (13) is equivalent to (37), we also have, by Theorem 11, that $\xi_1/\xi_0 \sim \eta(\infty)/\eta(-\infty)$. Thus, $\xi_0 \sim -\eta(\infty)^{-1}$.

Conversely, suppose $\xi_0 \sim -\eta(\infty)^{-1}$ and $\xi_1 \sim |\eta(-\infty)|^{-1}$. Theorem 10 gives us (11) and (10). Theorem 11 gives us (37) and (36), which are equivalent to (13) and (12), respectively. □

In this section, we have established that (11) implies (12) implies (10), and (13) implies (12). Example 13 shows that it is possible to have all four conditions holding. The remaining examples illustrate that there are no implications besides those already mentioned.

Example Let $\xi_0 = -f(\varepsilon)\eta(\infty)^{-1}$ and $\xi_1 = g(\varepsilon)|\eta(-\infty)|^{-1}$, where $0 < f, g \leq 1$ with $f, g \to 1$ as $\varepsilon \to 0$. By Lemma 11 and Theorem 12, this is the most general family of choices such that the resulting coupling sequence satisfies (10)-(13). □

In the remaining examples, let

$$L(\varepsilon) = \sqrt{\frac{F''(x_1)}{F''(x_0)}}\, e^{-(F(x_1)-F(x_0))/\varepsilon},$$

so that by Proposition 7, we have $\eta(\infty)/\eta(-\infty) \sim -L(\varepsilon)^{-1}$. Choose $0 < f \le 1$ and $h \ge L$, and let $g = L/h$, so that $0 < g \le 1$. Let $\xi_0 = -f(\varepsilon)\eta(\infty)^{-1}$ and $\xi_1 = g(\varepsilon)|\eta(-\infty)|^{-1}$. By Lemma 11, these are admissible choices for ξ_0 and ξ_1.

Note that $\xi_0 \sim -f(\varepsilon)L(\varepsilon)|\eta(-\infty)|^{-1}$, so that by Theorem 9, we have (10) in all these examples. Also note that by Theorem 10, we have (11) if and only if $h \sim L$. For applying Theorem 11, note that $\xi_1/\xi_0 \sim -g/(fL) = -1/(fh)$. Thus, (12) holds if and only if $fh \to 0$ and (13) holds if and only $fh \sim L$.

Example Let $f = h = 1$. Then none of (11), (12), or (13) hold, so we see that (10) does not imply any of the other conditions. □

Example Let $f = 1$ and $h = \sqrt{L}$. In this case, we have (12), but neither (11) nor (13) hold. Hence, (12) implies neither (11) nor (13). □

Example Let $f = h = L$. In this case, (11) and (12) hold, but (13) does not, showing that (11) does not imply (13). □

Example Let $f = h = \sqrt{L}$. Here we have (12) and (13), but not (11), showing that (13) does not imply (11). □

Acknowledgements This paper was completed while the first author was visiting the University of California, San Diego with the support of the Charles Lee Powell Foundation. The hospitality of that institution, particularly that of Professor Ruth Williams, was greatly appreciated.

Appendix 1

Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be Lipschitz and let $\varphi_{x,b}$ be the unique solution to $\varphi'_{x,b} = b(\varphi_{x,b})$ with $\varphi_{x,b}(0) = x$. For $\varepsilon > 0$, let $X_{\varepsilon,b}$ be defined by

$$X_{\varepsilon,b}(t) = X_{\varepsilon,b}(0) + \int_0^t b(X_{\varepsilon,b}(s))\,ds + \sqrt{2\varepsilon}W(t),$$

where W is a standard d-dimensional Brownian motion. As in Section 1, if $F : \mathbb{R}^d \to \mathbb{R}$ is given, then $\varphi_x = \varphi_{x,-\nabla F}$ and $X_\varepsilon = X_{\varepsilon,-\nabla F}$. For the F we use later, $-\nabla F$ is *not* Lipschitz. This will cause no difficulty, however, since it will be locally Lipschitz, and we will only apply these theorems on compact sets.

This first theorem is [11, Theorem 2.40]. It describes the asymptotic mean time to leave a domain of attraction.

Theorem Let $F : \mathbb{R}^d \to \mathbb{R}$ have continuous and bounded derivatives up to second order. Let D be a bounded open domain in $\mathbb{R}^d$ with boundary ∂D of class C^2 and $\langle -\nabla F(x), n(x)\rangle < 0$ for all $x \in \partial D$, where $n(x)$ is the outward unit normal vector to ∂D at x.

Let $x_0 \in D$. Assume that if G is a neighborhood of x_0, then there exists a neighborhood $\widetilde{G}$ of x_0 such that $\widetilde{G} \subset G$ and, for all $x \in \widetilde{G}$, we have $\varphi_x([0,\infty)) \subset G$ and $\varphi_x(t) \to x_0$ as $t \to \infty$. Further assume that, for each $x \in \overline{D}$, we have $\varphi_x((0,\infty)) \subset D$.

Then for any $x \in D$,

(i) $\lim_{\varepsilon\to 0} 2\varepsilon \log E^x[\tau(D^c)] = \inf_{y\in\partial D} 2(F(y) - F(x_0)) =: V_0$, and
(ii) for all $\zeta > 0$, we have $\lim_{\varepsilon\to 0} P^x(e^{(V_0-\zeta)/(2\varepsilon)} < \tau(D^c) < e^{(V_0+\zeta)/(2\varepsilon)}) = 1$.

Moreover, both convergences hold uniformly in x on each compact subset of D. □

This next theorem is [11, Lemma 2.34(b)]. It asserts that the diffusion cannot linger for long inside the domain of attraction without quickly coming into a small neighborhood of the associated minimum.

Theorem Assume the hypotheses of Theorem 18. Fix $\delta > 0$. Then there exists $C > 0$, $T > 0$, and $\varepsilon_0 > 0$ such that

$$P^x(\tau(D^c \cup B_\delta(x_0)) > t) \le e^{-C(t-T)/(2\varepsilon)},$$

for all $x \in \overline{D} \setminus B_\delta(x_0)$, all $t > T$, and all $\varepsilon < \varepsilon_0$. □

The last result we need gives the probability of leaving the domain of attraction through a given point. To state this result, we need some preliminary notation and definitions. See [11, Section 5.3] for more details.

Let $u : [0, T] \to \mathbb{R}^d$. If u is absolutely continuous, define

$$I_T(u) = \frac{1}{2}\int_0^T |u'(s) - b(u(s))|^2\, ds,$$

and define $I_T(u) = \infty$ otherwise.

Let G be a bounded domain in $\mathbb{R}^d$ with ∂G of class C^2 and define

$$V(x, y) = \inf\{I_T(u) \mid T > 0, u : [0, T] \to \mathbb{R}^d, u(0) = x, u(T) = y\}$$
$$V_G(x, y) = \inf\{I_T(u) \mid T > 0, u : [0, T] \to G \cup \partial G, u(0) = x, u(T) = y\}.$$

The functions V and V_G are continuous on $\mathbb{R}^d \times \mathbb{R}^d$ and $(G \cup \partial G) \times (G \cup \partial G)$, respectively. We have $V_G(x, y) \ge V(x, y)$ for all $x, y \in G \cup \partial G$. Also, for all $x, y \in G$, if $V_G(x, y) \le \min_{z\in\partial G} V(x, z)$, then $V_G(x, y) = V(x, y)$.

Note that if $\varphi_{x,b}(t) = y$ for some $t > 0$ and $\varphi_{x,b}([0, t]) \subset G \cup \partial G$, then $V_G(x, y) = 0$. An equivalence relation on $G \cup \partial G$ is defined by $x \sim_G y$ if and only if $V_G(x, y) = V_G(y, x) = 0$. It can be shown that if the equivalence class of y is nontrivial, then $\varphi_{y,b}([0, \infty))$ is contained in that equivalence class.

The ω-limit set of a point $y \in \mathbb{R}^d$ is denoted by $\omega(y)$ and defined as the set of accumulation points of $\varphi_{y,b}([0, \infty))$. Assume that G contains a finite number of compact sets $K_1, \ldots, K_\ell$ such that each K_i is an equivalence class of $\sim_G$. Assume further that, for all $y \in \mathbb{R}^d$, if $\omega(y) \subset G \cup \partial G$, then $\omega(y) \subset K_i$ for some i.

The function V_G is constant on $K_i \times K_j$, so we let $V_G(K_i, K_j)$, $V_G(x, K_i)$, and $V_G(K_i, x)$ denote this common value. Also, $V_G(K_i, \partial G) = \inf_{y\in\partial G} V_G(K_i, y)$.

Given a finite set $\mathcal{L}$ and a nonempty, proper subset $Q \subset \mathcal{L}$, let $\mathbb{G}(Q)$ denote the set of directed graphs on $\mathcal{L}$ with arrows $i \to j$, $i \in \mathcal{L} \setminus Q$, $j \in \mathcal{L}$, $j \ne i$, such that:

(i) from each $i \in \mathcal{L} \setminus Q$ exactly one arrow is issued; (ii) for each $i \in \mathcal{L} \setminus Q$ there is a chain of arrows starting at i and finishing at some point in Q. If j is such a point we say that the graph leads i to j. For $i \in \mathcal{L} \setminus Q$ and $j \in Q$, the set of graphs in $\mathbb{G}(Q)$ leading i to j is denoted by $\mathbb{G}_{i,j}(Q)$.

With $\mathcal{L} = \{K_1, \ldots, K_\ell, \partial G\}$, let

$$M_G = \min_{g \in \mathbb{G}(\partial G)} \sum_{(\alpha \to \beta) \in g} V_G(\alpha, \beta).$$

If $x \in G$ and $y \in \partial G$, then with $\mathcal{L} = \{K_1, \ldots, K_\ell, x, y, \partial G\}$, let

$$M_G(x, y) = \min_{g \in \mathbb{G}_{x,y}(\{y, \partial G\})} \sum_{(\alpha \to \beta) \in g} V_G(\alpha, \beta).$$

The following theorem is [11, Theorem 5.19].

Theorem Under the above assumptions and notation, for any compact set $K \subset G$, $\gamma > 0$, and $\delta > 0$, there exists $\varepsilon_0 > 0$ and $\delta_0 \in (0, \delta)$ so that for any $x \in K$, $y \in \partial G$, and $\varepsilon \in (0, \varepsilon_0)$, we have

$$\exp\left(-\frac{M_G(x, y) - M_G + 2\gamma}{2\varepsilon}\right) \le P^x(X_{\varepsilon,b}(\tau) \in B_{\delta_0}(y)) \le \exp\left(-\frac{M_G(x, y) - M_G - 2\gamma}{2\varepsilon}\right),$$

where $\tau = \tau^{X_{\varepsilon,b}}(\mathbb{R}^d \setminus G)$. □

The next two results are auxiliary results which are needed to apply Theorem 20. The first is [11, Proposition 2.37].

Proposition Under the assumptions of Theorem 18, we have

$$V(x_0, y) = 2(F(y) - F(x_0)),$$

for all $y \in \overline{D}$. □

Lemma 12 *Let $b = -\nabla F$, where F is as in Theorem 18. If there exists $T_0 > 0$ such that $\varphi_x(T_0) = y$, then $V(x, y) = 0$ and $V(y, x) = 2(F(x) - F(y))$.*

Proof Since $\varphi_x' = b(\varphi_x)$, we have $I_{T_0}(\varphi_x) = 0$, which implies $V(x, y) = 0$. Let $T > 0$ and let $\varphi : [0, T] \to \mathbb{R}^d$ satisfy $\varphi(0) = y$ and $\varphi(T) = x$. Then

$$\begin{aligned}
I_T(\varphi) &= \frac{1}{2}\int_0^T |\varphi'(s) - b(\varphi(s))|^2\,ds \\
&= \frac{1}{2}\int_0^T |\varphi'(s) + b(\varphi(s))|^2\,ds - 2\int_0^T \langle \varphi'(s), b(\varphi(s))\rangle\,ds \\
&= \frac{1}{2}\int_0^T |\varphi'(s) + b(\varphi(s))|^2\,ds + 2\int_0^T \langle \varphi'(s), \nabla F(\varphi(s))\rangle\,ds \\
&= \frac{1}{2}\int_0^T |\varphi'(s) + b(\varphi(s))|^2\,ds + 2(F(x) - F(y)).
\end{aligned}$$

This shows that $V(y,x) \geq 2(F(x)-F(y))$. Now let $\psi(t) = \varphi_x(T_0 - t)$. Then $\psi(0) = y$, $\psi(T_0) = x$, and $\psi' = -b(\psi)$. Hence, $V(y,x) \leq I_{T_0}(\psi) = 2(F(x)-F(y))$. □

Appendix 2

Finally, we need two classical results of Laplace that allow us to estimate exponential integrals. The following two results can be found in [5, pp. 36–37]. The notation $a \sim b$ means that $a/b \to 1$.

Theorem Let $I \subset \mathbb{R}$ be a (possibly infinite) open interval, $F \in C^2(I)$, and $x_0 \in I$. Suppose g is continuous at x_0. If $F(x_0)$ is the unique global minimum of F on I, and $F''(x_0) > 0$, then

$$\int_I g(x)e^{-F(x)/\varepsilon}\,dx \sim g(x_0)\sqrt{\frac{2\pi\varepsilon}{F''(x_0)}}\,e^{-F(x_0)/\varepsilon}, \tag{39}$$

as $\varepsilon \to 0$, provided the left-hand side exists for sufficiently small ε. □

Proposition Let $-\infty < a < x_0 < b \leq \infty$ and $F \in C^1(a,b)$. Suppose g is continuous at x_0. If $F(x_0)$ is the unique global minimum of F on $[x_0, b)$ and $F'(x_0) > 0$, then

$$\int_{x_0}^b g(x)e^{-F(x)/\varepsilon}\,dx \sim g(x_0)\frac{\varepsilon}{F'(x_0)}\,e^{-F(x_0)/\varepsilon}, \tag{40}$$

as $\varepsilon \to 0$, provided the left-hand side exists for sufficiently small ε. □

References

1. Anton Bovier, Michael Eckhoff, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.
2. Anton Bovier, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.

3. R. Courant and D. Hilbert. *Methods of mathematical physics. Vol. I.* Interscience Publishers, Inc., New York, N.Y., 1953.
4. Michael Eckhoff. Precise asymptotics of small eigenvalues of reversible diffusions in the metastable regime. *Ann. Probab.*, 33(1):244–299, 2005.
5. A. Erdélyi. *Asymptotic expansions*. Dover Publications, Inc., New York, 1956.
6. Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
7. M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.
8. Wilhelm Huisinga, Sean Meyn, and Christof Schütte. Phase transitions and metastability in Markovian and molecular systems. *Ann. Appl. Probab.*, 14(1):419–458, 2004.
9. Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
10. Thomas G. Kurtz and Jason Swanson. Finite Markov chains coupled to general Markov processes and an application to metastability I. Preprint, 2020.
11. Enzo Olivieri and Maria Eulália Vares. *Large deviations and metastability*, volume 100 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.
12. Makoto Sugiura. Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Japan*, 47(4):755–788, 1995.
13. Makoto Sugiura. Asymptotic behaviors on the small parameter exit problems and the singularly perturbation problems. *Ryukyu Math. J.*, 14:79–118, 2001.

Maximally Distributed Random Fields under Sublinear Expectation

Xinpeng Li and Shige Peng

Abstract This paper focuses on the maximal distribution on sublinear expectation space and introduces a new type of random fields with the maximally distributed finite-dimensional distribution. The corresponding spatial maximally distributed white noise is constructed, which includes the temporal-spatial situation as a special case due to the symmetrical independence property of maximal distribution. In addition, the stochastic integrals with respect to the spatial or temporal-spatial maximally distributed white noises are established in a quite direct way without the usual assumption of adaptability for integrand.

1 Introduction

In mathematics and physics, a random field is a type of parameterized family of random variables. When the parameter is time $t \in \mathbb{R}^+$, we call it a stochastic process, or a temporal random field. Quite often the parameter is space $x \in \mathbb{R}^d$, or time-space $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^d$. In this case, we call it a spatial or temporal-spatial random field. A typical example is the electromagnetic wave dynamically spread everywhere in our $\mathbb{R}^3$-space or more exactly, in $\mathbb{R}^+ \times \mathbb{R}^3$-time-space. In principle, it is impossible to know the exact state of the electromagnetic wave of our real world , namely, it is a nontrivial random field parameterized by the time-space $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^3$.

Classically, a random field is defined on a given probability space $(\Omega, \mathcal{F}, P)$. But for the above problem, can we really get to know the probability P? This involves the so called problem of uncertainty of probabilities.

Xinpeng Li
Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, 266237, Qingdao, China. e-mail: lixinpeng@sdu.edu.cn

Shige Peng (✉)
School of Mathematics, Shandong University, 250100, Jinan, China. e-mail: peng@sdu.edu.cn

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_14

Over the past few decades, non-additive probabilities or nonlinear expectations have become active domains for studying uncertainties, and received more and more attention in many research fields, such as mathematical economics, mathematical finance, statistics, quantum mechanics. A typical example of nonlinear expectation is sublinear one, which is used to model the uncertainty phenomenon characterized by a family of probability measures $\{P_\theta\}_{\theta\in\Theta}$ in which the true measure is unknown, and such sublinear expectation is usually defined by

$$\mathbb{E}[X] := \sup_{\theta\in\Theta} E_{P_\theta}[X].$$

This notion is also known as the upper expectation in robust statistics (see Huber [9]), or the upper prevision in the theory of imprecise probabilities (see Walley [20]), and has the closed relation with coherent risk measures (see Artzner et al. [1], Delbaen [4], Föllmer and Schied [6]). A first dynamical nonlinear expectation, called g-expectation was initiated by Peng [12].

The foundation of sublinear expectation theory with a new type of G-Brownian motion and the corresponding Itô's stochastic calculus was laid in Peng [13], which keeps the rich and elegant properties of classical probability theory except linearity of expectation. Peng [15] initially defined the notion of independence and identical distribution (i.i.d.) based on the notion of nonlinear expectation instead of the capacity. Based on the notion of new notions, the most important distribution called G-normal distribution introduced, which can be characterized by the so-called G-heat equation. The notions of G-expectation and G-Brownian motion can be regarded as a nonlinear generalization of Wiener measure and classical Brownian motion. The corresponding limit theorems as well as stochastic calculus of Itô's type under G-expectation are systematically developed in Peng [18]. Besides that, there is also another important distribution, called maximal distribution. The distribution of maximally distributed random variable X can be calculated simply by

$$\mathbb{E}[\varphi(X)] = \max_{v\in[-\mathbb{E}[-X],\mathbb{E}[X]]} \varphi(v), \quad \varphi \in C_b(\mathbb{R}).$$

The law of large numbers under sublinear expectation (see Peng [18]) shows that if $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent and identical distributed random variables with $\lim_{c\to\infty} \mathbb{E}[(|X_1| - c)^+] = 0$, then the sample average converges to maximal distribution in law, i.e.,

$$\lim_{n\to\infty} \mathbb{E}[\varphi(\frac{X_1 + \cdots + X_n}{n})] = \max_{v\in[-\mathbb{E}[-X_1],\mathbb{E}[X_1]]} \varphi(v), \quad \forall\varphi \in C_b(\mathbb{R}).$$

We note that the finite-dimensional distribution for quadratic variation process of G-Brownian motion is also maximal distributed.

Recently, Ji and Peng [10] introduced a new G-Gaussian random fields, which contains a type of spatial white noise as a special case. Such white noise is a natural generalization of the classical Gaussian white noise (for example, see Walsh [21], Dalang [2] and Da Prato and Zabczyk [3]). As pointed in [10], the space-indexed increments do not satisfy the property of independence. Once the sublinear G-expectation degenerates to linear case, the property of independence for the space-indexed part turns out to be true as in the classical probability theory.

In this paper, we introduce a very special but also typical random field, called maximally distributed random field, in which the finite-dimensional distribution is maximally distributed. The corresponding space-indexed white noise is also constructed. It is worth mentioning that the space-indexed increments of maximal white noise is independent, which is essentially different from the case of G-Gaussian white noise. Thanks to the symmetrical independence of maximally distributed white noise, it is natural to view the temporal-spatial maximally distributed white noise as a special case of the space-indexed maximally distributed white noise. The stochastic integrals with respect to spatial and temporal-spatial maximally distributed white noises can be constructed in a quite simple way, which generalize the stochastic integral with respect to quadratic variation process of G-Brownian motion introduced in Peng [18]. Furthermore, due to the boundedness of maximally distributed random field, the usual assumption of adaptability for integrand can be dropped. We emphasize that the structure of maximally distributed white noise is quite simple, it can be determined by only two parameters $\underline{\mu}$ and $\overline{\mu}$, and the calculation of the corresponding finite-dimensional distribution is taking the maximum of continuous function on the domain determined by $\underline{\mu}$ and $\overline{\mu}$. The use of maximally distributed random fields for modelling purposes in applications can be explained mainly by the simplicity of their construction and analytic tractability combined with the maximal distributions of marginal which describe many real phenomena due to the law of large numbers with uncertainty.

This paper is organized as follows. In Section 2, we review basic notions and results of nonlinear expectation theory and the notion and properties of maximal distribution. In Section 3, we first recall the general setting of random fields under nonlinear expectations, and then introduce the maximally distributed random fields. In Section 4, we construct the spatial maximally distributed white noise and study the corresponding properties. The properties of spatial as well as temporal-spatial maximally distributed white noise and the related stochastic integrals are established in Section 5.

2 Preliminaries

In this section, we recall some basic notions and properties in the nonlinear expectation theory. More details can be found in Denis et al. [5], Hu and Peng [8] and Peng [13, 14, 15, 16, 18, 19].

Let Ω be a given nonempty set and $\mathcal{H}$ be a linear space of real-valued functions on Ω such that if $X \in \mathcal{H}$, then $|X| \in \mathcal{H}$. $\mathcal{H}$ can be regarded as the space of random variables. In this paper, we consider a more convenient assumption: if random variables $X_1,\cdots,X_d \in \mathcal{H}$, then $\varphi(X_1, X_2, \cdots, X_d) \in \mathcal{H}$ for each $\varphi \in C_{b.Lip}(\mathbb{R}^d)$. Here $C_{b.Lip}(\mathbb{R}^d)$ is the space of all bounded and Lipschitz functions on $\mathbb{R}^d$.

We call $X = (X_1, \cdots, X_n)$, $X_i \in \mathcal{H}$, $1 \leq i \leq n$, an n-dimensional random vector, denoted by $X \in \mathcal{H}^n$.

Definition 1 A nonlinear expectation $\hat{E}$ on $\mathcal{H}$ is a functional $\hat{E} : \mathcal{H} \to \mathbb{R}$ satisfying the following properties: for each $X, Y \in \mathcal{H}$,

(i) Monotonicity: $\hat{E}[X] \geq \hat{E}[Y]$ if $X \geq Y$;
(ii) Constant preserving: $\hat{E}[c] = c$ for $c \in \mathbb{R}$;

The triplet $(\Omega, \mathcal{H}, \hat{E})$ is called a nonlinear expectation space. If we further assume that

(iii) Sub-additivity: $\hat{E}[X + Y] \leq \hat{E}[X] + \hat{E}[Y]$;
(iv) Positive homogeneity: $\hat{E}[\lambda X] = \lambda \hat{E}[X]$ for $\lambda \geq 0$.

Then $\hat{E}$ is called a sublinear expectation, and the corresponding triplet $(\Omega, \mathcal{H}, \hat{E})$ is called a sublinear expectation space.

Let $(\Omega, \mathcal{H}, \hat{E})$ be a nonlinear (resp., sublinear) expectation space. For each given n-dimensional random vector X, we define a functional on $C_{b.Lip}(\mathbb{R}^n)$ by

$$\mathbb{F}_X[\varphi] := \hat{E}[\varphi(X)], \text{ for each } \varphi \in C_{b.Lip}(\mathbb{R}^n).$$

$\mathbb{F}_X$ is called the distribution of X. It is easily seen that $(\mathbb{R}^n, C_{b.Lip}(\mathbb{R}^n), \mathbb{F}_X)$ forms a nonlinear (resp., sublinear) expectation space. If $\mathbb{F}_X$ is not a linear functional on $C_{b.Lip}(\mathbb{R}^n)$, we say X has distributional uncertainty.

Definition 2 Two n-dimensional random vectors X_1 and X_2 defined on nonlinear expectation spaces $(\Omega_1, \mathcal{H}_1, \hat{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \hat{E}_2)$ respectively, are called identically distributed, denoted by $X_1 \overset{d}{=} X_2$, if $\mathbb{F}_{X_1} = \mathbb{F}_{X_2}$, i.e.,

$$\hat{E}_1[\varphi(X_1)] = \hat{E}_2[\varphi(X_2)], \ \forall \varphi \in C_{b.Lip}(\mathbb{R}^n).$$

Definition 3 Let $(\Omega, \mathcal{H}, \hat{E})$ be a nonlinear expectation space. An n-dimensional random vector Y is said to be independent from another m-dimensional random vector X under the expectation $\hat{E}$ if, for each test function $\varphi \in C_{b.Lip}(\mathbb{R}^{m+n})$, we have

$$\hat{E}[\varphi(X, Y)] = \hat{E}[\hat{E}[\varphi(x, Y)]_{x=X}].$$

Remark 1 Peng [15] (see also Peng [18]) introduced the notions of the distribution and the independence of random variables under a nonlinear expectation, which play a crucially important role in the nonlinear expectation theory.

For simplicity, the sequence $\{X_i\}_{i=1}^n$ is called independence if X_{i+1} is independent from $(X_1, \cdots, X_i)$ for $i = 1, 2, \cdots, n-1$. Let $\bar{X}$ and X be two n-dimensional random vectors on $(\Omega, \mathcal{H}, \hat{E})$. $\bar{X}$ is called an independent copy of X, if $\bar{X} \overset{d}{=} X$ and $\bar{X}$ is independent from X.

Remark 2 It is important to note that "Y is independent from X" does not imply that "X is independent from Y" (see Peng [18]).

In this paper, we focus on an important distribution on sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$, called maximal distribution.

Definition 4 An n-dimensional random vector $X = (X_1, \cdots, X_n)$ on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$ is said to be maximally distributed, if there exists a bounded and closed convex subset $\Lambda \subset \mathbb{R}^n$ such that, for every continuous function $\varphi \in C(\mathbb{R}^n)$,

$$\hat{E}[\varphi(X)] = \max_{x \in \Lambda} \varphi(x).$$

Remark 3 Here Λ characterizes the uncertainty of X. It is easy to check that this maximally distributed random vector X satisfies

$$X + \bar{X} \overset{d}{=} 2X,$$

where $\bar{X}$ is an independent copy of X. Conversely, suppose a random variable X satisfying $X + \bar{X} \overset{d}{=} 2X$, if we further assume the uniform convergence condition $\lim_{c \to \infty} \hat{E}[(|X| - c)^+] = 0$ holds, then we can deduce that X is maximally distributed by the law of large numbers (see Peng [18]). An interesting problem is that is X still maximally distributed without such uniform convergence condition? We emphasize that the law of large numbers does not hold in this case, a counterexample can be found in Li and Zong [11].

Proposition 1 *Let $g(p) = \max_{v \in \Lambda} v \cdot p$ be given. Then an n-dimensional random variable is maximally distributed if and only if for each $\varphi \in C(\mathbb{R}^n)$, the following function*

$$u(t, x) := \hat{E}[\varphi(x + tX)] = \max_{v \in \Lambda} \varphi(x + tv), \ (t, x) \in [0, \infty) \times \mathbb{R}^n \tag{1}$$

is the unique viscosity solution of the the following nonlinear partial differential equation

$$\partial_t u - g(D_x u) = 0, \ \ u|_{t=0} = \varphi(x). \tag{2}$$

This property implies that, each sublinear function g on $\mathbb{R}^n$ determines uniquely a maximal distribution. The following property is easy to check.

Proposition 2 *Let X be an n-dimensional maximally distributed random vector characterized by its generating function*

$$g(p) := \hat{E}[X \cdot p], \ \ p \in \mathbb{R}^n.$$

Then, for any function $\psi \in C(\mathbb{R}^n)$, $Y = \psi(X)$ is also an $\mathbb{R}$-valued maximally distributed random variable:

$$\mathbb{E}[\varphi(Y)] = \max_{v \in [\underline{\rho}, \overline{\rho}]} \varphi(v), \qquad \overline{\rho} = \max_{\gamma \in \Lambda} \psi(\gamma), \quad \underline{\rho} = \min_{\gamma \in \Lambda} \psi(\gamma).$$

Proposition 3 *Let $X = (X_1, \cdots, X_n)$ be an n-dimensional maximal distribution on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$. If the corresponding generating function satisfies, for all $p = (p_1, \cdots, p_n) \in \mathbb{R}^n$,*

$$g(p) = \hat{E}[X_1 p_1 + \cdots + X_n p_n] = \hat{E}[X_1 p_1] + \cdots + \hat{E}[X_n p_n],$$

then $\{X_i\}_{i=1}^n$ is a sequence of independent maximally distributed random variables.

Moreover, for any permutation π of $\{1, 2, \cdots, n\}$, the sequence $\{X_{\pi(i)}\}_{i=1}^n$ is also independent.

Proof For $i = 1, \cdots, n$, we denote $\overline{\mu}_i = \hat{E}[X_i]$ and $\underline{\mu}_i = -\hat{E}[-X_i]$. Since

$$\begin{aligned} g(p) &= \hat{E}[X_1 \cdot p_1 + \cdots + X_n \cdot p_n] = \hat{E}[X_1 \cdot p_1] + \hat{E}[X_2 \cdot p_2] + \cdots + \hat{E}[X_n \cdot p_n] \\ &= \sum_{i=1}^n \max_{v_i \in [\underline{\mu}_i, \overline{\mu}_i]} p_i v_i = \max_{(v_1, \cdots, v_n) \in \otimes_{i=1}^n [\underline{\mu}_i, \overline{\mu}_i]} (p_1 v_1 + \cdots + p_n v_n), \end{aligned}$$

it follows Proposition 1 that $(X_1, \cdots, X_n)$ is an n-dimensional maximally distributed random vector such that, $\forall \varphi \in C(\mathbb{R}^n)$,

$$\hat{E}[\varphi(X_1, \cdots, X_n)] = \max_{(v_1, \cdots, v_n) \in \otimes_{i=1}^n [\underline{\mu}_i, \overline{\mu}_i]} \varphi(v_1, \cdots, v_n).$$

It is easy to check that $\{X_i\}_{i=1}^n$ is independent, and so does the permuted sequence $\{X_{\pi(i)}\}_{i=1}^n$. □

Remark 4 The independence of maximally distributed random variables is symmetrical. But, as discussed in Remark 2, under a sublinear expectation, X is independent from Y does not automatically imply that Y is also independent from X. In fact, Hu and Li [7] proved that, if X is independent from Y, and Y is also independent from X, and both of X and Y have distributional uncertainty, then (X, Y) must be maximally distributed.

3 Maximally distributed random fields

In this section, we first recall the general setting of random fields defined on a nonlinear expectation space introduced by Ji and Peng [10].

Definition 5 Under a given nonlinear expectation space $(\Omega, \mathcal{H}, \hat{E})$, a collection of m-dimensional random vectors $W = (W_\gamma)_{\gamma \in \Gamma}$ is called an m-dimensional random field indexed by Γ, if for each $\gamma \in \Gamma$, $W_\gamma \in \mathcal{H}^m$.

In order to introduce the notion of finite-dimensional distribution of a random field W, we denote the family of all sets of finite indices by

$$\mathcal{J}_\Gamma := \{\underline{\gamma} = (\gamma_1, \cdots, \gamma_n) : \ \forall n \in \mathbb{N}, \ \gamma_1, \cdots, \gamma_n \in \Gamma, \ \gamma_i \neq \gamma_j \text{ if } i \neq j\}.$$

Definition 6 Let $(W_\gamma)_{\gamma\in\Gamma}$ be an m-dimensional random field defined on a nonlinear expectation space $(\Omega, \mathcal{H}, \hat{E})$. For each $\underline{\gamma} = (\gamma_1, \cdots, \gamma_n) \in \mathcal{J}_\Gamma$ and the corresponding random vector $W_{\underline{\gamma}} = (W_{\gamma_1}, \cdots, W_{\gamma_n})$, we define a functional on $C_{b.Lip}(\mathbb{R}^{n\times m})$ by

$$\mathbb{F}^W_{\underline{\gamma}}[\varphi] = \hat{E}[\varphi(W_{\underline{\gamma}})]$$

The collection $(\mathbb{F}^W_{\underline{\gamma}}[\varphi])_{\underline{\gamma}\in\mathcal{J}_\Gamma}$ is called the family of finite-dimensional distributions of $(W_\gamma)_{\gamma\in\Gamma}$.

It is clear that, for each $\underline{\gamma} \in \mathcal{J}_\Gamma$, the triple $(\mathbb{R}^{n\times m}, C_{b.Lip}(\mathbb{R}^{n\times m}), \mathbb{F}^W_{\underline{\gamma}})$ constitutes a nonlinear expectation space.

Let $(W^{(1)}_\gamma)_{\gamma\in\Gamma}$ and $(W^{(2)}_\gamma)_{\gamma\in\Gamma}$ be two m-dimensional random fields defined on nonlinear expectation spaces $(\Omega_1, \mathcal{H}_1, \hat{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \hat{E}_2)$ respectively. They are said to be identically distributed, denoted by $(W^{(1)}_\gamma)_{\gamma\in\Gamma} \overset{d}{=} (W^{(2)}_\gamma)_{\gamma\in\Gamma}$, or simply $W^{(1)} \overset{d}{=} W^{(2)}$, if for each $\underline{\gamma} = (\gamma_1, \cdots, \gamma_n) \in \mathcal{J}_\Gamma$,

$$\hat{E}_1[\varphi(W^{(1)}_{\underline{\gamma}})] = \hat{E}_2[\varphi(W^{(2)}_{\underline{\gamma}})], \ \ \forall \varphi \in C_{b.Lip}(\mathbb{R}^{n\times m}).$$

For any given m-dimensional random field $W = (W_\gamma)_{\gamma\in\Gamma}$, the family of its finite-dimensional distributions satisfies the following properties of consistency:

(1) Compatibility: For each $(\gamma_1, \cdots, \gamma_n, \gamma_{n+1}) \in \mathcal{J}_\Gamma$ and $\varphi \in C_{b.Lip}(\mathbb{R}^{n\times m})$,

$$\mathbb{F}^W_{\gamma_1,\cdots,\gamma_n}[\varphi] = \mathbb{F}^W_{\gamma_1,\cdots,\gamma_n,\gamma_{n+1}}[\widetilde{\varphi}], \tag{3}$$

where the function $\widetilde{\varphi}$ is a function on $\mathbb{R}^{(n+1)\times m}$ defined for any $y_1, \cdots, y_n, y_{n+1} \in \mathbb{R}^m$,

$$\widetilde{\varphi}(y_1, \cdots, y_n, y_{n+1}) = \varphi(y_1, \cdots, y_n);$$

(2) Symmetry: For each $(\gamma_1, \cdots, \gamma_n) \in \mathcal{J}_\Gamma$, $\varphi \in C_{b.Lip}(\mathbb{R}^{n\times m})$ and each permutation π of $\{1, \cdots, n\}$,

$$\mathbb{F}^W_{\gamma_{\pi(1)},\cdots,\gamma_{\pi(n)}}[\varphi] = \mathbb{F}^W_{\gamma_1,\cdots,\gamma_n}[\varphi_\pi] \tag{4}$$

where we denote $\varphi_\pi(y_1, \cdots, y_n) = \varphi(y_{\pi(1)}, \cdots, y_{\pi(n)})$, for $y_1, \cdots, y_n \in \mathbb{R}^m$.

The following theorem generalizes the classical Kolmogorov's existence theorem to the situation of sublinear expectation space, which is a variant of Theorem 3.8 in Peng [17]. The proof can be founded in Ji and Peng [10].

Theorem 1 *Let $\{\mathbb{F}_{\underline{\gamma}}, \underline{\gamma} \in \mathcal{J}_\Gamma\}$ be a family of finite-dimensional distributions satisfying the compatibility condition* (3) *and the symmetry condition* (4)*. Then there exists an m-dimensional random field $W = (W_\gamma)_{\gamma\in\Gamma}$ defined on a nonlinear expectation space $(\Omega, \mathcal{H}, \hat{E})$ whose family of finite-dimensional distributions coincides with $\{\mathbb{F}_{\underline{\gamma}}, \underline{\gamma} \in \mathcal{J}_\Gamma\}$. Moreover, if we assume that each $\mathbb{F}_{\underline{\gamma}}$ in $\{\mathbb{F}_{\underline{\gamma}}, \underline{\gamma} \in \mathcal{J}_\Gamma\}$ is sublinear, then the corresponding expectation $\hat{E}$ on the space of random variables $(\Omega, \mathcal{H})$ is also sublinear.*

Now we consider a new random fields under a sublinear expectation space.

Definition 7 Let $(W_\gamma)_{\gamma\in\Gamma}$ be an m-dimensional random field, indexed by Γ, defined on a sublinear expectation space $(\Omega,\mathcal{H},\hat{E})$. $(W_\gamma)_{\gamma\in\Gamma}$ is called a maximally distributed random field if for each $\underline{\gamma}=(\gamma_1,\cdots,\gamma_n)\in\mathcal{J}_\Gamma$, the following $(n\times m)$-dimensional random vector

$$\begin{aligned} W_{\underline{\gamma}} =&(W_{\gamma_1},\cdots,W_{\gamma_n})\\ =&(W_{\gamma_1}^{(1)},\cdots W_{\gamma_1}^{(m)},\cdots,W_{\gamma_n}^{(1)},\cdots,W_{\gamma_n}^{(m)}),\ \ W_{\gamma_i}^{(j)}\in\mathcal{H}, \end{aligned}$$

is maximally distributed.

For each $\underline{\gamma}=(\gamma_1,\cdots,\gamma_n)\in\mathcal{J}_\Gamma$, we define

$$g_{\underline{\gamma}}^W(p)=\hat{E}[W_{\underline{\gamma}}\cdot p],\quad p\in\mathbb{R}^{n\times m},$$

Then $(g_{\underline{\gamma}}^W)_{\underline{\gamma}\in\mathcal{J}_\Gamma}$ constitutes a family of sublinear functions:

$$g_{\underline{\gamma}}^W:\mathbb{R}^{n\times m}\mapsto\mathbb{R},\quad \underline{\gamma}=(\gamma_1,\cdots,\gamma_n),\ \gamma_i\in\Gamma,\ 1\le i\le n,\ n\in\mathbb{N},$$

which satisfies the properties of consistency in the following sense:

(1) Compatibility: For any $(\gamma_1,\cdots,\gamma_n,\gamma_{n+1})\in\mathcal{J}_\Gamma$ and $p=(p_i)_{i=1}^{n\times m}\in\mathbb{R}^{n\times m}$,

$$g^W_{\gamma_1,\cdots,\gamma_n,\gamma_{n+1}}(\bar{p})=g^W_{\gamma_1,\cdots,W_{\gamma_n}}(p), \tag{5}$$

where $\bar{p}=\begin{pmatrix}p\\0\end{pmatrix}\in\mathbb{R}^{(n+1)\times m}$;

(2) Symmetry: For any permutation π of $\{1,\cdots,n\}$,

$$g^W_{\gamma_{\pi(1)},\cdots,\gamma_{\pi(n)}}(p)=g^W_{\gamma_1,\cdots,\gamma_n}(\pi^{-1}(p)), \tag{6}$$

where $\pi^{-1}(p)=(p^{(1)},\ldots,p^{(n)})$,

$$p^{(i)}=(p_{(\pi^{-1}(i)-1)m+1},\ldots,p_{(\pi^{-1}(i)-1)m+m}),\ 1\le i\le n.$$

If the above type of family of sublinear functions $(g_\gamma)_{\underline{\gamma}\in\mathcal{J}_\Gamma}$ is given, following the construction procedure in the proof of Theorem 3.5 in Ji and Peng [10], we can construct a maximally distributed random field on sublinear expectation space.

Theorem 2 *Let* $(g_{\underline{\gamma}})_{\underline{\gamma}\in\mathcal{J}_\Gamma}$ *be a family of real-valued functions such that, for each* $\underline{\gamma}=(\gamma_1,\cdots,\gamma_n)\in\mathcal{J}_\Gamma$*, the real function* $g_{\underline{\gamma}}$ *is defined on* $\mathbb{R}^{n\times m}\mapsto\mathbb{R}$ *and satisfies the sub-linearity. Moreover, this family* $(g_{\underline{\gamma}})_{\underline{\gamma}\in\mathcal{J}_\Gamma}$ *satisfies the compatibility condition* (5) *and symmetry condition* (6)*. Then there exists an m-dimensional maximally distributed random field* $(W_\gamma)_{\gamma\in\Gamma}$ *on a sublinear expectation space* $(\Omega,\mathcal{H},\hat{E})$ *such that for each* $\underline{\gamma}=(\gamma_1,\cdots,\gamma_n)\in\mathcal{J}_\Gamma$*,* $W_{\underline{\gamma}}=(W_{\gamma_1},\cdots,W_{\gamma_n})$ *is maximally distributed with generating function*

$$g_{\underline{\gamma}}^{W}(p) = \hat{E}[W_{\underline{\gamma}} \cdot p] = g_{\underline{\gamma}}(p), \quad for\ any\ p \in \mathbb{R}^{n\times m}.$$

Furthermore, if there exists another maximally distributed random field $(\bar{W}_\gamma)_{\gamma\in\Gamma}$, with the same index set Γ, defined on a sublinear expectation space $(\bar{\Omega}, \bar{\mathcal{H}}, \bar{E})$ such that for each $\underline{\gamma} = (\gamma_1, \cdots, \gamma_n) \in \mathcal{J}_\Gamma$, $\bar{W}_{\underline{\gamma}}$ is maximally distributed with the same generating function $g_{\underline{\gamma}}$, namely,

$$\bar{E}[\bar{W}_{\underline{\gamma}} \cdot p] = g_{\underline{\gamma}}(p) \quad for\ any\ p \in \mathbb{R}^{n\times m},$$

then we have $W \overset{d}{=} \bar{W}$.

4 Maximally distributed white noise

In this section, we formulate a new type of maximally distributed white noise on $\mathbb{R}^d$.

Given sublinear expectation space $\Omega, \mathcal{H}, \hat{E}$, let $\mathbb{L}^p(\Omega)$ be the completion of $\mathcal{H}$ under the Banach norm $\|X\| := \hat{E}[|X|^p]^{\frac{1}{p}}$. For any $X, Y \in \mathbb{L}^1(\Omega)$, we say that $X = Y$ if $\hat{E}[|X - Y|] = 0$. As shown in Chapter 1 of Peng [18], $\hat{E}$ can be continuously extended to the mapping from $\mathbb{L}^1(\Omega)$ to $\mathbb{R}$ and properties (i)-(iv) of Definition 1 still hold. Moreover, $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$ also forms a sublinear expectation space, which is called the complete sublinear expectation space.

Definition 8 Let $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$ be a complete sublinear expectation space and $\Gamma = \mathcal{B}_0(\mathbb{R}^d) := \{A \in \mathcal{B}(\mathbb{R}^d), \lambda_A < \infty\}$, where λ_A denotes the Lebesgue measure of $A \in \mathcal{B}(\mathbb{R}^d)$. Let $g : \mathbb{R} \mapsto \mathbb{R}$ be a given sublinear function, i.e.,

$$g(p) = \overline{\mu}p^+ - \underline{\mu}p^-, \quad -\infty < \underline{\mu} \le \overline{\mu} < +\infty.$$

A random field $W = \{W_A\}_{A\in\Gamma}$ is called a one-dimensional maximally distributed white noise if

(i) For each $A_1, \cdots, A_n \in \Gamma$, $(W_{A_1}, \cdots, W_{A_n})$ is a $\mathbb{R}^n$-maximally distributed random vector under $\hat{E}$, and for each $A \in \Gamma$,

$$\hat{E}[W_A \cdot p] = g(p)\lambda_A, \quad p \in \mathbb{R}. \tag{7}$$

(ii) Let $A_1, A_2, \cdots, A_n$ be in Γ and mutually disjoint, then $\{W_{A_i}\}_{i=1}^n$ are independent sequence, and

$$W_{A_1\cup A_2\cup\cdots\cup A_n} = W_{A_1} + W_{A_2} + \cdots + W_{A_n}. \tag{8}$$

Remark 5 For each $A \in \Gamma$, we can restrict that W_A takes values in $[\lambda_A\underline{\mu}, \lambda_A\overline{\mu}]$. Indeed, let

$$d_A(x) := \min_{y\in[\lambda_A\underline{\mu}, \lambda_A\overline{\mu}]}\{|x - y|\},$$

by the definition of maximal distribution,

$$\hat{E}[d_A(W_A)] = \max_{v\in[\lambda_A\underline{\mu},\lambda_A\overline{\mu}]} \min_{y\in[\lambda_A\underline{\mu},\lambda_A\overline{\mu}]} \{|v-y|\} = 0,$$

which implies that $d_A(W_A) = 0$.

We can construct a spatial maximal white noise satisfying Definition 8 in the following way.

For each $\underline{\gamma} = (A_1, \cdots, A_n) \in \mathcal{J}_\Gamma$, $\Gamma = \mathcal{B}_0(\mathbb{R}^d)$, consider the mapping $g_{\underline{\gamma}}(\cdot) : \mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$g_{\underline{\gamma}}(p) := \sum_{k\in\{0,1\}^n} g(k\cdot p)\lambda_{B(k)}, \quad p \in \mathbb{R}^n, \tag{9}$$

where $k = (k_1, \cdots, k_n) \in \{0,1\}^n$, and $B(k) = \cap_{j=1}^n B_j$, with

$$B_j = \begin{cases} A_j & \text{if } k_j = 1, \\ A_j^c & \text{if } k_j = 0. \end{cases}$$

For example, given $A_1, A_2, A_3 \in \Gamma$ and $p = (p_1, p_2, p_3) \in \mathbb{R}^3$,

$$\begin{aligned} &g_{A_1,A_2,A_3}(p) = g(p_1+p_2+p_3)\lambda_{A_1\cap A_2\cap A_3} \\ &+ g(p_1+p_2)\lambda_{A_1\cap A_2\cap A_3^c} + g(p_2+p_3)\lambda_{A_1^c\cap A_2\cap A_3} + g(p_1+p_3)\lambda_{A_1\cap A_2^c\cap A_3} \\ &+ g(p_1)\lambda_{A_1\cap A_2^c\cap A_3^c} + g(p_2)\lambda_{A_1^c\cap A_2\cap A_3^c} + g(p_3)\lambda_{A_1^c\cap A_2^c\cap A_3}. \end{aligned}$$

Obviously, for each $\underline{\gamma} = (A_1, \cdots, A_n) \subset \Gamma$, $g_{\underline{\gamma}}(\cdot)$ defined by (9) is a sublinear function defined on $\mathbb{R}^n$ due to the sub-linearity of function $g(\cdot)$. The following property shows that the consistency conditions (5) and (6) also hold for $\{g_{\underline{\gamma}}\}_{\gamma\in\mathcal{J}_\Gamma}$.

Proposition 4 *The family* $\{g_{\underline{\gamma}}\}_{\gamma\in\mathcal{J}_\Gamma}$ *defined by (9) satisfies the consistency conditions (5) and (6).*

Proof For compatibility (5), given $A_1, \cdots, A_n, A_{n+1} \in \Gamma$ and $\bar{p}^T = (p^T, 0) \in \mathbb{R}^{n+1}$, we have

$$\begin{aligned} g_{A_1,\cdots,A_{n+1}}(\bar{p}) &= \sum_{k\in\{0,1\}^{n+1}} g(k\cdot\bar{p})\lambda_{B(k)} \\ &= \sum_{k'\in\{0,1\}^n} g(k'\cdot p)(\lambda_{B(k')\cap A_{n+1}} + \lambda_{B(k')\cap A_{n+1}^c}) \\ &= \sum_{k'\in\{0,1\}^n} g(k'\cdot p)\lambda_{B(k')} = g_{A_1,\cdots,A_n}(p). \end{aligned}$$

The symmetry (6) can be easily verified since the operators $k\cdot p$ and $B(k) = \cap_{j=1}^n B_j$ are also symmetry. □

Now we present the existence of the maximally distributed white noises under the sublinear expectation.

Theorem 3 *For each given sublinear function*

$$g(p) = \max_{\mu \in [\underline{\mu}, \overline{\mu}]} (\mu \cdot p) = \overline{\mu} p^+ - \underline{\mu} p^-, \ p \in \mathbb{R},$$

there exists a one-dimensional maximally distributed random field $(W_\gamma)_{\gamma \in \Gamma}$ on a sublinear expectation space $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$ such that, for each $\underline{\gamma} = (A_1, \cdots, A_n) \in \mathcal{J}_\Gamma$, $W_{\underline{\gamma}} = (W_{A_1}, \cdots, W_{A_n})$ is maximally distributed.

Furthermore, $(W_\gamma)_{\gamma \in \Gamma}$ is a spatial maximally distributed white noise under $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$, namely, conditions (i) and (ii) of Definition 8 are satisfied.

If $(\bar{W}_\gamma)_{\gamma \in \Gamma}$ is another maximally distributed white noise with the same sublinear function g in (9)*, then $\bar{W} \overset{d}{=} W$.*

Proof Thanks to Proposition 4 and Theorem 2, the existence and uniqueness of the maximally distributed random field W in a sublinear expectation space $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$ with the family of generating functions defined by (9) hold. We only need to verify that the maximally distributed random field W satisfies conditions (i) and (ii) of Definition 8.

For each $A \in \Gamma$, $\hat{E}[W_A \cdot p] = g(p)\lambda_A$ by Theorem 2 and (9), thus (i) of Definition 8 holds.

We note that if $\{A_i\}_{i=1}^n$ are mutually disjoint, then for $p = (p_1, \cdots, p_n) \in \mathbb{R}^n$, by (9), we have

$$\hat{E}[p_1 W_{A_1} + \cdots + p_n W_{A_n}] = g(p_1)\lambda_{A_1} + \cdots + g(p_n)\lambda_{A_n},$$

thus the independence of $\{W_{A_i}\}_{i=1}^n$ can be implied by Proposition 3.

In order to prove (8), we only consider the case of two disjoint sets. Suppose that

$$A_1 \cap A_2 = \emptyset, \quad A_3 = A_1 \cup A_2,$$

an easy computation of (9) shows that

$$\begin{aligned} g_{A_1, A_2, A_3}(p) =& g(p_1 + p_3)\lambda_{A_1} + g(p_2 + p_3)\lambda_{A_2} \\ =& \max_{v_1 \in [\underline{\mu}\lambda_{A_1}, \overline{\mu}\lambda_{A_1}]} \ \max_{v_2 \in [\underline{\mu}\lambda_{A_2}, \overline{\mu}\lambda_{A_2}]} \ \max_{v_3 = v_1 + v_2} (p_1 \cdot v_1 + p_2 \cdot v_2 + p_3 \cdot v_3). \end{aligned}$$

Thus, for each $\varphi \in C(\mathbb{R}^3)$,

$$\hat{E}[\varphi(W_{A_1}, W_{A_2}, W_{A_3})] = \max_{v_1 \in [\underline{\mu}\lambda_{A_1}, \overline{\mu}\lambda_{A_1}]} \ \max_{v_2 \in [\underline{\mu}\lambda_{A_2}, \overline{\mu}\lambda_{A_2}]} \ \max_{v_3 = v_1 + v_2} \varphi(v_1, v_2, v_3).$$

In particular, we set $\varphi(v_1, v_2, v_3) = |v_1 + v_2 - v_3|$, it follows that

$$\hat{E}[|W_{A_1} + W_{A_2} - W_{A_1 \cup A_2}|] = 0.$$

which implies that

$$W_{A_1 \cup A_2} = W_{A_1} + W_{A_2}.$$

Finally, (ii) of Definition 8 holds. □

Remark 6 The finite-dimensional distribution of maximally distributed whiten noise can be uniquely determined by two parameters $\overline{\mu}$ and $\underline{\mu}$, which can be simply calculated by taking the maximum of the continuous function over the domain determined by $\overline{\mu}$ and $\underline{\mu}$.

Similar to the invariant property of G-Gaussian white noise introduced in Ji and Peng [10], it also holds for maximally distributed white noise due to the well-known invariance of the Lebesgue measure under rotation and translation.

Proposition 5 *For each $p \in \mathbb{R}^d$ and $O \in \mathbb{O}(d) := \{O \in \mathbb{R}^{d\times d} : O^T = O^{-1}\}$, we set*

$$T_{p,O}(A) = O \cdot A + p, \quad A \in \Gamma.$$

Then, for each $A_1, \cdots, A_n \in \Gamma$,

$$(W_{A_1}, \cdots, W_{A_n}) \overset{d}{=} (W_{T_{p,O}(A_1)}, \cdots, W_{T_{p,O}(A_n)}).$$

5 Spatial and temporal maximally distributed white noise and related stochastic integral

In Ji and Peng [10], we see that a spatial G-white noise is essentially different from the temporal case or the temporal-spatial case, since there is no independence property for the spatial G-white noise. But for the maximally distributed white noise, spatial or temporal-spatial maximally distributed white noise has the independence property due to the symmetrical independence for maximal distribution.

Combining symmetrical independence and boundedness properties of maximal distribution, the integrand random fields can be largely extended when we consider the stochastic integral with respect to spatial maximally distributed white noise. For stochastic integral with respect to temporal-spatial case, the integrand random fields can even contain the "non-adapted" situation.

5.1 Stochastic integral with respect to the spatial maximally distributed white noise

We firstly define the stochastic integral with respect to the spatial maximally distributed white noise in a quite direct way.

Let $\{W_\gamma\}_{\gamma\in\Gamma}$, $\Gamma = \mathcal{B}_0(\mathbb{R}^d)$, be a one-dimensional maximally distributed white noise defined on a complete sublinear expectation space $(\Omega, \mathbb{L}^1(\Omega), \hat{E})$, with $g(p) =$

$\overline{\mu}p^+ - \underline{\mu}p^-$, $-\infty < \underline{\mu} \le \overline{\mu} < \infty$. We introduce the following type of random fields, called simple random fields.

Given $p \ge 1$, set

$$M_g^{p,0}(\Omega) = \{\eta(x,\omega) = \sum_{i=1}^{n} \xi_i(\omega) 1_{A_i}(x),\ A_1, \cdots, A_n \in \Gamma \text{ are mutually disjoint}$$
$$i = 1, 2, \ldots, n,\ \ \xi_1, \cdots, \xi_n \in \mathbb{L}^p(\Omega), \quad n = 1, 2, \cdots, \}.$$

For each simple random fields $\eta \in M_g^{p,0}(\Omega)$ of the form

$$\eta(x,\omega) = \sum_{i=1}^{n} \xi_i(\omega) \mathbf{1}_{A_i}(x), \tag{10}$$

the related Bohner's integral for η with respect to the Lebesgue measure λ is

$$I_B(\eta) = \int_{\mathbb{R}^d} \eta(x,\omega)\lambda(dx) := \sum_{i=1}^{n} \xi_i(\omega)\lambda_{A_i}.$$

It is immediate that $I_B(\eta) : M_g^{p,0}(\Omega) \mapsto \mathbb{L}^p(\Omega)$ is a linear and continuous mapping under the norm for η, defined by,

$$\|\eta\|_{M^p} = \hat{E}[\int_{\mathbb{R}^d} |\eta(x,\omega)|^p \lambda(dx)]^{\frac{1}{p}}.$$

The completion of $M_g^{p,0}(\Omega)$ under this norm is denoted by $M_g^p(\Omega)$ which is a Banach space. The unique extension of the mapping I_B is denoted by

$$\int_{\mathbb{R}^d} \eta(x,\omega)\lambda(dx) := I_B(\eta),\ \ \eta \in M_g^p(\Omega).$$

Now for a simple random field $\eta \in M_g^{p,0}(\Omega)$ of form (10), we define its stochastic integral with respect to W as

$$I_W(\eta) := \int_{\mathbb{R}^d} \eta(x,\omega)W(dx) = \sum_{i=1}^{n} \xi_i(\omega)W_{A_i}.$$

With this formulation, we have the following estimation.

Lemma 1 *For each $\eta \in M_g^{1,0}(\Omega)$ of form (10), we have*

$$\hat{E}\left[\left|\int_{\mathbb{R}^d} \eta(x,\omega)W(dx)\right|\right] \le \kappa \hat{E}\left[\int_{\mathbb{R}^d} |\eta(x,\omega)|\lambda(dx)\right] \tag{11}$$

where $\kappa = \max\{|\underline{\mu}|, |\overline{\mu}|\}$.

Proof We have

$$\hat{E}[|\int_{\mathbb{R}^d}\eta(x,\omega)W(dx)|] = \hat{E}[|\sum_{i=1}^{N}\xi_i(\omega)W_{A_i}|] \leq \hat{E}[\sum_{i=1}^{N}|\xi_i(\omega)|\cdot|W_{A_i}|]$$
$$\leq \kappa\hat{E}[\sum_{i=1}^{N}|\xi_i(\omega)|\cdot\lambda_{A_i}] = \kappa\hat{E}[\|\eta\|_{M_g^1(\Omega)}].$$

The last inequality is due to the boundedness of maximal distribution (see Remark 5). □

This lemma shows that $I_W : M_g^{1,0}(\Omega) \mapsto \mathbb{L}^1(\Omega)$ is a linear continuous mapping. Consequently, I_W can be uniquely extended to the whole domain $M_g^1(\Omega)$. We still denote this extended mapping by

$$\int_{\mathbb{R}^d}\eta W(dx) := I_W(\eta).$$

Remark 7 Different from the stochastic integrals with respect to G-white noise in Ji and Peng [10] which is only defined for the deterministic integrand, here the integrand can be a random field.

5.2 Maximally distributed random fields of temporal-spatial types and related stochastic integral

It is well-known that the framework of the classical white noise defined in a probability space $(\Omega,\mathcal{F},P)$ with 1-dimensional temporal and d-dimensional spatial parameters is in fact a $\mathbb{R}^{1+d}$-indexed space type white noise. But Peng [17] and then Ji and Peng [10] observed a new phenomenon: Unlike the classical Gaussian white noise, the d-dimensional space-indexed G-white noise cannot have the property of incremental independence, thus spatial G-white noise is essentially different from temporal-spatial or temporal one. Things will become much direct for the case of maximally distributed white noise due to the incremental independence property of maximal distributions. This means that a time-space maximally distributed $(1+d)$-white noise is essentially a $(1+d)$-spatial white noise. The corresponding stochastic integral is also the same. But in order to make clear the dynamic properties, we still provide the description of the temporal-spatial white-noise on the time-space framework:

$$\mathbb{R}^+\times\mathbb{R}^d = \{(t,x_1,\ldots,x_d)\in\mathbb{R}^+\times\mathbb{R}^d\},$$

where the index $t\in[0,\infty)$ is specially preserved to be the index for time.

Let $\Gamma = \{A\in\mathcal{B}(\mathbb{R}^+\times\mathbb{R}^d),\lambda_A<\infty\}$, the maximally distributed white noise $\{W_A\}_{A\in\Gamma}$ is just like in the spatial case with dimension $1+d$.

More precisely, let

$$\Omega = \{\omega \in \mathbb{R}^{\Gamma} : \ \omega(A \cup B) = \omega(A) + \omega(B),\\ \forall A, B \in \Gamma, \quad A \cup B = \emptyset\},$$

and $W = (W_\gamma(\omega) = \omega_\gamma)_{\gamma \in \Gamma}$ the canonical random field.

For $T > 0$, denote the temporal-spatial sets before time T by

$$\Gamma_T := \{A \in \Gamma : (s, x) \in A \Rightarrow 0 \leq s < T\}.$$

Set $\mathcal{F}_T = \sigma\{W_A, A \in \Gamma_T\}$, $\mathcal{F} = \bigvee_{T \geq 0} \mathcal{F}_T$, and

$$L_{ip}(\Omega_T) = \{\varphi(W_{A_1}, \ldots, W_{A_n}), \ \forall n \in \mathbb{N},\\ A_i \in \Gamma_T, i = 1, \ldots, n, \varphi \in C_{b.Lip}(\mathbb{R}^n)\}.$$

We denote

$$L_{ip}(\Omega) = \bigcup_{n=1}^{\infty} L_{ip}(\Omega_n).$$

For each $X \in L_{ip}(\Omega)$, without loss of generality, we assume X has the form

$$X = \varphi(W_{A_{11}}, \cdots, W_{A_{1m}}, \cdots, W_{A_{n1}}, \cdots, W_{A_{nm}}),$$

where $A_{ij} = [t_{i-1}, t_i) \times A_j$, $1 \leq i \leq n, 1 \leq j \leq m$, $0 = t_0 < t_1 < \cdots < t_n < \infty$, $\{A_1, \cdots, A_m\} \subset \mathcal{B}_0(\mathbb{R}^d)$ are mutually disjoint and $\varphi \in C_{b.Lip}(\mathbb{R}^{n \times m})$. Then the corresponding sublinear expectation for X can be defined by

$$\begin{aligned} \hat{E}[X] &= \hat{E}[\varphi(W_{A_{11}}, \cdots, W_{A_{1m}}, \cdots, W_{A_{n1}}, \cdots, W_{A_{nm}}) \\ &= \max_{v_{ij} \in [\underline{\mu}, \overline{\mu}]} \varphi(\lambda_{A_{11}} v_{11}, \cdots, \lambda_{A_{1m}} v_{1m}, \cdots, \lambda_{A_{n1}} v_{v_{n1}}, \cdots, \lambda_{A_{nm}} v_{nm}), \\ &\qquad 1 \leq i \leq m, 1 \leq j \leq n \end{aligned}$$

and the related conditional expectation of X under $\mathcal{F}_t$, where $t_j \leq t < t_{j+1}$, denoted by $\hat{E}[X|\mathcal{F}_t]$, is defined by

$$\hat{E}[\varphi(W_{A_{11}}, \cdots, W_{A_{1m}}, \cdots, W_{A_{n1}}, \cdots, W_{A_{nm}})|\mathcal{F}_t]\\ = \psi(W_{A_{11}}, \cdots, W_{A_{1m}}, \cdots, W_{A_{j1}}, \cdots, W_{A_{jm}}),$$

where

$$\psi(x_{11}, \cdots, x_{1m}, \cdots, x_{j1}, \cdots, x_{jm}) = \hat{E}[\varphi(x_{11}, \cdots, x_{1m}, \cdots, x_{j1}, \cdots, x_{jm}, \tilde{W})].$$

Here

$$\tilde{W} = (W_{A_{(j+1)1}}, \cdots, W_{A_{(j+1)m}}, \cdots, W_{A_{n1}}, \cdots, W_{A_{nm}}).$$

It is easy to verify that $\hat{E}[\cdot]$ defines a sublinear expectation on $L_{ip}(\Omega)$ and the canonical process $(W_\gamma)_{\gamma\in\Gamma}$ is a one-dimensional temporal-spatial maximally distributed white noise on $(\Omega, L_{ip}(\Omega), \hat{E})$.

For each $p \geq 1$, $T \geq 0$, we denote by $L_g^p(\Omega_T)$(resp., $L_g^p(\Omega)$) the completion of $L_{ip}(\Omega_T)$(resp., $L_{ip}(\Omega)$) under the norm $\|X\|_p := (\hat{E}[|X|^p])^{1/p}$. The conditional expectation $\hat{E}[\cdot\,|\mathcal{F}_t] : L_{ip}(\Omega) \to L_{ip}(\Omega_t)$ is a continuous mapping under $\|\cdot\|_p$ and can be extended continuously to the mapping $L_g^p(\Omega) \to L_g^p(\Omega_t)$ by

$$|\hat{E}[X\,|\mathcal{F}_t] - \hat{E}[Y\,|\mathcal{F}_t]\,| \leq \hat{E}[|X-Y|\,|\mathcal{F}_t] \quad \text{for } X, Y \in L_{ip}(\Omega).$$

It is easy to verify that the conditional expectation $\hat{E}[\cdot|\mathcal{F}_t]$ satisfies the following properties, and the proof is very similar to the corresponding one of Proposition 5.3 in Ji and Peng [10].

Proposition 6 *For each $t \geq 0$, the conditional expectation $\hat{E}[\cdot\,|\mathcal{F}_t] : L_g^p(\Omega) \to L_g^p(\Omega_t)$ satisfies the following properties: for any $X, Y \in L_g^p(\Omega)$, $\eta \in L_g^p(\Omega_t)$,*

(i) $\hat{E}[X\,|\mathcal{F}_t] \geq \hat{E}[Y\,|\mathcal{F}_t]$ *for* $X \geq Y$.
(ii) $\hat{E}[\eta\,|\mathcal{F}_t] = \eta$.
(iii) $\hat{E}[X+Y\,|\mathcal{F}_t] \leq \hat{E}[X\,|\mathcal{F}_t] + \hat{E}[Y\,|\mathcal{F}_t]$.
(iv) $\hat{E}[\eta X\,|\mathcal{F}_t] = \eta^+\hat{E}[X\,|\mathcal{F}_t] + \eta^-\hat{E}[-X\,|\mathcal{F}_t]$ *if η is bounded.*
(v) $\hat{E}[\hat{E}[X\,|\mathcal{F}_t]\,|\mathcal{F}_s] = \hat{E}[X|\mathcal{F}_{t\wedge s}]$ *for* $s \geq 0$.

Now we define the stochastic integral with respect to the spatial-temporal maximally distributed white noise W, which is similar to the spatial situation.

For each given $p \geq 1$, let $M^{p,0}(\Omega_T)$ be the collection of simple processes with the form:

$$f(s,x;\omega) = \sum_{i=0}^{n-1}\sum_{j=1}^{m} X_{ij}(\omega)\mathbf{1}_{A_j}(x)\mathbf{1}_{[t_i,t_{i+1})}(s), \tag{12}$$

where $X_{ij} \in L_g^p(\Omega_T)$, $i = 0, \cdots, n-1$, $j = 1, \cdots, m$, $0 = t_0 < t_1 < \cdots < t_n = T$, and $\{A_j\}_{j=1}^m \subset \Gamma$ is mutually disjoint.

Remark 8 Since we only require $X_{ij} \in L_g^p(\Omega_T)$, the integrand may "non-adapted". This issue is essentially different from the requirement of adaptability in the definition of stochastic integral with respect to temporal-spatial G-white noise in Ji and Peng [10].

The completion of $M^{p,0}(\Omega_T)$ under the norm $\|\cdot\|_{M^p}$, denoted by $M_g^p(\Omega_T)$, is a Banach space, where the Banach norm $\|\cdot\|_{M^p}$ is defined by

$$\begin{aligned}\|f\|_{M^p} :=& \left(\hat{E}\left[\int_0^T\int_{\mathbb{R}^d}|f(s,x)|^p ds\lambda(dx)\right]\right)^{\frac{1}{p}}\\ =& \left\{\hat{E}\left[\sum_{i=0}^{n-1}\sum_{j=1}^{m}|X_{ij}|^p(t_{i+1}-t_i)\lambda_{A_j}\right]\right\}^{\frac{1}{p}}.\end{aligned}$$

For $f \in M^{p,0}(\Omega_T)$ with the form as (12), the related stochastic integral with respect to the temporal-spatial maximally distributed white noise W can be defined as follows:

$$I_W(f) = \int_0^T \int_{\mathbb{R}^d} f(s,x)W(ds,dx) := \sum_{i=0}^{n-1} \sum_{j=1}^{m} X_{ij} W([t_j, t_{j+1}) \times A_j). \tag{13}$$

Similar to Lemma 1, we have

Lemma 2 *For each* $f \in M^{1,0}([0,T] \times \mathbb{R}^d)$,

$$\hat{E}\left[\left|\int_0^T \int_{\mathbb{R}^d} f(s,x)W(ds,dx)\right|\right] \le \kappa \hat{E}\left[\int_0^T \int_{\mathbb{R}^d} |f(s,x)| ds dx\right], \tag{14}$$

where $\kappa = \max\{|\overline{\mu}|, |\underline{\mu}|\}$.

Thus $I_W : M^{1,0}(\Omega_T) \mapsto L_g^1(\Omega_T)$ is a continuous linear mapping. Consequently, I_W can be uniquely extend to the domain $M_g^1(\Omega_T)$. We still denote this mapping by

$$\int_0^T \int_{\mathbb{R}^d} f(s,x)W(ds,dx) := I_W(f) \;\text{ for } f \in M_g^1(\Omega_T).$$

Remark 9 Thanks to the boundedness of maximally distributed white noise, the domain of integrand $M_g^1(\Omega_T)$ is much larger since the usual requirement of adaptability for integrand can be dropped.

It is easy to check that the stochastic integral has the following properties.

Proposition 7 *For each* $f, g \in M_g^1(\Omega_T)$, $0 \le s \le r \le t \le T$,
(i) $\int_s^t \int_{\mathbb{R}^d} f(u,x)W(du,dx) = \int_s^r \int_{\mathbb{R}^d} f(u,x)W(du,dx) + \int_r^t \int_{\mathbb{R}^d} f(u,x)W(du,dx)$.
(ii) $\int_s^t \int_{\mathbb{R}^d} (\alpha f(u,x) + g(u,x))W(du,dx)$
$= \alpha \int_s^t \int_{\mathbb{R}^d} f(u,x)W(du,dx) + \int_s^t \int_{\mathbb{R}^d} g(u,x)W(du,dx)$, *where* $\alpha \in L_g^1(\Omega_T)$ *is bounded.*

Remark 10 In particular, if we only consider temporal maximally distributed white noise and further assume that $\underline{\mu} \ge 0$. In this case, the index set $\Gamma = \{[s,t) : 0 \le s < t < \infty\}$. The canonical process $W([0,t))$ is the quadratic variation process of G-Brownian motion, more details about the quadratic variation process can be found in Peng [18].

Acknowledgements This research is partially supported by NSF of China (No.L1624032, No.11526205, No.11601281), 111 Project of Chinese SAFEA (No.B12023), National Key R&D Program of China (No.2018YFA0703900).

References

1. Artzner, P., Delbaen, F., Eber, J-M., Heath, D.: Coherent measures of risk. Math. Financ. **9**, 203–228 (1999)
2. Dalang, R. C.: Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. Electron. J. Probab. **4**, 1–29 (1999)
3. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (1992)
4. Delbaen, F.: Coherent Risk Measures. Edizioni della Normale, Pisa (2000)
5. Denis, L., Hu, M., Peng, S.: Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. Potential Anal. **34**, 139–161 (2011)
6. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Financ. Stoch. **6**, 429-447 (2002)
7. Hu, M., Li, X.: Independence under the G-expectation framework. J. Theor. Probab. **27**, 1011-1020 (2014)
8. Hu, M., Peng, S.: On representation theorem of G-expectations and paths of G-Brownian motion. Acta. Math. Sin.-English Ser. **25**, 539–546 (2009)
9. Huber, P. J.: Robust Statistics. John Wiley & Sons, New York (1981)
10. Ji, X., Peng, S.: Spatial and temporal white noises under sublinear G-expectation. Sci. China-Math. **63**(1), 61–82 (2020)
11. Li, X., Zong, G.: On the necessary and sufficient conditions for Peng's law of large numbers under sublinear expectations, arXiv:2106.00902 (2021)
12. Peng, S.: Backward SDE and related g-expectations. In: El Karoui, N., Mazliak L. (eds.) Backward Stochastic Differential Equations, pp. 141–159. Longman, Harlow (1997)
13. Peng, S.: G-expectation, G-Brownian motion and related stochastic calculus of Itô's type. In: Benth, F. E. et al. (eds.) Stochastic Analysis and Applications, pp. 541–567, Springer, Berlin (2007)
14. Peng, S.: Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. Stoch. Process. Their Appl. **118**, 2223–2253 (2008)
15. Peng, S.: A new central limit theorem under sublinear expectations. arXiv:0803.2656 (2008)
16. Peng, S.: Tightness, weak compactness of nonlinear expectations and application to CLT. arXiv:1006.2541 (2010)
17. Peng, S.: G-Gaussian processes under sublinear expectations and q-Brownian motion in quantum mechanics. arXiv:1105.1055 (2011)
18. Peng, S.: Nonlinear Expectations and Stochastic Calculus under Uncertainty. Springer, Berlin (2019)
19. Peng, S.: Law of large numbers and central limit theorem under nonlinear expectations. Probab. Uncertaint. Quant. Risk **4** (4), 1–8 (2019)
20. Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)
21. Walsh, J. B.: An Introduction to Stochastic Partial Differential Equations. In: Hennequin, P. L. (ed.) École d'été de Probabilités de Saint-Flour, XIV–1984, pp.265–439, Springer, Berlin (1986)

Pairs Trading under Geometric Brownian Motion Models

Phong Luu, Jingzhi Tie, and Qing Zhang

Abstract This survey paper is concerned with pairs trading strategies under geometric Brownian motion models. Pairs trading is about trading simultaneously a pair of securities, typically stocks. The idea is to monitor the spread of their price movements over time. A pairs trade is triggered by their price divergence (e.g., one stock moves up a significant amount relative to the other) and consists of a short position in the strong stock and a long position in the weak one. Such a strategy bets on the reversal of their price strengths and the eventual convergence of the price spread. Pairs trading is popular among trading institutions because its risk neutral nature. In practice, the trader needs to decide when to initiate a pairs position (how much divergence is enough) and when to close the position (how to take profits or cut losses). It is the main goals of this paper to address these issues and theoretical findings along with related practical considerations.

Key words: pairs trading, optimal trading strategy, geometric Brownian motions

Phong Luu
Department of Mathematics, University of North Georgia, Oakwood, GA 30566,
e-mail: phong.luu@ung.edu

Jingzhi Tie
Department of Mathematics, University of Georgia, Athens, GA 30602,
e-mail: jtie@uga.edu

Qing Zhang
Department of Mathematics, University of Georgia, Athens, GA 30602,
e-mail: qz@uga.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_15

1 Introduction

This paper is about strategies for simultaneously trading a pair of stocks. The idea is to track the price movements of these two securities over some period of time and compare their relative price strength. A pairs trade is triggered when their prices diverge, e.g., one stock moves up substantially relative to the other. A pairs trade is entered and consists of a short position in the stronger stock and a long position in the weaker one. Such a strategy bets on the reversal of their price strength and eventual convergence of their price spread.

Pairs trading was introduced by Garry Bamberger and followed by Nunzio Tartaglia's quantitative group at Morgan Stanley in the 1980s. Tartaglia's group used advanced statistical tools and developed high tech trading systems by incorporating trader's intuition and disciplined filtering rules. They were able to identify pairs of stocks and trade them with a great success. See Gatev et al. [7] for related background details. In addition, there are studies addressing why pairs trading works. For related in-depth discussions in connection with the cause of the price divergence and subsequent convergence, we refer the reader to the books by Vidyamurthy [21] and Whistler [22].

Empirical studies and related considerations can be found in papers by Do and Faff [4, 5], Gatev et al. [7], and books by Vidyamurthy [21] and Whistler [22]. Issues involved in these works include statistical characterization of the spread process, performance of pairs trading with various trading thresholds, and the impact of trading costs in connection with pairs trading.

A major advantage of pairs trading is its 'market neutral' nature in the sense that it helps to hedge market risks. For example, if the market crashes and takes both stocks with it, the trade would result in a gain on the short side and a loss on the long side of the position. The gain and loss cancel out each other and to some extent, reduce the market risk.

In pairs trading, a crucial step is to determine when to initiate a pairs trade (i.e., how much spread divergence is sufficient to trigger a trade) and when to close the position (when to lock in profits). Following empirical developments documented in Gatev et al. [7], increasing efforts were made addressing theoretical aspects of pairs trading. The main focus was devoted to development of mathematical models that capture the spread movements, filtering techniques, optimal entry and exit timings, money management and risk control. For example, in Elliott et al. [6], the price spread is assumed to be a mean reversion process with additive noise. Several filtering techniques were explored to identify entry points. One exit rule with a fixed holding period was discussed in detail. In Deshpande and Barmish [3], a general (mean-reversion based) framework was developed. Using a 'spread' function, they were able to determine the numbers of shares of each stock every moment and how to adjust them over time. They showed that such an algorithm leads to positive expected returns.

Some recent efforts on pairs trading have been devoted to in-depth analysis based on mean reversion models. For example, Kuo et al. [11] considered an optimal selling rule. The objective is to determine the time of closing an existing pairs position in order to maximize an expected return or to cut losses short. In particular, given a fixed cut-loss level, the optimal target level can be determined under a mean reversion model. Further results on mean reversion models can be found in Song and Zhang [18]. They have developed a complete system with both entry and exit signals. They have shown that the optimal trading rule can be determined by threshold levels. The calculation of these levels only involves algebraic equations.

We would like to point out that almost all literature on pairs trading is mean reversion based one way or the other. On the one hand, this makes the trading more intuitive. On the other, such constraint adds a severe limitation on its potential applications. In order to meet the mean-reversion requirement, tradable pairs are typically selected among stocks from the same industrial sector. From a practical viewpoint, it is highly desirable to have a broad range of stock selections for pairs trading. Mathematically speaking, this amounts to the possibility of treating pairs trading under models other than mean reversion. In Tie et al. [19], they have developed a new method to treat the pairs-trading problem under general geometric Brownian motions. In particular, under a two-dimensional geometric Brownian motion model, they were able to fully characterize the optimal policy in terms of two threshold lines obtained by solving the associated variational inequalities. The principal idea of pairs trading is that one builds the position of a pair when the cost is low and closes the position when the pairs' value is high. These two threshold switching lines quantify exactly how low is low and how high is high. These policies are easy to compute and implement. The most striking feature of these results is the simplicity of the solution: Clean-cut assumptions and closed-form trading policies.

One important consideration in trading has yet received deserved attention: How to trade with cutting losses. There are many scenarios when cutting losses may arise. A typical one is a margin call. This is often proceeded with heavy losses leading to an enforced closure of part or the entire pairs position. Often in practice, a pairs trader chooses a pre-determined stop-loss level due to a money management consideration. From a modeling point of view, the prices of the pairs may cease to behave as the model prescribes due to undesirable events such as acquisition (or bankruptcy) of one stock in the pairs position. It is necessary to modify the trading rule accordingly in order to accommodate a pre-determined stop loss level. On the other hand, from a control theoretical viewpoint, forcing a stop loss amounts to imposing a hard state constraint. This often poses substantial challenges when solving the problem. Such issues were addressed in Liu et al. [13] recently. They were able to establish regions in terms of threshold lines to characterize trading rules. They also obtained sufficient conditions that guarantee the optimality of these trading rules.

In this paper, we mainly involve stocks. Nevertheless, the idea of pairs trading is not limited to stock trading. For example, the optimal timing of investments in irreversible projects can also be considered as a pairs-trading problem. Back in 1986, McDonald and Siegel [15] considered optimal timing of investment in an irreversible project. Two factors are included in their model: The value of the project and its cost. Greater project value growth potential and lesser future project cost will postpone the transaction. See also Hu and Øksendal [9] for more rigorous mathematical treatment. In terms of pairs trading, their results are about a pairs trading selling rule. Extension along this line can be found in Tie and Zhang [20]. They treated the pairs selling rule under a regime-switching model. They were also able to show threshold type selling policies.

The problem under consideration is closely related to traditional portfolio selection problems. Following Merton's work in the late 60's, much progress along this direction has been made. A thorough treatment of the problem can be found in Davis and Norman [2] in which they studied Merton's investment/consumption problem with the transaction costs and established wedge-shaped regions for the pair of bank and stock holdings. To some extent, pairs trading resembles portfolio selection. Rather than balancing between bank and stock holdings, pairs trading involves positions consisting of two stocks. In portfolio selection, risk control is achieved through adjusting proportion of stock holdings; while, in pairs trading, the risk is limited by focusing on highly correlated stocks that are traded in opposite directions. Early theoretical development along portfolio selection with transaction costs using viscosity solutions can be found in Zariphopoulou [23]. Further in-depth studies and a complete solution to investment and consumption problem with transaction costs can be found in Shreve and Soner [17].

Mathematical trading rules have been studied for many years. In addition to the work by Hu and Øksendal [9] and Song and Zhang [18], Zhang [25] considered a selling rule determined by two threshold levels, a target price and a stop-loss limit. In [25], such optimal threshold levels are obtained by solving a set of two-point boundary value problems. Guo and Zhang [8] studied the optimal selling rule under a model with switching Geometric Brownian motion. Using a smooth-fit technique, they obtained the optimal threshold levels by solving a set of algebraic equations. These papers are concerned with the selling side of trading in which the underlying price models are of GBM type. Some subsequent efforts were devoted to strategies on complete trading systems including buying and selling decision making. For example, Dai et al. [1] developed a trend-following rule based on a conditional probability indicator. They showed that the optimal trading rule can be determined by two threshold curves which can be obtained by solving the associated Hamilton-Jacobi-Bellman (HJB) equations. A similar idea was developed following a confidence interval approach by Iwarere and Barmish [10]. In addition, Merhi and Zervos [16] studied an investment capacity expansion/reduction problem following a dynamic programming approach under a geometric Brownian motion market model. In connection with mean reversion trading, Zhang and Zhang [24] obtained a buy-low and sell-high policy by characterizing the 'low' and 'high' levels in terms of the mean reversion parameters.

In this paper, we focus on the mathematical aspects of pairs trading. We present key ideas used in derivation of solutions to the associated HJB equations and summarize the main results. In §2, we consider pairs trading under geometric Brownian motions. It can be seen that pairs trading ideas are more general and they do not have to be cast under a mean reversion framework. In §3, we address pairs trading with a stop-loss constraint. We establish threshold type trading policies and provide sufficient conditions that guarantee the optimality of these policies. In §4, we consider a two-dimensional geometric Brownian model with regime-switching. We focus on related optimal pairs selling rules. Proofs of these results are omitted and can be found in [13, 19, 20]. Finally, some concluding remarks are given in §5.

2 Pairs Trading under a GBM

In this section, we consider pairs trading under a two-dimensional geometric Brownian motion model. A share of pairs position $\mathbf{Z}$ consists of one share long position in stocks $\mathbf{X}^1$ and one share short position in $\mathbf{X}^2$. Let (X_t^1, X_t^2) denote their prices at t satisfying the following stochastic differential equation:

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} X_t^1 & \\ & X_t^2 \end{pmatrix} \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} \right], \tag{1}$$

where μ_i, $i = 1,2$, are the return rates, σ_{ij}, $i,j = 1,2$, the volatility constants, and (W_t^1, W_t^2) a 2-dimensional standard Brownian motion.

We consider the case that the net position at any time can be either long (with one share of $\mathbf{Z}$) or flat (no stock position of either $\mathbf{X}^1$ or $\mathbf{X}^2$). Let $i = 0,1$ denote the initial net position and let $\tau_0 \le \tau_1 \le \tau_2 \le \cdots$ denote a sequence of buying and selling (stopping) times. If initially the net position is long ($i = 1$), then one should sell $\mathbf{Z}$ before acquiring any shares in the future. That is, to first sell the pair at τ_0, then buy at τ_1, sell at τ_2, buy at τ_3, and so on. The corresponding trading sequence is denoted by $\Lambda_1 = (\tau_0, \tau_1, \tau_2, \ldots)$. Likewise, if initially the net position is flat ($i = 0$), then one should start to buy a share of $\mathbf{Z}$. That is, to first buy at τ_1, sell at τ_2, then buy at τ_3, an so forth. The corresponding sequence of stopping times is denoted by $\Lambda_0 = (\tau_1, \tau_2, \ldots)$.

Let K denote the fixed percentage of transaction costs associated with buying or selling of stocks $\mathbf{X}^i$, $i = 1,2$. For example, the cost to establish the pairs position $\mathbf{Z}$ at $t = t_1$ is $(1+K)X_{t_1}^1 - (1-K)X_{t_1}^2$ and the proceeds to close it at a later time $t = t_2$ is $(1-K)X_{t_2}^1 - (1+K)X_{t_2}^2$. For ease of notation, let $\beta_{\mathrm{b}} = 1+K$ and $\beta_{\mathrm{s}} = 1-K$.

Given the initial state (x_1, x_2), net position $i = 0,1$, and the decision sequences Λ_0 and Λ_1, the corresponding reward functions

$$
\begin{aligned}
J_0(x_1,x_2,\Lambda_0) = E\Big\{ & [e^{-\rho\tau_2}(\beta_s X^1_{\tau_2} - \beta_b X^2_{\tau_2}) I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}(\beta_b X^1_{\tau_1} - \beta_s X^2_{\tau_1}) I_{\{\tau_1<\infty\}}] \\
& + [e^{-\rho\tau_4}(\beta_s X^1_{\tau_4} - \beta_b X^2_{\tau_4}) I_{\{\tau_4<\infty\}} - e^{-\rho\tau_3}(\beta_b X^1_{\tau_3} - \beta_s X^2_{\tau_3}) I_{\{\tau_3<\infty\}}] + \cdots \Big\}, \\
J_1(x_1,x_2,\Lambda_1) = E\Big\{ & e^{-\rho\tau_0}(\beta_s X^1_{\tau_0} - \beta_b X^2_{\tau_0}) I_{\{\tau_0<\infty\}} \\
& + [e^{-\rho\tau_2}(\beta_s X^1_{\tau_2} - \beta_b X^2_{\tau_2}) I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}(\beta_b X^1_{\tau_1} - \beta_s X^2_{\tau_1}) I_{\{\tau_1<\infty\}}] \\
& + [e^{-\rho\tau_4}(\beta_s X^1_{\tau_4} - \beta_b X^2_{\tau_4}) I_{\{\tau_4<\infty\}} - e^{-\rho\tau_3}(\beta_b X^1_{\tau_3} - \beta_s X^2_{\tau_3}) I_{\{\tau_3<\infty\}}] + \cdots \Big\},
\end{aligned}
\tag{2}
$$

where $\rho > 0$ is a given discount factor and I_A is the indicator function of an event A.

For $i = 0, 1$, let $V_i(x_1, x_2)$ denote the value functions with $(X^1_0, X^2_0) = (x_1, x_2)$ and initial net positions $i = 0, 1$. That is, $V_i(x_1, x_2) = \sup_{\Lambda_i} J_i(x_1, x_2, \Lambda_i)$, $i = 0, 1$.

Remark. Note that the 'one-share' assumption can be easily relaxed. For example, one can consider any pairs $\mathbf{Z}$ consisting of n_1 shares of long position in $\mathbf{X}^1$ and n_2 shares of short position in $\mathbf{X}^2$. This case can be treated by changing of the state variables $(X^1_t, X^2_t) \to (n_1 X^1_t, n_2 X^2_t)$. Due to the nature of GBMs, the corresponding system equation in (1) will stay the same. The new allocations will only affect the reward function in (2) implicitly. In addition, we only focus on the 'long' side of pairs trading and note that the 'short' side of trading can also be treated by simply switching the roles of the two stocks $\mathbf{X}^1$ and $\mathbf{X}^2$. □

Example. In this example, we consider stock prices of Target Corp. (TGT) and Wal-Mart Stores Inc. (WMT). In Figure 1, daily closing prices of both stocks from 1985 to 2014 are plotted. The data is divided into two parts. The first part (1985-1999) will be used to calibrate the model and the second part (2000-2014) to backtest the performance of our results. Using the prices (1985-1999) and following the traditional least squares method, we obtain $\mu_1 = 0.2059$, $\mu_2 = 0.2459$, $\sigma_{11} = 0.3112$, $\sigma_{12} = 0.0729$, $\sigma_{21} = 0.0729$, $\sigma_{22} = 0.2943$.

We assume **(A1):** $\rho > \mu_1$ and $\rho > \mu_2$. Under these conditions, we can show that, for all $x_1, x_2 > 0$,

$$
0 \le V_0(x_1,x_2) \le x_2, \quad \text{and} \quad \beta_s x_1 - \beta_b x_2 \le V_1(x_1,x_2) \le \beta_b x_1 + K x_2. \tag{3}
$$

Formally, the associated HJB equations have the form: For $x_1, x_2 > 0$,

$$
\begin{aligned}
&\min\Big\{\rho v_0(x_1,x_2) - \mathcal{A} v_0(x_1,x_2),\ v_0(x_1,x_2) - v_1(x_1,x_2) + \beta_b x_1 - \beta_s x_2\Big\} = 0, \\
&\min\Big\{\rho v_1(x_1,x_2) - \mathcal{A} v_1(x_1,x_2),\ v_1(x_1,x_2) - v_0(x_1,x_2) - \beta_s x_1 + \beta_b x_2\Big\} = 0,
\end{aligned}
\tag{4}
$$

where

$$
\mathcal{A} = \frac{1}{2}\left\{ a_{11} x_1^2 \frac{\partial^2}{\partial x_1^2} + 2 a_{12} x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} + a_{22} x_2^2 \frac{\partial^2}{\partial x_2^2} \right\} + \mu_1 x_1 \frac{\partial}{\partial x_1} + \mu_2 x_2 \frac{\partial}{\partial x_2},
$$

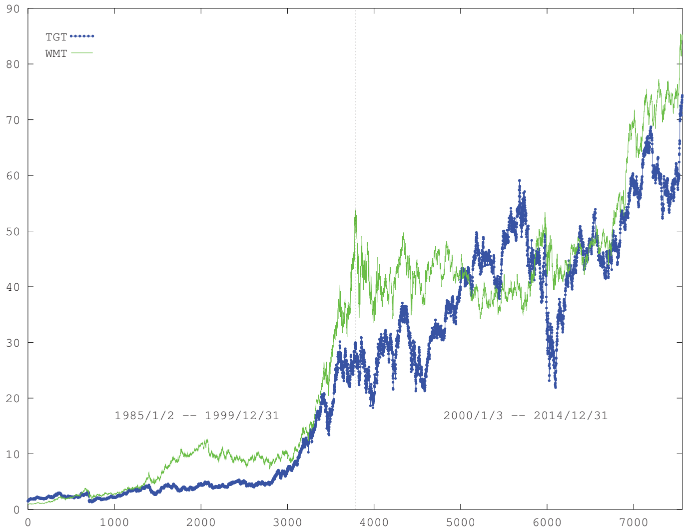

Fig. 1 Daily Closing Prices of TGT and WMT from 1985 to 2014.

and $a_{11} = \sigma_{11}^2 + \sigma_{12}^2$, $a_{12} = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}$, and $a_{22} = \sigma_{21}^2 + \sigma_{22}^2$.

We convert these HJB equations into single variable equations. Let $y = x_2/x_1$ and $v_i(x_1, x_2) = x_1 w_i(x_2/x_1)$, for some function $w_i(y)$ and $i = 0, 1$. By direct calculation, we have

$$\frac{\partial v_i}{\partial x_1} = w_i(y) - y w_i'(y), \quad \frac{\partial v_i}{\partial x_2} = w_i'(y),$$
$$\frac{\partial^2 v_i}{\partial x_1^2} = \frac{y^2 w_i''(y)}{x_1}, \quad \frac{\partial^2 v_i}{\partial x_2^2} = \frac{w_i''(y)}{x_1}, \text{ and } \frac{\partial^2 v_1}{\partial x_1 \partial x_2} = -\frac{y w_i''(y)}{x_1}.$$

We can write $\mathcal{A} v_i$ in terms of w_i and obtain

$$\mathcal{A} v_i = x_1 \left\{ \frac{1}{2} [a_{11} - 2a_{12} + a_{22}] y^2 w_i''(y) + (\mu_2 - \mu_1) y w_i'(y) + \mu_1 w_i(y) \right\}.$$

Let $\mathcal{L}[w_i(y)] = \lambda y^2 w_i''(y) + (\mu_2 - \mu_1) y w_i'(y) + \mu_1 w_i(y)$ with $\lambda = (a_{11} - 2a_{12} + a_{22})/2$. Then, the above HJB equations can be given as follows:

$$\begin{aligned} &\min \left\{ \rho w_0(y) - \mathcal{L} w_0(y),\ w_0(y) - w_1(y) + \beta_{\mathrm{b}} - \beta_{\mathrm{s}} y \right\} = 0, \\ &\min \left\{ \rho w_1(y) - \mathcal{L} w_1(y),\ w_1(y) - w_0(y) - \beta_{\mathrm{s}} + \beta_{\mathrm{b}} y \right\} = 0. \end{aligned} \tag{5}$$

In this paper, we only consider the case when $\lambda \neq 0$. If $\lambda = 0$, the problem reduces to a first order case and can be similarly treated. To solve these equations, we first focus on $(\rho-\mathcal{L})w_i(y)=0$, $i=0,1$. These are the Euler equations and their solutions are of the form y^δ, for some δ. We substitute this into the equation $(\rho-\mathcal{L})w_i=0$ and obtain the corresponding characteristic equation $\delta^2-(1+(\mu_1-\mu_2)/\lambda)\delta-(\rho-\mu_1)/\lambda=0$. There are two real roots

$$\begin{aligned}\delta_1&=\frac{1}{2}\left(1+\frac{\mu_1-\mu_2}{\lambda}+\sqrt{\left(1+\frac{\mu_1-\mu_2}{\lambda}\right)^2+\frac{4\rho-4\mu_1}{\lambda}}\right)>1,\\ \delta_2&=\frac{1}{2}\left(1+\frac{\mu_1-\mu_2}{\lambda}-\sqrt{\left(1+\frac{\mu_1-\mu_2}{\lambda}\right)^2+\frac{4\rho-4\mu_1}{\lambda}}\right)<0.\end{aligned}\tag{6}$$

The general solution of $(\rho-\mathcal{L})w_i(y)=0$ should be of the form: $w_i(y)=c_{i1}y^{\delta_1}+c_{i2}y^{\delta_2}$, for some c_{i1} and c_{i2}, $i=1,2$.

Intuitively, if X_t^1 is small and X_t^2 is large, then one should buy $\mathbf{X}^1$ and sell (short) $\mathbf{X}^2$. I.e., to open a pairs position $\mathbf{Z}$. If, on the other hand, X_t^1 is large and X_t^2 is small, then one should close the pairs position $\mathbf{Z}$ by selling $\mathbf{X}^1$ and buying back $\mathbf{X}^2$. In view of this, the first quadrant $P=\{(x_1,x_2):\ x_1>0 \text{ and } x_2>0\}$ into three regions Γ_1, Γ_2, and Γ_3 where $\Gamma_1=\{(x_1,x_2)\in P:\ x_2\le k_1x_1\}$, $\Gamma_2=\{(x_1,x_2)\in P:\ k_1x_1<x_2<k_2x_1\}$, and $\Gamma_3=\{(x_1,x_2)\in P:\ x_2\ge k_2x_1\}$. This is illustrated in Figure 2.

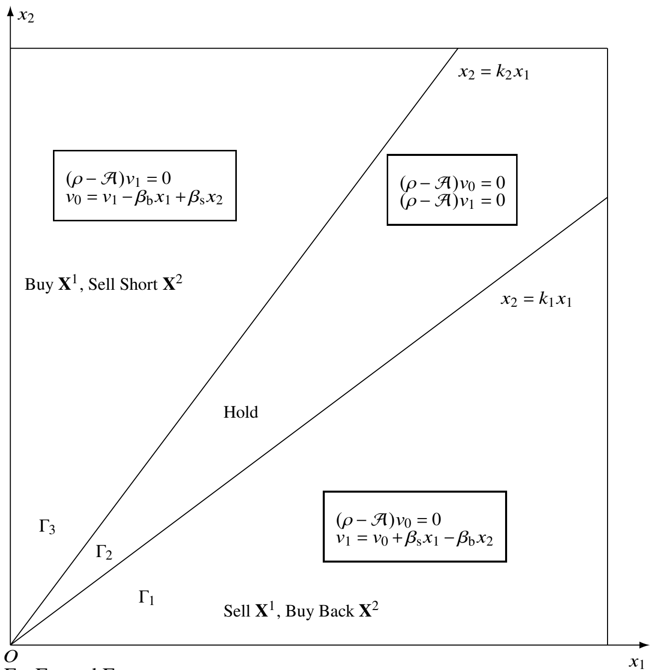

Fig. 2 Regions Γ_1, Γ_2, and Γ_3

With a little bit abuse of notation, we can write the corresponding Γ_i, $i = 1,2,3$, in terms of $y(= x_2/x_1)$: $\Gamma_1 = \{y:\ 0 < y \le k_1\}$, $\Gamma_2 = \{y:\ k_1 < y < k_2\}$, and $\Gamma_3 = \{y : y \ge k_2\}$. Here $0 < k_1 < k_2$ are slopes (thresholds) to be determined so that on

$$\begin{aligned}
&\Gamma_1 : (\rho - \mathcal{L})w_0 = 0, && w_1 = w_0 + \beta_s - \beta_b y;\\
&\Gamma_2 : (\rho - \mathcal{L})w_0 = 0, && (\rho - \mathcal{L})w_1 = 0;\\
&\Gamma_3 : w_0 = w_1 - \beta_b + \beta_s y, && (\rho - \mathcal{L})w_1 = 0.
\end{aligned} \tag{7}$$

Recall the boundedness of the value function in (3) and $\delta_2 < 0$. The coefficient of the term y^{δ_2} in w_0 on Γ_1 has to be zero. Thus, $w_0 = C_0 y^{\delta_1}$ for some C_0 on Γ_1. Likewise, on Γ_3, the coefficient of y^{δ_1} must be zero because $\delta_1 > 1$. The solution is $w_1 = C_1 y^{\delta_2}$ for some C_1 on Γ_3. Finally, these functions are extended to Γ_2 and are given by $w_0 = C_0 y^{\delta_1}$ and $w_1 = C_1 y^{\delta_2}$. The solutions on each region should have the form:

$$\begin{aligned}
&\Gamma_1 : w_0 = C_0 y^{\delta_1}, && w_1 = C_0 y^{\delta_1} + \beta_s - \beta_b y;\\
&\Gamma_2 : w_0 = C_0 y^{\delta_1}, && w_1 = C_1 y^{\delta_2};\\
&\Gamma_3 : w_0 = C_1 y^{\delta_2} - \beta_b + \beta_s y, && w_1 = C_1 y^{\delta_2}.
\end{aligned}$$

Next we use smooth-fit conditions to determine the values for parameters: k_1, k_2, C_0, and C_1. Necessarily, the continuity of w_1 and its first order derivative at $y = k_1$ imply $C_1 k_1^{\delta_2} = C_0 k_1^{\delta_1} + \beta_s - \beta_b k_1$ and $C_1 \delta_2 k_1^{\delta_2 - 1} = C_0 \delta_1 k_1^{\delta_1 - 1} - \beta_b$. These equations can be written in matrix form:

$$\begin{pmatrix} k_1^{\delta_1} & -k_1^{\delta_2} \\ \delta_1 k_1^{\delta_1 - 1} & -\delta_2 k_1^{\delta_2 - 1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_b k_1 - \beta_s \\ \beta_b \end{pmatrix}. \tag{8}$$

Similarly, the smooth-fit conditions for w_0 at $y = k_2$ lead to equations:

$$\begin{pmatrix} k_2^{\delta_1} & -k_2^{\delta_2} \\ \delta_1 k_2^{\delta_1 - 1} & -\delta_2 k_2^{\delta_2 - 1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_s k_2 - \beta_b \\ \beta_s \end{pmatrix}. \tag{9}$$

Solve for C_0 an C_1 and express the corresponding inverse matrices in terms of k_1 and k_2 to obtain

$$\begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \frac{1}{\delta_1-\delta_2}\begin{pmatrix} \beta_{\rm b}(1-\delta_2)k_1^{1-\delta_1}+\beta_{\rm s}\delta_2 k_1^{-\delta_1} \\ \beta_{\rm b}(1-\delta_1)k_1^{1-\delta_2}+\beta_{\rm s}\delta_1 k_1^{-\delta_2} \end{pmatrix} = \frac{1}{\delta_1-\delta_2}\begin{pmatrix} \beta_{\rm s}(1-\delta_2)k_2^{1-\delta_1}+\beta_{\rm b}\delta_2 k_2^{-\delta_1} \\ \beta_{\rm s}(1-\delta_1)k_2^{1-\delta_2}+\beta_{\rm b}\delta_1 k_2^{-\delta_2} \end{pmatrix}. \tag{10}$$

The second equality yields two equations of k_1 and k_2. We can simplify them and write

$$(1-\delta_2)(\beta_{\rm b}k_1^{1-\delta_1}-\beta_{\rm s}k_2^{1-\delta_1}) = \delta_2(\beta_{\rm b}k_2^{-\delta_1}-\beta_{\rm s}k_1^{-\delta_1}),$$
$$(1-\delta_1)(\beta_{\rm b}k_1^{1-\delta_2}-\beta_{\rm s}k_2^{1-\delta_2}) = \delta_1(\beta_{\rm b}k_2^{-\delta_2}-\beta_{\rm s}k_1^{-\delta_2}).$$

To solve these equations, let $r = k_2/k_1$ and replace k_2 by rk_1 to obtain

$$(1-\delta_2)(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_1})k_1 = \delta_2(\beta_{\rm b}r^{-\delta_1}-\beta_{\rm s})$$

and

$$(1-\delta_1)(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_2})k_1 = \delta_1(\beta_{\rm b}r^{-\delta_2}-\beta_{\rm s}).$$

We have

$$k_1 = \frac{\delta_2(\beta_{\rm b}r^{-\delta_1}-\beta_{\rm s})}{[(1-\delta_2)(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_1})]} = \frac{\delta_1(\beta_{\rm b}r^{-\delta_2}-\beta_{\rm s})}{[(1-\delta_1)(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_2})]}.$$

Using the second equality and write the difference of both sides, we have

$$\begin{aligned} f(r) &:= \delta_1(1-\delta_2)(\beta_{\rm b}r^{-\delta_2}-\beta_{\rm s})(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_1}) \\ &\quad -\delta_2(1-\delta_1)(\beta_{\rm b}r^{-\delta_1}-\beta_{\rm s})(\beta_{\rm b}-\beta_{\rm s}r^{1-\delta_2}) = 0. \end{aligned}$$

where $\beta = \beta_b/\beta_s (> 1)$. Then we can show $f(\beta^2) > 0$ and $f(r) \to -\infty$, as $r \to \infty$. Therefore, there exists $r_0 > \beta^2$ so that $f(r_0) = 0$. Using this r_0, we write k_1 and k_2:

$$\begin{aligned} k_1 &= \frac{\delta_2(\beta_{\rm b}r_0^{-\delta_1}-\beta_{\rm s})}{(1-\delta_2)(\beta_{\rm b}-\beta_{\rm s}r_0^{1-\delta_1})} = \frac{\delta_1(\beta_{\rm b}r_0^{-\delta_2}-\beta_{\rm s})}{(1-\delta_1)(\beta_{\rm b}-\beta_{\rm s}r_0^{1-\delta_2})}, \\ k_2 &= \frac{\delta_2(\beta_{\rm b}r_0^{1-\delta_1}-\beta_{\rm s}r_0)}{(1-\delta_2)(\beta_{\rm b}-\beta_{\rm s}r_0^{1-\delta_1})} = \frac{\delta_1(\beta_{\rm b}r_0^{1-\delta_2}-\beta_{\rm s}r_0)}{(1-\delta_1)(\beta_{\rm b}-\beta_{\rm s}r_0^{1-\delta_2})}. \end{aligned} \tag{11}$$

Finally, we can use these k_1 and k_2 to express C_0 and C_1 given in (10).

Theorem. Assume (A1). Then the solutions of the HJB equations (4) can be given as $v_0(x_1,x_2) = x_1 w_0(x_2/x_1)$ and $v_1(x_1,x_2) = x_1 w_1(x_2/x_1)$ where

$$w_0(y) = \begin{cases} \left(\dfrac{\beta_b(1-\delta_2)k_1^{1-\delta_1} + \beta_s\delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}\right) y^{\delta_1}, & \text{if } 0 < y < k_2, \\ \left(\dfrac{\beta_b(1-\delta_1)k_1^{1-\delta_2} + \beta_s\delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}\right) y^{\delta_2} + \beta_s y - \beta_b, & \text{if } y \geq k_2, \end{cases}$$

$$w_1(y) = \begin{cases} \left(\dfrac{\beta_b(1-\delta_2)k_1^{1-\delta_1} + \beta_s\delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}\right) y^{\delta_1} + \beta_s - \beta_b y, & \text{if } 0 < y \leq k_1, \\ \left(\dfrac{\beta_b(1-\delta_1)k_1^{1-\delta_2} + \beta_s\delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}\right) y^{\delta_2}, & \text{if } y > k_1. \end{cases}$$

The optimal trading rule can be determined by two threshold lines ($x_2 = k_1 x_1$ and $x_2 = k_2 x_1$) as follows:

Theorem. Assume (A1). Then, $v_i(x_1, x_2) = x_1 w_i(x_2/x_1) = V_i(x_1, x_2)$, $i = 0, 1$. Moreover, if initially $i = 0$, let $\Lambda_0^* = (\tau_1^*, \tau_2^*, \tau_3^*, \dots)$ such that $\tau_1^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_3^* = \inf\{t \geq \tau_2^* : (X_t^1, X_t^2) \in \Gamma_3\}$, and so on. Similarly, if initially $i = 1$, let $\Lambda_1^* = (\tau_0^*, \tau_1^*, \tau_2^*, \dots)$ such that $\tau_0^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_1^* = \inf\{t \geq \tau_0^* : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$, and so on. Then Λ_0^* and Λ_1^* are optimal. □

Example 1 (cont.) We backtest our pairs trading rule using the stock prices of TGT and WMT from 2000 to 2014. Using the parameters mentioned earlier, based on the historical prices from 1985 to 1999, we obtain the pair $(k_1, k_2) = (1.03905, 1.28219)$. A pairs trading (long $\mathbf{X}^1$ and short $\mathbf{X}^2$) is triggered when (X_t^1, X_t^2) enters Γ_3. The position is closed when (X_t^1, X_t^2) enters Γ_1. Initially, we allocate trading the capital \$100K. When the first long signal is triggered, buy \$50K TGT stocks and short the same amount of WMT. Such half-and-half capital allocation between long and short applies to all trades. In addition, each pairs transaction is charged \$5 commission. In Figure 3, the corresponding ratio X_t^2/X_t^1, the threshold levels k_1 and k_2, and the corresponding equity curve are plotted. There are total 3 trades and the end balance is \$155.914K.

We can also switch the roles of $\mathbf{X}^1$ and $\mathbf{X}^2$, i.e., to long WMT and short TGT by taking $\mathbf{X}^1$=WMT and $\mathbf{X}^2$=TGT. In this case, the new $(\tilde{k}_1, \tilde{k}_2) = (1/k_2, 1/k_1) = (1/1.28219, 1/1.03905)$. These levels and the corresponding equity curve is given in Figure 4. Such trade leads to the end balance \$132.340K. Note that both types of trades have no overlap, i.e., they do not compete for the same capital. The grand total profit is \$88254 which is a 88.25% gain.

Note also that there are only 5 trades in the fifteen year period leaving the capital in cash most of the time. This is desirable because the cash sitting in the account can be used for other types of shorter term trading in between, at least drawing interest over time.

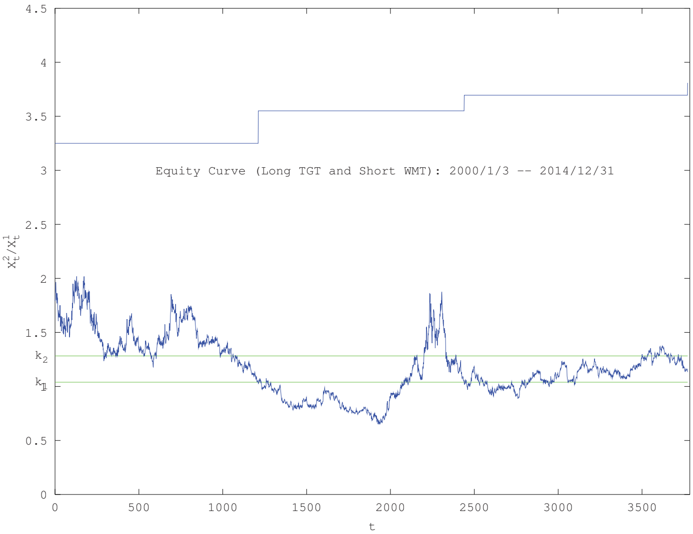

Fig. 3 $\mathbf{X}^1$=TGT, $\mathbf{X}^2$=WMT: The threshold levels k_1, k_2 and the corresponding equity curve

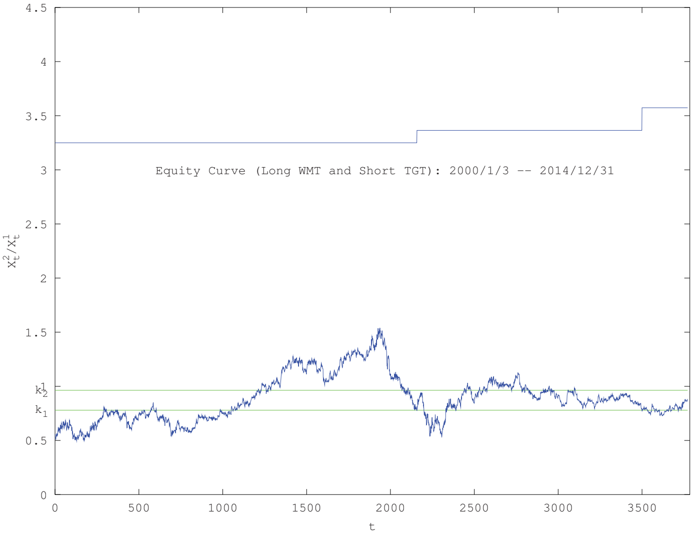

Fig. 4 $\mathbf{X}^1$=WMT, $\mathbf{X}^2$=TGT: The threshold levels k_1, k_2 and the corresponding equity curve

3 Pairs Trading with Cutting Losses

In this section, we consider our above-mentioned pairs trading rule with cutting losses. Recall that a pairs position consists of a long position in stock $\mathbf{X}^1$ and a short position in $\mathbf{X}^2$. The objective is to open (buy) and close (sell) the pairs positions sequentially to maximize the discounted reward function J_0 and J_1 in (2). In practice, unexpected events could cause substantial losses. This normally occurs when the long side X_t^1 shrinks while the short side X_t^2 rises. To limit the downside risk of the pairs position, we impose a hard cut loss level and require $X_t^2/X_t^1 \leq M$. Here M is a constant representing a stop-loss level to account for market reaction to undesirable events. The introduction of such stop-loss level amounts to imposing a hard state constraint which makes the corresponding optimal control problem much more difficult.

Let τ_M denote the corresponding exit time, i.e., $\tau_M = \{t : X_t^2/X_t^1 \geq M\}$. Then, $\tau_n \leq \tau_M$, for all n.

Our goal is to find Λ_0 and Λ_1 so as to maximize the reward functions $J_0(x_1, x_2, \Lambda_0)$ and $J_1(x_1, x_2, \Lambda_1)$ under such state constraints. For $i = 0, 1$, let $V_i(x_1, x_2)$ denote the corresponding value functions with the initial state $(X_0^1, X_0^2) = (x_1, x_2)$ and net positions $i = 0, 1$.

Example The main purpose of imposing a hard stop-loss level M is to limit losses to an acceptable level to account for undesirable market moves to unforeseeable events. The stock prices of Ford Motor (F) and General Motors (GM) are highly correlated historically. They make good candidates for pairs trading. In Figure 5, the daily closing price ratio (F/GM) from 1977 to 2009 is plotted. It can be seen that the ratio remains ‘normal’ for most of the time during this period of time. The ratio starts to rise when approaching the subprime crisis. This would normally trigger a pairs position longing GM and shorting F. Finally, it spikes prior to GM’s chapter 11 filing on June 1, 2009 causing heavy losses to any F/GM pair positions. Such hypothetical losses can be limited if one had a hard limit M in place to begin with to force close the position before prices getting out of control.

The choice of M depends on the investor’s risk preference. Smaller M (tighter stop-loss control) will cause frequent stop outs and limit profit potential. Larger M (loose stop-loss), on the other hand, will leave more room for the position to run with higher risks. □

Let H denote the feasible region under the hard state constraint $x_2/x_1 \leq M$. Then, $H = \{(x_1, x_2) : 0 < x_1, 0 < x_2 \leq M x_1\}$. We can show the same inequalities in (3) hold on H. The associated HJB equations on H can be given as follows:

$$\begin{cases} \min\{\rho v_0(x_1,x_2) - \mathcal{A} v_0(x_1,x_2),\ v_0(x_1,x_2) - v_1(x_1,x_2) + \beta_{\text{b}} x_1 - \beta_{\text{s}} x_2\} = 0, \\ \min\{\rho v_1(x_1,x_2) - \mathcal{A} v_1(x_1,x_2),\ v_1(x_1,x_2) - v_0(x_1,x_2) - \beta_{\text{s}} x_1 + \beta_{\text{b}} x_2\} = 0, \end{cases} \tag{12}$$

with the boundary conditions $v_0(x_1, M x_1) = 0$ and $v_1(x_1, M x_1) = \beta_{\text{s}} x_1 - \beta_{\text{b}} M x_1$.

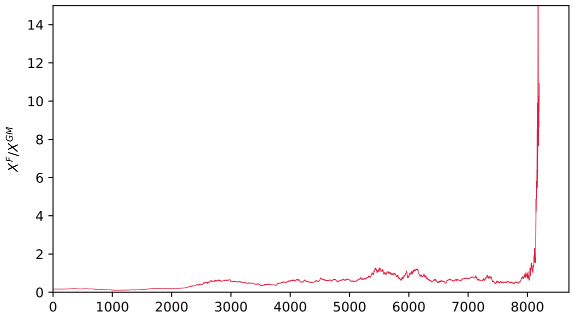

Fig. 5 Daily closing ratio of F/GM from 1977 to 2009

Following similar approach as in the previous section, we divide the feasible region H into four regions $\Gamma_1 = \{(x_1,x_2) \in H : 0 < x_2 \le k_1x_1\}$, $\Gamma_2 = \{(x_1,x_2) \in H : k_1x_1 < x_2 \le k_2x_1\}$, $\Gamma_3 = \{(x_1,x_2) \in H : k_2x_1 < x_2 \le k_3x_1\}$, and $\Gamma_4 = \{(x_1,x_2) \in H : k_3x_1 < x_2 \le Mx_1\}$, where $0 < k_1 < k_2 < k_3 < M$ are threshold slopes to be determined. The control actions on Γ_1, Γ_2, and Γ_3 are similar as before. Γ_4 is the hold and see region due to possible cut-loss at $x_2 = Mx_1$. This is illustrated in Figure 6.

Using the smooth-fit approach, we can show that the k_1 and k_2 are identical as the ones given in (11) with δ_1 and δ_2 in (6).

To determine k_3, let

$$f_1(x) = \frac{M^{\delta_1}\beta_s(x(1-\delta_2)+\beta\delta_2)}{x^{\delta_1}} + \frac{M^{\delta_2}\beta_s(x(\delta_1-1)-\beta\delta_1)}{x^{\delta_2}} + \beta_s(1-M\beta)(\delta_1-\delta_2).$$

We assume **(A2)**: There is a k_3 in (k_2, M) such that $f_1(k_3) = 0$.

A sufficient condition for this can be given as **(A2')** $f_1(k_3) > 0$.

On each of the regions Γ_i, $i = 1,2,3,4$, we can write the solutions of the HJB equations in terms of δ_i, $i = 1,2$, with coefficients C_j, $j = 0,1,\ldots,4$. Then using smooth-fit conditions, we can specify these constants as follows:

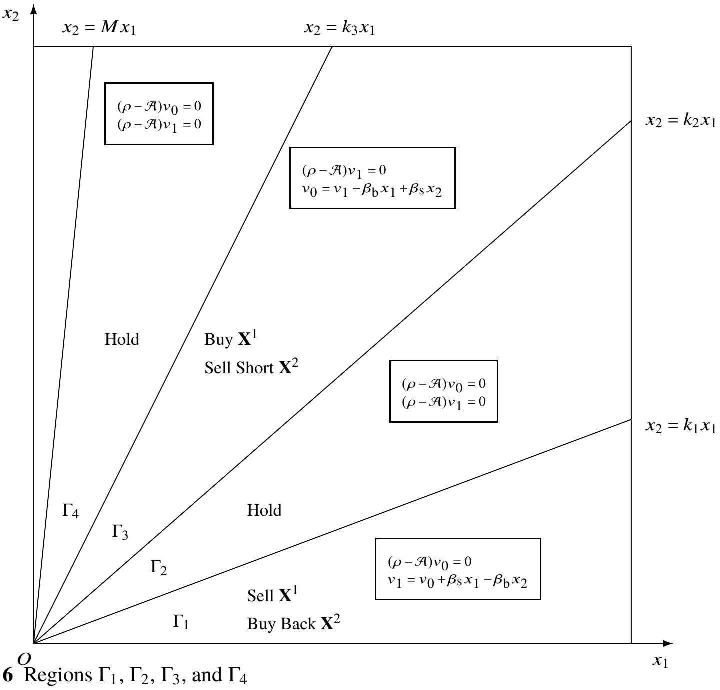

Fig. 6 Regions Γ_1, Γ_2, Γ_3, and Γ_4

$$\begin{cases} C_0 = \dfrac{1}{M^{\delta_1}}\left[(M^{\delta_1}, M^{\delta_2})K^0(k_1)\begin{pmatrix} \beta_s k_1 - \beta_b \\ \beta_b \end{pmatrix} + \beta_s - M\beta_b\right], \\ C_1 = C_0 + \dfrac{(\delta_2 - 1)\beta_b k_1^{1-\delta_1} - \beta_s \delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}, \\ C_2 = \dfrac{(1-\delta_1)\beta_b k_1^{1-\delta_2} + \beta_s \delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}, \\ C_3 = C_1 + \dfrac{(1-\delta_2)\beta_s k_3^{1-\delta_1} + \beta_b \delta_2 k_3^{-\delta_1}}{\delta_1 - \delta_2}, \\ C_4 = C_2 + \dfrac{(\delta_1 - 1)\beta_s k_3^{1-\delta_2} - \beta_b \delta_1 k_3^{-\delta_2}}{\delta_1 - \delta_2}, \end{cases} \tag{13}$$

where

$$K^0(x) = \frac{1}{\delta_1 - \delta_2}\begin{pmatrix} -\delta_2 x^{-\delta_1} & x^{1-\delta_1} \\ \delta_1 x^{-\delta_2} & -x^{1-\delta_2} \end{pmatrix}.$$

Finally, we need an additional condition to guarantee all inequalities in the HJB equations to hold. We assume (**A3**): Either $f_2'(M) < 0$, or $f_2''(M) < 0$, where,

$$f_2(y) = (C_1 y^{\delta_1} + C_2 y^{\delta_2}) - (C_3 y^{\delta_1} + C_4 y^{\delta_2}) + \beta_b y - \beta_s.$$

A sufficient condition fo (A3) can be given as **(A3')**: $\mu_1 \geq \mu_2$. Under these conditions, we have the following theorems.

Theorem Assume (A1), (A2), and (A3). Then the following functions $v_i(x_1, x_2) = x_1 w_i(x_2/x_1)$, $i = 0, 1$, satisfy the HJB equations (12) where

$$w_0(y) = \begin{cases} C_0 y^{\delta_1}, & 0 < y < k_2, \\ C_1 y^{\delta_1} + C_2 y^{\delta_2} + \beta_s y - \beta_b, & k_2 \leqslant y \leqslant k_3, \\ C_3 y^{\delta_1} + C_4 y^{\delta_2}, & k_3 < y \leqslant M; \end{cases}$$
$$w_1(y) = \begin{cases} C_0 y^{\delta_1} + \beta_s - \beta_b y, & 0 < y < k_1, \\ C_1 y^{\delta_1} + C_2 y^{\delta_2}, & k_1 \leqslant y \leqslant M. \end{cases}$$

Theorem Assume (A1), (A2), and (A3) and $v_0(x_1, x_2) \geq 0$. Then, $v_i(x_1, x_2) = x_1 w_i(x_2/x_1) = V_i(x_1, x_2)$, $i = 0, 1$. Moreover, if $i = 0$, let $\Lambda_0^* = (\tau_1^*, \tau_2^*, \tau_3^*, \ldots) = (\tau_1^0, \tau_2^0, \tau_3^0, \ldots) \wedge \tau_M$ where $\tau_1^0 = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^0 = \inf\{t \geq \tau_1^0 : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_3^0 = \inf\{t \geq \tau_2^0 : (X_t^1, X_t^2) \in \Gamma_3\}, \ldots$.

Similarly, if $i = 1$, let $\Lambda_1^* = (\tau_0^*, \tau_1^*, \tau_2^*, \ldots) = (\tau_0^0, \tau_1^0, \tau_2^0, \ldots) \wedge \tau_M$ where $\tau_0^0 = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_1^0 = \inf\{t \geq \tau_0^0 : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^0 = \inf\{t \geq \tau_1^0 : (X_t^1, X_t^2) \in \Gamma_1\}$,.... Then Λ_0^* and Λ_1^* are optimal. □

Next, we consider the daily closing prices of Target Corp. (TGT) and Wal-Mart Stores Inc. (WMT) from 1985 to 2019. The data are divided into two parts. The first part (1985-1999) is used to calibrate the model and the second part (2000-2014) to backtest the performance of our results. Let $\mathbf{X}^1$=WMT and $\mathbf{X}^2$=TGT. Using the traditional least squares method, we have

$$\mu_1 = 0.2459, \mu_2 = 0.2059, \sigma_{11} = 0.2943, \sigma_{12} = 0.0729, \sigma_{21} = 0.0729, \sigma_{22} = 0.3112. \tag{14}$$

And also, we take $K = 0.001$ and $\rho = 0.5$. Using these parameters, we obtain $(k_1, k_2, k_3) = (0.780, 0.963, 1.913)$.

Backtesting 1: (WMT-TGT): We backtest our pairs-trading rule using the daily closing prices of WMT and TGT from 2000/1/2 to 2019/3/15. Use $(k_1, k_2, k_3) = (0.780, 0.963, 1.913)$. Assume initial capital \$100K. We keep the 50:50 allocations in longs and shorts. In Figure 7, the ratio of $\mathbf{X}_t^{TGT}/\mathbf{X}_t^{WMT}$, the threshold levels (k_1, k_2, k_3), and the equity curve are plotted with the x-axis representing the number of trading days. Also, when there is no pairs position, we factor in a 3% interest for the cash position. The overall end balance is \$195.46K. For comparison purpose, a money market return with 3% interest rate is also plotted in Figure 7. In this example, the stop loss with $M = 2$ was not triggered and there was no forced stops.

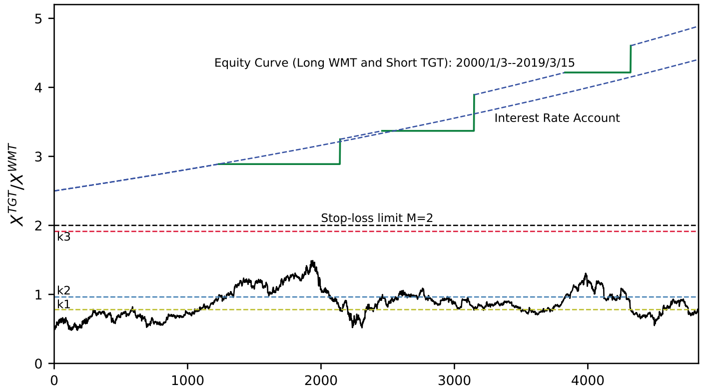

Fig. 7 The threshold levels k_1, k_2, k_3 and the equity curve

Backtesting 2: (GM-F). Next, we backtest using the daily closing prices of GM and F from 1998/1/2 to 2009/6/30. We take $M = 2$ and follow similar calculation (with 2:1 ratio of F/GM) to obtain $(k_1, k_2, k_3) = (0.760, 0.892, 1.946)$. Also assume the initial capital \$100K. We keep the 50:50 distribution in dollar amount between longs and shorts. In Figure 8, the ratio $2\mathbf{X}_t^F/\mathbf{X}_t^{GM}$, the threshold levels (k_1, k_2, k_3), and the equity curve are plotted. Similarly as in the previous example, when there is no pairs position, a 3% interest was factored in for the cash position. The overall end balance is \$149.52K after hitting stop-loss limit $M = 2$ on 2009/3/6.

On the other hand, without cutting losses, the initial \$100K will end up with \$86.38K in debt when the last pairs closed on GM's bankruptcy (2009/6/1). A pure money market return with 3% interest rate is also provided in Figure 8.

4 A Pairs Selling Rule with Regime Switching

Market models with regime-switching are important in market analysis. In this paper, we consider a geometric Brownian motion with regime-switching. The market mode is represented by a two-state Markov chain. We focus on the selling part of pairs trading and generalize the results of Hu and Øksendal [9] by incorporating models with regime switching. We show that the optimal selling rule can be determined by two threshold curves and establish a set of sufficient conditions that guarantee the optimality of the policy. We also include several numerical examples under a different set of parameter values.

We consider two stocks $\mathbf{X}^1$ and $\mathbf{X}^2$. Let $\{X_t^1, t \geq 0\}$ denote the prices of stock $\mathbf{X}^1$ and $\{X_t^2, t \geq 0\}$ that of stock $\mathbf{X}^2$. Let also α_t be a two-state Markov chain representing

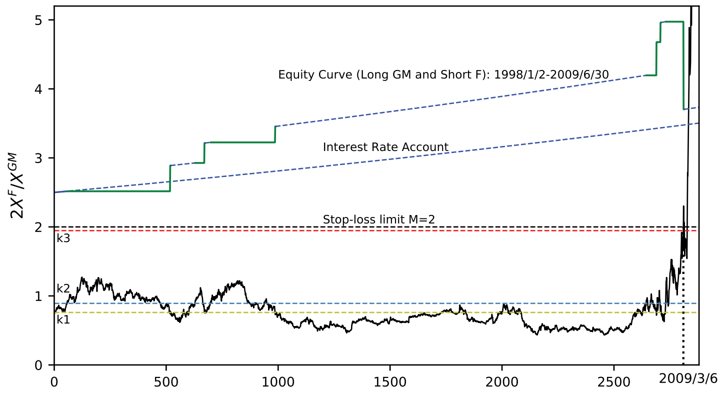

Fig. 8 The threshold levels k_1, k_2, k_3 and the equity curve

regime mode. They satisfy the following stochastic differential equation:

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} X_t^1 & \\ & X_t^2 \end{pmatrix} \left[\begin{pmatrix} \mu_1(\alpha_t) \\ \mu_2(\alpha_t) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11}(\alpha_t) & \sigma_{12}(\alpha_t) \\ \sigma_{21}(\alpha_t) & \sigma_{22}(\alpha_t) \end{pmatrix} d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} \right], \tag{15}$$

where μ_i, $i = 1,2$, are the return rates, σ_{ij}, $i,j = 1,2$, the volatility constants, and (W_t^1, W_t^2) a 2-dimensional standard Brownian motion.

Let $\mathcal{M} = \{1,2\}$ denote the state space for α_t and let $Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$, with $\lambda_1 > 0$ and $\lambda_2 > 0$, be its generator. We assume α_t and (W_t^1, W_t^2) are independent.

In this section, we consider a pairs selling rule under the regime switching model. Again, we assume the corresponding pairs position consists of a one-share long position in stock $\mathbf{X}^1$ and a one-share short position in stock $\mathbf{X}^2$. The problem is to determine an optimal stopping time τ to close the pairs position by selling $\mathbf{X}^1$ and buying back $\mathbf{X}^2$.

Given the initial state $(X_0^1, X_0^2) = (x_1, x_2)$, $\alpha_0 = i = 1,2$, and the selling time τ, the corresponding reward function

$$J(x_1, x_2, i, \tau) = E\left[e^{-\rho\tau}(\beta_s X_\tau^1 - \beta_b X_\tau^2)\right], \tag{16}$$

where $\rho > 0$ is a given discount factor, $\beta_b = 1 + K$, $\beta_s = 1 - K$, and K is the transaction cost in percentage.

The problem is to find an $\{\mathcal{F}_t\} = \sigma\{(X_r^1, X_r^2, \alpha_r) : \ r \le t\}$ stopping time τ to maximize J. Let $V(x_1, x_2, i)$ denote the corresponding value functions:

$$V(x_1,x_2,i) = \sup_{\tau} J(x_1,x_2,i,\tau). \tag{17}$$

As in the previous sections, we impose the following conditions: **(B1)** For $i = 1,2$, $\rho > \mu_1(i)$ and $\rho > \mu_2(i)$.

Under these conditions, we can obtain

$$\beta_{\rm s}x_1 - \beta_{\rm b}x_2 \le V(x_1,x_2,i) \le \beta_{\rm s}x_1. \tag{18}$$

To consider the associated HJB equations, for $i = 1,2$, let

$$\begin{aligned}\mathcal{A}_i = \frac{1}{2}\left[a_{11}(i)x_1^2\frac{\partial^2}{\partial x_1^2} + 2a_{12}(i)x_1x_2\frac{\partial^2}{\partial x_1\partial x_2} + a_{22}(i)x_2^2\frac{\partial^2}{\partial x_2^2}\right]\\ +\mu_1(i)x_1\frac{\partial}{\partial x_1} + \mu_2(i)x_2\frac{\partial}{\partial x_2}\end{aligned} \tag{19}$$

where

$$\begin{cases} a_{11}(i) = \sigma_{11}^2(i) + \sigma_{12}^2(i),\\ a_{12}(i) = \sigma_{11}(i)\sigma_{21}(i) + \sigma_{12}(i)\sigma_{22}(i),\\ a_{22}(i) = \sigma_{21}^2(i) + \sigma_{22}^2(i).\end{cases}$$

Using these generators, the associated HJB equations have the form:

$$\begin{cases} \min\{(\rho - \mathcal{A}_1)v(x_1,x_2,1) - \lambda_1(v(x_1,x_2,2) - v(x_1,x_2,1)),\\ \qquad\qquad v(x_1,x_2,1) - \beta_{\rm s}x_1 + \beta_{\rm b}x_2\} = 0,\\ \min\{(\rho - \mathcal{A}_2)v(x_1,x_2,2) - \lambda_2(v(x_1,x_2,1) - v(x_1,x_2,2)),\\ \qquad\qquad v(x_1,x_2,2) - \beta_{\rm s}x_1 + \beta_{\rm b}x_2\} = 0.\end{cases} \tag{20}$$

To solve the HJB equations (20), we can introduce change of variables: $y = x_2/x_1$ and $v(x_1,x_2,i) = x_1w_i(x_2/x_1)$, for some functions $w_i(y)$ and $i = 1,2$.

Consider characteristic equations for $(\rho - \mathcal{A}_1)v_1 - \lambda_1(v_2 - v_1) = 0$ and $(\rho - \mathcal{A}_2)v_2 - \lambda_2(v_1 - v_2) = 0$:

$$[\rho + \lambda_1 - \theta_1(\delta)][\rho + \lambda_2 - \theta_2(\delta)] - \lambda_1\lambda_2 = 0, \tag{21}$$

where, for $i = 1,2$. $\theta_i(\delta) = \sigma_i\delta(\delta - 1) + [(\mu_2(i) - \mu_1(i)]\delta + \mu_1(i)$ and $\sigma_i = [a_{11}(i) - 2a_{12}(i) + a_{22}(i)]/2$.

It can be seen the above equation has four zeros: $\delta_1 \ge \delta_2 > 1 > 0 > \delta_3 \ge \delta_4$.

Heuristically, one should close the pairs position when X_t^1 is large and X_t^2 is small. In view of this, we introduce $H_1 = \{(x_1,x_2) : x_2 \le k_1x_1\}$ and $H_2 = \{(x_1,x_2) : x_2 \le k_2x_1\}$, for some k_1 and k_2 (to be determined) so that one should sell when (X_t^1, X_t^2)

enters H_i provided $\alpha_t = i$, $i = 1,2$. In this paper, we only consider the case: $k_1 < k_2$. Other cases can be treated similarly.

To represent the solutions to the HJB equations on each of these regimes, we apply the smooth-fit approach and obtain:

$$\begin{cases} C_1 = \dfrac{-\delta_4\beta_s + (\delta_4 - 1)\beta_b k_2}{\eta_3(\delta_3 - \delta_4)k_2^{\delta_3}}, \\ C_2 = \dfrac{\delta_3\beta_s + (1-\delta_3)\beta_b k_2}{\eta_4(\delta_3 - \delta_4)k_2^{\delta_4}}, \\ C_3 = \dfrac{\gamma_2(\beta_s - a_1) + (1-\gamma_2)(\beta_b + a_2)k_1}{(\gamma_2 - \gamma_1)k_1^{\gamma_1}}, \\ C_4 = \dfrac{-\gamma_1(\beta_s - a_1) + (\gamma_1 - 1)(\beta_b + a_2)k_1}{(\gamma_2 - \gamma_1)k_1^{\gamma_2}}, \end{cases}$$

where $\eta_j = (\rho + \lambda_1 - \theta_1(\delta_j))/\lambda_1$, for $j = 1,2,3,4$, and

$$\begin{cases} \gamma_1 = \dfrac{1}{2} + \dfrac{\mu_1(1) - \mu_2(1)}{2\sigma_1} + \sqrt{\left(\dfrac{1}{2} + \dfrac{\mu_1(1) - \mu_2(1)}{2\sigma_1}\right)^2 + \dfrac{\rho + \lambda_1 - \mu_1(1)}{\sigma_1}}, \\ \gamma_2 = \dfrac{1}{2} + \dfrac{\mu_1(1) - \mu_2(1)}{2\sigma_1} - \sqrt{\left(\dfrac{1}{2} + \dfrac{\mu_1(1) - \mu_2(1)}{2\sigma_1}\right)^2 + \dfrac{\rho + \lambda_1 - \mu_1(1)}{\sigma_1}}. \end{cases} \tag{22}$$

Let

$$g(r) = \frac{A_1 - \gamma_2(\beta_s - a_1)r^{\gamma_1}}{(1-\gamma_2)(\beta_b + a_2)r^{\gamma_1} - B_1 r} - \frac{A_2 + \gamma_1(\beta_s - a_1)r^{\gamma_2}}{(\gamma_1 - 1)(\beta_b + a_2)r^{\gamma_2} - B_2 r}, \tag{23}$$

where

$$\begin{cases} A_1 = \dfrac{-\delta_4\beta_s(\gamma_2 - \delta_3)}{\eta_3(\delta_3 - \delta_4)} + \dfrac{\delta_3\beta_s(\gamma_2 - \delta_4)}{\eta_4(\delta_3 - \delta_4)} - \gamma_2 a_1, \\ A_2 = \dfrac{-\delta_4\beta_s(\delta_3 - \gamma_1)}{\eta_3(\delta_3 - \delta_4)} + \dfrac{\delta_3\beta_s(\delta_4 - \gamma_1)}{\eta_4(\delta_3 - \delta_4)} + \gamma_1 a_1, \\ B_1 = \dfrac{(\delta_4 - 1)(\gamma_2 - \delta_3)\beta_b}{\eta_3(\delta_3 - \delta_4)} + \dfrac{(1-\delta_3)\beta_b(\gamma_2 - \delta_4)}{\eta_4(\delta_3 - \delta_4)} - (\gamma_2 - 1)a_2, \\ B_2 = \dfrac{(\delta_4 - 1)(\delta_3 - \gamma_1)\beta_b}{\eta_3(\delta_3 - \delta_4)} + \dfrac{(1-\delta_3)\beta_b(\delta_4 - \gamma_1)}{\eta_4(\delta_3 - \delta_4)} - (1-\gamma_1)a_2, \end{cases}$$

with $a_1 = \lambda_1\beta_s/(\rho + \lambda_1 - \mu_1(1))$ and $a_2 = -\lambda_1\beta_b/(\rho + \lambda_1 - \mu_2(1))$.

We assume **(B2)**: $g(r)$ has a zero $r_0 > 1$.

Using this r_0, we can obtain

$$\begin{cases} k_1 = \dfrac{A_1 - \gamma_2(\beta_{\mathrm{s}} - a_1)r_0^{\gamma_1}}{(1-\gamma_2)(\beta_{\mathrm{b}} + a_2)r_0^{\gamma_1} - B_1 r_0}, \\ k_2 = r_0 k_1 = \dfrac{A_1 r_0 - \gamma_2(\beta_{\mathrm{s}} - a_1)r_0^{\gamma_1+1}}{(1-\gamma_2)(\beta_{\mathrm{b}} + a_2)r_0^{\gamma_1} - B_1 r_0}. \end{cases} \tag{24}$$

We can express C_1, C_2, C_3, and C_4 in terms of k_1 and k_2. The solutions to the HJB equations have the form $v(x_1, x_2, \alpha) = x_1 w_\alpha(x_2/x_1)$, $\alpha = 1, 2$, with

$$w_1(y) = \begin{cases} \beta_{\mathrm{s}} - \beta_{\mathrm{b}} y & \text{for } y \in \Gamma_1, \\ C_3 y^{\gamma_1} + C_4 y^{\gamma_2} + a_1 + a_2 y & \text{for } y \in \Gamma_2, \\ C_1 y^{\delta_3} + C_2 y^{\delta_4} & \text{for } y \in \Gamma_3; \end{cases}$$

$$w_2(y) = \begin{cases} \beta_{\mathrm{s}} - \beta_{\mathrm{b}} y & \text{for } y \in \Gamma_1 \cup \Gamma_2, \\ C_1 \eta_3 y^{\delta_3} + C_2 \eta_4 y^{\delta_4} & \text{for } y \in \Gamma_3, \end{cases}$$

where

$$\Gamma_1 = (0, k_1], \quad \Gamma_2 = (k_1, k_2), \quad \text{and} \quad \Gamma_3 = [k_2, \infty).$$

To guarantee the variational inequalities in the HJB equations, we need the following conditions:

$$k_1 \le \min\left\{ \frac{(\rho - \mu_1(1))\beta_{\mathrm{s}}}{(\rho - \mu_2(1))\beta_{\mathrm{b}}}, \frac{(\rho - \mu_1(2))\beta_{\mathrm{s}}}{(\rho - \mu_2(2))\beta_{\mathrm{b}}} \right\}; \tag{25}$$

$$w_1(y) \le \beta_{\mathrm{s}} - \beta_{\mathrm{b}} y + \frac{1}{\lambda_2}[(\rho - \mu_1(2))\beta_{\mathrm{s}} - (\rho - \mu_2(2))\beta_{\mathrm{b}} y] \text{ on } \Gamma_2. \tag{26}$$

In addition, let $\phi(y) = w_1(y) - \beta_{\mathrm{s}} + \beta_{\mathrm{b}} y$. Then

$$\begin{cases} \phi''(k_1) = C_3 \gamma_1 (\gamma_1 - 1) k_1^{\gamma_1 - 2} + C_4 \gamma_2 (\gamma_2 - 1) k_1^{\gamma_2 - 2}, \text{ and} \\ \phi(k_2) = C_3 k_2^{\gamma_1} + C_4 k_2^{\gamma_2} + a_1 + a_2 k_2 - \beta_{\mathrm{s}} + \beta_{\mathrm{b}} y. \end{cases}$$

We need conditions

$$\phi''(k_1) \ge 0 \text{ and } \phi(k_2) \ge 0. \tag{27}$$

Finally, on Γ_3, let $\psi(y) = w_2(y) - \beta_{\mathrm{s}} + \beta_{\mathrm{b}} y$. Then,

$$\psi''(k_2) = C_1 \eta_3 \delta_3 (\delta_3 - 1) k_2^{\delta_3 - 2} + C_2 \eta_4 \delta_4 (\delta_4 - 1) k_2^{\delta_4 - 2}.$$

We need

$$\psi''(k_2) \ge 0 \text{ and } C_1 y^{\delta_3} + C_2 y^{\delta_4} \ge \beta_{\mathrm{s}} - \beta_{\mathrm{b}} y \text{ on } \Gamma_3. \tag{28}$$

Theorem. Assume (B1) and (B2). Assume also (25), (26), (27), and (28) hold. Then, $v(x_1,x_2,\alpha)=x_1 w_\alpha(x_2/x_1)=V(x_1,x_2,\alpha)$, $\alpha=1,2$. Let $D=\{(x_1,x_2,1): x_2>k_1x_1\}\cup\{(x_1,x_2,2): x_2>k_2x_1\}$. Let $\tau^*=\inf\{t: (X_t^1,X_t^2,\alpha_t)\notin D\}$. Then τ^* is optimal. □

Finally, we give an example to illustrate the results.

Example In this example, we take

$$\mu_1(1)=0.20,\quad \mu_2(1)=0.25,\quad \mu_1(2)=-0.30,\ \mu_2(2)=-0.35,$$
$$\sigma_{11}(1)=0.30,\ \sigma_{12}(1)=0.10,\ \sigma_{21}(1)=0.10,\ \sigma_{22}(1)=0.35,$$
$$\sigma_{11}(2)=0.40,\ \sigma_{12}(2)=0.20,\ \sigma_{21}(2)=0.20,\ \sigma_{22}(2)=0.45,$$
$$\lambda_1=6.0,\qquad \lambda_2=10.0,\qquad K=0.001,\qquad \rho=0.50.$$

Then, we use the function $g(r)$ in (23) and find the unique zero $r_0=1.020254>1$. Using this r_0 and (24), we obtain $k_1=0.723270$ and $k_2=0.737920$. Then, we calculate and get $C_1=0.11442$, $C_2=-0.00001$, $C_3=0.29121$, $C_4=0.00029$, $\eta_3=0.985919$, and $\eta_4=-1.541271$. With these numbers, we verify all variational inequalities required in (B2). The graphs of the value functions are given in Figure 9. □

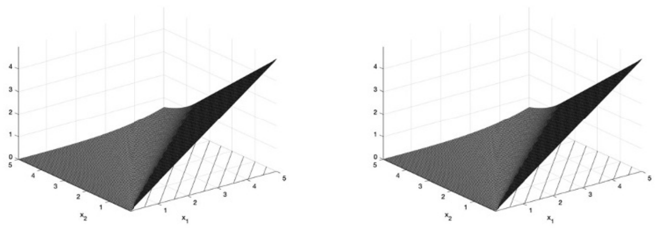

Fig. 9 Value Functions $V(x_1,x_2,1)$ and $V(x_1,x_2,2)$

5 Conclusions

In this paper, we have surveyed pairs trading under geometric Brownian motion models. We were able to obtain closed-form solutions. The trading rules are given in terms of threshold levels and are simple and easy to implement. The major advantage of pairs trading is its risk-neutral nature, i.e., it can be profitable regardless of the general market directions. Pairs trading is a natural extension to McDonald and Siegel's [15] irreversible project investment decision making. We were able to obtain similar results under suitable conditions.

Some initial efforts in connection with numerical computations and implementation have been done in Luu [14]. In particular, stochastic approximation techniques (see Kushner and Yin [12]) can be used to effectively estimate these threshold levels directly. Finally, it would be interesting to examine how these methods work through backtests for a larger selection of stocks.

It would be interesting to extend the results to incorporate more involved models (e.g., models with incomplete observation in market mode α_t). In this case, nonlinear filtering methods such as the Wonham filter can be used for calculation of the conditional probabilities of $\alpha = i$ given the stock prices up to time t. Some ideas along this line have been used in Dai et al. [1] in connection with the trend-following trading.

References

1. Dai, M., Zhang, Q., Zhu, Q.: Trend following trading under a regime switching model. SIAM J. Fin. Math. **1**(1), 780–810 (2010). DOI 10.1137/090770552. URL http://dx.doi.org/10.1137/090770552
2. Davis, M.H.A., Norman, A.R.: Portfolio selection with transaction costs. Mathematics of Operations Research **15**(4), 676–713 (1990). DOI 10.1287/moor.15.4.676. URL https://doi.org/10.1287/moor.15.4.676
3. Deshpande, A., Barmish, B.: A general framework for pairs trading with a control-theoretic point of view. pp. 761–766 (2016). DOI 10.1109/CCA.2016.7587910
4. Do, B., Faff, R.: Does simple pairs trading still work? Financial Analysts Journal **66**(4), 83–95 (2010). DOI 10.2469/faj.v66.n4.1. URL https://doi.org/10.2469/faj.v66.n4.1
5. Do, B., Faff, R.: Are pairs trading profits robust to trading costs? Journal of Financial Research **35**(2), 261–287 (2012). DOI 10.1111/j.1475-6803.2012.01317.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6803.2012.01317.x
6. Elliott, R.J., Hoek, J.V.D., Malcolm, W.P.: Pairs trading. Quantitative Finance **5**(3), 271–276 (2005). DOI 10.1080/14697680500149370. URL https://doi.org/10.1080/14697680500149370
7. Gatev, E., Goetzmann, W., Rouwenhorst, K.: Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies **19**(3), 797–827 (2006). URL https://EconPapers.repec.org/RePEc:oup:rfinst:v:19:y:2006:i:3:p:797-827
8. Guo, X., Zhang, Q.: Optimal selling rules in a regime switching model. IEEE Transactions on Automatic Control **50**(9), 1450–1455 (2005). DOI 10.1109/TAC.2005.854657
9. Hu, Y., Øksendal, B.: Optimal time to invest when the price processes are geometric brownian motions. Finance and Stochastics **2**, 295–310 (1998). DOI 10.1007/s007800050042. URL https://doi.org/10.1007/s007800050042
10. Iwarere, S., Barmish, B.R.: A confidence interval triggering method for stock trading via feedback control. In: Proceedings of the 2010 American Control Conference, pp. 6910–6916 (2010). DOI 10.1109/ACC.2010.5531311
11. Kuo, K., Luu, P., Nguyen, D., Perkerson, E., Thompson, K., Zhang, Q.: Pairs trading: An optimal selling rule. Mathematical Control & Related Fields **5**, 489–499 (2015). DOI 10.3934/mcrf.2015.5.489. URL http://aimsciences.org/article/id/27a4b8fb-5a99-46b4-939e-9a12c9b9bca2

12. Kushner, H., Yin, G.: Stochastic Approximation Algorithms and Applications. Springer, New York (1997)
13. Liu, R., Wu, Z., Zhang, Q.: Pairs-trading under geometric brownian motions: An optimal strategy with cutting losses. Automatica **115**(C) (2020). DOI 10.1016/j.automatica.2020.108912. URL https://doi.org/10.1016/j.automatica.2020.108912
14. Luu, P.: Optimal pairs trading rules and numerical methods. Ph.D. dissertation, University of Georgia (2016)
15. McDonald, R., Siegel, D.: The value of waiting to invest. The Quarterly Journal of Economics **101**(4), 707–727 (1986). URL https://EconPapers.repec.org/RePEc:oup:qjecon:v:101:y:1986:i:4:p:707-727.
16. Merhi, A., Zervos, M.: A model for reversible investment capacity expansion. SIAM Journal on Control and Optimization **46**(3), 839–876 (2007). DOI 10.1137/050640758. URL https://doi.org/10.1137/050640758
17. Shreve, S.E., Soner, H.M.: Optimal investment and consumption with transaction costs. Ann. Appl. Probab. **4**(3), 609–692 (1994). DOI 10.1214/aoap/1177004966. URL https://doi.org/10.1214/aoap/1177004966
18. Song, Q., Zhang, Q.: An optimal pairs-trading rule. Automatica **49**, 3007–3014 (2013)
19. Tie, J., Zhang, H., Zhang, Q.: An optimal strategy for pairs trading under geometric brownian motions. Journal of Optimization Theory and Applications **179**(2), 654–675 (2018)
20. Tie, J., Zhang, Q.: Pairs trading: An optimal selling rule under a regime switching model. Stochastic Modeling and Control (2020)
21. Vidyamurthy, G.: Pairs Trading: Quantitative Methods and Analysis. Wiley Finance. Wiley (2011). URL https://books.google.com/books?id=SQLU014X7JwC
22. Whistler, M.: Trading Pairs – Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies. Wiley, New York (2004)
23. Zariphopoulou, T.: Investment-consumption models with transaction fees and markov-chain parameters. SIAM Journal on Control and Optimization **30**(3), 613–636 (1992). DOI 10.1137/0330035. URL https://doi.org/10.1137/0330035
24. Zhang, H., Zhang, Q.: Trading a mean-reverting asset: Buy low and sell high. Automatica **44**(6), 1511 – 1518 (2008). DOI https://doi.org/10.1016/j.automatica.2007.11.003. URL http://www.sciencedirect.com/science/article/pii/S0005109808000216. Stochastic Modelling, Control, and Robust Optimization at the Crossroads of Engineering, Environmental Economics, and Finance
25. Zhang, Q.: Stock trading: An optimal selling rule. SIAM Journal on Control and Optimization **40**(1), 64–87 (2001). DOI 10.1137/S0363012999356325. URL https://doi.org/10.1137/S0363012999356325

Equilibrium Model of Limit Order Books: A Mean-Field Game View

Jin Ma and Eunjung Noh

Abstract In this paper, we propose a continuous time equilibrium model of the (sell-side) limit order book (LOB) in which the liquidity dynamics follows a non-local, reflected mean-field stochastic differential equation (SDE) with state-dependent intensity. To motivate the model we first study an N-seller static mean-field type Bertrand game among the liquidity providers. We shall then formulate the continuous time model as the limiting mean-field dynamics of the representative seller, and argue that the frontier of the LOB (e.g., the best ask price) is the value function of a mean-field stochastic control problem by the representative seller. Using a dynamic programming approach, we show that the value function is a viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation, which can be used to determine the equilibrium density function of the LOB, in the spirit of [32].

1 Introduction

With the rapid growth of electronic trading, the study of order-driven markets has become an increasingly prominent focus in quantitative finance. Indeed, in the current financial world more than half of the markets use a limit order book (LOB) mechanism to facilitate trade. There has been a large amount of literature studying LOB from various angles, combined with some associated optimization problems such as placement, liquidation, executions, etc. (see, e.g. [1, 3, 4, 5, 7, 14, 18, 20, 30, 35, 36, 39, 40] to mention a few). Among many important structural issues of LOB, one of the focuses has been the dynamic move-

Jin Ma
Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA., e-mail: jinma@usc.edu

Eunjung Noh
Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA., e-mail: en221@math.rutgers.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_16

ment of the LOB, both its frontier and its “density" (or “shape"). The latter was shown to be a determining factor of the “liquidity cost" (cf. [32]), an important aspect that impacts the pricing of the asset. We refer to, e.g., [2, 19, 26, 32] for the study of LOB particularly concerning its shape formation.

In this paper, we assume that all sellers are *patient* and all buyers are *impatient*. We extend dynamic model of LOB proposed in [32] in two major aspects. The guiding idea is to specify the *expected equilibrium utility function*, which plays an essential role in the modeling of the shape of the LOB in that it endogenously determines both the dynamic density of the LOB and its frontier. More precisely, instead of assuming, more or less in an ad hoc manner, that the equilibrium price behaves like an “utility function", we shall consider it as the consequence of a Bertrand-type game among a large number of liquidity providers (sellers who set limit orders). Following the argument of [13], we first study an N-seller static Bertrand game, where a profit function of each seller involves not only the limit order price less the waiting cost, but also the average of the other sellers' limit order prices observed. We show that the Nash equilibrium exists in such a game. With an easy randomization argument, we can then show that, as $N \to \infty$, the Nash equilibrium converges to an optimal strategy of a single player's optimization problem with a mean-field nature, as expected.

We note that the Bertrand game in finance can be traced back to as early as 1800s, when Cournot [15] and Bertrand [8] first studied oligopoly models of markets with a small number of competitive players. We refer to [17] and [41] for background and references. Since Cournot's model uses quantity as a strategic variable to determine the price, while Bertrand model does the opposition, we choose to use the Bertrand game as it fits our problem better. We shall assume that the sellers use the same marginal profit function, but with different choices of the price-waiting cost preference to achieve the optimal outcome (see Sect. 3 for more detailed formulation).

We would like to point out that our study of Bertrand game is in a sense “motivational" for the second main feature of this paper, that is, the continuous time, mean-field type dynamic liquidity model. More precisely, we assume that the liquidity dynamics is a pure-jump Markov process, with a mean-field type state dependent jump intensity. Such a dynamic game is rather complicated, and is expected to involve systems of nonlinear, mean-field type partial differential equations (see, e.g., [23, 27]). We therefore consider the limiting case as the number of sellers tends to infinity, and argue that the dynamics of the total liquidity follows a pure jump SDE with reflecting boundary conditions and mean-field-type state-dependent jump intensity. We note that such SDE is itself new and therefore interesting in its own right.

We should point out that the special features of our underlying liquidity dynamics (mean-field type; state-dependent intensity; and reflecting boundary conditions) require the combined technical tools in mean-field games, McKean-Vlasov SDEs with state-dependent jump intensity, and SDEs with discontinuous paths and reflecting boundary conditions. In particular, we refer to the works [10, 11, 12, 21, 22, 24, 25, 29, 31, 33, 37] (and the references cited therein) for the technical foundation of this paper. Furthermore, apart from justifying the underlying liquidity dynamics, another main task of this paper is to substantiate the corresponding stochastic control problem, including validating the dynamic programming principle (DPP) and showing that the value function is a viscosity solution to the corresponding Hamilton-Jacobi-Bellman (HJB) equation.

This paper is organized as follows. In Sect. 2, we introduce necessary notations and preliminary concepts, and study the well-posedness of a reflected mean-field SDEs with jumps that will be essential in our study. We shall also provide an Itô's formula involving reflected mean-field SDEs with jumps for ready reference. In Sect. 3 we investigate a static Bertrand game with N sellers, and its limiting behavior as N tends to infinity. Based on the results, we then propose in Sect. 4 a continuous time mean-field type dynamics of the representative seller, as well as a mean-field stochastic control problem as the limiting version of dynamic Bertrand game when the number of sellers becomes sufficiently large. In Sect. 5 and Sect. 6 we validate the *Dynamic Programming Principle* (DPP), derive the HJB equation, and show that the value function is a viscosity solution to the corresponding HJB equation.

2 Preliminaries

Throughout this paper we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which is defined two standard Brownian motions $W = \{W_t : t \geq 0\}$ and $B = \{B_t : t \geq 0\}$. Let $(\mathcal{A}, \mathscr{B}_{\mathcal{A}})$ and $(\mathcal{B}, \mathscr{B}_{\mathcal{B}})$ be two measurable spaces. We assume that there are two Poisson random measures $\mathcal{N}^s$ and $\mathcal{N}^b$, defined on $\mathbb{R}_+ \times \mathcal{A} \times \mathbb{R}_+$ and $\mathbb{R}_+ \times \mathcal{B}$, and with Lévy measures $\nu^s(\cdot)$ and $\nu^b(\cdot)$, respectively. In other words, we assume that the Poisson measures $\mathcal{N}^s$ and $\mathcal{N}^b$ have mean measures $\widehat{\mathcal{N}}^s(\cdot) := (m \times \nu^s \times m)(\cdot)$ and $\widehat{\mathcal{N}}^b(\cdot) := (m \times \nu^b)(\cdot)$, respectively, where $m(\cdot)$ denotes the Lebesgue measure on $\mathbb{R}_+$, and we denote the *compensated* random measures $\widetilde{\mathcal{N}}^s(A) := (\mathcal{N}^s - \widehat{\mathcal{N}}^s)(A) = \mathcal{N}^s(A) - (m \times \nu^s \times m)(A)$ and $\widetilde{\mathcal{N}}^b(B) := (\mathcal{N}^b - \widehat{\mathcal{N}}^b)(B) = \mathcal{N}^b(B) - (m \times \nu^b)(B)$, for any $A \in \mathscr{B}(\mathbb{R}_+ \times \mathcal{A} \times \mathbb{R}_+)$ and $B \in \mathscr{B}(\mathbb{R}_+ \times \mathcal{B})$. For simplicity, throughout this paper we assume that both ν^s and ν^b are finite, that is, $\nu^s(\mathcal{A}), \nu^b(\mathcal{B}) < \infty$, and we assume the Brownian motions and Poisson random measures are mutually independent. We note that for any $A \in \mathscr{B}(\mathcal{A} \times \mathbb{R}_+)$ and $B \in \mathscr{B}(\mathcal{B})$, the processes $(t, \omega) \mapsto \widetilde{\mathcal{N}}^s([0,t] \times A, \omega)$, $\widetilde{\mathcal{N}}^b([0,t] \times B, \omega)$ are both $\mathbb{F}^{\mathcal{N}^s, \mathcal{N}^b}$-martingales. Here $\mathbb{F}^{\mathcal{N}^s, \mathcal{N}^b}$ denotes the filtration generated by $\mathcal{N}^s$ and $\mathcal{N}^b$.

For a generic Euclidean space E and for $T > 0$, we denote $C([0,T]; E)$ and $\mathbb{D}([0,T]; E)$ to be the spaces of continuous and càdlàg functions, respectively. We endow both spaces with "sup-norms", so that both of them are complete metric spaces. Next, for $p \geq 1$ we denote $L^p(\mathcal{F}; E)$ to be the space of all E-valued $\mathcal{F}$-measurable random variable ξ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}[|\xi|^p] < \infty$. Also, for $T \geq 0$, we denote $L^p_{\mathbb{F}}([t,T]; E)$ to be all E-valued $\mathbb{F}$-adapted process η on $[t,T]$, such that $\|\eta\|_{p,T} := \mathbb{E}[\int_t^T |\eta_s|^p ds]^{1/p} < \infty$. We often use the notations $L^p(\mathbb{F}; C([0,T]; E))$ and $L^p(\mathbb{F}; \mathbb{D}([0,T]; E))$ when we need to specify the path properties for elements in $L^p_{\mathbb{F}}([0,T]; E)$.

For $p \geq 1$ we denote by $\mathscr{P}_p(E)$ the space of probability measures μ on $(E, \mathscr{B}(E))$ with finite p-th moment, i.e. $\|\mu\|_p^p := \int_E |x|^p \mu(dx) < \infty$. Clearly, for $\xi \in L^p(\mathcal{F}; E)$, its law $\mathcal{L}(\xi) = \mathbb{P}_\xi := \mathbb{P} \circ \xi^{-1} \in \mathscr{P}_p(E)$. We endow $\mathscr{P}_p(E)$ with the following p-Wasserstein metric:

$$
\begin{aligned}
W_p(\mu,\nu) &:= \inf\Big\{\Big(\int_{E\times E}|x-y|^p\pi(dx,dy)\Big)^{\frac{1}{p}} : \pi \in \mathscr{P}_p(E\times E) \\
&\qquad\qquad \text{with marginals } \mu \text{ and } \nu\Big\} \qquad (1)\\
&= \inf\Big\{\|\xi-\xi'\|_{L^p(\Omega)} : \xi,\xi' \in L^p(\mathcal{F};E) \text{ with } \mathbb{P}_\xi = \mu,\ \mathbb{P}_{\xi'} = \nu\Big\}.
\end{aligned}
$$

Furthermore, we suppose that there is a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$ such that (i) the Brownian motion W and Poisson random measures $\mathcal{N}^s, \mathcal{N}^b$ are independent of $\mathcal{G}$; and (ii) $\mathcal{G}$ is "rich enough" in the sense that for every $\mu \in \mathscr{P}_2(\mathbb{R})$, there is a random variable $\xi \in L^2(\mathcal{G}; E)$ such that $\mu = \mathbb{P}_\xi$. Let $\mathbb{F} = \mathbb{F}^{W,B,\mathcal{N}^s,\mathcal{N}^b \vee \mathcal{G}} = \{\mathcal{F}_t\}_{t\geq 0}$, where $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^B \vee \mathcal{F}_t^{\mathcal{N}^s} \vee \mathcal{F}_t^{\mathcal{N}^b} \vee \mathcal{G}$, $t \geq 0$, be the filtration generated by W, B, $\mathcal{N}^s$, $\mathcal{N}^b$, and $\mathcal{G}$, augmented by all the $\mathbb{P}$-null sets so that it satisfies the *usual hypotheses* (cf. [38]).

Let us introduce two spaces that are useful for our analysis later. We write $C_b^{1,1}(\mathscr{P}_2(\mathbb{R}))$ to denote the space of all differentiable functions $f : \mathscr{P}_2(\mathbb{R}) \to \mathbb{R}$ such that $\partial_\mu f$ exists, and is bounded and Lipschitz continuous. That is, for some constant $C > 0$, it holds

(i) $|\partial_\mu f(\mu, x)| \leq C$, $\mu \in \mathscr{P}_2(\mathbb{R})$, $x \in \mathbb{R}$;

(ii) $|\partial_\mu f(\mu, x) - \partial_\mu f(\mu', x')| \leq C\{|x - x'| + W_2(\mu, \mu')\}$, $\mu, \mu' \in \mathscr{P}_2(\mathbb{R})$, $x, x' \in \mathbb{R}$.

We shall denote $C_b^{2,1}(\mathscr{P}_2(\mathbb{R}))$ to be the space of all functions $f \in C_b^{1,1}(\mathscr{P}_2(\mathbb{R}))$ such that

(i) $\partial_\mu f(\cdot, x) \in C_b^{1,1}(\mathscr{P}_2(\mathbb{R}))$ for all $x \in \mathbb{R}$;

(ii) $\partial_\mu^2 f : \mathscr{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \otimes \mathbb{R}$ is bounded and Lipschitz continuous;

(iii) $\partial_\mu f(\mu, \cdot) : \mathbb{R} \to \mathbb{R}$ is differentiable for every $\mu \in \mathscr{P}_2(\mathbb{R})$, and its derivative $\partial_y \partial_\mu f : \mathscr{P}_2(\mathbb{R}) \times \mathbb{R} \to \mathbb{R} \otimes \mathbb{R}$ is bounded and Lipschitz continuous.

2.1 Mean-field SDEs with reflecting boundary conditions

In this subsection we consider the following (discontinuous) SDE with reflection, which will be a key element of our discussion: for $t \in [0, T]$,

$$
\begin{aligned}
X_s &= x + \int_t^s \int_{A\times\mathbb{R}_+} \theta(r, X_{r-}, \mathbb{P}_{X_r}, z)\mathbf{1}_{[0,\lambda(r,X_{r-},\mathbb{P}_{X_r})]}(y)\widetilde{\mathcal{N}}^s(drdzdy) \qquad (2)\\
&\quad + \int_t^s b(r, X_r, \mathbb{P}_{X_r})dr + \int_t^s \sigma(r, X_r, \mathbb{P}_{X_r})dB_r + \beta_s + K_s, \qquad s \in [t, T],
\end{aligned}
$$

where θ, λ, b, σ are measurable functions defined on appropriate subspaces of $[0,T]\times\Omega\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R})\times\mathbb{R}$, β is an $\mathbb{F}$-adapted process with càdlàg paths, and K is a "reflecting process". That is, it is an $\mathbb{F}$-adapted, non-decreasing, and càdlàg process so that

(i) $X_\cdot \geq 0$, $\mathbb{P}$-a.s.;

(ii) $\int_0^T \mathbf{1}_{\{X_r>0\}} dK_r^c = 0$, $\mathbb{P}$-a.s. (K^c denotes the continuous part of K); and

(iii) $\Delta K_t = (X_{t-}+\Delta Y_t)^-$ for all $t\in[0,T]$, where $Y = X - K$.

We call SDE (2) a *mean-field SDE with discontinuous paths and reflections* (MFSDEDR), and we denote the solution by $(X^{t,x}, K^{t,x})$, although the superscript is often omitted when context is clear. If $b,\sigma = 0$ and β is pure jump, then the solution (X,K) becomes pure jump as well (i.e., $dK^c \equiv 0$). We note that the main feature of this SDE is that the jump intensity $\lambda(\cdots)$ of the solution X is "state-dependent" with mean-field nature. Its well-posedness thus requires some attention since, to the best of our knowledge, it has not been studied in the literature.

We shall make use of the following *Standing Assumptions*.

Assumption 2.1 *The mappings* $\lambda : [0,T]\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R}) \mapsto \mathbb{R}_+$, $b : [0,T]\times\Omega\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R}) \mapsto \mathbb{R}$, $\sigma : [0,T]\times\Omega\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R}) \mapsto \mathbb{R}$, *and* $\theta : [0,T]\times\Omega\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R})\times\mathbb{R} \mapsto \mathbb{R}$ *are all uniformly bounded and continuous in* (t,x), *and satisfy the following conditions, respectively:*

(i) For fixed $\mu\in\mathscr{P}_2(\mathbb{R})$ *and* $x,z\in\mathbb{R}$, *the mappings* $(t,\omega)\mapsto\theta(t,\omega,x,\mu,z)$, $(b,\sigma)(t,\omega,x,\mu)$ *are* $\mathbb{F}$*-predictable;*

(ii) For fixed $\mu\in\mathscr{P}_2(\mathbb{R})$, $(t,z)\in[0,T]\times\mathbb{R}$, *and* $\mathbb{P}$*-a.e.* $\omega\in\Omega$, *the functions* $\lambda(t,\cdot,\mu)$, $b(t,\omega,\cdot,\mu)$, $\sigma(t,\omega,\cdot,\mu)$, $\theta(t,\omega,\cdot,\mu,z)\in C_b^1(\mathbb{R})$;

(iii) For fixed $(t,x,z)\in[0,T]\times\mathbb{R}^2$, *and* $\mathbb{P}$*-a.e.* $\omega\in\Omega$, *the functions* $\lambda(t,x,\cdot)$, $b(t,\omega,x,\cdot)$, $\sigma(t,\omega,x,\cdot)$, $\theta(t,\omega,x,\cdot,z)\in C_b^{1,1}(\mathscr{P}_2(\mathbb{R}))$;

(iv) There exists $L>0$, *such that for* $\mathbb{P}$*-a.e.* $\omega\in\Omega$, *it holds that*

$$
\begin{aligned}
&|\lambda(t,x,\mu)-\lambda(t,x',\mu')|+|b(t,\omega,x,\mu)-b(t,\omega,x',\mu')| \\
&+|\sigma(t,\omega,x,\mu)-\sigma(t,\omega,x',\mu')|+|\theta(t,\omega,x,\mu,z)-\theta(t,\omega,x',\mu',z)| \\
&\qquad\leq L\left(|x-x'|+W_1(\mu,\mu')\right), \qquad t\in[0,T],\ x,x',z\in\mathbb{R},\ \mu,\mu'\in\mathscr{P}_2(\mathbb{R}).
\end{aligned}
$$

Remark 2.2 (i) The requirements on the coefficients in Assumption 2.1 (such as boundedness) are stronger than necessary, only to simplify the arguments. More general (but standard) assumptions are easily extendable without substantial difficulties.

(ii) Throughout this paper, unless specified, we shall denote $C>0$ to be a generic constant depending only on T and the bounds in Assumption 2.1. Furthermore, we shall allow it to vary from line to line.

It is well-known that (see, e.g., [9]), as a mean-field SDE, the solution to (2) may not satisfy the so-called "flow property", in the sense that $X_r^{t,x} \neq X_r^{s,X_s^{t,x}}$, $0\leq t\leq s\leq r\leq T$. It is also noted in [9] that if we consider the following accompanying SDE of (2): for $s\in[t,T]$,

$$X_s^{t,\xi} = \xi + \int_t^s b(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}})dr + \int_t^s \sigma(r, X_r^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}})dB_r + \beta_s + K_s^{t,\xi}$$
$$+ \int_t^s \int_{A\times\mathbb{R}_+} \theta(r, X_{r-}^{t,\xi}, \mathbb{P}_{X_r^{t,\xi}}, z)\mathbf{1}_{[0,\lambda(r,X_{r-}^{t,\xi},\mathbb{P}_{X_r^{t,\xi}})]}(y)\widetilde{\mathcal{N}}^s(drdzdy) \quad (3)$$

and then using the law $\mathbb{P}_{X^{t,\xi}}$ to consider a slight variation of (3):

$$X_s^{t,x,\xi} = x + \int_t^s b(r, X_r^{t,x,\xi}, \mathbb{P}_{X_r^{t,\xi}})dr + \int_t^s \sigma(r, X_r^{t,x,\xi}, \mathbb{P}_{X_r^{t,\xi}})dB_r + \beta_s + K_s^{t,x,\xi}$$
$$+ \int_t^s \int_{A\times\mathbb{R}_+} \theta(r, X_{r-}^{t,x,\xi}, \mathbb{P}_{X_r^{t,\xi}}, z)\mathbf{1}_{[0,\lambda(r,X_{r-}^{t,x,\xi},\mathbb{P}_{X_r^{t,\xi}})]}(y)\widetilde{\mathcal{N}}^s(drdzdy), \quad (4)$$

where $\xi \in L^2(\mathcal{F}_t;\mathbb{R})$, then we shall argue below that the following flow property holds:

$$\left(X_r^{s,X_s^{t,x,\xi},X_s^{t,\xi}}, X_r^{s,X_s^{t,\xi}}\right) = (X_r^{t,x,\xi}, X_r^{t,\xi}), \qquad 0 \le t \le s \le r \le T, \qquad (5)$$

for all $(x,\xi) \in \mathbb{R}\times L^2(\mathcal{F}_t;\mathbb{R})$. We should note that although both SDEs (3) and (4) resemble the original equation (2), the process $X^{t,x,\xi}$ has the full information of the solution given the initial data (x,ξ), where ξ provides the initial distribution $\mathbb{P}_\xi$, and x is the actual initial state.

To prove the well-posedness of SDEs (3) and (4), we first recall the so-called "Discontinuous Skorohod Problem" (DSP) (see, e.g., [16, 31]). Let $Y \in \mathbb{D}([0,T])$, $Y_0 \ge 0$. We say that a pair $(X,K) \in \mathbb{D}([0,T])^2$ is a solution to the DSP(Y) if

(i) $X = Y + K$;

(ii) $X_t \ge 0$, $t \ge 0$; and

(iii) K is nondecreasing, $K_0 = 0$, and $K_t = \int_0^t \mathbf{1}_{\{X_{s-}=0\}}dK_s$, $t \ge 0$.

It is well-known that the solution to DSP exists and is unique, and it can be shown (see [31]) that the condition (iii) amounts to saying that $\int_0^t \mathbf{1}_{\{X_{s-}>0\}}dK_s^c = 0$, where K^c denotes the continuous part of K, and $\Delta K_t = (X_{t-} + \Delta Y_t)^-$. Furthermore, it is shown in [16] that solution mapping of the DSP, $\Gamma : \mathbb{D}([0,T]) \mapsto \mathbb{D}([0,T])$, defined by $\Gamma(Y) = X$, is Lipschitz continuous under uniform topology. That is, there exists a constant $L > 0$ such that

$$\sup_{t\in[0,T]} |\Gamma(Y^1)_t - \Gamma(Y^2)_t| \le L \sup_{t\in[0,T]} |Y_t^1 - Y_t^2|, \qquad Y^1, Y^2 \in \mathbb{D}([0,T]). \qquad (6)$$

Before we proceed to prove the well-posedness of (3) and (4), we note that the two SDEs can be argued separately. Moreover, while (3) is a mean-field (or McKean-Vlasov)-type of SDE, (4) is actually a standard SDE (although with state-dependent intensity) with discontinuous paths and reflection, given the law of the solution to (3), $\mathbb{P}_{X^{t,\xi}}$, and it can be argued similarly but much simpler. Therefore, in what follows we shall focus only on the well-posedness of SDE (3). Furthermore, for simplicity we shall assume $b \equiv 0$, as the general case can be argued similarly without substantial difficulty.

The scheme of solving the SDE (3) is more or less standard (see, e.g., [31]). We shall first consider an SDE without reflection: for $\xi \in L^2(\mathcal{F}_t;\mathbb{R})$ and $s \in [t,T]$,

$$\begin{aligned} Y_s^{t,\xi} &= \xi + \int_t^s \int_{A\times\mathbb{R}_+} \theta(r,\Gamma(Y^{t,\xi})_{r-},\mathbb{P}_{\Gamma(Y^{t,\xi})_r},z)\mathbf{1}_{[0,\lambda_{r-}^{\Gamma(t,\xi)}]}(y)\widetilde{\mathcal{N}}^s(drdzdy) \\ &\quad + \int_t^s \sigma(r,\Gamma(Y^{t,\xi})_r,\mathbb{P}_{\Gamma(Y^{t,\xi})_r})dB_r + \beta_s, \end{aligned} \tag{7}$$

where $\lambda_{r-}^{\Gamma(t,\xi)} := \lambda(r,\Gamma(Y^{t,\xi})_{r-},\mathbb{P}_{\Gamma(Y^{t,\xi})_r})$. Clearly, if we can show that (7) is well-posed, then by simply setting $X_s^{t,\xi} = \Gamma(Y^{t,\xi})_s$ and $K_s^{t,\xi} = X_s^{t,\xi} - Y_s^{t,\xi}$, $s \in [t,T]$, we see that $(X^{t,\xi},K^{t,\xi})$ would solve SDE (3)(!). We should note that a technical difficulty caused by the presence of the state-dependent intensity is that the usual L^2-norm does not work as naturally as expected, as we shall see below. We nevertheless have the following result.

Theorem 2.3 *Assume that Assumptions 2.1 is in force. Then, there exists a solution $Y^{t,\xi} \in L^2_{\mathbb{F}}(\mathbb{D}([t,T]))$ to SDE (7). Furthermore, such solution is pathwisely unique.*

Proof Assume $t = 0$. For a given $T_0 > 0$, and $y \in L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$, consider a mapping $\mathscr{T}$:

$$\begin{aligned} \mathscr{T}(y)_s &:= \xi + \int_0^s \int_{A\times\mathbb{R}_+} \theta(r,\Gamma(y)_{r-},\mathbb{P}_{\Gamma(y)_r},z)\mathbf{1}_{[0,\lambda(r,\Gamma(y)_{r-},\mathbb{P}_{\Gamma(y)_r})]}(u)\widetilde{\mathcal{N}}^s(drdzdu) \\ &\quad + \int_0^s \sigma(r,\Gamma(y)_r,\mathbb{P}_{\Gamma(y)_r})dB_r + \beta_s, \qquad s \geq 0. \end{aligned} \tag{8}$$

We shall argue that $\mathscr{T}$ is a contraction mapping on $L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$ for $T_0 > 0$ small enough.

To see this, denote, for $\eta \in \mathbb{D}([0,T_0])$, $|\eta|_s^* := \sup_{0\leq r\leq s}|\eta_r|$, and define $\theta_s(z) := \theta(s,\Gamma(y)_s,\mathbb{P}_{\Gamma(y)_s},z)$, $\lambda_s := \lambda(s,\Gamma(y)_s,\mathbb{P}_{\Gamma(y)_s})$, $\sigma_s := \sigma(s,\Gamma(y)_s,\mathbb{P}_{\Gamma(y)_s})$, $s \in [0,T_0]$. Then, we have

$$\begin{aligned} \mathbb{E}[|\mathscr{T}(y)|_{T_0}^*] &\leq C\Big\{\mathbb{E}|\xi| + \mathbb{E}\Big[\int_0^{T_0}\int_{A\times\mathbb{R}_+} |\theta_r(z)\mathbf{1}_{[0,\lambda_r]}(u)|\nu^s(dz)dudr\Big] \\ &\qquad + \mathbb{E}\Big[\Big(\int_0^{T_0}|\sigma_r|^2dr\Big)^{1/2}\Big]\Big\} \\ &\leq C\mathbb{E}|\xi| + C\mathbb{E}\Big[\int_0^{T_0}\int_A |\theta_r(z)\lambda_r|\nu^s(dz)dr\Big] \\ &\qquad + C\mathbb{E}\Big[\Big(\int_0^{T_0}|\sigma_r|^2dr\Big)^{1/2}\Big] < \infty, \end{aligned}$$

thanks to Assumption 2.1. Hence, $\mathscr{T}(y) \in L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$.

We now show that $\mathscr{T}$ is a contraction mapping on $L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$. For $y_1, y_2 \in L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$, we denote θ^i, λ^i, and σ^i, respectively, as before, and denote $\Delta\varphi :=$

$\varphi^1 - \varphi^2$, for $\varphi = \theta, \lambda, \sigma$, and $\Delta\mathscr{T}(s) := \mathscr{T}(y_1)_s - \mathscr{T}(y_2)_s$, $s \geq 0$. Then, we have, for $s \in [0, T_0]$,

$$\Delta\mathscr{T}(s) = \int_0^s \Delta\sigma_r dB_r + \int_0^s \int_{A\times\mathbb{R}_+} \Big[\Delta\theta_r(z)\mathbf{1}_{[0,\lambda_r^1]}(y) \\ +\theta_r^2(z)\big(\mathbf{1}_{[0,\lambda_r^1]}(y) - \mathbf{1}_{[0,\lambda_r^2]}(y)\big)\Big]\widetilde{\mathcal{N}}^s(drdzdy).$$

Clearly, $\Delta\mathscr{T} = \mathscr{T}(y_1) - \mathscr{T}(y_2)$ is a martingale on $[0, T_0]$. Since $\widetilde{\mathcal{N}} = \mathcal{N} - \widehat{\mathcal{N}}$ and $|\mathbf{1}_{[0,a]}(\cdot) - \mathbf{1}_{[0,b]}(\cdot)| \leq \mathbf{1}_{[a\wedge b, a\vee b]}(\cdot)$ for any $a, b \in \mathbb{R}$, we have, for $0 \leq s \leq T_0$,

$$\mathbb{E}|\Delta\mathscr{T}|_s^* \leq \mathbb{E}\Big[\Big(\int_0^s |\Delta\sigma_r|^2 dr\Big)^{\frac{1}{2}}\Big] + 2\mathbb{E}\Big[\int_0^s \int_{A\times\mathbb{R}_+} |\theta_r^2(z)\big(\mathbf{1}_{[0,\lambda_r^1]}(y) - \mathbf{1}_{[0,\lambda_r^2]}(y)\big) \\ +\Delta\theta_r(z)\mathbf{1}_{[0,\lambda_r^1]}(y)|\nu^s(dz)dydr\Big] := I_1 + I_2. \tag{9}$$

Recalling from Remark 2.2-(ii) for the generic constant $C > 0$, and by Assumption 2.1-(iv), (6), and the definition of $W_1(\cdot,\cdot)$ (see (1)), we have

$$\begin{aligned} I_1 &\leq C\mathbb{E}\Big[\Big(\int_0^s \{|y_1 - y_2|_r^{*,2} + W_1(\mathbb{P}_{\Gamma(y_1)_r}, \mathbb{P}_{\Gamma(y_2)_r})^2\}dr\Big)^{1/2}\Big] \\ &\leq C\mathbb{E}[\sqrt{s}(|y_1 - y_2|_s^* + \mathbb{E}|y_1 - y_2|_s^*)] \leq C\sqrt{T_0}\|y_1 - y_2\|_{L^1(\mathbb{D}([0,T_0]))} \\ I_2 &\leq C\Big(\mathbb{E}\Big[\int_0^s \int_A |\Delta\theta_r(z)|\nu^s(dz)dr\Big] + \mathbb{E}\Big[\int_0^s |\Delta\lambda_r|dr\Big]\Big) \\ &\leq C\mathbb{E}\Big[\int_0^s \Big(|\Gamma(y_1)_r - \Gamma(y_2)_r| + W_1(\mathbb{P}_{\Gamma(y_1)_r}, \mathbb{P}_{\Gamma(y_2)_r})\Big)dr\Big] \\ &\leq C\mathbb{E}\Big[\int_0^s |\Gamma(y_1) - \Gamma(y_2)|_r^* dr\Big] \leq CT_0\|y_1 - y_2\|_{L^1(\mathbb{D}([0,T_0]))}. \end{aligned} \tag{10}$$

Combining (9) and (10), we deduce that

$$\|\Delta\mathscr{T}\|_{L^1(\mathbb{D}([0,T_0]))} \leq C(T_0 + \sqrt{T_0})\|y_1 - y_2\|_{L^1(\mathbb{D}([0,T_0]))}, \qquad s \in [0, T_0]. \tag{11}$$

Therefore, by choosing T_0 such that $C(T_0 + \sqrt{T_0}) < 1$, we see that the mapping $\mathscr{T}$ is a contraction on $L^1(\mathbb{D}([0,T_0]))$, which implies that (7) has a unique solution in $L^1_{\mathbb{F}}(\mathbb{D}([0,T_0]))$. Moreover, we note that T_0 depends only on the universal constant in Assumption 2.1. We can repeat the argument for the time interval $[T_0, 2T_0], [2T_0, 3T_0], \cdots$, and conclude that (7) has a unique solution in $L^1_{\mathbb{F}}(\mathbb{D}([0,T]))$ for any given $T > 0$.

Finally, we claim that the solution $Y \in L^2_{\mathbb{F}}(\mathbb{D}([0,T]))$. Indeed, by Burkholder-Davis-Gundy's inequality and Assumption 2.1, we have

$$
\begin{aligned}
\mathbb{E}[|Y|_s^{*,2}] &\le C\Big\{\mathbb{E}|\xi|^2+\mathbb{E}\Big[\int_0^s\int_{A\times\mathbb{R}_+}|\theta_r(z)\mathbf{1}_{[0,\lambda_r]}(y))|^2\nu^s(dz)dydr\Big] \\
&\quad +\mathbb{E}[\int_0^s|\sigma_r|^2dr]+\mathbb{E}|\beta|_T^{*,2}\Big\} \qquad (12)\\
&\le C\Big\{\mathbb{E}|\xi|^2+\mathbb{E}\Big[\int_0^s[1+|Y_r|^2+W_1(0,\Gamma(Y)_r)]^2dr\Big]+\mathbb{E}|\beta|_T^{*,2}\Big\}\\
&\le C\Big\{\mathbb{E}|\xi|^2+\int_0^s(1+\mathbb{E}[|Y|_r^{*,2}])dr+\mathbb{E}|\beta|_T^{*,2}\Big\}, \qquad s\in[0,T].
\end{aligned}
$$

Here, in the last inequality above we used the fact that

$$W_1(0,\Gamma(Y)_r)^2\le(\|\Gamma(Y)_r\|_{L^1(\Omega)})^2\le(\mathbb{E}|\Gamma(Y)|_r^*)^2\le C\mathbb{E}[|Y|_r^{*,2}], \qquad r\in[0,s].$$

Applying the Gronwall inequality, we obtain that $\mathbb{E}[|Y|_T^{*,2}]<\infty$. The proof is now complete. □

Remark 2.4 (i) It is worth noting that once we solved $X^{t,\xi}$, then we know $\mathbb{P}_{X^{t,\xi}}$, and (4) can be viewed as a standard SDEDR with coefficient $\tilde\lambda(s,x):=\lambda(s,x,\mathbb{P}_{X_s^{t,\xi}})$, which is Lipschitz in x. This guarantees the existence and uniqueness of the solution $(X^{t,x,\xi},K^{t,x,\xi})$ to (4).

(ii) The uniqueness of the solutions to (3) and (4) implies that $X_s^{t,x,\xi}\,|_{x=\xi}=X_s^{t,\xi}$, $s\in[t,T]$. That is, $X_s^{t,x,\xi}\,|_{x=\xi}$ solves the same SDE as $X_s^{t,\xi}$, $s\in[t,T]$. (See more detail in [34].)

(iii) Given $(t,x)\in[0,T]\times\mathbb{R}$, if $\mathbb{P}_{\xi_1}=\mathbb{P}_{\xi_2}$ for $\xi_1,\xi_2\in L^2(\mathcal{F}_t;\mathbb{R})$, then X^{t,x,ξ_1} and X^{t,x,ξ_2} are indistinguishable. So, $X^{t,x,\mathbb{P}_\xi}:=X^{t,x,\xi}$, i.e. $X^{t,x,\xi}$ depends on ξ only through its law.

2.2 An Itô's formula

We shall now present an Itô's formula that will be frequently used in our future discussion. We note that a similar formula for mean-field SDE can be found in [9], and the one involving jumps was given in the recent work [22]. The one presented below is a slight modification of that of [22], taking the particular state-dependent intensity feature of the dynamics into account. Since the proof is more or less standard but quite lengthy, we refer to [34] for the details.

In what follows we let $(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ be a copy of the probability space $(\Omega,\mathcal{F},\mathbb{P})$, and denote $\tilde{\mathbb{E}}[\cdot]$ to be the expectation under $\tilde{\mathbb{P}}$. For any random variable ϑ defined on $(\Omega,\mathcal{F},\mathbb{P})$, we denote, when there is no danger of confusion, $\tilde\vartheta\in(\tilde\Omega,\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ to be a copy of ϑ such that $\tilde{\mathbb{P}}_{\tilde\vartheta}=\mathbb{P}_\vartheta$. We note that that $\tilde{\mathbb{E}}[\cdot]$ acts only on the variables of the form $\tilde\vartheta$.

We first define the following classes of functions.

Definition 2.5 *We say that* $F\in C_b^{1,2,(2,1)}([0,T]\times\mathbb{R}\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R}))$*, if* *(i)* $F(t,v,\cdot,\cdot) \in C_b^{2,1}(\mathbb{R}\times\mathscr{P}_2(\mathbb{R}))$*, for all* $t\in[0,T]$ *and* $v\in\mathbb{R}$*;*

(ii) $F(\cdot,v,x,\mu)\in C_b^1([0,T])$*, for all* $(v,x,\mu)\in\mathbb{R}\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R})$*;*

(iii) $F(t,\cdot,x,\mu)\in C_b^2(\mathbb{R})$*, for all* $(t,x,\mu)\in[0,T]\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R})$*;*

(iv) All derivatives involved in the definitions above are uniformly bounded over $[0,T]\times\mathbb{R}\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R})$ *and Lipschitz continuous in* (x,μ)*, uniformly with respect to* t*.*

We are now ready to state the Itô's formula. Let $V^{t,v}$ be an Itô process given by

$$V_s^{t,v} = v + \int_t^s b^V(r,V_r^{t,v})dr + \int_t^s \sigma^V(r,V_r^{t,v})dB_r^V, \tag{13}$$

where $v\in\mathbb{R}$ and $(B_t^V)_{t\in[0,T]}$ is a standard Brownian motion independent of $(B_t)_{t\in[0,T]}$. For notational simplicity, in what follows for the coefficients $\varphi = b,\sigma,\beta,\lambda$, we denote $\varphi_s^{t,x,\xi} := \varphi(s,X_s^{t,x,\xi},\mathbb{P}_{X_s^{t,\xi}})$, $\theta_s^{t,x,\xi}(z) := \theta(s,X_s^{t,x,\xi},\mathbb{P}_{X_s^{t,\xi}},z)$, $\tilde{\varphi}_s^{t,\xi} := \varphi(s,\tilde{X}_s^{t,\tilde{\xi}},\mathbb{P}_{X_s^{t,\xi}})$, and $\tilde{\theta}_s^{t,\xi}(z) := \theta(s,\tilde{X}_s^{t,\tilde{\xi}},\mathbb{P}_{X_s^{t,\xi}},z)$. Similarly, denote $b_s^{t,v} := b^V(s,V_s^{t,v})$ and $\sigma_s^{t,v} := \sigma^V(s,V_s^{t,v})$. Also, let us write $\Theta_s^t := (s,V_s^{t,v},X_s^{t,x,\xi},\mathbb{P}_{X_s^{t,\xi}})$, from which we have $\Theta_t^t = (t,v,x,\mathbb{P}_\xi)$.

Proposition 2.6 (Itô's Formula) *Let* $\Phi\in C_b^{1,2,(2,1)}([0,T]\times\mathbb{R}\times\mathbb{R}\times\mathscr{P}_2(\mathbb{R}))$*, and* $(X^{t,\xi},X^{t,x,\xi},V^{t,v})$ *be the solutions to (3), (4) and (13), respectively, on* $[t,T]$*. Then, for* $0\le t\le s\le T$*, it holds*

$$\begin{aligned}
&\Phi(\Theta_s^t)-\Phi(\Theta_t^t)\\
&=\int_t^s \Big(\partial_t\Phi(\Theta_r^t)+\partial_x\Phi(\Theta_r^t)b_r^{t,x,\xi}+\frac{1}{2}\partial_{xx}^2\Phi(\Theta_r^t)(\sigma_r^{t,x,\xi})^2+\partial_v\Phi(\Theta_r^t)b_r^{t,v}\\
&\qquad+\frac{1}{2}\partial_{vv}^2\Phi(\Theta_r^t)(\sigma_r^{t,v})^2\Big)dr\\
&\quad+\int_t^s\partial_x\Phi(\Theta_r^t)\sigma_r^{t,x,\xi}dB_r+\int_t^s\partial_v\Phi(\Theta_r^t)\sigma_r^{t,v}dB_r^V+\int_t^s\partial_x\Phi(\Theta_{r-}^t)\mathbf{1}_{\{X_{r-}=0\}}dK_r\\
&\quad+\int_t^s\int_A\Big(\Phi(r,V_r^{t,v},X_{r-}^{t,x,\xi}+\theta_r^{t,x,\xi}(z),\mathbb{P}_{X_r^{t,\xi}})-\Phi(\Theta_{r-}^t)\\
&\qquad\qquad-\partial_x\Phi(\Theta_{r-}^t)\theta_{r-}^{t,x,\xi}(z)\Big)\lambda_r^{t,x,\xi}\nu^s(dz)dr \qquad\qquad (14)\\
&\quad+\int_t^s\int_{A\times\mathbb{R}_+}\Big(\Phi(r,V_r^{t,v},X_{r-}^{t,x,\xi}+\theta_r^{t,x,\xi}(z),\mathbb{P}_{X_r^{t,\xi}})-\Phi(\Theta_{r-}^t)\Big)\mathbf{1}_{[0,\lambda_r^{t,x,\xi}]}(y)\widetilde{\mathcal{N}}^s(drdzdy)\\
&\quad+\int_t^s\widetilde{\mathbb{E}}\Big[\partial_\mu\Phi(\Theta_r^t,\tilde{X}_r^{t,\tilde{\xi}})\tilde{b}_r^{t,\xi}+\frac{1}{2}\partial_y(\partial_\mu\Phi)(\Theta_r^t,\tilde{X}_r^{t,\tilde{\xi}})(\tilde{\sigma}_r^{t,\xi})^2\\
&\qquad+\int_0^1\int_A\Big(\partial_\mu\Phi(\Theta_r^t,\tilde{X}_r^{t,\tilde{\xi}}+\rho\tilde{\theta}_r^{t,x,\xi}(z))-\partial_\mu\Phi(\Theta_r^t,\tilde{X}_s^{t,\tilde{\xi}})\Big)\tilde{\theta}_r^{t,\xi}(z)\tilde{\lambda}_r^{t,\xi}\nu^s(dz)d\rho\Big]dr.
\end{aligned}$$

3 A Bertrand game among the sellers (static case)

In this section we analyze a price setting mechanism among liquidity providers (investors placing limit orders), and use it as the basis for our continuous time model in the rest of the paper. To begin with, recall that in this paper we assume all sellers are patient and all buyers are impatient. We therefore consider only the sell-side LOB. Following the ideas of [13, 27, 28], we shall consider the process of the (static) price setting as a Bertrand-type of game among the sellers, each placing a certain number of sell limit orders at a specific price, and trying to maximize her expected utility. To be more precise, we assume that sellers use the price at which they place limit orders as their *strategic variable*, and the number of shares submitted would be determined accordingly. Furthermore, we assume that there is a *waiting cost*, also as a function of the price. Intuitively, a higher price will lead to a longer execution time, hence a higher waiting cost. Thus, there is a competitive game among the sellers for better total reward. Finally, we assume that the sellers are homogeneous in the sense that they have the same subjective probability measure, so that they share the same degree of risk aversion.

We now give a brief description of the problem. We assume that there are N sellers, and the jth seller places limit orders at price $p_j = X + l_j$, $j = 1, 2, \cdots, N$, where X is the fundamental price. Without loss of generality, we may assume $X = 0$. As a main element in an *oligopolistic competitions* (cf. e.g., [28]), we assume that each seller i is equipped with a *demand function*, denoted by $h_i^N(p_1, p_2, \cdots, p_N)$, for a given price vector $p = (p_1, p_2, \cdots, p_N)$, reflecting the seller's perceived demand from the buyers. The seller i will determine the number of shares of limit orders to be placed in the LOB based on the values of his/her demand function, given the price vector. Hence this is a *Bertrand-type game*[1]. More specifically, we assume that the demand functions h_i^N, $i = 1, 2, \cdots, N$, are smooth and satisfy the following properties:

$$\frac{\partial h_i^N}{\partial p_i} < 0, \quad \text{and} \quad \frac{\partial h_i^N}{\partial p_j} > 0, \quad \text{for } j \neq i. \tag{15}$$

We note that (15) simply amounts to saying that each seller expects less demand (for her orders) when her own price increases, and more demand when other seller(s) increase their prices. Furthermore, we shall assume that the demand functions are invariant under permutations of the other sellers' prices, in the sense that, for fixed $p_1, \cdots, p_N$ and all $i, j \in \{1, \cdots, N\}$,

$$h_i^N(p_1, \cdots, p_i, \cdots, p_j, \cdots, p_N) = h_j^N(p_1, \cdots, p_j, \cdots, p_i, \cdots, p_N). \tag{16}$$

It is worth noting that the combination of (15) and (16) is the following fact: if a price vector p is ordered by $p_1 \leq p_2 \leq \cdots \leq p_N$, then for any $i < j$, it holds that

[1] A *Cournot game* is one such that the price p_i is the function of the numbers of shares $q = (q_1, \cdots, q_N)$ through a demand function. The two games are often exchangeable if the demand functions are invertible (see, e.g., [28]).

$$h_j^N(p) = h_j^N(p_1,\cdot\cdot,p_j,\cdot\cdot,p_i,\cdot\cdot,p_N) \le h_i^N(p_1,\cdot\cdot,p_i,\cdot\cdot,p_i,\cdot\cdot,p_N) \le h_i^N(p). \quad (17)$$

That is, the demand functions are ordered in a reversed way, which in a sense indicates that the shape of LOB should be a non-increasing function of prices in the LOB. Finally, for each i, we denote the price vector for "other" prices for seller i by p_{-i}. For seller i, the "least favorable" price for given p_{-i} is one that would generate zero demand, which is often referred to as the *choke price*. We shall assume such a price exists, and denote it by $\hat{p}_i(p_{-i}) < \infty$, namely,

$$h_i^N(p_1,\cdots,p_{i-1},\hat{p}_i,p_{i+1},\cdots,p_N) = 0. \quad (18)$$

We note that the existence of the choke price, together with the monotonicity property (15), indicates the possibility that $h_j^N(p) < 0$, for some j and some price vector p. But since the size of order placement cannot be negative, such scenario becomes unpractical. To amend this, we introduce the notion of *actual demand*, denoted by $\{\widehat{h}_i(p)\}$, which we now describe.

Consider an ordered price vector $p = (p_1,\cdots,p_N)$, with $p_i \le p_j$, $i \le j$, and we look at $h_N^N(p)$. If $h_N^N(p) \ge 0$, then by (17) we have $h_i^N(p) \ge 0$ for all $i = 1,\cdots,N$. In this case, we denote $\widehat{h}_i(p) = h_i^N(p)$ for all $i = 1,\cdots,N$. If $h_N^N(p) < 0$, then we set $\widehat{h}_N(p) = 0$. That is, the N-th seller does not act at all. We assume that the remaining $N-1$ sellers will observe this fact and modify their strategy as if there are only $N-1$ sellers. More precisely, we first choose a choke price $\widehat{p}_N$ so that $h_N^N(p_1,\cdots,p_{N-1},\widehat{p}_N) = 0$, and define

$$h_i^{N-1}(p_1,p_2,\cdots,p_{N-1}) := h_i^N(p_1,p_2,\cdots,p_{N-1},\widehat{p}_N), \qquad i = 1,\cdots,N-1,$$

and continue the game among the $N-1$ sellers.

In general, for $1 \le n \le N-1$, assume the $(n+1)$-th demand functions $\{h_i^{n+1}\}_{i=1}^{n+1}$ are defined. If $h_{n+1}^{n+1}(p_1,\cdots,p_{n+1}) < 0$, then other n sellers will assume $(n+1)$-th seller sets a price at $\hat{p}_{n+1}$ with zero demand (i.e., $h_{n+1}^{n+1}(p_1,p_2,\cdots,p_n,\hat{p}_{n+1}) = 0$), and modify their demand functions to

$$h_i^n(p_1,p_2,\cdots,p_n) := h_i^{n+1}(p_1,p_2,\cdots,p_n,\hat{p}_{n+1}), \qquad i = 1,\cdots,n. \quad (19)$$

We can now define the actual demand function $\{\widehat{h}_i\}_{i=1}^N$.

Definition 3.1 (*Actual demand function*) *Assume that $\{h_i^N\}_{i=1}^N$ is a family of demand functions. The family of "actual demand functions", denoted by $\{\widehat{h}_i\}_{i=1}^N$, are defined in the following steps: for a given ordered price vector p,*

(i) if $h_N^N(p) \ge 0$, then we set $\widehat{h}_i(p) = h_i^N(p)$ for all $i = 1,\cdots,N$;

(ii) if $h_N^N(p) < 0$, then we define recursively for $n = N-1,\cdots 1$ the demand functions $\{h_i^n\}_{i=1}^n$ as in (19). In particular, if there exists an $n < N$ such that $h_{n+1}^{n+1}(p_1,p_2,\cdots,p_n,p_{n+1}) < 0$ and $h_n^n(p_1,p_2,\cdots,p_n) \ge 0$, then we set

$$\widehat{h}_i(p) = \begin{cases} h_i^n(p_1, p_2, \cdots, p_n) & i = 1, \cdots, n \\ 0 & i = n+1, \cdots, N; \end{cases} \tag{20}$$

(iii) if there is no such n, then $\widehat{h}_i(p) = 0$ *for all* $i = 1, \cdots, N$.

We note that the actual demand function will always be non-negative, but for each price vector p, the number $\#\{i : \widehat{h}_i(p) > 0\} \leq N$, and could even be zero.

3.1 The Bertrand game and its Nash equilibrium

Besides the demand function, a key ingredient in the placement decision making process is the "waiting cost" for the time it takes for the limit order to be executed. We shall assume that each seller has her own waiting cost function $c_i^N \triangleq c_i^N(p_1, p_2, \cdots, p_N, Q)$, where Q is the total number of shares available in the LOB. Similar to the demand function, we shall assume the following assumptions for the waiting cost.

Assumption 3.2 *For each seller* $i \in \{1, \cdots, n\}$ *with* $n \in [1, N]$, *each* c_i^N *is smooth in all variables such that*

(i) (Monotonicity) $\frac{\partial c_i^N}{\partial p_i} > 0$, *and* $\frac{\partial c_i^N}{\partial p_j} < 0$, *for* $j \neq i$;

(ii) (Exchangeability) $c_i^N(p_1, \cdot\cdot, p_i, \cdot\cdot, p_i, \cdot\cdot, p_N) = c_j^N(p_1, \cdot\cdot, p_j, \cdot\cdot, p_i, \cdot\cdot, p_N)$;

(iii) $c_i^N(p)\big|_{p_i=0} = 0$, *and* $\frac{\partial c_i^N}{\partial p_i}\Big|_{p_i=0+} \in (0,1)$;

(iv) $\lim_{p_i \to \infty} \frac{p_i}{c_i^N(p)} = 0$, $i = 1, \cdots, N$.

Remark 3.3 (a) Assumption 3.2-(i), (ii) ensure that the price ordering leads to the same ordering for waiting cost functions, similar to what we argued before for demand functions. In particular, the second part of Assumption 3.2-(i) is due to (15). That is, if other seller submits an order at a higher price, the demand for seller i increases, which would lead to faster execution, hence shorter waiting time and lower waiting cost.

(b) Consider the function $J_i(p, Q) = p_i - c_i^N(p, Q)$. Assumption 3.2 amounts to saying that $J_i(p, Q)\big|_{p_i=0} = 0$, $\frac{\partial J_i(p,Q)}{\partial p_i}\Big|_{p_i=0+} > 0$, and $\lim_{p_i \to \infty} J_i(p, Q) < 0$. Thus, there exists $p_i^0 = p_i^0(p_{-i}, Q) > 0$ such that $\frac{\partial J_i(p,Q)}{\partial p_i}\Big|_{p_i=p_i^0} = 0$, and $\frac{\partial J_i(p,Q)}{\partial p_i}\Big|_{p_i>p_i^0} < 0$.

(c) Since $J_i(0, Q) = 0$, and $\frac{\partial J_i(p,Q)}{\partial p_i}\Big|_{p_i=0+} > 0$, one can easily check that $J_i(p_i^0, Q) > 0$. This, together with Assumption 3.2-(iv), shows that there exists $\tilde{p}_i = \tilde{p}_i(p_{-i}, Q) > p_i^0$, such that $J_i(p_i, Q)\big|_{p_i=\tilde{p}_i} = 0$ (or, equivalently $c_i^N(p_1, \cdots, p_{i-1}, \tilde{p}_i, p_{i+1}, \cdots, p_N, Q) = \tilde{p}_i$). Furthermore, remark above implies that $J_i(p_i, Q) < 0$ for all $p_i > \tilde{p}_i(p_{-i}, Q)$. In other words, any selling price higher than $\tilde{p}_i(p_{-i}, Q)$ would yield a negative profit, and therefore should be prevented.

The Bertrand game among sellers can now be formally introduced: each seller chooses its price to maximize profit in a non-cooperative manner, and their decision will be based not only on her own price, but also on the actions of all other sellers. We denote the profit of each seller by

$$\Pi_i(p_1,p_2,\cdots,p_N,Q) := \widehat{h}_i(p_1,p_2,\cdots,p_N)\times\Big(p_i - c_i^N(p_1,p_2,\cdots,p_N,Q)\Big), \qquad (21)$$

and each seller tries to maximize her profit Π. For each fixed Q, we are looking for a Nash equilibrium price vector $p^{*,N}(Q) = (p_1^{*,N}(Q),\cdots,p_N^{*,N}(Q))$. We note that in the case when $\widehat{h}_i(p^{*,N}) = 0$ for some i, the i-th seller will not participate in the game (with zero profit), so we shall modify the price

$$p_i^{*,N}(Q) \triangleq c_i^N(p_1^{*,N},\cdots,p_N^{*,N},Q) = c_i^N(p^{*,N},Q), \qquad (22)$$

and consider a subgame involving the $N-1$ sellers, and so on. That is, for a subgame with n sellers, they solve

$$p_i^{*,n} = \arg\max_{p\geq 0}\Pi_i^n(p_1^{*,n},p_2^{*,n},\cdots,p_{i-1}^{*,n},p,p_{i+1}^{*,n},\cdots,p_n^{*,n},Q), \quad i=1,\cdots,n \qquad (23)$$

to get $p^{*,n} = (p_1^{*,n},\cdots,p_n^{*,n},c_{n+1}^{*,n+1},\cdots,c_N^{*,N})$. More precisely, we define a Nash Equilibrium as follows.

Definition 3.4 *A vector of prices $p^* = p^*(Q) = (p_1^*,p_2^*,\cdots,p_N^*)$ is called a Nash equilibrium if*

$$p_i^* = \arg\max_{p\geq c_i}\Pi_i(p_1^*,p_2^*,\cdots,p_{i-1}^*,p,p_{i+1}^*,\cdots,p_N^*,Q), \qquad (24)$$

and $p_i^* = c_i^{*,i}(p^*,Q)$ whenever $\widehat{h}_i(p^*) = 0$, $i = 1,2,\cdots,N$.

We assume the following on a subgame for our discussion.

Assumption 3.5 *For $n = 1,\cdots,N$, we assume that there exists a unique solution to the system of maximization problems in equation (23).*

Remark 3.6 We observe from Definition of the Nash Equilibrium that, in equilibrium, a seller is actually participating in the Bertrand game only when her actual demand function is positive, and those with zero actual demand function will be ignored in the subsequent subgames. However, a participating seller does not necessarily have positive profit unless she sets the price higher than the waiting cost. In other words, it is possible that $\widehat{h}_i(p^*) > 0$, but $p_i^* = c_i(p^*,Q)$, so that $\Pi_i(p^*,Q) = 0$. We refer to such a case the *boundary* case, and denote the price to be $c_i^{*,b}$.

The following result details the procedure of finding the Nash equilibrium for the Bertrand competition. The idea is quite similar to that in [13], except for the general form of the waiting cost. We sketch the proof for completeness.

Proposition 3.7 *Assume that Assumption 3.2 is in force. Then there exists a Nash equilibrium to the Bertrand game (21) and (24).*

Moreover, the equilibrium point p^*, after modifications, should take the following form:

$$p^* = (p_1^*, \cdots, p_k^*, c_{k+1}^{*,b}, \cdots, c_n^{*,b}, c_{n+1}^*, \cdots, c_N^*), \tag{25}$$

from which we can immediately read: $\widehat{h}_i(p^*) > 0$ and $p_i^* > c_i^*$, $i = 1, \cdots, k$; $\widehat{h}_i(p^*) > 0$ but $p_i^* \le c_i^*$, $i = k+1, \cdots, n$; and $\widehat{h}_i(p^*) \le 0$, $i = n+1, \cdots, N$.

Proof We start with N sellers, and we shall drop the superscript N from all the notations, for simplicity. Let $p^* = (p_1^*, p_2^*, \cdots, p_N^*)$ be the candidate equilibrium prices (obtained by, for example, the first-order condition). By exchangeability, we can assume without loss of generality that the prices are ordered: $p_1^* \le p_2^* \le \cdots \le p_N^*$, and so are the corresponding cost functions $c_1^* \le c_2^* \le \cdots \le c_N^*$, where $c_i^* = c_i(p^*, Q)$ for $i = 1, \cdots, N$.

We first compare $p_N^{*,N}$ and $c_N^{*,N}$.

Case 1. $p_N^* > c_N^*$. We consider the following cases:

(a) If $h_N^N(p^*) > 0$, then by Definition 3.1 we have $\widehat{h}_i(p^*) = h_i^N(p^*) > 0$, for all i, and $p^* = (p_1^*, p_2^*, \cdots, p_N^*)$ is an equilibrium point.

(b) If $h_N^N(p^*) \le 0$, then in light of the definition of actual demand function (Definition 3.1), we have $\widehat{h}_N(p^*) = 0$. Thus, the N-th seller will have zero profit regardless where she sets the price. We shall require in this case that the N-th seller reduces her price to c_N^*, and we shall consider remaining $(N-1)$-sellers' candidate equilibrium prices $p^{*,N-1} = (p_1^{*,N-1}, \cdots, p_{N-1}^{*,N-1})$.

Case 2. $p_N^* \le c_N^*$. In this case the N-th seller would have a non-positive profit at the best. Thus, she sets $p_N^* = c_N^*$, and quits the game, and again the problem is reduced to a subgame with $(N-1)$ sellers, and to Case 1-(b). We should note that in the "boundary case" described in Remark 3.6, we will write $p_N^* = c_N^{*,b}$.

Repeating the same procedure for the subgames (for $n = N-1, \cdots, 2$), we see that eventually we will get a modified equilibrium point p^* of the form (25), proving the proposition. □

3.2 A linear mean-field case

In this subsection, we consider a special case, studied in [27], but with the modified waiting cost functions. More precisely, we assume that there are N sellers, each with demand function

$$h_i^N(p_1, \cdots, p_N) \triangleq A - Bp_i + C\bar{p}_i^N, \tag{26}$$

where $A, B, C > 0$, and $B > C$, and $\bar{p}_i^N = \frac{1}{N-1}\sum_{j\neq i}^{N} p_j$. We note that the structure of the demand function (26) obviously reflects a mean-field nature, and one can easily check that it satisfies all the assumptions mentioned in the previous subsection. Furthermore, as was shown in [27, Proposition 2.4], the actual demand function takes the form: for each $n \in \{1, \cdots, N-1\}$,

$$h_i^n(p_1, \cdots, p_n) = a_n - b_n p_i + c_n \bar{p}_i^n, \qquad \text{for } i = 1, \cdots, n,$$

where $\bar{p}_i^n = \frac{1}{n-1}\sum_{j\neq i}^{n} p_j$, and the parameters (a_n, b_n, c_n) can be calculated recursively for $n = N, \cdots, 1$, with $a_N = A$, $b_N = B$ and $c_N = C$. We note that in these works the (waiting) costs are assumed to be constant.

Let us now assume further that the waiting cost is also linear. For example, for $n = 1, \cdots, N$,

$$c_i^n = c_i^n(p_i, \bar{p}_i^n, Q) \triangleq x_n(Q)p_i - y_n(Q)\bar{p}_i^n, \qquad x_n(Q), y_n(Q) > 0.$$

Note that the profit function for seller i is

$$\Pi_i(p_1, \cdots, p_n, Q) = (a_n - b_n p_i + c_n \bar{p}_i^n) \cdot \left(p_i - (x_n p_i - y_n \bar{p}_i^n)\right). \tag{27}$$

An easy calculation shows that the critical point for the maximizer is

$$p_i^{*,n} = \frac{a_n}{2b_n} + \left(\frac{c_n}{2b_n} - \frac{y_n}{2(1-x_n)}\right)\bar{p}_i^n, \tag{28}$$

which is the optimal choice of seller i if the other sellers set prices with average $\bar{p}_i^n = \frac{1}{n-1}\sum_{j\neq i}^{n} p_j^*$. Now, let us define

$$\bar{p}^n := \frac{1}{n}\sum_{i=1}^{n} p_i^* = \frac{a_n(1-x_n)}{2b_n(1-x_n) - c_n(1-x_n) + b_n y_n}. \tag{29}$$

Then, it is readily seen that $\bar{p}_i^n = \frac{n}{n-1}\bar{p}^n - \frac{1}{n-1}p_i^*$, which means (plugging back into (28))

$$p_i^{*,n} = \frac{a_n}{2b_n + \frac{c_n}{n-1} - \frac{1}{n-1}\frac{b_n y_n}{1-x_n}} + \frac{1}{\frac{n-1}{n}\frac{2b_n(1-x_n)}{c_n(1-x_n)-b_n y_n} + \frac{1}{n}}\bar{p}^n. \tag{30}$$

For the sake of argument, let us assume that the coefficients $(a_n, b_n, c_n, x_n(Q), y_n(Q))$ converge to $(a, b, c, x(Q), y(Q))$ as $n \to \infty$. Then, we see from (29) and (30) that

$$\begin{cases} \lim_{n\to\infty} \bar{p}^n = \dfrac{a(1-x)}{2b(1-x)-c(1-x)+by} =: \bar{p}; \\ \lim_{n\to\infty} p_i^{*,n} = \dfrac{a}{2b} + \dfrac{c(1-x)-by}{2b(1-x)} \lim_{n\to\infty} \bar{p}^n = \dfrac{a(1-x)}{(2b-c)(1-x)+by} =: p^*. \end{cases} \tag{31}$$

It is worth noting that if we assume that there is a "representative seller" who randomly sets prices $p = p_i$ with equal probability $\frac{1}{n}$, then we can randomize the profit function (27):

$$\Pi_n(p,\bar{p}) = (a_n - b_n p + c_n \bar{p})\Big(p - (x_n p - y_n \bar{p})\Big), \tag{32}$$

where p is a random variable taking value $\{p_i\}$ with equal probability, and $\bar{p} \sim \mathbb{E}[p]$, thanks to the Law of Large Numbers, when n is large enough. In particular, in the limiting case as $n \to \infty$, we can replace the randomized profit function Π_n in (32) by:

$$\Pi_\infty = \Pi(p,\mathbb{E}[p]) := (a - bp + c\mathbb{E}[p])\Big(p - (xp - y\mathbb{E}[p])\Big). \tag{33}$$

A similar calculation as (28) shows that $(p^*, \mathbb{E}[p^*]) \in \text{argmax}\Pi(p,\mathbb{E}[p])$ will take the form

$$p^* = \frac{c(1-x)-by}{2b(1-x)}\mathbb{E}[p^*] + \frac{a}{2b} \qquad \text{and} \qquad \mathbb{E}[p^*] = \frac{a(1-x)}{2b(1-x)-c(1-x)+by}.$$

Consequently, we see that $p^* = \frac{a(1-x)}{(2b-c)(1-x)+by}$, as we see in (31).

Remark 3.8 The analysis above indicates the following facts: (i) If we consider the sellers in a "homogeneous" way, and as the number of sellers becomes large enough, all of them will actually choose the same strategy, as if there is a "representative seller" that sets the prices uniformly; (ii) The limit of equilibrium prices actually coincides with the optimal strategy of the representative seller under a limiting profit function. These facts are quite standard in mean-field theory, and will be used as the basis for our dynamic model for the (sell) LOB in the next section.

4 Mean-field type liquidity dynamics in continuous time

In this section we extend the idea of Bertrand game to the continuous time setting. To begin with, we assume that the contribution of each individual seller to the LOB is measured by the "liquidity" (i.e., the number of shares of the given asset) she provides, which is the function of the selling price she chooses, hence under the Bertrand game framework.

4.1 A general description

We begin by assuming that there are N sellers, and denote the liquidity that the i-th seller “adds" to the LOB at time t by Q^i_t. We shall assume that it is a pure jump Markov process, with the following generator: for any $f \in C([0,T]\times\mathbb{R}^N)$, and $(t,q) \in [0,T]\times\mathbb{R}^N$,

$$\mathscr{A}^{i,N}[f](t,q) := \int_{\mathbb{R}} \lambda^i(t,q,\theta)\Big(f(t,q_{-i}(q_i + h^i(t,\theta,z)) - f(t,q) \\ - \langle \partial_{x_i} f, h^i(t,\theta,z)\rangle\Big)\nu^i(dz), \tag{34}$$

where $q \in \mathbb{R}^N$, and $q_{-i}(y) = (q_1,\cdots,q_{i-1},y,q_{i+1},\cdots,q_N)$. Furthermore, h^i denotes the demand function for the i-th seller, and $\theta \in \mathbb{R}^k$ is a certain market parameter which will be specified later. Roughly speaking, (34) indicates that the i-th seller would act (or “jump") at stopping times $\{\tau^i_j\}_{j=1}^\infty$ with the waiting times $\tau^i_{j+1} - \tau^i_j$ having exponential distribution with intensity $\lambda^i(\cdot)$, and jump size being determined by the demand function $h^i(\cdots)$. The total liquidity provided by all the sellers is then a pure jump process with the generator

$$\mathscr{A}^N[f](t,q,\theta) = \sum_{i=1}^N \mathscr{A}^{i,N}[f](t,q), \qquad q \in \mathbb{R}^N,\ N \in \mathbb{N}, \quad t \in [0,T]. \tag{35}$$

We now specify the functions λ^i and h^i further. Recalling the demand function introduced in the previous section, we assume that there are two functions λ and h, such that for each i, and for $(t,x,q,p) \in [0,T]\times\mathbb{R}\times\mathbb{R}^{2N}$,

$$\lambda^i(t,q,\theta) = \lambda(t,q^i,p^i,\mu^N), \quad h^i(t,\theta,z) = h(t,x,q^i,p^i,z), \tag{36}$$

where $\mu^N := \frac{1}{N}\sum_{i=1}^N \delta_{p^i}$, x denotes the fundamental price at time t, and p^i is the sell price. We shall consider $p = (p^1,\cdots,p^N)$ as the control variable, as the Bertrand game suggests. Now, if we assume $\nu^i = \nu$ for all i, then we have a *pure jump Markov game of mean-field-type*, similar to the one considered in [6], in which each seller adds liquidity (in terms of number of shares) dynamically as a pure jump Markov process, denoted by Q^i_t, $t \geq 0$, with the kernel

$$\nu(t,q^i,\mu^N,p^i,dz) = \lambda(t,q^i,p^i,\mu^N)[\nu\circ h^{-1}(t,x,q^i,p^i,\cdot)](dz). \tag{37}$$

Furthermore, in light of the static case studied in the previous section, we shall assume that the seller's instantaneous profit at time $t > 0$ takes the form $(p^i_t - c^i_t)\Delta Q^i_t$, where c^i_t is the “waiting cost" for i-th seller at time t. We observe that the submitted sell price p^i can be written as $p^i = x + l^i$, where x is the fundamental price and l^i is the distance from x that the i-th seller chooses to set. Now let us assume that there is an invertible relationship between the selling prices p and the corresponding number of shares q, e.g., $p = \varphi(q)$ (such a relation is often used to convert the Bertrand game to Cournot game, see, e.g., [27]), and consider l as the *control variable*. We can then rewrite the functions λ and h of (36) in the following form:

$$\lambda^i(t,q,\theta) = \lambda(t,q^i,l^i,\tilde{\mu}^N(\varphi(q))), \quad h^i(t,q,\theta) = h(t,x,q^i,l^i,z). \tag{38}$$

To simplify the presentation, in what follows, we shall assume that λ does not depend on the control variable l^i, and that both λ and h are time-homogeneous. In other words, we assume that each Q^i follows a pure jump SDE studied in §2:

$$Q^i_t = q^i + \int_0^t \int_{A\times\mathbb{R}_+} h(X_r, Q^i_{r-}, l^i_r, z)\mathbf{1}_{[0,\lambda(Q^i_{r-},\mu^N_{\varphi(\mathbf{Q}_r)})]}(y)\mathcal{N}^s(drdzdy), \tag{39}$$

where $\mathbf{Q}_t = (Q^1_t,\cdots,Q^N_t)$, $\mathcal{N}^s$ is a Poisson random measure on $\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$, and $\{X_t\}_{t\geq 0}$ is the fundamental price process of the underlying asset which we assume to satisfy the SDE (cf. [32]):

$$X^{t,x}_s = x + \int_t^s b(X^{t,x}_r)dr + \int_t^s \sigma(X^{t,x}_r)dW_r, \tag{40}$$

where b and σ are deterministic functions satisfying some standard conditions. We shall assume that the i-th seller is aiming at maximizing the expected total accumulated profit:

$$\begin{aligned} &\mathbb{E}\Big\{\sum_{t\geq 0}(p^i_t - c^i_t)\Delta Q^i_t\Big\} \\ &= \mathbb{E}\Big\{\int_0^\infty \int_A h(X_t, Q^i_t, l^i_t, z)(X_t + l^i_t - c^i_t)\lambda(Q^i_t, \mu^N_{\varphi(\mathbf{Q}_t)})\nu^s(dz)dt\Big\}. \end{aligned} \tag{41}$$

We remark that in (41) the time horizon is allowed to be infinity, which can be easily converted to finite horizon by setting $h(X_t,\cdots) = 0$ for $t\geq T$, for a given time horizon $T > 0$, which we do not want to specify at this point. Instead, our focus will be mainly on the limiting behavior of the equilibrium when $N\to\infty$. In fact, given the “symmetric" nature of the problem (i.e., all seller’s having the same λ and h), as well as the results in the previous section, we envision a “representative seller" in a limiting mean-field type control problem whose optimal strategy coincides with the limit of N-seller Nash equilibrium as $N\to\infty$, just as the well-known continuous diffusion cases (see, e.g., [29] and [9, 12]). We should note such a result for pure jump cases has been substantiated in a recent work [6], in which it was shown that, under reasonable conditions, in the limit the total liquidity $Q_t = \sum_{i=1}^N Q^i_t$ will converge to a pure jump Markovian process with a mean-field type generator. Based on this result, as well as the individual optimization problem (39) and (41), it is reasonable to consider the following (limiting) mean-filed type pure jump stochastic control problem for a representative seller, whose total liquidity has a dynamics that can be characterized by the following mean-field type pure jump SDE:

$$Q_t = q + \int_0^t \int_{A\times\mathbb{R}_+} h(X_r, Q_{r-}, l_r, z)\mathbf{1}_{[0,\lambda(Q_{r-},\mathbb{P}_{Q_r})]}(y)\mathcal{N}^s(drdzdy), \tag{42}$$

where $\lambda(Q,\mathbb{P}_Q) := \lambda(Q,\mathbb{E}[\varphi(Q)])$ by a slight abuse of notation, and with the cost functional:

$$\Pi(q,l) = \mathbb{E}\Big\{\int_0^\infty \int_A h(X_t,Q_t,l_t,z)(X_t + l_t - c_t)\lambda(Q_t,\mathbb{P}_{Q_t})\nu^s(dz)dt\Big\}. \quad (43)$$

4.2 Problem formulation

With the general description in mind, we now give the formulation of our problem. First, we note that the liquidity of the LOB will not only be affected by the liquidity providers (i.e., the sellers), but also by liquidity consumer, that is, the market buy orders as well as the cancellations of sell orders (which we assume is free of charge). We shall describe its collective movement (in terms of number of shares) of all such consumptional orders as a compound Poisson process, denoted by $\beta_t = \sum_{i=1}^{N_t} \Lambda_t$, $t \geq 0$, where $\{N_t\}$ is a standard Poisson process with parameter λ, and $\{\Lambda_i\}$ is a sequence of i.i.d. random variables taking values in a set $B \subseteq \mathbb{R}$, with distribution ν. Without loss of generality, we assume that counting measure of β coincides with the canonical Poisson random measure $\mathcal{N}^b$, so that the Lévy measure is $\nu^b = \lambda\nu$. In other words, $\beta_t := \int_0^t \int_B z\, \widetilde{\mathcal{N}}^b(drdz)$, and the total liquidity satisfies the SDE:

$$Q_t^0 = q + \int_0^t \int_{A\times\mathbb{R}_+} h(X_r,Q_{r-}^0,l_r,z)\mathbf{1}_{[0,\lambda(Q_{r-}^0,\mathbb{P}_{Q_r^0})]}(y)\mathcal{N}^s(drdzdy) - \beta_t. \quad (44)$$

We remark that there are two technical issues for the dynamics (44). First, the presence of the buy order process β brings in the possibility that $Q_t^0 < 0$, which should never happen in reality. We shall therefore assume that the buy order has a natural upper limit: the total available liquidity Q_t^0. That is, if we denote $\mathcal{S}_\beta = \{t : \Delta\beta_t \neq 0\}$, then for all $t \in \mathcal{S}_\beta$, we have $Q_t^0 = (Q_{t-}^0 - \Delta\beta_t)^+$. Consequently, we can assume that there exists a process $K = \{K_t\}$, where K is a non-decreasing, pure jump process such that (i) $\mathcal{S}_K = \mathcal{S}_\beta$; (ii) $\Delta K_t := (Q_{t-}^0 - \Delta\beta_t)^-$, $t \in \mathcal{S}_K$; and (iii) the Q^0-dynamics (44) can be written as, for $t \geq 0$,

$$\begin{aligned}
Q_t &= q + \int_0^t \int_{A\times\mathbb{R}_+} h(X_r,Q_{r-},l_r,z)\mathbf{1}_{[0,\lambda(Q_{r-},\mathbb{P}_{Q_r})]}(y)\mathcal{N}^s(drdzdy) - \beta_t + K_t \\
&= q + \int_0^t \int_{A\times\mathbb{R}_+} h(X_r,Q_{r-},l_r,z)\mathbf{1}_{[0,\lambda(Q_{r-},\mathbb{P}_{Q_r})]}(y)\widetilde{\mathcal{N}}^s(drdzdy) \\
&\quad - \int_0^t \int_B z\,\widetilde{\mathcal{N}}^b(drdz) + \int_0^t \int_A h(X_r,Q_r,l_r,z)\lambda(Q_r,\mathbb{P}_{Q_r})\nu^s(dz)dr + K_t.
\end{aligned} \quad (45)$$

where K is a “reflecting process", and $\widetilde{\mathcal{N}}^s(drdzdy)$ is the compensated Poisson martingale measure of $\mathcal{N}^s$. That is, (45) is a (pure jump) mean-field SDE with reflection as was studied in §2.

Now, in light of the discussion of MFSDEDR in §2, we shall consider the following two MFSDEDRs that are slightly more general than (45): for $\xi \in L^2(\mathcal{F}_t;\mathbb{R})$, $q \in \mathbb{R}$, and $0 \le s \le t$,

$$
\begin{aligned}
Q_s^{t,\xi} &= \xi + \int_t^s \int_{A\times\mathbb{R}_+} h(X_r^{t,x}, Q_{r-}^{t,\xi}, l_r, z)\, \mathbf{1}_{[0,\lambda(Q_{r-}^{t,\xi}, \mathbb{P}_{Q_r^{t,\xi}})]}(y) \widetilde{\mathcal{N}}^s(drdzdy) \\
&\quad - \int_t^s \int_B z\, \widetilde{\mathcal{N}}^b(drdz) + \int_t^s a(X_r^{t,x}, Q_r^{t,\xi}, \mathbb{P}_{Q_r^{t,\xi}}, l_r)dr + K_s^{t,\xi}, \qquad (46) \\
Q_s^{t,q,\xi} &= q + \int_t^s \int_{A\times\mathbb{R}_+} h(X_r^{t,x}, Q_{r-}^{t,q,\xi}, l_r, z)\, \mathbf{1}_{[0,\lambda(Q_{r-}^{t,q,\xi}, \mathbb{P}_{Q_r^{t,\xi}})]}(y) \widetilde{\mathcal{N}}^s(drdzdy) \\
&\quad - \int_t^s \int_B z\, \widetilde{\mathcal{N}}^b(drdz) + \int_t^s a(X_r^{t,x}, Q_r^{t,q,\xi}, \mathbb{P}_{Q_r^{t,\xi}}, l_r)dr + K_s^{t,q,\xi}, \qquad (47)
\end{aligned}
$$

where $l = \{l_s\}$ is the control process for the representative seller, and $Q_s = Q_s^{t,q,\xi}$, $s \ge t$, is the total liquidity of the sell-side LOB. We shall consider the following set of *admissible strategies*:

$$
\mathcal{U}_{ad} := \{l \in L^1_{\mathbb{F}}([0,\infty);\mathbb{R}_+) : l \text{ is } \mathbb{F}\text{-predictable}\}. \qquad (48)
$$

The objective of the seller is to solve the following mean-field stochastic control problem:

$$
v(x,q,\mathbb{P}_\xi) = \sup_{l\in\mathcal{U}_{ad}} \Pi(x,q,\mathbb{P}_\xi,l) = \sup_{l\in\mathcal{U}_{ad}} \mathbb{E}\Big[\int_0^\infty e^{-\rho r} L(X_r^x, Q_r^{q,\xi}, \mathbb{P}_{Q_r^\xi}, l_r)dr\Big] \qquad (49)
$$

where $L(x,q,\mu,l) := \int_A h(x,q,l,z)c(x,q,l)\lambda(q,\mu)\nu^s(dz)$, and $\mathcal{U}_{ad}$ is defined in (48). Here we denote $X^x := X^{0,x}$, $Q^{q,\xi} := Q^{0,q,\xi}$.

Remark 4.1 (i) In (46) and (47), we allow a slightly more general drift function a, which in particular could be $a(x,q,\mu,l) = \lambda(q,\mu)\int_A h(x,q,l,z)\nu^s(dz)$, as is in (45).

(ii) In (49), the pricing function $c(x,q,l)$ is a more general expression of the original form $x + l - c$ in (43), taking into account the possible dependence of the waiting cost c_t on the sell position l and the total liquidity q at time t.

(iii) Compared to (43), we see that a *discounting factor* $e^{-\rho t}$ is added to the cost functional $\Pi(\cdots)$ in (49), reflecting its nature as the "present value".

In the rest of the paper we shall assume that the market parameters b, σ, λ, h, the pricing function c in (46) – (49), and the discounting factor ρ satisfy the following assumptions.

Assumption 4.2 *All functions* $b, \sigma \in C^0(\mathbb{R})$, $\lambda \in L^0(\mathbb{R}\times\mathscr{P}_2(\mathbb{R});\mathbb{R}_+)$, $h \in L^0(\mathbb{R}^2 \times \mathbb{R}_+ \times A)$, *and* $c \in L^0(\mathbb{R}\times\mathbb{R}_+\times\mathbb{R}_+)$ *are bounded, and satisfy the following conditions, respectively:*

(i) b *and* σ *are uniformly Lipschitz continuous in* x *with Lipschitz constant* $L > 0$*;*

(ii) $\sigma(0) = 0$ *and* $b(0) \ge 0$*;*

(iii) λ *and* h *satisfy Assumption 2.1;*

(iv) For $l \in \mathbb{R}_+$, $c(x,q,l)$ *is Lipschitz continuous in* (x,q)*, with Lipschitz constant* $L > 0$*;*

(v) h *is non-increasing, and* c *is non-decreasing in the variable* l*;*

(vi) $\rho > L + \frac{1}{2}L^2$*, where* $L > 0$ *is the Lipschitz constant in Assumption 2.1;*

(vii) For $(x,\mu,l) \in \mathbb{R}_+ \times \mathscr{P}_2(\mathbb{R}) \times \mathbb{R}_+$, $\Pi(x,q,\mu,l)$ *is convex in* q.

Remark 4.3 (i) The monotonicity assumptions in Assumption 4.2-(v) are inherited from §3. Specifically, they are the assumption (15) for h, and Assumption 3.1-(i) for c, respectively.

(ii) Under Assumption 4.2, one can easily check that the SDEs (40) as well as (46) and (47) all have pathwisely unique strong solutions in $L^2_{\mathbb{F}}(\mathbb{D}([0,T]))$, thanks to Theorem 2.3; and Assumption 4.2-(ii) implies that $X^{t,x}_s \geq 0$, $s \in [t,\infty)$, $\mathbb{P}$-a.s., whenever $x \geq 0$.

5 Dynamic programming principle

In this section we substantiate the *dynamic programming principle* (DPP) for the stochastic control problem (46)–(49). We begin by examining some basic properties of the value function.

Proposition 5.1 *Under the Assumptions 2.1 and 4.2, the value function* $v(x,q,\mathbb{P}_\xi)$ *is Lipschitz continuous in* $(x,q,\mathbb{P}_\xi)$*, non-decreasing in* x*, and decreasing in* q.

Proof We first check the Lipschitz property in x. For $x,x' \in \mathbb{R}$, denote $X^x = X^{0,x}$ and $X^{x'} = X^{0,x'}$ as the corresponding solutions to (40), respectively. Denote $\Delta X_t = X^x_t - X^{x'}_t$, and $\Delta x = x - x'$. Then, applying Itô's formula to $|\Delta X_t|^2$ and by some standard arguments, one has

$$|\Delta X_t|^2 = |\Delta x|^2 + \int_0^t (2\alpha_s + \beta_s^2)|\Delta X_s|^2 ds + \int_0^t 2\beta_s |\Delta X_s|^2 dW_s,$$

where α, β are two processes bounded by the Lipschitz constants L in Assumption 2.1, thanks to Assumption 4.2. Thus, one can easily check, by taking expectation and applying Burkholder-Davis-Gundy and Gronwall inequalities, that

$$\mathbb{E}[|\Delta X|^{*,2}_t] \leq |\Delta x|^2 e^{(2L+L^2)t}, \qquad t \geq 0. \tag{50}$$

Furthermore, it is clear that, under Assumption 4.2, the function $L(x,q,\mu,l)$ is uniformly Lipschitz in x, uniformly in (q,μ,l). That is, for some generic constant $C > 0$ we have

$$|\Pi(x,q,\mathbb{P}_\xi,l)-\Pi(x',q,\mathbb{P}_\xi,l)|\le C\mathbb{E}\Big[\int_0^\infty\int_A e^{-\rho t}|\Delta X_t|\nu^s(dz)dt\Big]$$
$$\le C\mathbb{E}\Big[\int_0^\infty e^{-\rho t}\sqrt{\mathbb{E}[|\Delta X|_t^{*,2}]}dt\Big]\le C|\Delta x|\int_0^\infty e^{-\rho t}e^{(L+\frac{1}{2}L^2)t}dt\le C|x-x'|.$$

Here the last inequality is due to Assumption 4.2-(vi). Consequently, we obtain

$$|v(x,q,\mathbb{P}_\xi)-v(x',q,\mathbb{P}_\xi)|\le C|x-x'|,\quad \forall x,x'\in\mathbb{R}. \tag{51}$$

To check the Lipschitz properties for q and $\mathbb{P}_\xi$, we denote, for $(q,\mathbb{P}_\xi)\in\mathbb{R}_+\times\mathscr{P}_2(\mathbb{R})$, $h_s^{q,\xi}\equiv h(X_s,Q_s^{t,q,\xi},l_s,z)$, $\lambda_s^{q,\xi}\equiv\lambda(Q_s^{t,q,\xi},\mathbb{P}_{Q_s^{t,\xi}})$, and $c_s^{q,\xi}\equiv c(X_s,Q_s^{t,q,\xi},l_s)$, $s\ge t$. Furthermore, for $q,q'\in\mathbb{R}_+$ and $\mathbb{P}_\xi,\mathbb{P}_{\xi'}\in\mathscr{P}_2(\mathbb{R})$, we denote $\Delta\psi_r\equiv\psi_r^{q,\xi}-\psi_r^{q',\xi'}$ for $\psi=h,\lambda,c$. Now, by Assumptions 2.1 and 4.2, and following a similar argument of Theorem 2.3, one shows that

$$|\Pi(x,q,\mathbb{P}_\xi,l)-\Pi(x,q',\mathbb{P}_{\xi'},l)|$$
$$\le\mathbb{E}\Big\{\int_0^\infty\int_A e^{-\rho r}\,(h_r^{q,\xi}c_r^{q,\xi}|\Delta\lambda_r|+c_r^{q,\xi}\lambda_r^{q',\xi'}|\Delta h_r|+h_r^{q',\xi'}\lambda_r^{q',\xi'}|\Delta c_r|)\,\nu^s(dz)dr\Big\}$$
$$\le\mathbb{E}\Big\{\int_0^\infty\int_A e^{-\rho r}|Q_r^{t,q,\xi}-Q_r^{t,q',\xi'}|\nu^s(dz)dr\Big\}\le C\Big(|q-q'|+W_1(\mathbb{P}_\xi,\mathbb{P}_{\xi'})\Big),$$

which implies that

$$|v(x,q,\mathbb{P}_\xi)-v(x,q',\mathbb{P}'_\xi)|\le C\Big(|q-q'|+W_1(\mathbb{P}_\xi,\mathbb{P}_{\xi'})\Big). \tag{52}$$

Finally, the respective monotonicity of the value function on x and q follows from the comparison theorem of the corresponding SDEs and Assumption 4.2. This completes the proof. □

We now turn our attention to the DPP. The argument will be very similar to that of [32], except for some adjustments to deal with the mean-field terms. But, by using the flow-property (5) we can carry out the argument without substantial difficulty.

Theorem 5.2 *Assume that Assumptions 2.1 and 4.2 are in force. Then, for any* $(x,q,\mathbb{P}_\xi)\in\mathbb{R}^2\times\mathscr{P}_2(\mathbb{R})$ *and for any* $t\in(0,\infty)$*,*

$$v(x,q,\mathbb{P}_\xi)=\sup_{l\in\mathscr{U}_{ad}}\mathbb{E}\Big[\int_0^t e^{-\rho s}L(X_s^x,Q_s^{q,\xi;l},\mathbb{P}_{Q_s^{\xi;l}},l_s)ds$$
$$+e^{-\rho t}v(X_t^x,Q_t^{q,\xi;l},\mathbb{P}_{Q_t^{\xi;l}})\Big]. \tag{53}$$

Proof Let us denote the right side of (53) by $\tilde{v}(x,q,\mathbb{P}_\xi)=\sup_l\tilde{\Pi}(x,q,\mathbb{P}_\xi;l)$. We first note that X_r and $(Q_r^{t,\xi},Q_r^{t,q,\xi})$ have the flow property. So, for any $l\in\mathscr{U}_{ad}$,

$$
\begin{aligned}
&\Pi(x,q,\mathbb{P}_\xi;l) = \mathbb{E}\Big[\int_0^\infty e^{-\rho s} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds\Big]\\
&= \mathbb{E}\Big[\int_0^t e^{-\rho s} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds\\
&\qquad + e^{-\rho t}\mathbb{E}\Big\{\int_t^\infty e^{-\rho(s-t)} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds\Big|\mathcal{F}_t\Big\}\Big] \qquad (54)\\
&= \mathbb{E}\Big[\int_0^t e^{-\rho s} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds + e^{-\rho t}\Pi(X_t^x, Q_t^{q,\xi;l}, \mathbb{P}_{Q_t^{\xi;l}}; l)\Big]\\
&\le \mathbb{E}\Big[\int_0^t e^{-\rho s} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds + e^{-\rho t} v(X_t^x, Q_t^{q,\xi;l}, \mathbb{P}_{Q_t^{\xi;l}})\Big]\\
&= \tilde{\Pi}(x,q,\mathbb{P}_\xi;l).
\end{aligned}
$$

This implies that $v(x,q,\mathbb{P}_\xi) \le \tilde{v}(x,q,\mathbb{P}_\xi)$.

To prove the other direction, let us denote $\Gamma = \mathbb{R}_+ \times \mathbb{R} \times \mathscr{P}_2(\mathbb{R})$, and consider, at each time $t \in (0,\infty)$, a countable partition $\{\Gamma_i\}_{i=1}^\infty$ of Γ and $(x_i, q_i, \mathbb{P}_{\xi_i}) \in \Gamma_i$, $\xi_i \in L^2(\mathcal{F}_t)$, $i = 1, 2, \cdots$, such that for any $(x,q,\mu) \in \Gamma_i$ and for fixed $\varepsilon > 0$, it holds $|x - x_i| \le \varepsilon$, $q_i - \varepsilon \le q \le q_i$, and $W_2(\mu, \mathbb{P}_{\xi_i}) \le \varepsilon$. Now, for each i, choose an ε-optimal strategy $l^i \in \mathscr{U}_{ad}$, such that $v(t,x_i,q_i,\mathbb{P}_{\xi_i}) \le \Pi(t,x_i,q_i,\mathbb{P}_{\xi_i};l^i) + \varepsilon$, where $\Pi(t,x_i,q_i,\mathbb{P}_{\xi_i};l^i) := \mathbb{E}[\int_t^\infty e^{-\rho(s-t)} L(X_s^{t,x_i}, Q_s^{t,q_i,\xi_i}, \mathbb{P}_{Q_s^{t,\xi_i}}, l_s^i)ds]$ and $v(t,x_i,q_i,\mathbb{P}_{\xi_i}) = \sup_{l^i \in \mathscr{U}_{ad}} \Pi(t,x_i,q_i,\mathbb{P}_{\xi_i};l^i)$.

Then, by definition of the value function and the Lipschitz properties (Proposition 5.1) with some constant $C > 0$, for any $(x,q,\mu) \in \Gamma_i$, it holds that

$$
\begin{aligned}
\Pi(t,x,q,\mu;l^i) &\ge \Pi(t,x_i,q_i,\mathbb{P}_{\xi_i};l^i) - C\varepsilon \ge v(t,x_i,q_i,\mathbb{P}_{\xi_i}) - (C+1)\varepsilon\\
&\ge v(t,x,q,\mu) - (2C+1)\varepsilon. \qquad (55)
\end{aligned}
$$

Now, for any $l \in \mathscr{U}_{ad}$, we define a new strategy $\tilde{l}$ as follows:

$$
\tilde{l}_s := l_s \mathbf{1}_{[0,t]}(s) + \Big[\sum_i l_s^i \mathbf{1}_{\Gamma_i}(X_t^x, Q_t^{q,\xi;l}, \mathbb{P}_{Q_t^{\xi;l}})\Big]\mathbf{1}_{(t,\infty)}(s). \qquad (56)
$$

Then, clearly $\tilde{l} \in \mathscr{U}_{ad}$. To simplify notation, let us denote

$$
I_1 = \int_0^t e^{-\rho s} L(X_s^x, Q_s^{q,\xi;l}, \mathbb{P}_{Q_s^{\xi;l}}, l_s)ds. \qquad (57)
$$

By applying (55) and flow property, we have

$$
\begin{aligned}
&v(x,q,\mu)\geq \Pi(x,q,\mu;\tilde{l})\\
=&\mathbb{E}\Big[I_1+e^{-\rho t}\mathbb{E}\Big\{\int_t^\infty e^{-\rho(s-t)}L(X_s^x,Q_s^{q,\xi;l},\mathbb{P}_{Q_s^{\xi;l}},l_s)ds\Big|\mathcal{F}_t\Big\}\Big]\\
=&\mathbb{E}\Big[I_1+e^{-\rho t}\Pi(t,X_t^x,Q_t^{q,\xi},\mathbb{P}_{Q_t^{\xi}};\tilde{l})\Big]\\
=&\mathbb{E}\Big[I_1+e^{-\rho t}\sum_i\Pi(t,X_t^x,Q_t^{q,\xi},\mathbb{P}_{Q_t^{\xi}};l^i)\mathbf{1}_{\Gamma_i}(X_t^x,Q_t^{q,\xi},\mathbb{P}_{Q_t^{\xi}})\Big]\\
\geq&\mathbb{E}\Big[I_1+e^{-\rho t}v(X_t^x,Q_t^{q,\xi},\mathbb{P}_{Q_t^{\xi}})\Big]-(2C+1)\varepsilon=\tilde{\Pi}(x,q,\mathbb{P}_{\xi};l)-(2C+1)\varepsilon.
\end{aligned}
$$

Since $\varepsilon>0$ is arbitrary, we get $v(x,q,\mathbb{P}_\xi)\geq\tilde{v}(x,q,\mathbb{P}_\xi)$, proving (53). □

Remark 5.3 We should note that while it is difficult to specify all the boundary conditions for the value function, the case when $q=0$ is relatively clear. Note that $q=0$ means there is zero liquidity for the asset. Then by definition of the liquidity dynamics (45) we see that Q_t will stay at zero until the first positive jump happens. During that period of time there would be no trade, thus by DPP (53) we should have

$$v(x,0,\mu)\equiv 0. \tag{58}$$

Furthermore, since the value function v is non-increasing in q, thanks to Proposition 5.1, and is always non-negative, we can easily see that the following boundary condition is also natural

$$\partial_q v(x,0,\mu)\equiv 0. \tag{59}$$

We shall use (58) and (59) frequently in our future discussion.

6 HJB equation and its viscosity solutions

In this section, we shall formally derive the HJB equation associated to the stochastic control problem studied in the previous section, and show that the value function of the control problem is indeed a viscosity solution of the HJB equation.

To begin with, we first note that, given the DPP (53), as well as the boundary conditions (58) and (59), if the value function v is smooth, then by standard arguments with the help of the Itô's formula (14) and the fact that

$$\partial_q v(X_{t-},Q_{t-},\mathbb{P}_{X_{t-}})\mathbf{1}_{\{Q_{t-}=0\}}dK_t=\partial_q v(X_{t-},0,\mathbb{P}_{X_{t-}})\mathbf{1}_{\{Q_{t-}=0\}}dK_t\equiv 0,$$

it is not difficult to show that the value function should satisfy the following HJB equation: for $(x,q,\mu)\in\mathbb{R}\times\mathbb{R}_+\times\mathscr{P}_2(\mathbb{R})$,

$$
\begin{cases}
\rho v(x,q,\mu)=\sup_{l\in\mathbb{R}_+}[\mathscr{J}^l[v](x,q,\mu)+L(x,q,\mu,l)],\\
v(x,0,\mu)=0,\quad \partial_q v(x,0,\mu)=0,
\end{cases}
\tag{60}
$$

where $\mathscr{J}^l$ is an integro-differential operator defined by, for any $\phi \in \mathbb{C}_b^{2,(2,1)}(\mathbb{R}\times\mathbb{R}_+\times\mathscr{P}_2(\mathbb{R}))$,

$$
\begin{aligned}
&\mathscr{J}^l[\phi](x,q,\mu) \stackrel{\triangle}{=} \Big(b(x)\partial_x + \sigma^2(x)\frac{1}{2}\partial^2_{xx} + a(x,q,\mu,l)\partial_q\Big)\phi(x,q,\mu)\\
&+\int_A \Big(\phi(x,q+h(x,q,l,z),\mu)-\phi(x,q,\mu)-\partial_q\phi(x,q,\mu)h(x,q,l,z)\Big)\lambda(q,\mu)\nu^s(dz)\\
&-\int_B \Big(\phi(x,q-z,\mu)-\phi(x,q,\mu)-\partial_q\phi(x,q,\mu)z\Big)\nu^b(dz)\\
&+\tilde{\mathbb{E}}\Big[\partial_\mu\phi(x,q,\mu,\tilde{\xi})a(x,\tilde{\xi},\mu,l)\Big]+\tilde{\mathbb{E}}\Big[\int_0^1\int_A\Big(\partial_\mu\phi(x,q,\mu,\tilde{\xi}+\gamma h(x,\tilde{\xi},l,z))\\
&\qquad -\partial_\mu\phi(x,q,\mu,\tilde{\xi})\Big)h(x,\tilde{\xi},l,z)\lambda(\tilde{\xi},\mu)\nu^s(dz)d\gamma\Big]\\
&-\tilde{\mathbb{E}}\Big[\int_0^1\int_B\Big(\partial_\mu\phi(x,q,\mu,\tilde{\xi}-\gamma z)-\partial_\mu\phi(x,q,\mu,\tilde{\xi})\Big)\times z\nu^b(dz)d\gamma\Big]. \qquad (61)
\end{aligned}
$$

We note that in general, whether there exists smooth solutions to the HJB equation (60) is by no means clear. We therefore introduce the notion of *viscosity solution* for (60). To this end, write $\mathscr{D} := \mathbb{R}\times\mathbb{R}_+\times\mathscr{P}_2(\mathbb{R})$, and for $(x,q,\mu)\in\mathscr{D}$, we denote

$$
\begin{aligned}
\mathscr{U}(x,q,\mu) &:= \Big\{\varphi\in\mathbb{C}_b^{2,(2,1)}(\mathscr{D}) : v(x,q,\mu)=\varphi(x,q,\mu)\Big\};\\
\overline{\mathscr{U}}(x,q,\mu) &:= \Big\{\varphi\in\mathscr{U}(x,q,\mu) : v-\varphi \text{ has a strict maximum at } (x,q,\mu)\Big\};\\
\underline{\mathscr{U}}(x,q,\mu) &:= \Big\{\varphi\in\mathscr{U}(x,q,\mu) : v-\varphi \text{ has a strict minimum at } (x,q,\mu)\Big\}.
\end{aligned}
$$

Definition 6.1 *We say a continuous function $v : \mathscr{D}\mapsto\mathbb{R}_+$ is a viscosity subsolution (supersolution, resp.) of (60) in $\mathscr{D}$ if*

$$
\rho\varphi(x,q,\mu) - \sup_{l\in\mathbb{R}_+}[\mathscr{J}^l[\varphi](x,q,\mu) + L(x,q,\mu,l)] \le 0, (resp. \ge 0) \qquad (62)
$$

for every $\varphi\in\overline{\mathscr{U}}(x,q,\mu)$ (resp. $\varphi\in\underline{\mathscr{U}}(x,q,\mu)$).

A function $v : \mathscr{D}\mapsto\mathbb{R}_+$ is called a viscosity solution of (60) on $\mathscr{D}$ if it is both a viscosity subsolution and a viscosity supersolution of (60) on $\mathscr{D}$.

Our main result of this section is the following theorem.

Theorem 6.2 *Assume that the Assumptions 2.1 and 4.2 are in force. Then, the value function v, defined by (49), is a viscosity solution of the HJB equation (60).*

Proof For a fixed $\bar{\mathbb{x}} := (\bar{x},\bar{q},\bar{\mu})\in\mathscr{D}$ with $\bar{\mu}=\mathbb{P}_{\bar{\xi}}$ and $\bar{\xi}\in L^2(\mathcal{F};\mathbb{R})$, and any $\eta>0$, consider the set $\mathscr{D}_{\bar{\mathbb{x}},\eta} := \{\mathbb{x}=(x,q,\mu)\in\mathscr{D} : \|\mathbb{x}-\bar{\mathbb{x}}\|<\eta\}$, where $\|\mathbb{x}-\bar{\mathbb{x}}\| := \Big(|x-\bar{x}|^2+|q-\bar{q}|^2+W_2(\mu,\bar{\mu})\Big)^{1/2}$, and $\mu=\mathbb{P}_\xi$ with $\xi\in L^2(\mathcal{F};\mathbb{R})$.

We first prove that the value function v is a *subsolution* to the HJB equation (60). We proceed by contradiction. Suppose not. Then there exist some $\varphi\in\overline{\mathscr{U}}(\mathbb{x})$ and

$\varepsilon_0 > 0$ such that

$$\rho\varphi(\bar{\mathbb{x}}) - \sup_{l\in\mathbb{R}_+}[\mathscr{J}^l[\varphi](\bar{\mathbb{x}}) + L(\bar{\mathbb{x}}, l)] =: 2\varepsilon_0 > 0. \tag{63}$$

Since $A^l(\mathbb{x}) := \mathscr{J}^l[\varphi](\mathbb{x}) + L(\mathbb{x}, l)$ is uniformly continuous in $\mathbb{x}$, uniformly in l, thanks to Assumption 4.2, one shows that there exists $\eta > 0$ such that for any $\mathbb{x} \in \mathscr{D}_{\bar{\mathbb{x}},\eta}$, it holds that

$$\rho\varphi(\mathbb{x}) - \sup_{l\in\mathbb{R}_+}[\mathscr{J}^l[\varphi](\mathbb{x}) + L(\mathbb{x}, l)] \geq \varepsilon_0. \tag{64}$$

Furthermore, since $\varphi \in \overline{\mathscr{U}}(\mathbb{x})$, we assume without loss of generality that $0 = v(\bar{\mathbb{x}}) - \varphi(\bar{\mathbb{x}})$ is the strict maximum. Thus for the given $\eta > 0$, there exists $\delta > 0$, such that

$$\max\left\{v(\mathbb{x}) - \varphi(\mathbb{x}) : \mathbb{x} \notin \mathscr{D}_{\bar{\mathbb{x}},\eta}\right\} = -\delta < 0. \tag{65}$$

On the other hand, for a fixed $\varepsilon \in (0, \min(\varepsilon_0, \delta\rho))$, by the continuity of v we can assume, modifying $\eta > 0$ if necessary, that

$$|v(\mathbb{x}) - v(\bar{\mathbb{x}})| = |v(\mathbb{x}) - \varphi(\bar{\mathbb{x}})| < \varepsilon, \qquad \mathbb{x} \in \mathscr{D}_{\bar{\mathbb{x}},\eta}. \tag{66}$$

Next, for any $T > 0$ and any $l \in \mathscr{U}_{ad}$ we set $\tau^T := \inf\{t \geq 0 : \bar{\Theta}_t \notin \mathscr{D}_{\bar{\mathbb{x}},\eta}\} \wedge T$, where $\bar{\Theta}_t := (X_t^{\bar{x}}, Q_t^{\bar{q},\bar{\xi},l}, \mathbb{P}_{Q_t^{\bar{q},\bar{\xi}}})$. Applying Itô's formula (14) to $e^{-\rho t}\varphi(\bar{\Theta}_t)$ from 0 to τ^T and noting that $v(\bar{\mathbb{x}}) = \varphi(\bar{\mathbb{x}})$ we have

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{\tau^T} e^{-\rho t} L(\bar{\Theta}_t, l_t)dt + e^{-\rho\tau^T} v(\bar{\Theta}_{\tau^T})\Big] \\
&= \mathbb{E}\Big[\int_0^{\tau^T} e^{-\rho t} L(\bar{\Theta}_t, l_t)dt + e^{-\rho\tau^T} \varphi(\bar{\Theta}_{\tau^T}) + e^{-\rho\tau^T}[v - \varphi](\bar{\Theta}_{\tau^T})\Big] \qquad (67)\\
&= \mathbb{E}\Big[\int_0^{\tau^T} e^{-\rho t}\Big(L(\bar{\Theta}_t, l_t) + \mathscr{J}^l[\varphi](\bar{\Theta}_t) - \rho\varphi(\bar{\Theta}_t)\Big)dt + e^{-\rho\tau^T}[v - \varphi](\bar{\Theta}_{\tau^T})\Big] + v(\bar{\mathbb{x}}) \\
&\leq \mathbb{E}\Big[-\frac{\varepsilon}{\rho}(1 - e^{-\rho\tau^T}) + e^{-\rho\tau^T}[v - \varphi](\bar{\Theta}_{\tau^T})\Big] + v(\bar{\mathbb{x}}) \\
&= \mathbb{E}\Big[e^{-\rho\tau^T}\Big(\frac{\varepsilon}{\rho} + [v - \varphi](\bar{\Theta}_{\tau^T})\Big) : \tau^T < T\Big] \\
&\quad + \mathbb{E}\Big[e^{-\rho\tau^T}\Big(\frac{\varepsilon}{\rho} + [v - \varphi](\bar{\Theta}_{\tau^T})\Big) : \tau^T = T\Big] + v(\bar{\mathbb{x}}) - \frac{\varepsilon}{\rho}.
\end{aligned}$$

Now note that on the set $\{\tau^T < T\}$ we must have $\bar{\Theta}_{\tau^T} \notin \mathscr{D}_{\bar{\mathbb{x}},\eta}$, thus $[v - \varphi](\bar{\Theta}_{\tau^T}) \leq -\delta$, thanks to (65). On the other hand, on the set $\{\tau^T = T\}$ we have $\bar{\Theta}_{\tau^T} = \bar{\Theta}_T \in \mathscr{D}_{\bar{\mathbb{x}},\eta}$, and then (66) implies that $[v - \varphi](\bar{\Theta}_T) \leq v(\bar{\mathbb{x}}) - \varphi(\bar{\Theta}_T) + \varepsilon$. Plugging these facts in (67), we can easily obtain that

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{\tau^T} e^{-\rho t} L(\bar{\Theta}_t, l_t)dt + e^{-\rho\tau^T} v(\bar{\Theta}_{\tau^T})\Big] \\
&\le \Big(\frac{\varepsilon}{\rho} - \delta\Big)\mathbb{P}\{\tau^T < T\} + (\frac{\varepsilon}{\rho} + \varepsilon)e^{-\rho T} + v(\bar{\mathbb{x}}) - \frac{\varepsilon}{\rho} \\
&\le \Big(\frac{\varepsilon}{\rho} + \varepsilon\Big)e^{-\rho T} + v(\bar{\mathbb{x}}) - \frac{\varepsilon}{\rho}.
\end{aligned}$$

Here in the last inequality above we used the fact that $\varepsilon/\rho - \delta < 0$, by definition of ε. Letting $T \to \infty$ we have

$$\mathbb{E}\Big[\int_0^{\tau^T} e^{-\rho t} L(\bar{\Theta}_t, l_t)dt + e^{-\rho\tau^T} v(\bar{\Theta}_{\tau^T})\Big] \le v(\bar{\mathbb{x}}) - \frac{\varepsilon}{\rho}.$$

Since $l \in \mathscr{U}_{ad}$ is arbitrary, this contradicts the dynamic programming principle (53).

The proof that v is viscosity supersolution of (60) is more or less standard, again with the help of Itô's formula (14). We only give a sketch here.

Let $\bar{\mathbb{x}} \in \mathscr{D}$ and $\varphi \in \underline{\mathscr{U}}(\bar{\mathbb{x}})$. Without loss of generality we assume that $0 = v(\bar{\mathbb{x}}) - \varphi(\bar{\mathbb{x}})$ is a global minimum. That is, $v(\mathbb{x}) - \varphi(\mathbb{x}) \ge 0$ for all $\mathbb{x} \in \mathscr{D}$. For any $h > 0$ and $l \in \mathscr{U}_{ad}$, we apply DPP (53) to get

$$\begin{aligned}
0 &\ge \mathbb{E}\Big[\int_0^{h} e^{-\rho t} L(\Theta_t, l_t)dt + e^{-\rho h} v(\Theta_h)\Big] - v(\mathbb{x}) \\
&\ge \mathbb{E}\Big[\int_0^{h} e^{-\rho t} L(\Theta_t, l_t)dt + e^{-\rho h} \varphi(\Theta_h)\Big] - \varphi(\mathbb{x}).
\end{aligned} \tag{68}$$

Applying Itô's formula to $e^{-\rho t}\varphi(\Theta_t)$ from 0 to h we have

$$0 \ge \mathbb{E}\Big[\int_0^{h} e^{-\rho t}\Big(L(\Theta_t, l_t) + \mathscr{J}^l[\varphi](\Theta_t) - \rho\varphi(\Theta_t)\Big)dt\Big]. \tag{69}$$

Dividing both sides by h and sending h to 0, we obtain $\rho\varphi(x,q,\mathbb{P}_\xi) \ge \mathscr{J}^l[\varphi](x,q,\mathbb{P}_\xi) + L(x,q,\mathbb{P}_\xi,l)$. By taking supremum over $l \in \mathscr{U}_{ad}$ on both sides, we conclude

$$\rho\varphi(x,q,\mathbb{P}_\xi) \ge \sup_{l \in \mathscr{U}_{ad}} [\mathscr{J}^l[\varphi](x,q,\mathbb{P}_\xi) + L(x,q,\mathbb{P}_\xi,l)].$$

The proof is now complete. □

Finally, we remark that, as the limiting case of a Bertrand-type of game for a large number of sellers, the value function $v(x,q,\mathbb{P}_\xi)$ in (49) can be thought of as the discounted lifelong expected utility of a representative seller, and thus can be considered as "equilibrium" discounted expected utility for all sellers. Moreover, as one can see in Proposition 5.1, the value function $v(x,q,\mathbb{P}_\xi)$ is uniformly Lipschitz continuous, non-decreasing in x, and decreasing in q. Also, by Assumption 4.2-(vii), the value function is convex in q. Consequently, we see that the value function $v(x,q,\mathbb{P}_\xi)$ resembles the *expected utility function* $U(x,q)$ in [32] which was defined by the following properties:

(i) the mapping $x \mapsto U(x,q)$ is non-decreasing, and $\frac{\partial U(x,q)}{\partial q} < 0$, $\frac{\partial^2 U(x,q)}{\partial q^2} > 0$;
(ii) the mapping $(x,q) \mapsto U(x,q)$ is uniformly Lipschitz continuous.

In particular, we may identify the two functions by setting $U(x,q) = v(x,q,\mathbb{P}_\xi)|_{\xi\equiv q}$, which amounts to saying that the equilibrium density function of a LOB is fully described by the value function of a control problem of the representative seller's Bertrand-type game. This would enhance the notion of "endogenous dynamic equilibrium LOB model" of [32] in a rather significant way.

Acknowledgements The author Jin Ma is supported in part by US NSF grants #DMS-1908665.

References

1. Alfonsi, A., Fruth, A., & Schied, A. (2008). Constrained portfolio liquidation in a limit order book model. *Banach Center Publications*, *83*, 9–25.
2. Alfonsi, A., Fruth, A., & Schied, A. (2010). Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance, 10*(2), 143–157.
3. Alfonsi, A., & Schied, A. (2010). *Optimal Trade Execution and Absence of Price Manipulations in Limit Order Book Models*. SIAM J. Financial Math., **1**(1) , 490-522.
4. Alfonsi, A., Schied, A., & Slynko, A. (2012). *Order book resilience, price manipulation, and the positive portfolio problem*. SIAM J. Financial Math., **3**(1), 511-533.
5. Avellaneda, M., & Stoikov, S. (2008). *High-frequency trading in a limit order book*. Quant. Finance, **8**(3), 217-224.
6. Basna, R., Hilbert, A., & Kolokoltsov, V.N. (2017). *An Approximate Nash Equilibrium for Pure Jump Markov Games of Mean-field-type on Continuous State Space*. Stochastics **89**, no. 6-7, 967-993.
7. Bayraktar, E., & Ludkovski, M. (2014). *Liquidation in limit order books with controlled intensity*. Math. Finance, **24**(4), pp 627-650.
8. Bertrand, J. (1883). *Théorie mathématique de la richesse sociale*. Journal des Savants, **67**, 499-508.
9. Buckdahn, R., Li, J., Peng, S., & Rainer, C. (2017). *Mean-field stochastic differential equations and associated PDEs*. Ann. Probab. **45**(2), 824-878.
10. Cardaliaguet, P. (2013). *Notes on Mean Field Games* (from P-L Lions' lectures at Collège de France). https://www.ceremade.dauphine.fr/ cardalia/.
11. Carmona, R., & Lacker, D. (2015). *A probabilistic weak formulation of mean field games and applications*. Ann. Appl. Probab. **25**(3), 1189-1231.
12. Carmona, R., Delarue, F., & Lachapelle, A. (2013). *Control of McKean-Vlasov dynamics versus mean field games*. Math. Financ. Econ. **7**(2), 131-166.
13. Chan, P., & Sircar, R. (2015). *Bertrand and Cournot mean field games*. Appl. Math. Optim. 71 , no. 3, 533-569.
14. Cont, R., Stoikov, S., & Talreja, R. (2010). *A stochastic model for order book dynamics*. Oper. Res. **58**(3), 549-563.
15. Cournot, A. (1838). *Recherches sur les Principes Mathematiques de la Theorie des Richesses*. Hachette, Paris, . English translation by N. T. Bacon published in Economic Classics, Macmillan, 1897, and reprinted in 1960 by Augustus M. Kelly.
16. Dupuis, P., & Ishii, H. (1991). *On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications*. Stochastics and Stochastic Reports, **35**, 31-62.
17. Friedman, J. (1983). *Oligopoly Theory*. Cambridge University Press,
18. Gatheral, J., Schied, A., Slynko, A. (2011). *Exponential resilience and decay of market impact*. Econophysics of Order-driven Markets, Springer, pp 225-236.

19. Gayduk, R., & Nadtochiy, S. (2018). *Liquidity Effects of Trading Frequency*. Math. Finance 28 (3), 839-876
20. Guéant, O., Lehalle, C.-A., & Fernandez-Tapia, J. (2012). *Optimal portfolio liquidation with limit orders*. SIAM J. Financial Math. 13(1), , 740-764.
21. Guéant, O., Lasry, J.-M., & Lions, P.-L. (2011). *Mean field games and applications*. Paris-Princeton Lectures on Mathematical Finance 2010, Lecture Notes in Mathematics, vol. 2003, Springer Berlin / Heidelberg, pp. 205-266.
22. Hao, T., & Li, J. (2016). *Mean-field SDEs with jumps and nonlocal integral-PDEs*. NoDEA Nonlinear Differential Equations Appl. **23**(2)
23. Harris, C., Howison, S., & Sircar, R. (2010). *Games with Exhaustible Resources*. SIAM J. Appl. Math. 70(7), pages 2556-2581.
24. Ikeda, N., & Watanabe, W. (1981). *Stochastic Differential Equations and Diffusion Processes*. North-Holland Pub. Co.
25. Kingman, J. F. C. (1993). *Poisson processes*. Oxford University Press
26. Lachapelle, A., Lasry, J.-M., Lehalle, C.-A., & Lions, P.-L. (2016). *Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis*. Math. Financ. Econ. Vol.10, Issue 3, 223-262.
27. Ledvina, A., & Sircar, R. (2011). *Dynamic Bertrand Oligopoly*. App. Math. Optim. Vol. 63, Issue 1 (2011), 11-44.
28. Ledvina, A., & Sircar, R. (2012). *Oligopoly games under asymmetric costs and an application to energy production*. Math. Financ. Econ. 6, no. 4, 261-293.
29. Lasry, J.-M., Lions, P.L. (2007). *Mean field games*. Jpn. J. Math. 2, 229-260.
30. Lokka, A. (2014). *Optimal liquidation in a limit order book for a risk averse investor*. Math. Finance, Vol. 24, Issue 4, pp. 696-727.
31. Ma, J. (1993). *Discontinuous reflection and a class of singular stochastic control problems for diffusions*. Stochastics and Stochastic Reports, Vol.44, No.3-4 , 225-252.
32. Ma, J., Wang, X., & Zhang, J. (2015). *Dynamic equilibrium limit order book model and optimal execution problem*. Math. Control Relat. Fields, Vol.5, No.3
33. Ma, J., Yong, J., & Zhao, Y. (2010). *Four Step Scheme for General Markovian Forward-Backward SDEs*. J. Syst. Sci. Complex. 23, no. 3, 546-571.
34. Noh, E. (2018). *Equilibrium Model of Limit Order Book and Optimal Execution Problem*. Ph.D. dissertation, Department of Mathematics, University of Southern California
35. Obizhaeva, A., & Wang, J. (2013). *Optimal trading strategy and supply/demand dynamics*. Journal of Financial Markets, 71, No.4 , 1027-1063.
36. Potters, M., & Bouchaud, J.-P. (2003). *More statistical properties of order books and price impact*. Phys. A, 324, No. 1-2 , 133-140.
37. Protter, P. (1983). *Point process differentials with evolving intensities*. Nonlinear Stochastic Problems, Vol 104 of the series NATO ASI Series , 133-140.
38. Protter, P. (1990). *Stochastic Integrals and Stochastic Differential Equations*. , Springer.
39. Predoiu, S., Shaikhet, G., & Shreve, S. (2011). *Optimal execution of a general one-sided limit-order book*. SIAM Journal on Financial Mathematics, 2 , 183-212.
40. Rosu, I. (2009). *A dynamic model of the limit order book*. The Review of Financial Studies, 22, 4601-4641.
41. Vives, X. (1999). *Oligopoly pricing: old ideas and new tools*. MIT Press, Cambridge, MA

Bounded Regret for Finitely Parameterized Multi-Armed Bandits

Kishan Panaganti, Dileep Kalathil, and Pravin Varaiya

Abstract We consider multi-armed bandits where the model of the underlying stochastic environment is characterized by a common unknown parameter. The true parameter is unknown to the learning agent. However, the set of possible parameters, which is finite, is known a priori. We propose an algorithm that is simple and easy to implement, which we call Finitely Parameterized Upper Confidence Bound (FP-UCB) algorithm, which uses the information about the underlying parameter set for faster learning. In particular, we show that the FP-UCB algorithm achieves a bounded regret under a structural condition on the underlying parameter set. We also show that, if the underlying parameter set does not satisfy this structural condition, the FP-UCB algorithm achieves a logarithmic regret, but with a smaller preceding constant compared to the standard UCB algorithm. We also validate the superior performance of the FP-UCB algorithm through extensive numerical simulations.

1 Introduction

The Multi-Armed Bandit (MAB) problem is a canonical formalism for studying how an agent learns to take optimal actions through repeated interactions with a stochastic environment. The learning agent receives a reward at each time step which will depend on the action of the agent as well as the stochastic uncertainty of the environment. The goal of the agent is to act so as to maximize the cumulative reward. When the model of the environment is known, computing the optimal action is a standard optimization problem. The challenge in MAB is that the agent does not know the stochastic model of environment a priori. The agent needs to *explore*, i.e., take ac-

Kishan Panaganti and Dileep Kalathil are with the Department of Electrical and Computer Engineering at Texas A&M University, College Station, Texas, USA, e-mail: kpb, dileep.kalathil@tamu.edu · Pravin Varaiya is with the Department of Electrical Engineering and Computer Sciences at University of California, Berkeley, California, USA, e-mail: varaiya@berkeley.edu

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_17

tions to gather information and estimate the model of the system. At the same time the agent must *exploit* the available information to maximize the cumulative reward. This *exploration vs. exploitation* trade-off is at the core of the MAB problem.

Lai and Robbins in their seminal paper [19] formulated the non-Bayesian stochastic MAB problem and characterized the performance of a learning algorithm using the metric of *regret*. They showed that no learning algorithm can achieve a regret better than $\mathscr{O}(\log T)$. They also proposed a learning algorithm that achieves an asymptotic logarithmic regret, matching the fundamental lower bound. A simple index-based algorithm called UCB algorithm was introduced in [5] which achieves the order optimal regret in a non-asymptotic manner. This approach led to a number of interesting algorithms, among them linear bandits [13], contextual bandits [11], combinatorial bandits [10], and decentralized and multi-player bandits [15].

Thompson (Posterior) Sampling is another class of algorithms that give superior numerical performance for MAB problems. The posterior sampling heuristic was first introduced by Thompson [25], but the first rigorous performance guarantee, an $\mathscr{O}(\log T)$ regret, was given in [2]. The Thompson sampling idea has been used in algorithms for bandits with multiple plays [17], contextual bandits [3], general online learning problem [14], and reinforcement learning [23]. Both classes of algorithms have been used in a number of practical applications, like communication networks [24], smart grids [16], and recommendation systems [29].

Our contribution: We consider a class of multi-armed bandits problems where the reward corresponding to each arm is characterized by a common unknown parameter with a finite set of possible values. This restriction is inspired by real-world applications. For example, in recommendation systems and e-commerce applications (Amazon, Netflix), it is typical to assume that each user has a certain 'type' parameter (denoted by θ in our formulation), and the set of possible parameter values is finite. The preferences of the user is characterized by her type (for example, prefer science books over fiction books). The set of all possible types and the preferences of each type may be known a priori, but the type of a new user may be unknown. So, instead of learning the preferences of this user over all possible choices, it may be easier to learn the type parameter of this user from a few observations. In this work, we propose an algorithm that explicitly uses the availability of such structural information about the underlying parameter set which enables faster learning.

We propose an algorithm that is simple and easy to implement, which we call FP-UCB algorithm, which uses the structural information for faster learning. We show that the proposed FP-UCB algorithm can achieve a bounded regret ($\mathscr{O}(1)$) under some structural condition on the underlying parameter set. This is in sharp contrast to the increasing ($\mathscr{O}(\log T)$) regret of standard multi-armed bandits algorithms. We also show that, if the underlying parameter set does not satisfy the structural condition, the FP-UCB algorithm achieves a regret of $\mathscr{O}(\log T)$, but with a smaller preceding constant compared to the standard UCB algorithm. The regret achieved by our algorithm also matches with the fundamental lower bound given by [1]. One remarkable aspect of our algorithm is that, it is oblivious to whether the underlying parameter set satisfies the necessary condition or not, thereby avoiding re-tuning of the algorithm depending on the problem instance. Instead, it achieves the best possible performance given the problem instance.

Related work: Finitely parameterized multi-armed bandits problem were first studied by Agrawal et al. [1]. They proposed an algorithm for this setting, and proved that it achieves a bounded regret when the parameter set satisfies some necessary condition, and logarithmic regret otherwise. However, their algorithm is rather complicated, which limits practical implementations and extension to other settings. The regret analysis is also involved and asymptotic in nature, different from the recent simpler index-based bandits algorithms and their finite time analysis. [1] also provided a fundamental lower bound for this class of problems. Compared to this work, our FP-UCB algorithm is simple, easy to implement, and easy to analyze, while providing non-asymptotic performance guarantees that match the lower bound.

Some recent works exploit the available structure of the MAB problem to get tighter regret bounds. In particular, [4] [20] [22] [12] consider the problem setting similar to our paper where the mean reward of each arm is characterized by a single unknown parameter. [4] assumes that the reward functions are continuous in the global parameter and gives a bounded regret result. [20] gives specific conditions on the mean reward to achieve a bounded regret. [22] considers a latent bandit problem where the reward distributions are partitioned into a number of clusters and indexed by a latent parameter corresponding to the cluster. [12] characterizes the minimal rates at which sub-optimal arms have to be explored depending on the structural information, and proposes an algorithm that achieves these rates. [8] [7] [26] exploit a different structural information where it is shown that if the mean value of the best arm and the second best arm (but not the identity of the arms) are known, a bounded regret can be achieved. There also are bandit algorithms that exploit side information [28] [9], and recently in the context of contextual bandits [6]. Our problem formulation, algorithm, and analysis are different from these works. We also note that our problem formulation is fundamentally different from the system identification problems [21] [18] because the goal here is to learn an optimal policy online.

2 Problem Formulation

We consider the following sequential decision-making problem. In each time step $t \in \{1,2,\ldots,T\}$, the agent selects an arm (action) from the set of L possible arms, denoted $a(t) \in [L] = \{1,\ldots,L\}$. Each arm i, when selected, yields a random real-valued reward. Let $X_i(\tau)$ be the random reward from arm i in its τth selection. We assume that $X_i(\tau)$ is drawn according to a probability distribution $P_i(\cdot;\theta^o)$ with mean $\mu_i(\theta^o)$. Here θ^o is the (true) parameter that determines the distribution of the stochastic rewards. The agent does not know θ^o or the corresponding mean value $\mu_i(\theta^o)$. The random rewards obtained from playing an arm repeatedly are i.i.d. and independent of the plays of the other arms. The rewards are bounded with support in $[0,1]$. The goal of the agent is to select a sequence of actions that maximizes the expected cumulative reward, $\mathbb{E}[\sum_{t=1}^{T} \mu_{a(t)}(\theta^o))]$. The action $a(t)$ depends on the history of observations available to the agent until time t. So, $a(t)$ is stochastic and the expectation is with respect to its randomness.

Clearly, the optimal choice is to select the best arm (the arm with the highest mean value) all the time, i.e., $a(t) = a^*(\theta^o), \forall t$, where $a^*(\theta^o) = \arg\max_{i \in [L]} \mu_i(\theta^o)$. However, the agent will be able to make this optimal decision only if she knows the parameter θ^o or the corresponding mean values $\mu_i(\theta^o)$ for all i. The goal of a MAB algorithm is to learn to make the optimal sequence of decisions without knowing the true parameter θ^o.

We consider the setting where the agent knows the set of possible parameters Θ. We assume that Θ is finite. If the true parameter were $\theta \in \Theta$, then agent selecting arm i will get a random reward drawn according to a distribution $P_i(\cdot;\theta)$ with mean $\mu_i(\theta)$. We assume that for each $\theta \in \Theta$, the agent knows $P_i(\cdot;\theta)$ and $\mu_i(\theta)$ for all $i \in [L]$. The optimal arm corresponding to the parameter θ is $a^*(\theta) = \arg\max_{i \in [L]} \mu_i(\theta)$. We emphasize that the agent does not know the true parameter θ^o (and hence the optimal action $a^*(\theta^o)$) except that it is in the finite set Θ.

In the multi-armed bandits literature, it is standard to characterize the performance of an online learning algorithm using the metric of regret. Regret is defined as the performance loss of an algorithm as compared to the optimal algorithm with complete information. Since this is $b(t) = a^*(\theta^o)$, the expected cumulative regret of a multi-armed bandits algorithm after T time steps is defined as

$$\mathbb{E}[R(T)] := \mathbb{E}\left[\sum_{t=1}^{T} (\mu_{a^*(\theta^o)}(\theta^o) - \mu_{a(t)}(\theta^o))\right]. \tag{1}$$

The goal of a MAB learning algorithm is to select actions sequentially in order to minimize $\mathbb{E}[R(T)]$.

3 UCB Algorithm for Finitely Parameterized Multi-Armed Bandits

In this section, we present our algorithm for finitely parameterized multi-armed bandits and the main theorem. We first introduce a few notations for presenting the algorithm and the results succinctly.

Let $n_i(t)$ be the number of times arm i has been selected by the algorithm until time t, i.e., $n_i(t) = \sum_{\tau=1}^{t} \mathbb{1}\{a(\tau) = i\}$. Here $\mathbb{1}\{.\}$ is an indicator function. Define the empirical mean corresponding to arm i at time t as,

$$\hat{\mu}_i(t) := \frac{1}{n_i(t)} \sum_{\tau=1}^{n_i(t)} X_i(\tau). \tag{2}$$

Define the set $A := \{a^*(\theta) : \theta \in \Theta\}$, which is the collection of optimal arms corresponding to all parameters in Θ. Intuitively, a learning agent can restrict selection to arms from the set A. Clearly, $A \subset [L]$ and this reduction can be useful when $|A|$ is much smaller than L.

Our FP-UCB Algorithm is given in Algorithm 1. Figure 1 gives an illustration of the episodes and time slots of the FP-UCB algorithm.

For stating the main result, we introduce a few more notations. We define the *confusion set* $B(\theta^o)$ and $C(\theta^o)$ as,

$$B(\theta^o) := \{\theta \in \Theta : a^*(\theta) \neq a^*(\theta^o) \text{ and } \mu_{a^*(\theta^o)}(\theta^o) = \mu_{a^*(\theta^o)}(\theta)\},$$
$$C(\theta^o) := \{a^*(\theta) : \theta \in B(\theta^o)\}.$$

Intuitively, $B(\theta^o)$ is the set of parameters that can be confused with the true parameter θ^o. If $B(\theta^o)$ is non-empty, selecting $a^*(\theta^o)$ and estimating the empirical mean is not sufficient to identify the true parameter because the same mean reward can result from other parameters in $B(\theta^o)$. So, if $B(\theta^o)$ is non-empty, more exploration (i.e., selecting sub-optimal arms other than $a^*(\theta^o)$) is necessary to identify the true parameter. This exploration will contribute to the regret. On the other hand, if $B(\theta^o)$ is empty, the optimal parameter can be identified with much less exploration, which results in a bounded regret. $C(\theta^o)$ is the corresponding set of arms that needs to be explored sufficiently to identify the optimal parameter. So, whether $B(\theta^o)$ is empty or not is the structural condition that decides the performance of the algorithm.

We make the following assumption.

Assumption (Unique best action) For all $\theta \in \Theta$, the optimal action, $a^*(\theta)$, is unique.

We note that this is a standard assumption in the literature. This assumption can be removed at the expense of more notations. We define Δ_i as,

$$\Delta_i := \mu_{a^*(\theta^o)}(\theta^o) - \mu_i(\theta^o), \tag{3}$$

which is the difference between the mean value of the optimal arm and the mean value of arm i for the true parameter θ^o. This is the standard optimality gap notion used in the MAB literature [5]. Without loss of generality assume natural logarithms.

For each arm in $i \in C(\theta^o)$, we define,

$$\beta_i := \min_{\theta : \theta \in B(\theta^o), a^*(\theta) = i} |\mu_i(\theta^o) - \mu_i(\theta)|. \tag{4}$$

We use the following Lemma to compare our result with classical MAB result. The proof for this lemma is given in the appendix.

Lemma 1 *Let Δ_i and β_i be as defined in* (3) *and* (4) *respectively. Then, for each $i \in C(\theta^o)$, $\beta_i > 0$. Moreover, $\beta_i > \Delta_i$.*

We now present the finite time performance guarantee for our FP-UCB algorithm.

Algorithm 1 FP-UCB

```
Initialization: Select each arm in the set A once
Initialize episode number k = 1, time step t = |A| + 1
while t ≤ T do
    t_k = t − 1
    Compute the set
```

$$A_k = \left\{ a^*(\theta), \theta \in \Theta : \forall i \in A,\ |\hat{\mu}_i(t_k) - \mu_i(\theta)| \leq \sqrt{\frac{3\log(k)}{n_i(t_k)}} \right\}$$

```
    if |A_k| ≠ 0 then
        Select each arm in the set A_k once
        t ← t + |A_k|
    else
        Select each arm in the set A once
        t ← t + |A|
    end if
    k ← k + 1
end while
```

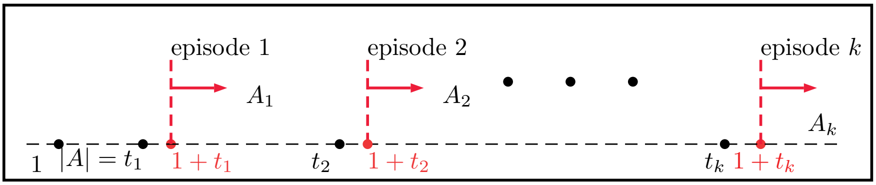

Fig. 1: An illustration of the episodes and time slots of the FP-UCB algorithm.

Theorem 1 *Under the FP-UCB algorithm,*

$$\mathbb{E}[R(T)] \leq D_1, \quad \textit{if } B(\theta^o) \textit{ empty, and}$$

$$\mathbb{E}[R(T)] \leq D_2 + 12\log(T) \sum_{i \in C(\theta^o)} \frac{\Delta_i}{\beta_i^2}, \quad \textit{if } B(\theta^o) \textit{ non-empty,} \tag{5}$$

where D_1 and D_2 are problem dependent constants that depend only on the problem parameters $|A|$ and $(\mu_i(\theta), \theta \in \Theta)$, but do not depend on T.

Remark 1 (Comparison with the classical MAB results) Both UCB type algorithms and Thompson Sampling type algorithms give a problem dependent regret bound $\mathscr{O}(\log T)$. More precisely, assuming that the optimal arm is arm 1, the regret of the UCB algorithm, $\mathbb{E}[R_{\text{UCB}}(T)]$, is given by [5]

$$\mathbb{E}[R_{\text{UCB}}(T)] = \mathscr{O}\left(\sum_{i=2}^{L} \frac{1}{\Delta_i} \log T \right).$$

On the other hand, the FP-UCB algorithm achieves the regret, $\mathbb{E}[R_{\text{FP-UCB}}(T)]$,

$$\mathscr{O}(1), \text{if } B(\theta^o) \text{ empty, and } \mathscr{O}\left(\sum_{i \in C(\theta^o)} \frac{\Delta_i}{\beta_i^2} \log T\right), \text{if } B(\theta^o) \text{ non-empty.}$$

Clearly, for some MAB problems, FP-UCB algorithm achieves a bounded regret ($\mathscr{O}(1)$) as opposed to the increasing regret ($\mathscr{O}(\log T)$) of the standard UCB algorithm. Even in the cases where FP-UCB algorithm incurs an increasing regret ($\mathscr{O}(\log T)$), the preceding constant (Δ_i/β_i^2) is smaller than the preceding constant ($1/\Delta_i$) of the standard UCB algorithm because $\beta_i > \Delta_i$.

We now give the asymptotic lower bound for the finitely parameterized multi-armed bandits problem from [1], for comparing the performance of our FP-UCB algorithm.

Theorem 2 (Lower bound [1])

For any uniformly good control scheme under the parameter θ^o,

$$\liminf_{T\to\infty} \frac{\mathbb{E}[R(T)]}{\log(T)} \geq \min_{h\in H} \max_{\theta\in B(\theta^o)} \frac{\sum_{u\in A\setminus\{a^*(\theta^o)\}} h_u(\mu_{a^*(\theta^o)}(\theta^o) - \mu_u(\theta^o))}{\sum_{u\in A\setminus\{a^*(\theta^o)\}} h_u D_u(\theta^o \| \theta)}.$$

where H is a probability simplex with $|A|-1$ vertices and, for any $u \in A \setminus \{a^*(\theta^o)\}$, $D_u(\theta^o\|\theta) = \int P_u(x;\theta^o)\log(P_u(x;\theta^o)/P_u(x;\theta))dx$ is the KL-divergence between the probability distributions $P_u(\cdot;\theta^o)$ and $P_u(\cdot;\theta)$.

Remark 2 (Optimality of the FP-UCB algorithm) From Theorem 2, the achievable regret of any multi-armed bandits learning algorithm is lower bounded by $\Omega(1)$ when $B(\theta^o)$ is empty, and $\Omega(\log T)$ when $B(\theta^o)$ is non-empty. Our FP-UCB algorithm achieves these bounds and hence achieves the order optimal performance.

4 Analysis of the FP-UCB Algorithm

In this section, we give the proof of Theorem 1. For reducing the notation, without loss of generality we assume that the true optimal arm is arm 1, i.e., $a^* = a^*(\theta^o) = 1$. We will also denote $\mu_j(\theta^o)$ as μ_j^o, for any $j \in A$.

Now, we can rewrite the expected regret from (1) as

$$\begin{aligned}\mathbb{E}[R(T)] &= \mathbb{E}\left[\sum_{t=1}^{T}(\mu_1^o - \mu_{a(t)}^o)\right] \\ &= \sum_{i=2}^{L}\Delta_i\,\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}\{a(t)=i\}\right] = \sum_{i=2}^{L}\Delta_i\,\mathbb{E}\left[n_i(T)\right].\end{aligned}$$

Since the algorithm selects arms only from the set A, this can be written as

$$\mathbb{E}[R(T)] = \sum_{i\in A}\Delta_i\,\mathbb{E}\left[n_i(T)\right]. \tag{6}$$

We first prove the following important propositions.

Proposition 1 *For all $i \in A \setminus C(\theta^o), i \neq 1$, under FP-UCB algorithm,*

$$\mathbb{E}\left[n_i(T)\right] \leq C_i, \tag{7}$$

where C_i is a problem dependent constant that does not depend on T.

Proof Consider an arm $i \in A \setminus C(\theta^o), i \neq 1$. Then, by definition, there exists a $\theta \in \Theta$ such that $a^*(\theta) = i$. Fix a θ which satisfies this condition. Define

$$\alpha_1(\theta) := |\mu_1(\theta^o) - \mu_1(\theta)|.$$

It is straightforward to note that when $i \in A \setminus C(\theta^o)$, then the θ which we considered above is not in $B(\theta^o)$. Hence, by definition, $\alpha_1(\theta) > 0$.

For notational convenience, we will denote $\mu_j(\theta)$ simply as μ_j, for any $j \in A$. Notice that the algorithm picks i^{th} arm once in $t \in \{1, \ldots, |A|\}$. Define K_T (note that this is a random variable) to be the total number of episodes in time horizon T for the FP-UCB algorithm. It is straightforward that $K_T \leq T$. Now,

$$\begin{aligned}
\mathbb{E}[n_i(T)] &= 1 + \mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t) = i\}\right] \\
&\overset{(a)}{=} 1 + \mathbb{E}\left[\sum_{k=1}^{K_T} (\mathbb{1}\{i \in A_k\} + \mathbb{1}\{A_k = \varnothing\})\right] \\
&\leq 1 + \sum_{k=1}^{T} \left[\mathbb{P}(\{i \in A_k\}) + \mathbb{P}(\{A_k = \varnothing\})\right] \qquad (8) \\
&= 1 + \sum_{k=1}^{T} \left[\mathbb{P}(\{i \in A_k, 1 \in A_k\}) + \mathbb{P}(\{i \in A_k, 1 \notin A_k\}) + \mathbb{P}(\{A_k = \varnothing\})\right] \\
&\leq 1 + \sum_{k=1}^{T} \left[\mathbb{P}(\{i \in A_k, 1 \in A_k\}) + \mathbb{P}(\{i \in A_k, 1 \notin A_k\}) + \mathbb{P}(\{i \notin A_k, 1 \notin A_k\})\right] \\
&\leq 1 + \sum_{k=1}^{T} [\mathbb{P}(\{i \in A_k, 1 \in A_k\}) + \mathbb{P}(\{1 \notin A_k\})]. \qquad (9)
\end{aligned}$$

Here (a) follows from the algorithm definition.

We will first analyze the second summation term in (9). First observe that, we can write $n_j(t_k) = 1 + \sum_{\tau=1}^{k-1} (\mathbb{1}\{j \in A_\tau\} + \mathbb{1}\{A_\tau = \varnothing\})$ for any $j \in A$ and episode k. Thus, $n_j(t_k)$ lies between 1 and k. Now,

$$
\begin{aligned}
\sum_{k=1}^{T} \mathbb{P}(\{1 \notin A_k\}) &\overset{(a)}{=} \sum_{k=1}^{T} \mathbb{P}\left(\bigcup_{j \in A} \left\{ |\hat{\mu}_j(t_k) - \mu_j^o| > \sqrt{\frac{3\log k}{n_j(t_k)}} \right\} \right) \\
&\overset{(b)}{\leq} \sum_{k=1}^{T} \sum_{j \in A} \mathbb{P}\left(|\hat{\mu}_j(t_k) - \mu_j^o| > \sqrt{\frac{3\log k}{n_j(t_k)}} \right) \\
&\overset{(c)}{=} \sum_{k=1}^{T} \sum_{j \in A} \mathbb{P}\left(\left| \frac{1}{n_j(t_k)} \sum_{\tau=1}^{n_j(t_k)} X_j(\tau) - \mu_j^o \right| > \sqrt{\frac{3\log k}{n_j(t_k)}} \right) \\
&\overset{(d)}{\leq} \sum_{k=1}^{T} \sum_{j \in A} \sum_{m=1}^{k} \mathbb{P}\left(\left| \frac{1}{m} \sum_{\tau=1}^{m} X_j(\tau) - \mu_j^o \right| > \sqrt{\frac{3\log k}{m}} \right) \\
&\overset{(e)}{\leq} \sum_{k=1}^{T} \sum_{j \in A} \sum_{m=1}^{k} 2\exp\left(-2m \frac{3\log k}{m} \right) = \sum_{k=1}^{T} \sum_{j \in A} 2k^{-5} \leq 4|A|. \quad (10)
\end{aligned}
$$

Here (a) follows from algorithm definition, (b) from the union bound, and (c) from the definition in (2). Inequality (d) follows by conditioning the random variable $n_j(t_k)$ that lies between 1 and k for any $j \in A$ and episode k. Inequality (e) follows from Hoeffding's inequality [27, Theorem 2.2.6].

For analyzing the first summation term in (9), define the event $E_k := \{n_1(t_k) < 12\log k/\alpha_1^2(\theta)\}$. Denote the complement of this event as E_k^c. Now the first summation term in (9) can be written as

$$
\sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k\}) = \underbrace{\sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k, E_k^c\})}_{= \, Term_1} + \underbrace{\sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k, E_k\})}_{= \, Term_2}. \quad (11)
$$

Analyzing $Term_1$ in (11), we get,

$$
\begin{aligned}
&\mathbb{P}(\{i \in A_k, 1 \in A_k, E_k^c\}) \\
&= \mathbb{P}\left(\bigcap_{j \in A} \{ |\hat{\mu}_j(t_k) - \mu_j^o| < \sqrt{\frac{3\log k}{n_j(t_k)}} \} \bigcap_{j \in A} \{ |\hat{\mu}_j(t_k) - \mu_j| < \sqrt{\frac{3\log k}{n_j(t_k)}} \} \bigcap E_k^c \right) \\
&\leq \mathbb{P}\left(\{ |\hat{\mu}_1(t_k) - \mu_1^o| < \sqrt{\frac{3\log k}{n_1(t_k)}} \}, |\{ \hat{\mu}_1(t_k) - \mu_1 | < \sqrt{\frac{3\log k}{n_1(t_k)}} \}, E_k^c \right) = 0. \quad (12)
\end{aligned}
$$

This is because the events $\{|\hat{\mu}_1(t_k) - \mu_1^o| < \sqrt{\frac{3\log k}{n_1(t_k)}}\}$ and $\{|\hat{\mu}_1(t_k) - \mu_1| < \sqrt{\frac{3\log k}{n_1(t_k)}}\}$ are disjoint under E_k^c, that is, when $n_1(t_k) \geq 12\log(k)/\alpha_1^2(\theta)$. To see this, notice that

$$\left\{|\hat{\mu}_1(t_k) - \mu_1^o| < \sqrt{\frac{3\log k}{n_1(t_k)}}\right\} \subseteq \left\{|\hat{\mu}_1(t_k) - \mu_1^o| < \frac{\alpha_1(\theta)}{2}\right\},$$

$$\left\{|\hat{\mu}_1(t_k) - \mu_1| < \sqrt{\frac{3\log k}{n_1(t_k)}}\right\} \subseteq \left\{|\hat{\mu}_1(t_k) - \mu_1| < \frac{\alpha_1(\theta)}{2}\right\},$$

for $n_1(t_k) \geq 12\log k/\alpha_1^2(\theta)$. Moreover, since $|\mu_1^o - \mu_1| = \alpha_1(\theta)$, $\{|\hat{\mu}_1(t_k) - \mu_1^o| < \alpha_1(\theta)/2\}$ and $\{|\hat{\mu}_1(t_k) - \mu_1| < \alpha_1(\theta)/2\}$ are disjoint sets. Hence, their subsets are also disjoint.

For analyzing $Term_2$ in (11), we start by setting up few notations. Define $n_1'(t_k) := 1 + \sum_{\tau=1}^{k-1} \mathbb{1}\{1 \in A_\tau\}$. Note that, according to the FP-UCB algorithm, arm 1 can be selected if A_τ is empty as well, so $n_1'(t_k) \leq n_1(t_k)$. Define $k_i(\theta)$ and $m(k)$ as,

$$k_i(\theta) := \min\left\{k : k \geq 3, k > \lceil 12\log(k)/\alpha_1^2(\theta)\rceil\right\}, \tag{13}$$

$$m(k) := \max\{1, k - \lceil 12\log(k)/\alpha_1^2(\theta)\rceil\}. \tag{14}$$

Note that $k_i(\theta)$ is a problem dependent constant and does not depend on T. Also, $m(k) = k - \lceil 12\log(k)/\alpha_1^2(\theta)\rceil$ for all $k \geq k_i(\theta)$. We claim that for all $k \geq k_i(\theta)$,

$$\left\{n_1'(t_k) < 12\log(k)/\alpha_1^2(\theta)\right\} \subseteq \left\{1 \notin A_\tau, \text{for some } \tau, m(k) \leq \tau \leq k-1\right\}. \tag{15}$$

To see this, suppose there exists no τ, $m(k) \leq \tau \leq k-1$, such that $1 \notin A_\tau$. Then, $1 \in A_\tau$ for all τ, where $m(k) \leq \tau \leq k-1$. So, by definition $n_1'(t_k) \geq (k - m(k)) = \lceil 12\log(k)/\alpha_1^2(\theta)\rceil$ for $k \geq k_i(\theta)$. So, the complement of the RHS of (15) is a subset of the complement of the LHS of (15). Hence the claim follows.

Now,

$$\begin{aligned}
\sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k, E_k\}) &\leq \sum_{k=1}^{T} \mathbb{P}(E_k) \\
&\stackrel{(a)}{\leq} \sum_{k=1}^{T} \mathbb{P}\left(n_1'(t_k) < 12\log(k)/\alpha_1^2(\theta)\right) \\
&\stackrel{(b)}{\leq} k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \mathbb{P}(n_1'(t_k) < 12\log(k)/\alpha_1^2(\theta)) \\
&\stackrel{(c)}{\leq} k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \mathbb{P}(\{1 \notin A_\tau, \text{for some } \tau, m(k) \leq \tau \leq k-1\}) \\
&\stackrel{(d)}{=} k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \mathbb{P}\left(\bigcup_{\tau=m(k)}^{k-1} \bigcup_{j \in A} |\hat{\mu}_j(\tau) - \mu_j^o| > \sqrt{\frac{3\log\tau}{n_j(t_\tau)}}\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \sum_{\tau=m(k)}^{k-1} \sum_{j \in A} \mathbb{P}\left(|\hat{\mu}_j(\tau) - \mu_j^o| > \sqrt{\frac{3\log\tau}{n_j(t_\tau)}} \right) \\
&\overset{(e)}{\leq} k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \sum_{\tau=m(k)}^{k-1} \frac{2|A|}{\tau^5} \qquad (16) \\
&\leq k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \frac{2|A|k}{(m(k))^5} \\
&= k_i(\theta) + \sum_{k=k_i(\theta)}^{T} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5} \overset{(f)}{=} k_i(\theta) + K_i(\theta), \qquad (17)
\end{aligned}
$$

where $K_i(\theta)$ is a problem dependent constant that does not depend on T.

In the above analysis, (a) follows from the definition of E_k and the observation that $n'_1(t_k) \leq n_1(t_k)$. Considering T to be greater than or equal to $k_i(\theta)|A|$, equality (b) follows; note that this is an artifact of the proof technique and does not affect the theorem statement since $\mathbb{E}[n_i(T')]$, for any T' less than $k_i(\theta)|A|$, can be trivially upper bounded by $\mathbb{E}[n_i(T)]$. Inequality (c) follows from (15), (d) by the FP-UCB algorithm, (e) is similar to the analysis in (10), and (f) follows from the fact that $k > \lceil 12\log(k)/\alpha_1^2(\theta) \rceil$ for all $k \geq k_i(\theta)$.

Now, using (17) and (12) in (11), we get,

$$\sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k\}) \leq k_i(\theta) + K_i(\theta). \qquad (18)$$

Using (18) and (10) in (9), we get,

$$\mathbb{E}[n_i(T)] \leq C_i,$$

where $C_i = 1 + 4|A| + \min_{\theta : a^*(\theta)=i}(k_i(\theta) + K_i(\theta))$, which is a problem dependent constant that does not depend on T. This concludes the proof. □

Proposition 2 *For any $i \in C(\theta^o)$, under the FP-UCB algorithm,*

$$\mathbb{E}\left[n_i(T)\right] \leq 2 + 4|A| + \frac{12\log(T)}{\beta_i^2}. \qquad (19)$$

Proof Fix an $i \in C(\theta^o)$. Then there exists $\theta \in B(\theta^o)$ such that $a^*(\theta) = i$. Fix θ which satisfies this condition. Define the event $F(t) := \left\{ n_i(t-1) < 12\log T/\beta_i^2 \right\}$. Now,

$$\mathbb{E}[n_i(T)] = 1 + \mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i\}\right]$$
$$= 1 + \mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i, F(t)\}\right] + \mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i, F^c(t)\}\right]. \tag{20}$$

Analyzing the first summation term in (20) we get,

$$\mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i, F(t)\}\right] = \mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i\}\mathbb{1}\left\{n_i(t-1) < 12\log T/\beta_i^2\right\}\right]$$
$$\leq 1 + 12\log T/\beta_i^2. \tag{21}$$

We use the same decomposition as in the proof of Proposition 1 for the second summation term in (20). Thus we get,

$$\mathbb{E}\left[\sum_{t=|A|+1}^{T} \mathbb{1}\{a(t)=i, F^c(t)\}\right] =$$
$$\mathbb{E}\left[\sum_{k=1}^{K_T} \mathbb{1}\{i \in A_k, F^c(t_k+1)\} + \mathbb{1}\{A_k = \varnothing, F^c(t_k+1)\}\right]$$
$$\leq \sum_{k=1}^{T} \mathbb{P}(\{i \in A_k, 1 \in A_k, F^c(t_k+1)\}) \tag{22}$$
$$+ \sum_{k=1}^{T} \mathbb{P}(\{1 \notin A_k, F^c(t_k+1)\}), \tag{23}$$

following the analysis in (9). First, consider (23). From the analysis in (10) we have

$$\sum_{k=1}^{T} \mathbb{P}(\{1 \notin A_k, F^c(t_k+1)\}) \leq \sum_{k=1}^{T} \mathbb{P}(\{1 \notin A_k\}) \leq 4|A|. \tag{24}$$

For any $i \in A$ and episode k under event $F^c(t_k+1)$, we have

$$n_i(t_k) \geq \frac{12\log T}{\beta_i^2} \geq \frac{12\log t_k}{\beta_i^2} \geq \frac{12\log k}{\beta_i^2}$$

since t_k satisfies $k \leq t_k \leq T$. From (4), it further follows that

$$\sqrt{\frac{3\log k}{n_j(t_k)}} \leq \frac{\beta_i}{2} \leq \frac{|\mu_i(\theta^o) - \mu_i(\theta)|}{2}.$$

So, following the analysis in (12) for (22), we get

$$\mathbb{P}(\{i \in A_k, 1 \in A_k, F^c(t_k+1)\})$$
$$= \mathbb{P}\left(\begin{array}{l} \bigcap_{j\in A}\{|\hat{\mu}_j(t_k) - \mu_j(\theta^o)| < \sqrt{\frac{3\log k}{n_j(t_k)}}\}, \\ \bigcap_{j\in A}\{|\hat{\mu}_j(t_k) - \mu_j(\theta)| < \sqrt{\frac{3\log k}{n_j(t_k)}}\}, F^c(t_k+1) \end{array} \right)$$
$$\leq \mathbb{P}\left(\begin{array}{l} \{|\hat{\mu}_i(t_k) - \mu_i(\theta^o)| < \sqrt{\frac{3\log k}{n_i(t_k)}}\}, \\ \{|\hat{\mu}_i(t_k) - \mu_i(\theta)| < \sqrt{\frac{3\log k}{n_i(t_k)}}\}, F^c(t_k+1) \end{array} \right) = 0. \quad (25)$$

Using equations (21), (24), and (25) in (20), we get

$$\mathbb{E}[n_i(T)] \leq 2 + 4|A| + \frac{12\log(T)}{\beta_i^2}.$$

This completes the proof. □

We now give the proof of our main theorem.

Proof **(of Theorem 1)**

From (6),

$$\mathbb{E}[R(T)] = \sum_{i\in A} \Delta_i \mathbb{E}[n_i(T)] = \sum_{i\in A\setminus C(\theta^o)} \Delta_i \mathbb{E}[n_i(T)] + \sum_{i\in C(\theta^o)} \Delta_i \mathbb{E}[n_i(T)]. \quad (26)$$

Whenever $B(\theta^o)$ is empty, notice that $C(\theta^o)$ is empty. So, using Proposition 1, (26) becomes

$$\mathbb{E}[R(T)] = \sum_{i\in A} \Delta_i \mathbb{E}[n_i(T)] \leq \sum_{i\in A} \Delta_i C_i \leq |A| \max_{i\in A} \Delta_i C_i.$$

Whenever $B(\theta^o)$ is non-empty, $C(\theta^o)$ is non-empty. Analyzing (26), we get,

$$\begin{aligned}\mathbb{E}[R(T)] &= \sum_{i\in A\setminus C(\theta^o)} \Delta_i \mathbb{E}[n_i(T)] + \sum_{i\in C(\theta^o)} \Delta_i \mathbb{E}[n_i(T)] \\ &\overset{(a)}{\leq} \sum_{i\in A\setminus C(\theta^o)} \Delta_i C_i + \sum_{i\in C(\theta^o)} \Delta_i \mathbb{E}[n_i(T)] \\ &\overset{(b)}{\leq} \sum_{i\in A\setminus C(\theta^o)} \Delta_i C_i + \sum_{i\in C(\theta^o)} \Delta_i \left(2 + 4|A| + \frac{12\log(T)}{\beta_i^2}\right) \\ &\leq |A| \max_{i\in A} \Delta_i (2 + C_i + 4|A|) + 12\log(T) \sum_{i\in C(\theta^o)} \frac{\Delta_i}{\beta_i^2}.\end{aligned}$$

Here (a) follows from Proposition 1 and (b) from Proposition 2. Setting

$$D_1 := |A| \max_{i\in A} \Delta_i C_i \text{ and } D_2 := |A| \max_{i\in A} \Delta_i (2 + C_i + 4|A|) \quad (27)$$

proves the regret bounds in (5) of the theorem. □

We now provide the following lemma to characterize the problem dependent constants C_i given in Proposition 1. The proof for this lemma is given in the appendix.

Lemma 2 *Under the hypotheses in Proposition 1, we have*

$$C_i \leq 1 + 4|A| + \min_{\theta : a^*(\theta) = i} (2E_i(\theta)(E_i(\theta)+1)|A| + 4|A|\alpha_1^{10}(\theta)),$$

where $E_i(\theta) = \max\{3, \lceil 144/\alpha_1^4(\theta) \rceil\}$ *and* $\alpha_1(\theta) = |\mu_1(\theta^o) - \mu_1(\theta)|$.

Now, using the above lemma with (27), we have a characterization of the problem dependent constants in Theorem 1.

5 Simulations

In this section, we present detailed numerical simulation to illustrate the performance of FP-UCB algorithm compared to the other standard multi-armed bandits algorithms.

We first consider a simple setting to illustrate intuition behind FP-UCB algorithm. Consider $\Theta = \{\theta^1, \theta^2\}$ with $[\mu_1(\theta^1)\mu_2(\theta^1)] = [0.9, 0.5]$ and $[\mu_1(\theta^2), \mu_2(\theta^2)] = [0.2, 0.5]$. Consider the reward distributions $P_i, i = 1, 2$ to be Bernoulli. Clearly, $a^*(\theta^1) = 1$ and $a^*(\theta^2) = 2$.

Suppose the true parameter is θ^1, i.e., $\theta^o = \theta^1$. Then, it is easy to note that, in this case $B(\theta^o)$ is empty, and hence $C(\theta^o)$ is empty. So, according to Theorem 1, FP-UCB will achieve an $\mathcal{O}(1)$ regret. The performance of the algorithm for this setting is shown in Fig. 2. Indeed, the regret doesn't increase after some time steps, which shows the bounded regret property. We note that in all the figures, the regret is averaged over 10 runs, with the thick line showing the average regret and the band around shows the ± 1 standard deviation.

Now, suppose the true parameter is θ^2, i.e., $\theta^o = \theta^2$. In this case $B(\theta^o)$ is non-empty. In fact, $B(\theta^o) = \theta^1$ and $C(\theta^o) = 1$. So, according to Theorem 1, FP-UCB will achieve an $\mathcal{O}(\log T)$ regret. The performance of the algorithm shown in Fig. 3 suggests the same. Fig. 4 plots the regret scaled by $\log t$, and the curve converges to a constant value, confirming the $\mathcal{O}(\log T)$ regret performance.

We consider a problem with 4 arms where the mean values for the arms (corresponding to the true parameter θ^o) are $\mu(\theta^o) = [0.6, 0.4, 0.3, 0.2]$. Consider the parameter set Θ such that $\mu(\theta)$ for any θ is a permutation of $\mu(\theta^o)$. Note that the cardinality of the parameter set, $|\Theta| = 24$, in this case. It is straightforward to show that $B(\theta^o)$ is empty for this case. We compare the performance of FP-UCB algorithm for this case with two standard multi-armed bandits algorithms. Fig. 5 shows the performance of standard UCB algorithm and that of FP-UCB algorithm. Fig. 6 compares the performance of standard Thompson sampling algorithm with that of FP-UCB algorithm. The standard bandits algorithm incurs an increasing regret, while FP-UCB achieves a bounded regret. For $\mu(\theta') = [0.4, 0.6, 0.3, 0.2]$, we have $a^*(\theta') = 2$. Now we give a typical value for the $k_2(\theta')$, defined in (13), used in the proof. For this θ' we have $k_2(\theta') = \min\{k : k \geq 3, k > \lceil 12\log(k)/\alpha_1^2(\theta')\rceil\} = \min\{k : k \geq 3, k > \lceil 12\log(k)/0.2^2\rceil\} = 2326$ since $\alpha_1(\theta') = 0.2$. When the reward distributions are not necessarily Bernoulli, note that $k_i(\theta)$ is 3 for any θ with $a^*(\theta) = i$ satisfying $\alpha_1(\theta) > 2\sqrt{3/e}$.

As before assume that $\mu(\theta^o) = [0.6, 0.4, 0.3, 0.2]$. But consider a larger parameter set Θ such that for any $\theta \in \Theta$, $\mu(\theta) \in \{0.6, 0.4, 0.3, 0.2\}^4$. Note that, due to repetitions in the mean rewards for the arms, definition of $a^*(\theta)$ needs to be updated, and the algorithmic way is to pick the minimum arm index out of which are having the same mean rewards. For example, consider $\mu(\theta) = [0.5, 0.6, 0.6, 0.2]$, and so as per our new definition, $a^*(\theta) = 2$. Even in this scenario, we have $B(\theta^o)$ to be empty. Thus, FP-UCB achieves an $\mathscr{O}(1)$ regret rather than $\mathscr{O}(\log(T))$ as opposed to standard UCB algorithm and Thompson sampling algorithm.

We now consider a case where FP-UCB incurs an increasing regret. We again consider a problem with 4 arms where the mean values for the arms are $\mu(\theta^o) = [0.4, 0.3, 0.2, 0.2]$. But consider a larger parameter set Θ such that for any $\theta \in \Theta$, $\mu(\theta) \in \{0.6, 0.4, 0.3, 0.2\}^4$. Note that the cardinality of Θ, $|\Theta| = 4^4$ in this case. It is easy to observe that $B(\theta^o)$ is non-empty, for instance θ with mean arm values $[0.4, 0.6, 0.3, 0.2]$ is in $B(\theta^o)$. Fig. 7 compares the performance of standard UCB and FP-UCB algorithms for this case. We see FP-UCB incurring $\mathscr{O}(\log(T))$ regret here. Also note that the performance of the FP-UCB in this case also is superior to the standard UCB algorithm.

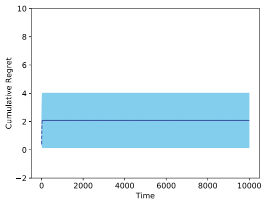

Fig. 2

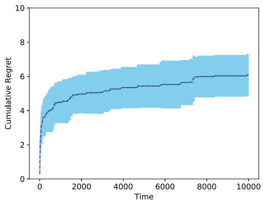

Fig. 3

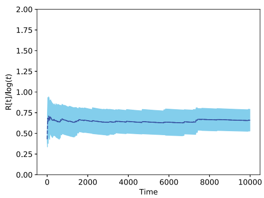

Fig. 4

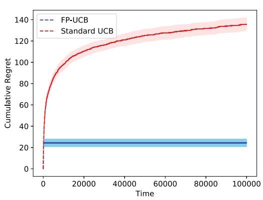

Fig. 5

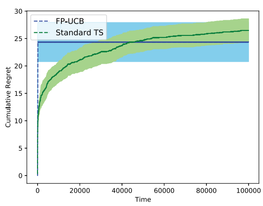

Fig. 6

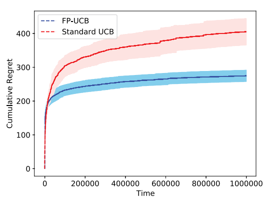

Fig. 7

6 Conclusion and Future Work

We proposed an algorithm for finitely parameterized multi-armed bandits. Our FP-UCB algorithm achieves bounded regret if the parameter set satisfies some necessary condition and logarithmic regret in other cases. In both cases, the theoretical performance guarantees for our algorithm are superior to the standard UCB algorithm for multi-armed bandits. Our algorithm also shows superior numerical performance.

In the future, we will extend this approach to linear bandits and contextual bandits. Reinforcement learning problems where the underlying MDP is finitely parameterized is another research direction we plan to explore. We will also develop similar algorithms using Thompson sampling approaches.

References

1. Rajeev Agrawal, Demosthenis Teneketzis, and Venkatachalam Anantharam. Asymptotically efficient adaptive allocation schemes for controlled iid processes: Finite parameter space. *IEEE Transactions on Automatic Control*, 34(3):258–267, 1989.
2. Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23, pages 39.1–39.26. PMLR, 2012.
3. Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135, 2013.
4. Onur Atan, Cem Tekin, and Mihaela Schaar. Global multi-armed bandits with hölder continuity. In *Artificial Intelligence and Statistics*, pages 28–36, 2015.
5. Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.
6. Hamsa Bastani, Mohsen Bayati, and Khashayar Khosravi. Mostly exploration-free algorithms for contextual bandits. *arXiv:1704.09011*, 2017.
7. Sébastien Bubeck and Che-Yu Liu. Prior-free and prior-dependent regret bounds for thompson sampling. In *Advances in Neural Information Processing Systems*, pages 638–646, 2013.
8. Sébastien Bubeck, Vianney Perchet, and Philippe Rigollet. Bounded regret in stochastic multi-armed bandits. In *Conference on Learning Theory*, pages 122–134, 2013.
9. Stéphane Caron, Branislav Kveton, Marc Lelarge, and Smriti Bhagat. Leveraging side observations in stochastic bandits. *Conference on Uncertainty in Artificial Intelligence*, 2012.
10. Nicolo Cesa-Bianchi and Gábor Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012.
11. Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214, 2011.
12. Richard Combes, Stefan Magureanu, and Alexandre Proutiere. Minimal exploration in structured stochastic bandits. In *Advances in Neural Information Processing Systems*, pages 1763–1771, 2017.
13. Varsha Dani, Thomas P Hayes, and Sham M Kakade. Stochastic linear optimization under bandit feedback. In *COLT*, 2008.
14. Aditya Gopalan, Shie Mannor, and Yishay Mansour. Thompson sampling for complex online problems. In *International Conference on Machine Learning*, pages 100–108, 2014.
15. Dileep Kalathil, Naumaan Nayyar, and Rahul Jain. Decentralized learning for multiplayer multiarmed bandits. *IEEE Transactions on Information Theory*, 60(4):2331–2345, 2014.

16. Dileep Kalathil and Ram Rajagopal. Online learning for demand response. In *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 218–222. IEEE, 2015.
17. Junpei Komiyama, Junya Honda, and Hiroshi Nakagawa. Optimal regret analysis of thompson sampling in stochastic multi-armed bandit problem with multiple plays. In *International Conference on Machine Learning*, pages 1152–1161, 2015.
18. Panqanamala Ramana Kumar and Pravin Varaiya. *Stochastic systems: Estimation, identification, and adaptive control*, volume 75. SIAM, 2015.
19. Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.
20. Tor Lattimore and Rémi Munos. Bounded regret for finite-armed structured bandits. In *Advances in Neural Information Processing Systems*, pages 550–558, 2014.
21. Lennart Ljung. *System Identification: Theory for the User*. Prentice Hall, New Jersey, 1998.
22. Odalric-Ambrym Maillard and Shie Mannor. Latent bandits. In *International Conference on Machine Learning*, pages 136–144, 2014.
23. Ian Osband, Daniel Russo, and Benjamin Van Roy. (more) efficient reinforcement learning via posterior sampling. In *Advances in Neural Information Processing Systems*, pages 3003–3011, 2013.
24. Cem Tekin and Mingyan Liu. Approximately optimal adaptive learning in opportunistic spectrum access. In *2012 Proceedings IEEE INFOCOM*, pages 1548–1556. IEEE, 2012.
25. William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.
26. Sattar Vakili and Qing Zhao. Achieving complete learning in multi-armed bandit problems. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 1778–1782. IEEE, 2013.
27. Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.
28. Chih-Chun Wang, Sanjeev R Kulkarni, and H Vincent Poor. Bandit problems with side observations. *IEEE Transactions on Automatic Control*, 50(3):338–355, 2005.
29. Shi Zong, Hao Ni, Kenny Sung, Nan Rosemary Ke, Zheng Wen, and Branislav Kveton. Cascading bandits for large-scale recommendation problems. In *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*, pages 835–844. AUAI Press, 2016.

Appendix

6.1 Proof of Lemma 1

Proof Fix an $i \in C(\theta^o)$. Then there exists a $\theta \in B(\theta^o)$ such that $a^*(\theta) = i$. For this θ, by the definition of $B(\theta^o)$, we have

$$\mu_1(\theta^o) = \mu_1(\theta). \tag{28}$$

Using Assumption 1, it follows that

$$\mu_i(\theta) = \mu_{a^*(\theta)}(\theta) > \mu_1(\theta) = \mu_1(\theta^o) = \mu_{a^*(\theta^o)}(\theta^o) > \mu_i(\theta^o).$$

Thus, $\beta_i = \min_{\theta:\theta \in B(\theta^o), a^*(\theta)=i} |\mu_i(\theta^o) - \mu_i(\theta)| > 0$.

Now, for any given θ considered above, suppose $|\mu_i(\theta) - \mu_i(\theta^o)| \leq \Delta_i$. Since $\Delta_i > 0$ by definition, this implies that

$$\mu_{a^*(\theta)}(\theta) = \mu_i(\theta) \leq \Delta_i + \mu_i(\theta^o) \overset{(a)}{=} \mu_1(\theta^o) - \mu_i(\theta^o) + \mu_i(\theta^o) = \mu_1(\theta^o) \overset{(b)}{=} \mu_1(\theta),$$

where (a) follows from definition of Δ_i and (b) from (28). This is a contradiction because $\mu_{a^*(\theta)}(\theta) > \mu_1(\theta)$.

Thus, $|\mu_i(\theta) - \mu_i(\theta^o)| > \Delta_i$ for any $\theta \in B(\theta^o)$ such that $a^*(\theta) = i$. So, $\beta_i > \Delta_i$.□

6.2 Proof of Lemma 2

Proof We have $C_i = 1 + 4|A| + \min_{\theta: a^*(\theta)=i}(k_i(\theta) + K_i(\theta))$.

First recall that $k_i(\theta) := \min\left\{k : k \geq 3, k > \lceil 12\log(k)/\alpha_1^2(\theta)\rceil\right\}$. Since $\log(x) \leq (x-1)/\sqrt{x}$ for all $1 \leq x < \infty$, we have

$$\left\{k : k \geq 3, k > \frac{12(k-1)}{\alpha_1^2(\theta)\sqrt{k}} + 1\right\} \subseteq \left\{k : k \geq 3, k > \lceil 12\log(k)/\alpha_1^2(\theta)\rceil\right\}.$$

The LHS of the above equation simplifies to $\left\{k : k \geq 3, k > 144/\alpha_1^4(\theta)\right\}$. Thus, we have $k_i(\theta) \leq \max\{3, \lceil 144/\alpha_1^4(\theta)\rceil\}$.

Now, recall that $K_i(\theta)$ is defined as

$$\begin{aligned} K_i(\theta) &= \sum_{k=k_i(\theta)}^{T} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5} \\ &\leq \sum_{k=k_i(\theta)}^{\infty} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5} \\ &= \sum_{k=k_i(\theta)}^{E_i(\theta)} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5} + \sum_{k=E_i(\theta)+1}^{\infty} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5}. \end{aligned} \tag{29}$$

We analyze the first summation in (29). Thus, we get,

$$\sum_{k=k_i(\theta)}^{E_i(\theta)} \frac{2|A|k}{(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil)^5} \leq \sum_{k=k_i(\theta)}^{E_i(\theta)} 2|A|k \leq \sum_{k=1}^{E_i(\theta)} 2|A|k = E_i(\theta)(E_i(\theta)+1)|A|. \tag{30}$$

Since $\log(x) \leq (x-1)/\sqrt{x}$ for all $1 \leq x < \infty$, we have

$$k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil \geq k - \frac{12\log(k)}{\alpha_1^2(\theta)} - 1 \geq \frac{(k-1)(\alpha_1^2(\theta)\sqrt{k} - 12)}{\alpha_1^2(\theta)\sqrt{k}}.$$

Using this, the second summation in (29) can be bounded as

$$\sum_{k=E_i(\theta)+1}^{\infty} \frac{2|A|k}{\left(k - \left\lceil \frac{12\log(k)}{\alpha_1^2(\theta)} \right\rceil\right)^5} \leq \sum_{k=E_i(\theta)+1}^{\infty} \frac{2|A|k^{7/2}\alpha_1^{10}(\theta)}{((k-1)(\alpha_1^2(\theta)\sqrt{k}-12))^5}$$
$$\stackrel{(a)}{\leq} \sum_{k=E_i(\theta)+1}^{\infty} \frac{2|A|k^{7/2}\alpha_1^{10}(\theta)}{(k-1)^5}$$
$$\leq 2|A|\alpha_1^{10}(\theta) \sum_{k=4}^{\infty} \frac{k^{7/2}}{(k-1)^5} \stackrel{(b)}{\leq} 4|A|\alpha_1^{10}(\theta) \qquad (31)$$

where (a) follows from the observation that $(\alpha_1^2(\theta)\sqrt{k}-12) > 1$ for $k \geq E_i(\theta)+1$ and (b) follows from calculus (an integral bound).

Thus using (30) and (31) in (29), we get $K_i(\theta) \leq E_i(\theta)(E_i(\theta)+1)|A| + 4|A|\alpha_1^{10}(\theta)$. This concludes the proof of this lemma. □

Developing the Path Signature Methodology and Its Application to Landmark-Based Human Action Recognition

Weixin Yang, Terry Lyons, Hao Ni, Cordelia Schmid, Lianwen Jin

Abstract Landmark-based human action recognition in videos is a challenging task in computer vision. One key step is to design a generic approach that generates discriminative features for the spatial structure and temporal dynamics. To this end, we regard the evolving landmark data as a high-dimensional path and apply path signature techniques to provide an expressive, robust, non-linear, and interpretable representation for the sequential events. We do not extract signature features from the raw path, rather we propose path disintegrations and path transformations as preprocessing steps. Path disintegrations turn a high-dimensional path linearly into a collection of lower-dimensional paths; some of these paths are in pose space while others are defined over a multi-scale collection of temporal intervals. Path transformations decorate the paths with additional coordinates in standard ways to allow the truncated signatures of transformed paths to expose additional features. For spatial representation, we apply the non-linear signature transform to vectorize the paths that arise out of pose disintegration, and for temporal representation, we apply it again to describe this evolving vectorization. Finally, all the features are joined together to constitute the input vector of a linear single-hidden-layer fully-connected network for classification. Experimental results on four diverse datasets demonstrated

Weixin Yang
Mathematical Institute, University of Oxford, UK, e-mail: weixin.yang@maths.ox.ac.uk

Terry Lyons
Mathematical Institute, University of Oxford, UK and Alan Turing Institute, UK, e-mail: tlyons@maths.ox.ac.uk

Hao Ni
Dept. of Mathematics, University College London and Alan Turing Institute, UK, e-mail: ucahhni@ucl.ac.uk

Cordelia Schmid
Inria, France, e-mail: cordelia.schmid@inria.fr

Lianwen Jin
College of Electronic and Information Engineering, South China University of Technology, China, e-mail: lianwen.jin@scut.edu.cn

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_18

that the proposed feature set with only a linear shallow network is effective and achieves comparable state-of-the-art results to the advanced deep networks, and meanwhile, is capable of interpretation.

1 Introduction

Human action recognition (HAR) is one of the most challenging tasks in computer vision with a wide range of applications, such as human-computer interaction, video surveillance, behavioral analysis, etc. A vast literature has been devoted to this task in recent years, among which are some informative surveys [1, 2, 3, 4, 5, 6, 7, 8]. An attractive option of HAR is Landmark-based HAR (LHAR) where the object is regarded as a system of correlated labelled landmarks. Johansson's classic moving light-spots experiment [9] demonstrated that people can detect motion patterns and recognize actions from several bright spots distributed on the body, which has stimulated research on pose estimation and LHAR [10, 11, 12]. Different from skeleton-based HAR (SHAR), LHAR, using no knowledge of skeletal structure, is flexible to extend to any landmark data streams with no explicit physical structures, e.g. traffic or people flow.

Although many solutions have been proposed to address the challenge of LHAR, the problem remains unsolved due to two main challenges. First, there is the problem of designing reliable discriminative features for spatial structural representation, and second of modelling the temporal dynamics of motion. In this paper, the path signature feature (PSF) is used and refined as an expressive, robust, non-linear, and interpretable feature set for spatial and temporal representation of LHAR.

The path signature, which was initially introduced in rough paths theory as a branch of stochastic analysis, has been successfully applied to many machine learning tasks. Most existing work can be devided into two categories: sliding-window-based and global-based. In the sliding temporal window approach [13, 14, 15, 16, 17, 18, 19], signatures of small paths are extracted and embedded into multi-channel feature maps as input of a CNN. The signatures herein are merely local descriptors from which the deep models are then trained to learn hierarchical representation. The global-based approaches combine all the cues into a high-dimensional path to compute high-level signatures over the whole time interval [20, 21] or low-level signatures over hierarchical intervals [22, 23]. They are straightforward but not efficient for high dimensional or spatio-temporal data.

To represent spatial pose, most methods [12, 24, 25, 26, 27, 28, 29, 30] used predefined skeletal structures. The connections distributed on a physical body are intuitive spatial constraints but not necessarily the crucial ones to distinguish actions. The connections discarded by imposing a skeletal structure could contain valuable non-local information. To solve this, hand-designed features [31, 32, 33, 34] were employed, but they are limited to encode non-linear dependencies. In this paper, we propose to localize a pose by disintegration into a collection of m-node subpaths. The signatures of these paths encode non-local and non-linear geometrical dependencies.

To model temporal dynamics, hand-designed local descriptors [31, 34] were popular, but it is difficult to encode complex spatio-temporal dependences in these. Recently, recurrent neural networks (RNN) [35], especially long short-term memory (LSTM) [36], have gained increasing popularity in handling sequential data, including human actions [37, 24, 38, 39]. In particular, a variation of LSTM [40, 25] succeeded in simultaneously exploring both spatial and temporal information. These deep models play a vital role in feature representation and achieve state-of-the-art performance, but the features learned by them are not as interpretable as hand-designed features. In this paper our temporal disintegration turns the original paths into hierarchical paths, from which the signatures encode multi-scale dynamical dependencies. Moreover, our path transformations decorates the paths with additional coordinates to allow signatures to expose additional valuable features.

To build the spatial and temporal representation, in each frame the spatial PSFs are extracted from the localized paths obtained by pose disintegration. In the clip, the evolution of each spatial feature along the time axis constitutes a spatio-temporal path. After path transformations and temporal disintegration, the temporal PSFs are then extracted from the spatio-temporal paths. Finally, the concatenation of all the features forms the input vector of a linear single-hidden-layer fully-connected network for classification. To extensively evaluate the effectiveness and flexibility of our method, several datasets (i.e., JHMDB [31], SBU [41], Berkeley MHAD [42], and NTURGB+D [39]) collected by different acquisition devices were used for experiments. Using our feature set and only a linear shallow net, we achieve comparable results to the advanced deep learning methods. Moreover, we took a further step toward understanding human actions by analyzing the PSFs and the linear classifier.

Our major contributions lie in four aspects:

1. PSFs are adopted and refined for LHAR with interpretations, proofs, experiments, and discussions of their properties and advantages.

2. Pose disintegration is proposed for non-local spatial dependencies, and temporal disintegration is proposed for multiscale temporal dependencies.

3. Path transformations, decorating the original paths with additional coordinates, are proposed to allow signatures to expose additional features.

4. Using signature-based spatio-temporal representation and only a linear shallow net, we achieve comparable state-of-the-art results to those with deep models. Meanwhile, this interpretable pipeline facilitates the understanding of HAR.

The authors are delighted to dedicate this paper to Mark H. A. Davis for many personal and professional reasons. Mark was wonderfully supportive friend. He was also an adventurous innovator who took mathematical ideas deep into commercial finance. In some sense this paper represents a similar pioneering spirit. It has a long history, and is the first effort to introduce path signature to the central area of action analysis and understanding in computer vision. This stream of research, as we report here, has developed these ideas into a viable methodology for analyzing evolving landmark style data in contexts where the datasets are too small to build effective deep learning approaches. We hope that by consolidating it here, we will recognize Mark with a paper he would have supported and approved of.

2 Related work

2.1 Path signature feature (PSF)

Rough path theory is concerned with capturing and making precise the interactions between highly oscillatory and non-linear systems [43]. The essential object in rough path theory, called the path signature, was first studied by Chen [44] whose work concentrates on piecewise regular paths. More recently, the path signature has been used by Lyons [45] to make sense of the solution to differential equations driven by very rough signals. It was extended by Lyons' theory from paths of bounded variation [45] to rough paths of finite p-variation for any $p \geq 1$ [46].

Some successful applications of the PSF have been made in the fields of machine learning, pattern recognition and data analysis. First of all, the most notable applications of using PSFs is handwriting understanding. Diehl [21] used iterated integrals of a handwritten curve for recognition and found that some linear functions of the PSF satisfy rotation invariance. Graham [19] used the sliding-window-based PSF as feature maps of a CNN for large-scale online handwritten character recognition, based on which he won the ICDAR2013 competition [47]. Inspired by this, Xie et al. [15, 16] extended the method to handwritten text recognition. Yang et al. [17, 18] explored the higher-level terms of the PSF for text-independent writer identification which requires subtle geometric features. For financial data, useful predictions can be made with only a small number of truncated PSFs [20, 48]. The truncated signature kernel for hand movement classification was presented in [49], and was further extended to an untruncated version [50]. Moreover, PSFs were used on self-reported mood data to distinguish psychiatric disorders [23]. In [51], path signature transform was applied to describe the behaviour of controlled differential equations for modelling temporal dynamics of irregular time series. To model topological data, a novel path signature feature based on the barcodes arising from persistent homology theory was proposed for classification tasks [52]. These applications demonstrate the value of the PSF as an effective and informative feature representation.

The paper has been a long time in development, and the preprints [53] on the ArXiv have already influenced other developments. To name a few, in [54, 55, 56], the extraction of the path signature feature was treated as a flexible intermediate layer in various end-to-end network architectures like CNNs, LSTMs, or Transformer Networks. Also, variants of our proposed feature set were successfully applied to tasks like Arabic handwriting recognition [57], writer identification [58], personal signature verification [59], sketch recognition [60], action/gesture recognition [61, 62], speech emotion recognition [63], etc., showing its generalization ability. The proposed invisibility-reset transformation was further analyzed in [64].

2.2 Landmark-based human action recognition

A human body can be regarded as an articulated system composed of joints that evolve in time [65]. For recent surveys of LHAR, we refer the reader to [8, 66, 67, 68].

Approaches based on hand-designed features for LHAR can be categorized into two classes: joint-based and part-based. The joint-based ones regard the human body as a set of points and attempt to capture the correlation among body joints by using the motion of 3D points [69], measuring the pairwise distances [31, 70, 26, 33, 34], or using the joint orientations [71]. On the other hand, the part-based approaches focus on connected segments of the human skeleton. They group the body into several parts and encode these parts separately [27, 28, 72, 73, 74, 29, 75]. Some methods in this category represent a pose by means of the geometric relations among body parts, for examples, [27, 28] employed quadruples of joints to form a new coordinate system for representation, and [12] considered measurements of the geometric transformation from one body part to another. Some methods assume that certain actions are usually associated with a subset of body parts, so they aim to identify and use the subsets of the most discriminative parts of the joints.

Given the recent success of deep learning frameworks, some works aim to capture correlation among joint positions using CNNs [76, 77, 78, 79]. In [76], the input feature maps of a CNN were joints colored according to their sequential orders, body parts, or velocity, while in [77] and [78], the CNN's inputs were the concatenation of hand-designed local features. Since human actions are usually recorded as video sequences, it is natural to apply RNNs or LSTMs. HBRNN [24] and Part-aware LSTM [39] contained multiple networks for different groups of joints. Zhu et al. [37] proposed a deep LSTM to learn the co-occurrence of discriminative joints using a mixed-norm regularization term in the cost function. By additional new gating to the LSTM, the Differential LSTM [38] is able to discover the salient motion patterns, and [40, 25] achieved robustness to noise. It is noteworthy that the spatio-temporal RNNs in [40, 25] concurrently encoded both spatial and temporal context of actions within a LSTM. Liu et al. [80] used an attention-based LSTM to iteratively select informative keypoints for recognition. Zhang et al. [81] used a multilayer LSTM to fuse several simple geometric features for recognition. By taking advantage of the graph structure of human skeleton, Graph Convolutional Networks (GCNs) were introduced into the action recognition task. Yan et al. [30] used spatial graph convolutions along with interleaving temporal convolutions. Concurrently, Li et al. [82] proposed a similar approach but introduced a multi-scale module for spatio-temporal modelling. DGNN [83] represented the skeleton as a directed acyclic graph to encode both joint and bone information. MV-IGNET [84] extracted multi-level spatial features and leveraged different skeleton topologies as multi-views to generate complementary action features. MMDGCN [85] proposed a dense graph convolution for local dependencies and used spatial-temporal attention module to reduce the redundancy. These deep learning methods achieved high accuracy on most large-scale action datasets, but they often require a lot of training data and suffer from a lack of interpretability.

3 Path Signature

3.1 Definition and geometric interpretation

The rigorous introduction of the path signature as a faithful description or feature set for unparameterized paths can be found in [43, 86, 87, 88], so in this paper we present it in a practical manner.

A d-dimensional path or stream of timestamped events P over the time interval $[0, T] \subset \mathbb{R}$ can be represented as a continuous map $P : [0, T] \rightarrow \mathbb{R}^d$. The coordinates of P at time τ are $P_\tau = \left(P_\tau^1, P_\tau^2, \ldots, P_\tau^d\right)$. To illustrate the idea, we consider the simplest case when $d = 1$. The path is a real-valued path for which the path integral is defined as

$$S(P)_{0,T}^1 = \int_{0<t\leq T} dP_\tau^1 = P_T^1 - P_0^1, \tag{1}$$

which is the increment of this 1-dimensional path over the whole time interval and is called the 1-fold iterated integral. We emphasize that $S(P)_{0,\tau}^1, 0 < \tau \leq T$ is also a real valued path w.r.t τ. The 2-fold iterated integral is

$$S(P)_{0,T}^{11} = \int_{0<\tau_2\leq T} S(P)_{0,\tau_2}^1 dP_{\tau_2}^1 = \frac{1}{2}\left(P_T^1 - P_0^1\right)^2, \tag{2}$$

which is proportional to the square of the increment. Again, $S(P)_{0,\tau}^{11}$ is a real-valued path, so if we continue recursively, the k-fold iterated integral of P is

$$\begin{aligned} S(P)_{0,T}^{11\ldots1} &= \int_{0<\tau_k\leq T} \cdots \int_{0<\tau_2\leq\tau_3} \int_{0<\tau_1\leq t_2} dP_{\tau_1}^1 dP_{\tau_2}^1 \ldots dP_{\tau_k}^1 \\ &= \frac{1}{k!}\left(P_T^1 - P_0^1\right)^k, \end{aligned} \tag{3}$$

which is proportional to the increment to the power of k.

Now, when $d = 2$, the 1-fold iterated integral of the path $\left\{P_\tau^1, P_\tau^2\right\}$ has 2 elements

$$S(P)_{0,T}^1 = \int_{0<t\leq T} dP_\tau^1 = P_T^1 - P_0^1, \tag{4}$$

$$S(P)_{0,T}^2 = \int_{0<t\leq T} dP_\tau^2 = P_T^2 - P_0^2. \tag{5}$$

Each element is the increment of the path on the corresponding axis over the time interval $[0, T]$. They denote the displacement of the given path. The 2-fold iterated integral of this 2D path contains $d^2 = 2^2$ elements

$$S(P)^{11}_{0,T} = \int_{0<\tau_2\leq T}\int_{0<\tau_1\leq t_2} dP^1_{\tau_1} dP^1_{t_2} = \frac{1}{2!}\left(P^1_T - P^1_0\right)^2, \tag{6}$$

$$S(P)^{22}_{0,T} = \int_{0<\tau_2\leq T}\int_{0<\tau_1\leq \tau_2} dP^2_{\tau_1} dP^2_{\tau_2} = \frac{1}{2!}\left(P^2_T - P^2_0\right)^2, \tag{7}$$

$$S(P)^{12}_{0,T} = \int_{0<\tau_2\leq T}\int_{0<\tau_1\leq \tau_2} dP^1_{\tau_1} dP^2_{\tau_2}, \tag{8}$$

$$S(P)^{21}_{0,T} = \int_{0<\tau_2\leq T}\int_{0<t_1\leq t_2} dP^2_{\tau_1} dP^1_{\tau_2}. \tag{9}$$

We note that the first two elements are the same as (2) in the 1-dimensional case. For the other two elements, the geometric intuitions are the areas shown in Fig. 1(a) and Fig. 1(b). Together they represent the Lévy area [86] depicted in Fig. 1(c). The Lévy area, which is a signed area enclosed by the path and the chord connecting the endpoints, can be expressed by

$$A_{0,T} = S(P)^{12}_{0,T} - S(P)^{21}_{0,T}. \tag{10}$$

The sign of the area depends on the sign of the winding number of the path moving around it [89]. The interpretation of the k-fold iterated integral ($k > 2$) of a 2D path is not trivial, so it is not included here. By analogy, for a 3D path, the 1-fold, 2-fold, and 3-fold iterated integrals are units of displacement, area, and volume respectively.

In general, for a path in $\mathbb{R}^d$, the superscript of the k-fold iterated integral, which describes the order of integration, is a multi-index $(i_1, i_2, \ldots, i_k) \in \{1, \ldots, d\}^k$. Therefore, the d^k elements of the k-fold iterated integral of a d-dimensional path can be generally expressed as

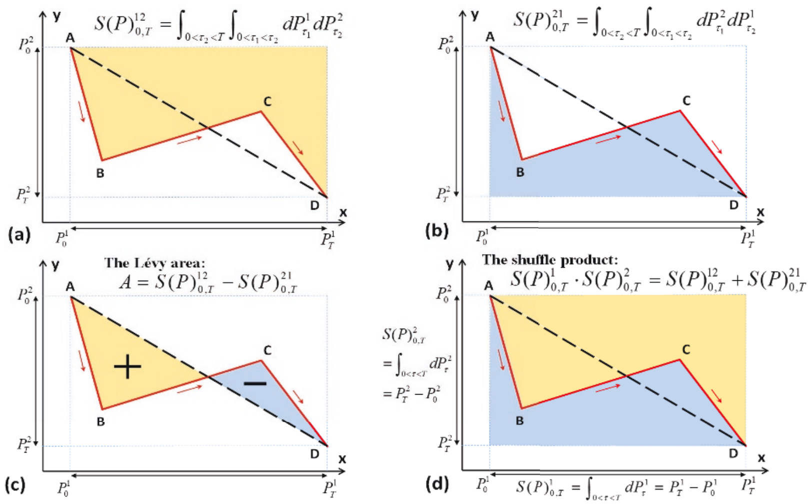

Fig. 1 The geometric intuition of the PSF of a 2D path. The path in red moves from A to D over the time interval [0, T]. The dashed line is the chord connecting the endpoints. Panels (a) and (b) depict two terms of the 2-fold iterated integrals of the path, (c) is the Lévy area enclosed by the path and its chord, and (d) is a demonstration of the shuffle product identity.

$$S(P)_{0,T}^{i_1,i_2,\ldots,i_k} = \int_{0<\tau_k\leq T} \cdots \int_{0<t_2\leq\tau_3} \int_{0<\tau_1\leq\tau_2} dP_{\tau_1}^{i_1} dP_{\tau_2}^{i_2} \ldots dP_{\tau_k}^{i_k}. \tag{11}$$

Then the signature of a path P over the time interval $[0,T]$ is the collection of all the iterated integrals of P:

$$\begin{aligned} S(P)_{0,T} = \Big(&1, S(P)_{0,T}^{1}, \ldots, S(P)_{0,T}^{d}, \\ &S(P)_{0,T}^{1,1}, \ldots, S(P)_{0,T}^{1,d}, S(P)_{0,T}^{2,1}, \ldots, S(P)_{0,T}^{d,1}, \ldots, S(P)_{0,T}^{d,d} \\ &\ldots, S(P)_{0,T}^{1,1,\ldots,1}, \ldots, S(P)_{0,T}^{i_1,i_2,\ldots,i_k}, \ldots, S(P)_{0,T}^{d,d,\ldots,d}, \ldots\Big), \end{aligned} \tag{12}$$

where the 0-th term is conventionally set to 1. Since the signature is defined on top of all the possible indices of finite length, the number of elements in the signature is infinite. In practical use we usually consider the signature truncated at a certain level n written as

$$S_n(P)_{0,T} = \Big(1, S(P)_{0,T}^{1}, \ldots, S(P)_{0,T}^{i_1,i_2,\ldots,i_n}, \ldots, S(P)_{0,T}^{d,d,\ldots,d}\Big) \tag{13}$$

of which the dimensionality is $\varphi(d,n) = \left(d^{n+1}-1\right)(d-1)^{-1}$. The elements of the truncated signature are taken as features (i.e., PSF) encoding the informative geometric properties of sequential data in applications in machine learning. For the feature set, the 0-th term (i.e., a constant value set to 1) is optional, so the dimension can be reduced to

$$\varphi'(d,n) = \left(d^{n+1}-d\right)(d-1)^{-1}. \tag{14}$$

For the 1-dimensional case ($d = 1$), the feature dimension is exactly equal to n (excluded the 0-th term) according to (1), (2), and (3).

3.2 Calculation of the signature for a discrete path

Although the path signature is initially defined for continuous paths with bounded variation, it is easily extended to discrete paths by linear interpolation [90]. The signature is canonical and does not depend on the choice of timescale used for the interpolation.

Computing the signature of a piecewise linear path does not require integrals. For each line segment of the path, the elements of its signature are given by

$$S(P)_{\tau,t+1}^{i_1,i_2,\ldots,i_k} = \frac{1}{k!} \prod_{j=1}^{k} \left(P_{\tau+1}^{i_j} - P_{\tau}^{i_j}\right), \tag{15}$$

where $P_\tau^{i_j}$ is the i_j-th coordinate value of path P at time τ. For the entire path, Chen's identity [44] states that for any time stamps (s,t,u) satisfying that $s < t < u$, we have

$$S(P)_{s,u}^{i,i_2,\ldots,i_k,\ldots,i_n} = \sum_{k=0}^{n} S(P)_{s,t}^{i_1,i_2,\ldots,i_k} S(P)_{t,u}^{i_{k+1},i_{k+2},\ldots,i_n}. \tag{16}$$

This implies that the signature of the entire path can be calculated from the signatures of its pieces.

We recommend the three open-source python software libraries, *esig* (on PyPi), derived from the *CoRoPa* C++ library *libalgebra* [91], *iisignature* [92], and *Signatory* [93] which has a dependency on PyTorch and works well on the CPU as well as the GPU. They all allow fast computation of the path signature.

3.3 Properties of the path signature

3.3.1 Uniqueness

It is proved that the path signature determines a path if and only if the path is not tree-like (this notion is introduced in [45]). A tree-like path is a trajectory containing a section where the path exactly retraces itself. Tree-like paths are common in real-world data streams, for instance, in some human actions, especially periodic ones, like clapping or jumping in place. An effective way to avoid the tree-like situation is adding an extra monotone dimension, such as time, to the original path.

3.3.2 Invariance under translation

The signature computed by (11) or (15) is invariant under translation of the paths, which is a practical advantage and avoids complex recentering normalization.

3.3.3 Invariance under time reparameterization

A time reparameterization of a path is a continuous, nondecreasing substitution for the time variable of a path. It changes the speed of recording of the path. Human actions are largely invariant under changing the speed of the action or viewing speed of the video. The ease with which the signature can completely filter out these changes in the representation is a major advantage for machine learning, substantially reducing the dimensionality of the feature set needed for action classification. The use of the path signature, with its fixed-dimensional feature set, can help the classifer to recognize the same action performed or sampled at different speeds. We refer the reader to [43, 88] for a detailed proof of the invariance of the path signature under time reparameterization.

3.3.4 Nonlinearity of the signature

The shuffle product identity [86] states that the product of two signatures of lower level can be expressed as a linear combination of some higher-level terms. For instance, for the two-dimensional case in section 3.1, we can easily derive the following equation from Fig. 1(d),

$$S(P)^{1}_{0,T} \cdot S(P)^{2}_{0,T} = S(P)^{12}_{0,T} + S(P)^{21}_{0,T}. \tag{17}$$

In other words, the nonlinear behavior in terms of lower level terms can be expressed by linear combination of higher-level terms. Therefore, when we incorporate the higher-level terms into the feature representation, we automatically include more nonlinear prior knowledge in our feature set. If the introduced nonlinearity is sufficient, we need only linear classifiers to distinguish the targets.

3.3.5 Fixed dimension under length variations

Another practical property of the path signature is that the dimension of the PSF extracted from the entire path depends on the truncation level of the signature and the intrinsic dimension of the path but is independent of the (sampled) length of the path, as described in 14. For human action recognition, the durations of actions are variable. The use of the path signature allows us to extract a fixed dimension of features and use them with classification methods which require a fixed-length input.

4 Path disintegrations and transformations

The principled and robust representation of unparameterized paths, along with the convenience of reducing polynomial functions on the space of paths to linear ones (which establishes their universality) provide the core motivations for using signatures as features. One can always take the signature of a raw path to remove any dependence on parameterization or translation, but sometimes it is prudent to apply path disintegrations or path transformations as preprocessing to improve the efficiency and effectiveness of PSFs. The disintegrations turn a path into a composition of subpaths while the transformations turn a path into a higher-dimensional path.

4.1 Path disintegrations

4.1.1 Pose Disintegration

In many cases, non-local clues are informative and straightforward, for instance, the non-local displacement between two hand points is a key feature for the action of clapping. To exploit both local and non-local clues in pose, we propose pose disintegration. Landmarks that are labelled with corresponding body parts have no inherent order, so a predefined priority order is randomly chosen and fixed – different random choices of initial order yield comparable results in preliminary experiments. The pose is then regarded as an ordered collection of points in $\mathbb{R}^d$. Our pose disintegration localizes the pose into all possible subposes containing m points. Connecting the m points in each subpose in the inherited order forms a unique m-node sub-path that visits each point once. We end up with a collection of sub-paths which do not need to be parts of physical body and are available for further path transformations or signature extractions.

We consider that functions on a pose can be approximated by functions on the piecewise linear localized paths of its subposes. For convenience, one can view these functions as linear functions on the signature of its localized paths. The terms of the first two levels of signatures cover the displacement and the area information similar to the traditional hand-designed features [31, 34], while the higher-level terms capture more non-linear features. For a pose with N joints, the dimension of the signatures of its localized paths is $C_N^m \cdot \varphi'(d, n)$, where m is the number of points in a subpose, d is the dimension of the path, and n is the truncated signature level. The selections of these parameter values are highly correlated and associated with the uniqueness of the paths. According to [94], any piecewise linear paths in $\mathbb{R}^d$, consisting of at most $m = d + 1$ points, can be uniquely recovered from the signature at the third level. A larger m brings semantically high-level components but requires a larger n for the path uniqueness [95], which exponentially increases the feature dimension according to 14, and means less shareability and more sub-paths. The number of m-node subpaths is in line with Pascal's triangle and increases along with $m(m \geq N/2)$. To avoid feature set of very large dimension, $m \leq 3, n = 3$ for $d = 2$ and $m \leq 4, n = 3$ for $d = 3$ are suggested. Beyond the signature level n required for the unique recovery of a path, the non-linearity (as described in 3.3.4) of the extra high-level terms may still contribute to facilitate the training of the model until the dimensionality of the feature set becomes impractical.

4.1.2 Temporal Disintegration

Temporal disintegration is based on the basic theory of the path signature which suggests that low-level terms of signatures on all intermediate length time intervals can be far more efficient than signatures of high levels over the whole time interval [86]. Therefore, instead of extracting the PSF over the whole time interval, the dyadic path signature features (DPSF) [22] split the entire interval into small intervals with a dyadic hierarchical structure and then extracts PSF over each small interval. Given a path over the whole time interval $[0, T] \subset \mathbb{R}$ the j-th dyadic level of the hierarchical structure is the collection of subintervals $\left[iT/2^j, (i+1)T/2^j\right], i \in \left[0, 2^j - 1\right], j \in \mathbb{N}$. Note that the 0-th dyadic level contains exactly the whole path. The DPSF over long, medium, and short time intervals describes multi-scale dynamical dependences more efficiently than the PSF over the entire interval, which requires higher-level terms to capture local dependencies.

Slightly different from the hierarchical structure in [22] which may break the events that occur near the conjunctive time stamps $\left\{iT/2^j \mid i \in \left[1, 2^j - 1\right], j \in \mathbb{N}^+\right\}$, we consider an overlapping version over the time intervals $\left[iT/2^{j+1}, (i+2)T/2^{j+1}\right]$, $i \in \left[0, 2 \cdot \left(2^j - 1\right)\right]$. The overlapping DPSF is expected to supplement the original DPSF with additional local details. Its dimension is

$$\hat{\varphi}(h, d, n) = \left(2^{h+1} - h - 2\right) \cdot \varphi'(d, n), \tag{18}$$

where $h \in \mathbb{N}^+$ is the number of the hierarchical level. The selection of h is a tradeoff between improving efficiency and introducing local noises over finer intervals.

4.2 Path transformations

4.2.1 Time-incorporated transformation

The signature is invariant under parameterization, but in many situations, one would like to keep the dependence on time. Adding a monotone increasing time dimension is adopted to encode motion speed. The signature of a time-incorporated path contains two parts: time-independent (TI) and time-dependent (TD). The TI part is exactly the signature of the original path, so its integration order is

$$i_1, i_2, \ldots, i_k \in \{1, \ldots, d\}. \tag{19}$$

The TD part is related with time. Its integration order is

$$i_1, i_2, \ldots, i_k \in \{1, \ldots, d+1\}, \exists m \in [1, k], i_m = d+1, \tag{20}$$

which means each term of the signature in TD is an integral along the time dimension at least once. Given the truncated signature level n, the dimensionality of the TD part is $\varphi'(d+1, n) - \varphi'(d, n)$. The signature of the original path filters out the information about the speed of motion and the sampling rate but the signature of the time-incorporated path allows us to select one and suppress the other according to significance to the classification.

4.2.2 Invisibility-reset transformation

The signature capturing relative position information is invariant under translation. Given that the absolute position may be essential for some scenarios (e.g., HAR under static CCTVs), we propose the invisibility-reset transformation of a path to retain the absolute position information in signatures. For a path P in $\mathbb{R}^d$ within the interval $[0, T]$, we add two time steps T+1 and T+2 with value P_T and 0 respectively at the end of P and add a visibility dimension v with values 1 in $[0, T]$ and 0 in $(T, T+2]$. In other words, the invisibility-reset path P_{IR} in $\mathbb{R}^{d+1}$ first becomes invisible at time T+1 and then is reset to the origin at T+2. According to (15) and (16), we have

$$S\,(P_{IR})^{i_1,i,\cdots,v,v}_{0,T+2} = -S(P)^{i,i_2\cdots i_k}_{0,T}, i_1, i_2, \ldots, i_k \in \{1, \ldots, d\} \tag{21}$$

which means certain terms in $S\,(P_{IR})$ encode the relative positions as in $S\,(P)$. Moreover, the terms of the first level of $S\,(P_{IR})$, given by

$$S\,(P_{IR})^{i_1}_{0,T+2} = -P^{i_1}_0, i_1 \in \{1, \ldots, d\}, \tag{22}$$

are the absolute position of the initial point. This simple transformation retains different position information in signatures and thus allows the network to select one and suppress the other according to significance to the task.

4.2.3 Multi-delayed lead-lag transformation

The lead-lag transformation proposed in [20, 87, 90] is designed to exploit the quadratic cross-variation between the original path and its delayed path. To extend its capability to describe long-term dependencies of sequential events, our modified lead-lag transformation, as shown in Fig. 2, is constructed by the original path and its multiple delayed paths (instead of one delayed path in [20]). We denote the dimension of a lead-lag transformed path as d_{LLT}. The signatures of lead-lag paths with smaller d_{LLT} encode short-term dependencies, while those with larger d_{LLT} explore more long-term dependencies.

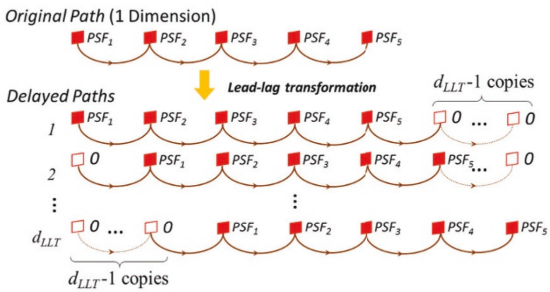

Fig. 2 The illustration of multi-delayed lead-lag transformation. The dimen-sion of lead-lag paths is d_{LLT}. The delayed paths are padded with zeros to ensure a fix length for each dimension.

Table 1 Proposed features for LHAR

# of joints	Spatial structural features (for each frame)	Temporal dynamical features (along the time axis)
1 (a single joint)	S-J: The d-dimensional coordinates of each of the predefined N joints are used.	T-J-PSF: The temporal PSF over the evolution of each joint up to signature level n_{TJ} is extracted.
2 (joint pair)	S-P-PSF: The PSF over each pair of joint up to signature level n_{SP} is extracted.	T-S-PSF: The evolution of each dimension of spatial PSF is treated as a path, over which the temporal PSF up to signature level n_{TS} is extracted.
3 (joint triple)	S-T-PSF: The PSF over each triple of joint up to signature level n_{ST} is extracted.	

5 Feature extraction for human action recognition

Our proposed feature set for LHAR, which we describe in this section, is outlined in Table 1. We note that an unofficial Python implementation of the feature set is available on GitHub [96].

5.1 Spatial structural features

First of all, the basic information describing the spatial structure is the d-dimensional coordinates of each of the N joints of the body. The keyword **S-J** denotes the spatial coordinate values of the joints. The dimension of this part is $D_{SJ} = N \cdot d$ for each sampled frame.

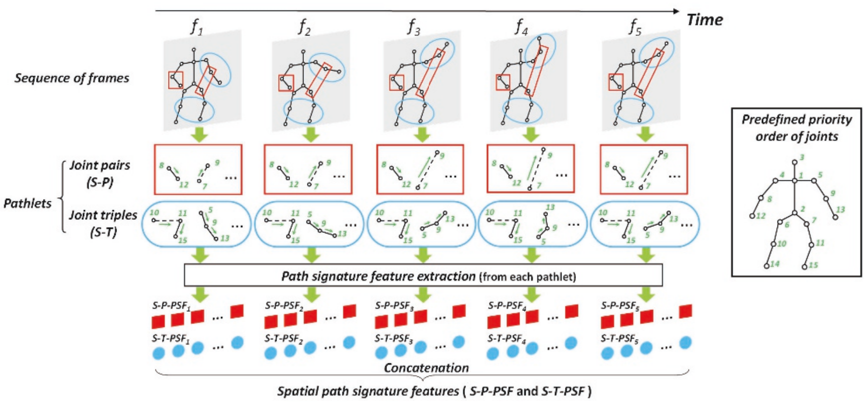

Fig. 3 The illustration of spatial feature (**S-P-PSF** and **S-T-PSF**) extraction. Note that we predefine the priority order of all the N joints (N = 15 in this figure). The red quadrangles denote the feature extraction of joint pairs, while the blue ellipses denote that of joint triples. All possible pairs and triples of joints are considered.

To encode the spatial context we use pose disintegration with $m = 2$ and $m = 3$, which means joint pairs and joint triples are used as illustrated in Fig. 3. The signatures of each of these subpaths are computed to model the spatial constraints in each frame. The spatial PSF of joint pairs and joint triples are denoted by **S-P-PSF** and **ST-PSF** respectively. If the truncation level of the signature of pairs and triples are n_{SP} and n_{ST} respectively, then the feature dimensions per frame are $D_{SP} = C_N^2 \cdot \varphi'(d, n_{SP})$ and $D_{ST} = C_N^3 \cdot \varphi'(d, n_{ST})$ respectively. Finally, the spatial features from all sampled frames are extracted and concatenated. The dimension of **S-P-PSF** and **S-T-PSF** per frame is denoted by $D_S = D_{SP} + D_{ST}$.

5.2 Temporal dynamical features

Inspired by the works in [40, 25] which jointly learned the spatial and temporal contexts in a variant of LSTM, we suggest that the dynamics of landmark-based human action can be described by the evolution of spatial context. The spatial context herein are the features we extracted in section 5.1, although other spatial features can be used alternatively. From these, we are going to extract two kinds of temporal features **T-J-PSF** and **T-S-PSF**.

The **T-J-PSF**, illustrated in Fig. 4, is the temporal PSF of the evolution of each joint along the time. The evolution of each joint is naturally a time-sequence, so we consider its time-incorporated transformation. For N-joint bodies in $\mathbb{R}^d$, the dimension of **T-J-PSF** is $D_{TJ} = N \cdot \varphi'(d+1, n_{TJ})$, where n_{TJ} is the truncation level of the signature.

Since each dimension of the spatial contextual features (**S-P-PSF** and **S-T-PSF**) characterizes one particular spatial constraint of a pose, the evolution of this spatial constraint along the time forms a spatio-temporal path which delivers temporal constraints of a motion. The temporal PSF of these spatio-temporal paths is denoted by **T-S-PSF** and illustrated in Fig. 5. Since the signature of a spatiotemporal path (i.e., a 1D path) is just the increments to a certain power, the multi-delayed lead-lag transformation is applied to each path to enrich the PSF with cross variations among

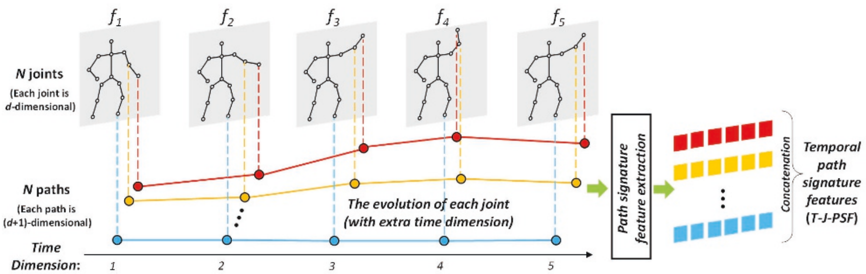

Fig. 4 Illustration of temporal features extracted from the evolution of each corresponding joint (**T-J-PSF**).

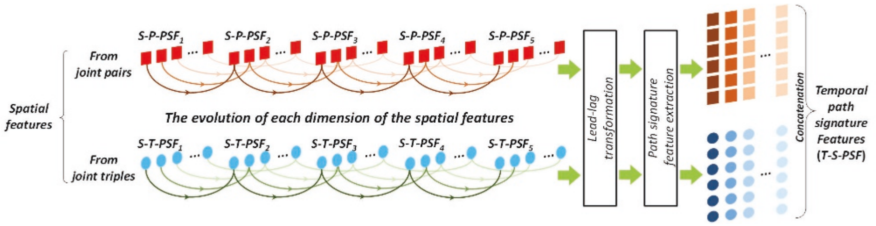

Fig. 5 Illustration of temporal features extracted from the evolution of spatial context (**T-S-PSF**). Each dimension of the spatial features is treated equally and individually.

events of the path. If the truncation level of the signature is n_{TS} and the dimension of the lead-lag paths is d_{LLT}, the dimension of **T-S-PSF** from all spatio-temporal paths is $D_{TS} = D_S \cdot \varphi' (d_{LLT}, n_{TS})$. Considering there might exist complicated or long-range actions, the temporal disintegration in section 4.1.2 can be applied. If so, the dimensions are $D_{TJ} = N \cdot \hat{\varphi} (h_{TJ}, d + 1, n_{TJ})$ and $D_{TS} = D_S \cdot \hat{\varphi} (h_{TS}, d_{LLT}, n_{TS})$ where h_{TJ} and h_{TS} are the corresponding hierarchical levels.

The dimension of all temporal PSFs is $D_T = D_{TJ} + D_{TS}$. Finally, the total dimension of spatial and temporal features per clip is $D = M \cdot (D_{SJ} + D_S) + D_T$, where M is the number of sampled frames. Moreover, the spatial components can be covered by the temporal PSFs extracted from invisibility-reset paths which require no sampling step.

6 Experimental results and analysis

6.1 Datasets

Monocular videos recorded by 2D cameras are capable of collecting spontaneous actions, but their sensitivity to viewpoint variations and occlusions makes recognition a difficult task [1]. Intuitively, human body is general in 3D space, so marker-based motion capture systems [97] were designed to collect highly accurate locations of human joints. However, they are often expensive and impractical for recording realistic action videos. Fortunately, costeffective depth cameras (e.g. Kinect sensor [98]) were designed to provide reliable joint locations via real-time pose estimation algorithms (e.g., [99]). Our method is general enough to be applied to various kinds of data. To extensively evaluate the proposed methods, we conducted experiments on four datasets containing examples of all three types of data: JHMDB [31], SBU [41], Berkeley MHAD [42], and NTURGB+D [39]. The information we used herein for action recognition is the locations of landmarks in all the frames. However, it is worth noting that our method is flexible and additional information such as visibility state or confidence score can be included.

The JHMDB dataset [31] is a 2D human action dataset. There are 928 clips, each clip containing between 15 and 40 frames. A clip captures only one person doing one of 21 actions. The 2D joint positions are manually annotated. There are 3 splits separating the whole dataset into training and testing set. The final result is the average of them. The sub-JHMDB is a subset of JHMDB with the full body inside the frame. The challenges are the spontaneity of the actions captured by the clips from YouTube and the loss of information due to 2D projection.

The SBU Interaction [41] is a 3D Kinect-based dataset. It has 282 clips categorized into 8 classes of two-actor interactions, and has 30 joints per frame. Its depth information suffers from self-occlusion, causing measurement errors in the estimated joint locations.

The third dataset is Berkeley MHAD dataset [42] captured by a marker-based motion capture system. It consists of 659 clips, of which 384 clips, performed by 7 actors, are used for training and 275 clips by 5 different actors are used for testing. The 3D locations of 43 joints captured by LED markers are very accurate.

The Kinect-based NTURGB+D [39] is one of the largest 3D action recognition datasets and contains 56 thousand clips of 60 classes. The large viewpoint variations and unconstrained number of actors pose considerable challenges for analysis of this dataset.

Note that the quantitative analysis was conducted on JHMDB, and all the parameters were determined by 5- fold cross validation on the training set of the first split.

6.2 Network configurations

Since PSFs are rich non-linear features, we adopted a single-hidden-layer linear neural network as our classifier (1- layer net also works well in preliminary experiments). The network is fully-connected and the activation of the hidden neurons is the identity function. The input dimension is decided by the PSF (i.e., **S-P-PSF**, **S-T-PSF**, **T-J-PSF**, **T-S-PSF**, or some combinations of them) and the output is a probability distribution given by a softmax layer over all the class labels in a dataset. The single hidden layer has 64 neurons. The networks are trained by stochastic gradient descent on the cross-entropy with momentum 0.7 and mini-batch size 30. The learning rate updates in accordance to $\alpha(t) = \alpha(0) \cdot \exp(-\lambda t)$ where $\alpha(0) = 0.005, \quad \lambda = 0.005$. The maximum epoch is 300 for all experiments.

Dropconnect [100], a generalization of Dropout [101], randomly omits a proportion of connections at each training iteration. It is applied to the connections between the input and the single hidden layer for regularization. A high ratio of Dropconnect is essential to mitigate overfitting because the features herein are of very high dimension. Additionally, since the actions of some joints are highly correlated with each other, a small proportion of joints or features may already be sufficient to distinguish some actions. Based on our preliminary experiments, the Dropconnect rate is set to 0.95.

6.3 Data preprocessing and benchmark

We used two kinds of data augmentation. One is horizontal flipping, and the other one is adding Gaussian noise (inspired by [31]) over joint coordinates to simulate the noisy positions caused by estimation or annotation.

To cope with translation variation, we normalized the joints from world coordinate system to person-centric coordinate system by placing the center point of the body at the origin. To compensate for the biometric differences, we normalized the coordinate values to the range of [-1,1] over the entire clip. For feature normalization, each feature was divided by the maximum absolute value of the corresponding dimension and normalized to [-1, 1].

The spatial components (**S-J**, **S-P-PSF**, and **S-T-PSF**) are calculated for each frame. To obtain a fixed-length input to the network, we uniformly sampled M (in this paper, $M = 10$) frames from each clip. As the signature has a fixed dimension under length variation, the temporal features (**T-J-PSF** and **T-S-PSF**) are extracted from all the frames without subsampling. Our baseline method is using **S-J**, the d-dimensional coordinate values of all N joints. This leads to MNd-dimensional feature set, for which we obtained a validation error rate of 57.54 ± 3.26%.

6.4 Investigation of the spatial features

As described in section 4.1.1 and 5.1, by pose disintegration with $m = 2$ and $m = 3$, we constructed all the joint pairs and triples as localized paths for **S-P-PSF** and **S-TPSF** respectively. The error rates on the validation set obtained by these feature sets are shown in Table 2 and Table 3. The performance improves when higher terms of the signature are considered, but the improvements tend to be negligible and the variance increases when the dimension of the feature grows exponentially with the signature level n. For the joint pairs, a suitable truncation level n_{SP} is 2 or 3, while for the joint triples, the level n_{ST} needs to be as high as 3 or 4, which suggests the choice of n should depend on m. We refer the reader to [95] which discusses the relationship among m, n, and the path dimension d from the view of path recovery. For the following experiments, we chose to fix $n_{SP} = 2$, $n_{ST} = 3$.

Table 2 Effect of different signature levels on the performance of S-P-PSF

Type of subpaths	Signature level n_{SP}	Feature dim.	Error rate (%)
Joint Pairs	1	2100	32.79 ± 4.49
	2	6300	25.41 ± 4.55
	3	14700	**24.10** ± **5.65**
	4	31500	24.10 ± 5.72

Table 3 Effect of different signature levels on the performance of S-T-PSF

Type of subpaths	Signature level n_{ST}	Feature dim.	Error rate (%)
Joint Triples	1	9100	43.93 ± 2.87
	2	27300	32.46 ± 3.26
	3	63700	26.39 ± 3.99
	4	136500	**24.75 ± 4.79**
	5	282100	23.77 ± 6.41
	6	573300	25.24 ± 6.44

6.5 Investigation of the temporal features

6.5.1 Investigation of T-J-PSF

First, we investigated the effect of the time-incorporated transformation and the truncation level n_{TJ} of the **T-J-PSF**. As shown in Fig. 6, if the truncation level n_{TJ} (the horizontal axis) is 1, adding a time dimension (the green plot) only improves the performance a little. This is because the first level term related to the time dimension is only the duration of the action. When n_{TJ} increases, the performance improvements of using time-incorporated PSF are more significant, showing the effectiveness of the time-incorporated path transformation. As to the truncation level, when n_{TJ} increases, the results have lower bias together with gradually higher variance, so a trade-off is required. Here, $n_{TJ} = 5$ is a good choice.

In addition, we investigated the effect of the signature of the time-incorporated path at different frame rates. We artificially increased the frame rate by interpolating additional frames at random time stamps of the original clips. Bodies of the additional frames were copied from those of their adjacent frames. On the other hand, we decreased the frame rate by random subsampling. The networks were trained using the training clips at original frame rate (30fps) and tested 10 times using artificial validation clips at each of the frame rates ranging from 6fps to 90fps in 6fps steps. As shown in Fig. 7, when the frame rate increases from 30fps to 90fps, the error rates of using the time-independent part (TI) stay the same, while those of using

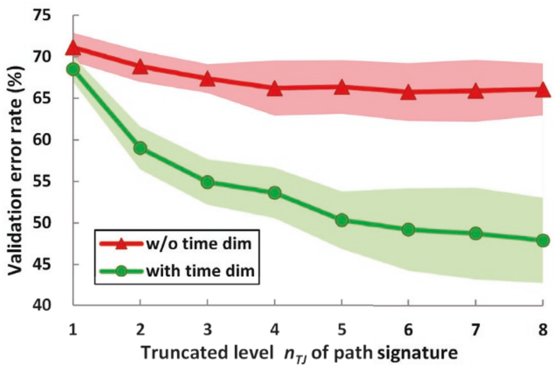

Fig. 6 Comparison of **T-J-PSF** w/ and w/o using time-incorporated paths. The colored areas are the error bands.

the time-dependent part (TD) raise rapidly. It demonstrates the TI (i.e. the signature of original path) is invariant under time reparameterization while the TD is very sensitive to speed variation. The larger the signature level n, the more sensitive the TD is to speed variation. Similarly, in the other direction, when the frame rate decreases from 30fps to 6fps, the influence to TD is far more significant than that to TI, showing the tolerance of TI under missing frames.

If we replace the PSF with the overlapping DPSF, then an appropriate hierarchical level h_{TJ} needs to be chosen. As shown in 8, in terms of performance, the low-level (e.g., $n_{TJ} = 2$) overlapping DPSFs over the hierarchical intervals (e.g., $h_{TJ} = 3$) often outperform the high-level (e.g., $n_{TJ} = 5$) PSFs over the whole interval ($h_{TJ} = 1$), which shows the efficiency of using temporal disintegration. However, when the disintegrated paths are too fragmented to avoid being dominated by local noises (e.g., when $h_{TJ} > 3$), the additional features are harmful. We thus fixed $h_{TJ} = 3$. Another observation is that the improvements from $h_{TJ} = 1$ to $h_{TJ} = 3$ become less significant along with the increasing n_{TJ}, demonstrating a trend that the high-level PSF and lowlevel DPSF yield similar information eventually.

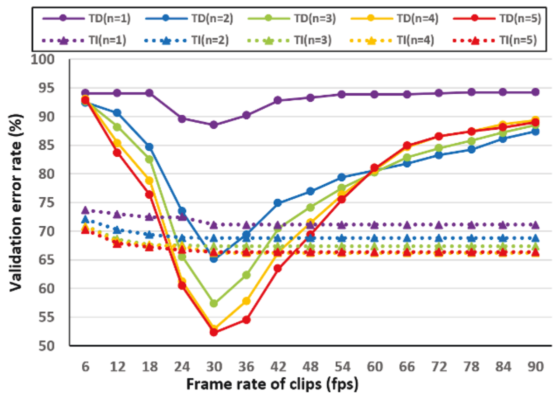

Fig. 7 Sensitivity of the time-dependent and time-independent part of the time-incorporated PSF to different frame rates.

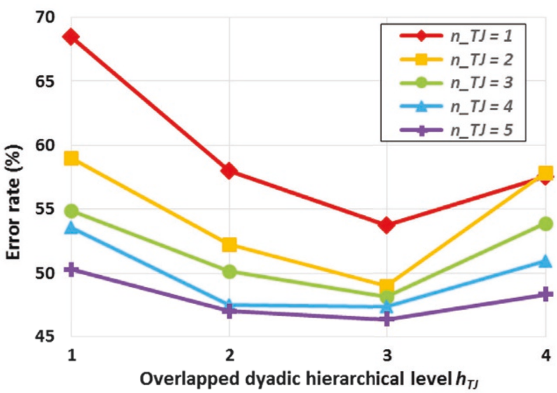

Fig. 8 Comparison of **T-J-PSF** with different dyadic hierarchical level h_{TJ} and different truncation level n_{TJ} of signature.

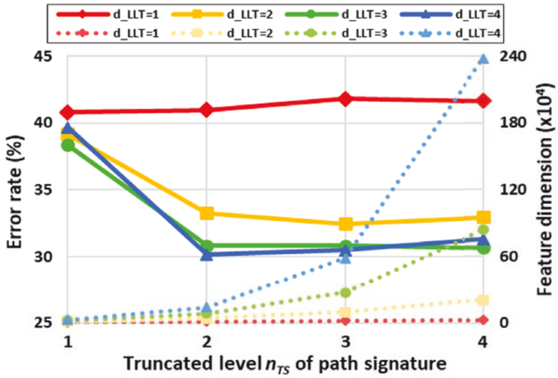

Fig. 9 Comparison of the error rates (solid) and the feature dimensions (dashed) of using **T-S-PSF** with different dimensions d_{LLT} of lead-lag paths and different truncation level n_{TS} of signature.

6.5.2 Investigation of T-S-PSF

Regarding the PSF derived from the evolution of the spatial context (**T-S-PSF**), two factors were evaluated: the dimension d_{LLT} of the lead-lag path and the truncation level n_{TS} of the signature. As shown in Fig. 9, the results improve when a higher dimension d_{LLT} of the lead-lag path is adopted, but the marginal improvement is less obvious when $d_{LLT} \geq 3$. For the truncation level n_{TS}, the improvements are significant from $n_{TS} = 1$ to $n_{TS} = 2$, but they are negligible when $n_{TS} > 2$. The dashed lines in Fig.9 show the trends of feature dimension under different parameters. By making a trade-off between model complexity and performance, we fixed $d_{LLT} = 3$, $n_{TS} = 2$.

By using the overlapping DPSF instead of PSF, the validation error rates are 30.82 ± 7.00%, 26.07 ± 6.12%, 26.39 ± 5.51%, and 26.07 ± 5.23%, when the hierarchical level h_{TS} is 1, 2, 3, and 4 respectively. Thus, we fixed h_{TS} to 3.

6.6 Ablation study

For the ablation study of our features on the JHMDB [31], we used the parameter setting for each feature based on the foregoing analysis. We retrained the network using the whole training set (including the validation set) and took the final result as the average of the three splits. The results are shown in Table 4. We can see that adding the spatial PSF (Ex. 4) to the baseline (Ex. 1) gives an improvement of about 20%, and further adding the temporal PSF (Ex. 9) contributes an additional 10%. The spatial context may be alternative between joint pairs and joint triples, for example Ex. 2 vs. Ex. 3, or Ex. 7 vs. Ex. 8, but they are complementary as shown in Ex. 4 and Ex. 9.

Applying the invisibility-reset transformation to all the paths before taking the temporal signatures allow us to remove all the spatial components **S-J**, **S-P-PSF**, and **S-TPSF**, while obtain the same accuracy as that of Ex. 9.

Also, we evaluated the method which directly takes all the evolving N landmarks in $\mathbb{R}^d$ as a Nd-dimensional path for signature extraction. Together with **S-J**, it achieves 55.0% in accuracy. The dimension of this PSF is $\varphi'(Nd, n) = 838,230$ when $n = 4$ and will be impractical when $n > 4$. This shows the cost-effectiveness of using pose and temporal disintegration.

Table 4 Effect of different signature levels on the performance of S-T-PSF

Ex.#	S-J	S-P-PSF	S-T-PSF	T-J-PSF	T-S-PSF (S-P)*	T-S-PSF (S-T)*	Accuracy (%)
1	o						48.9
2	o	o					68.4
3	o		o				68.7
4	o	o	o				69.2
5	o			o			62.0
6	o	o	o	o			73.5
7	o	o		o	o		79.1
8	o		o	o		o	78.3
9	o	o	o	o	o	o	80.4

∗ S-P (S-T) means the temporal features are only on the base of spatial joint pairs (joint triples).

6.7 Comparison with the state-of-the-art methods

To achieve our best results, we adopted the best settings of parameters from the foregoing analysis. For the JHMDB dataset [31], the results were given in the previous subsections. For the other three datasets, we followed the evaluation criteria in [40].

6.7.1 Comparison over small datasets

For the JHMDB dataset, previous state-of-the-art methods are high-level pose feature (HLPF) [31] and its modified version (i.e. Novel HLPF [34]), dense trajectory features [102] encoded by Fisher vectors [103], and the pose-based CNN features (P-CNN) [79]. As shown in Table 5, our method, which uses only the joint locations, achieve better performance than the P-CNN which requires additional RGB information. Further, our method manages the high degree of nonlinearity, and outperforms other methods using hand-designed features like HLPF. Also, the computation of our feature extraction is very fast. The average speed using *esig* [91] on a single thread of an Intel 2.4GHz Xeon Gold 6240R CPU is 85 fps on the JHMDB dataset.

Moreover, we used the off-the-shelf pose estimation called Alphapose (with Pose-flow) [104] to get a set of 17 estimated joints from the RGB videos of the sub-JHMDB dataset, and then trained and tested the network using the estimated poses. By using only location information, our test accuracy is 68.2%, which outperforms that of PCNN [79] (66.8%), PA-AP [105] (61.5%), JointAP [106] (61.2%), or HLPF [31] (54.1%). As an example of the flexibility of our method on additional clues, taking the confidence scores from the pose estimation as an additional dimension of landmarks raises the accuracy rate to 75.7%. However, a gap of accuracy still exists between using estimated poses and ground truth poses (84.23% by ours).

Table 5 Comparison of methods on JHMDB using ground-truth landmarks

Methods	Accuracy (%)
DT-FV [102]	65.9
P-CNN [79]	74.6
HLPF [31]	76.0
Novel HLPF [34]	79.6
Path Signature (Ours)	80.4

For the SBU Interaction dataset, the two human bodies are regarded as one united articulated system with a total of 30 joints in 3D. As shown in Table 6, the proposed method using PSF significantly outperforms the other skeleton-based methods including many RNN-based or LSTM-based ones. Aside from the accuracy, the interpretable PSF could facilitate further understanding of interactions.

Table 6 Comparison of methods on SBU dataset

Method	Accuracy (%)
Yun et al. [41]	80.3
Ji et al. [107]	86.9
CHARM [108]	83.9
HBRNN [24] (reported by [37])	80.4
Deep LSTM (reported by [37])	86.0
Co-occurrence LSTM [37]	90.4
STA-LSTM [109]	91.5
ST-LSTM-Trust Gate [40, 25]	93.3
SkeletonNet [110]	93.5
GC-Attention-LSTM [80]	94.1
Path Signature (Ours)	96.8

Table 7 Comparison of methods on MHAD dataset

Method	Accuracy (%)
Vantigodi et al. [111]	96.1
Ofli et al. [73]	95.4
Vantigodi et al. [112]	97.6
Kapsouras et al. [113]	98.2
HBRNN [24]	100
ST-LSTM-Trust Gate [40, 25]	100
Path Signature (Ours)	100

For the Berkeley MHAD dataset, we achieve the same accuracy (100%) as the state-of-the-art methods shown in Table 7, showing the effectiveness of PSF for recognizing actions with accurate joint locations.

These results show that the proposed hand-designed feature set with single-layer linear network can outperform most deep learning methods on small datasets.

6.7.2 Comparison over large-scale datasets

We also conducted experiments on the large-scale NTURGB+D data.

For normalization, we applied the same 3D rotation and scaling as those in [39], so the body in the first frame faces the camera directly and those in the following frames are compensated accordingly. Since in this dataset different actions contain different number of detected actors, we applied a two-stage classification. The first stage is a binary classifier separating the actions into two types: 1-body or 2-body actions, then the second stage is the corresponding classifier (1-body or 2-body classifier) for each type. The supervised label of the binary classification at the first stage can be found by going through all the training samples and calculating the average number of actors in each action class. The range of the numbers is [1.02, 1.06] for the first 49 classes which are annotated as 1-body actions, while the range is [1.87, 2.04] for the remaining 11 classes which are annotated as 2-body actions.

Before feature extraction, we ranked all the detected actors in each clip based on the magnitudes of their movements. Then, for the 1-body classifier, features were extracted from the most active actor. For the first-stage binary classifier and the 2-body classifier, the two most active actors were regarded as one evolving object; this means we ended up having twice the number of joints per frame (i.e., 50 joints per frame). If a body is missing in the entire clip, the coordinates of this body are set to 0; if a body is missing in some medial frames, its coordinates are filled in using cubic spline interpolation [114].

The final results were given by two-stage classification as shown in Table 8. Table 9 shows that our method outperforms many deep learning based methods. The GCN [30] and its variants [83, 84, 85] achieve the current state-of-the-art accuracy on NTURGB+D dataset by taking advantage of the human skeleton structure. To utilize this skeleton structure as a prior knowledge to reduce complexity in our feature set is worth further studying.

Table 8 Accuracy (%) of the two-stage classification on NTURGB+D dataset

Task	The 1st stage	The 2nd stage		Final
		1-body	2-body	
Cross-subject	99.2	75.7	91.9	78.3
Cross-view	99.3	82.5	94.4	86.1

Table 9 Comparison of methods on NTURGB+D dataset

Method	Deep networks?	Cross-subject	Cross-view
Dynamic Skeletons [115]	X(SVM)	60.2	65.2
HBRNN [24]	✓(RNN)	59.1	64.0
Part-aware LSTM [39]	✓(LSTM)	62.9	70.3
ST-LSTM-Trust Gate [40, 25]	✓(LSTM)	69.2	77.7
STA-LSTM [109]	✓(LSTM)	73.4	81.2
SkeletonNet [110]	✓(CNN)	75.9	81.2
Joint Distance Maps [116]	✓(CNN)	76.2	82.3
GC-Attention-LSTM [80]	✓(LSTM)	74.4	82.8
Deep STGC [82]	✓(GCN)	74.8	86.3
ST-GCN [30]	✓(GCN)	81.5	88.3
Path Signature (Ours)	X(Single-layer NN)	78.3	86.1

6.8 Toward understanding of human actions

The interpretable geometric properties of PSF facilitate the understanding of human actions. By using a linear classifier the importance of each feature to each action class can be evaluated by the product of the two-layer weight matrices. For each class of sub-JHMDB, we ranked the joint pairs/triples according to the average over the weights connecting the features of joint groups and the corresponding class label. The top-3 joint pairs/triples for spatial and temporal features are shown in Fig. 10. The spatial ones often emphasize static constraints while the temporal ones highlight dynamic variations. Notice that many top pairs/triples are physically non-local, which demonstrates the effectiveness of the pose disintegration method.

Moreover, by using temporal disintegration ($h = 3$), we can evaluate the importance of different timescales and time intervals. As shown in Fig. 11, discriminative motions often appear in various intervals of finer timescales, e.g., the start of “catch” or “pick”, the middle of “kick ball” or “swing ball”, and the end of “golf” or “jump”.

7 Conclusions

In this paper, we refined the path signature as a robust, nonlinear, and interpretable feature for landmark-based data. Path disintegrations and transformations are proposed to improve the effectiveness and efficiency of signature features. Based on these, we designed and built the signature-based spatio-temporal representation of action sequences. Experimental results show that using our feature set, a linear shallow fully-connected neural network achieves comparable results to advanced methods including CNN-based and RNN-based ones, especially on small datasets.

For future work, one could reduce the size of the representation of the body or feature set based on our analysis and understanding of human actions. It would also be interesting to integrate our landmark-based representation with other informative cues (e.g., appearance) to improve the performance of HAR. Moreover, our method

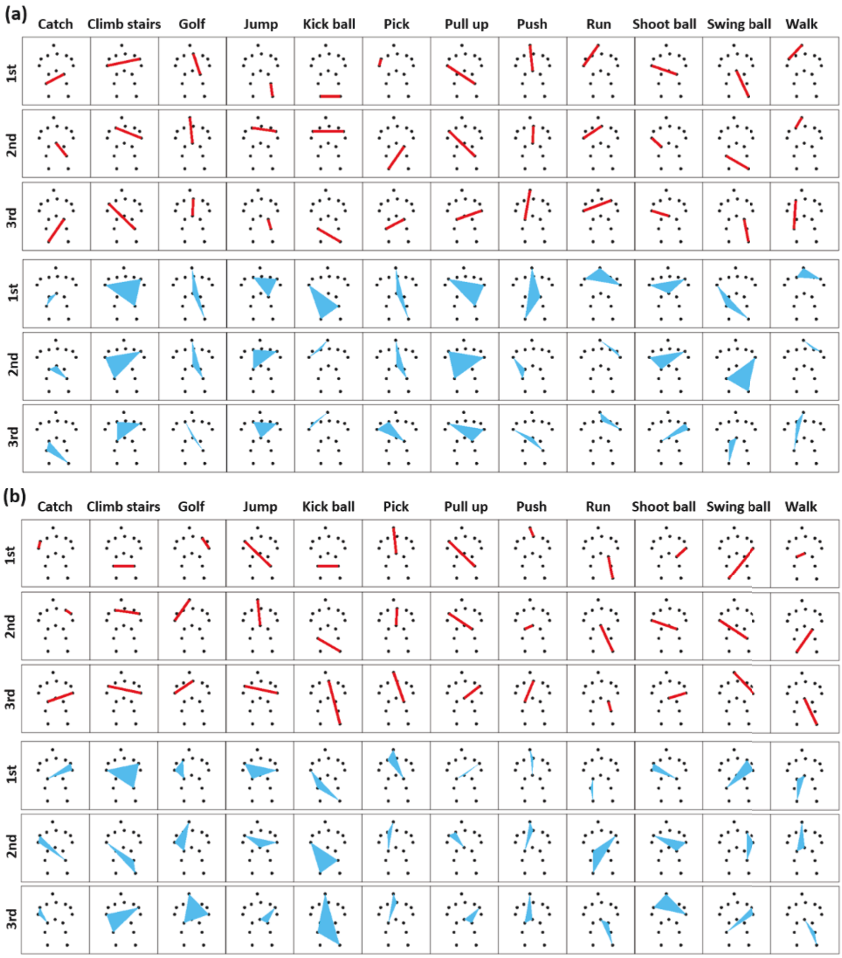

Fig. 10 Top-3 most important joint pairs/triples for (a) spatial features and (b) temporal features based on our linear network.

is general enough for other landmark-based objects where the given information in each landmark can be diverse.

Acknowledgements T. L. and H. N. are supported by the Alan Turing Institute under the EPSRC grant EP/N510129/1. This work was supported by ERC advanced grant ESig (no. 291244). C. S. was supported in part by ERC advanced grant ALLEGRO. This work was supported in part by the Alexander von Humbolt Foundation. W. Y. and L. J. were supported in part by NSFC (Grant no.: 61472144, 61673182), the National Key Research & Development Plan of China (no.2016YFB1001405), GD-NSF (no.2017A030312006), GDSTP (Grant no.:2015B010101004, 2015B010130003), GZSTP (no.201607010227). W. Y. was supported by Royal Society Newton International Fellowship (NIF/R1/181466).

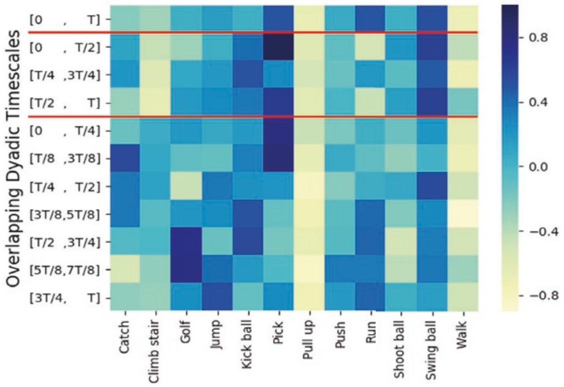

Fig. 11 Visualization of the important timescales and time periods for the actions in sub-JHMDB dataset. The darker in color, the more important it is.

References

[1] T. B. Moeslund and E. Granum. "A survey of computer vision-based human motion capture". In: *Computer vision and image understanding* 81.3 (2001), pp. 231–268.

[2] S. Mitra and T. Acharya. "Gesture recognition: A survey". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37.3 (2007), pp. 311–324.

[3] R. Poppe. "A survey on vision-based human action recognition". In: *Image and vision computing* 28.6 (2010), pp. 976–990.

[4] D. Weinland, R. Ronfard, and E. Boyer. "A survey of vision-based methods for action representation, segmentation and recognition". In: *Computer vision and image understanding* 115.2 (2011), pp. 224–241.

[5] M. Ziaeefard and R. Bergevin. "Semantic human activity recognition: A literature review". In: *Pattern Recognition* 48.8 (2015), pp. 2329–2345.

[6] G. Guo and A. Lai. "A survey on still image based human action recognition". In: *Pattern Recognition* 47.10 (2014), pp. 3343–3361.

[7] C. H. Lim, E. Vats, and C. S. Chan. "Fuzzy human motion analysis: A review". In: *Pattern Recognition* 48.5 (2015), pp. 1773–1796.

[8] L. L. Presti and M. La Cascia. "3D skeleton-based human action classification: A survey". In: *Pattern Recognition* 53 (2016), pp. 130–147.

[9] G. Johansson. "Visual perception of biological motion and a model for its analysis". In: *Perception & psychophysics* 14.2 (1973), pp. 201–211.

[10] S. Sadanand and J. J. Corso. "Action bank: A high-level representation of activity in video". In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2012, pp. 1234–1241.

[11] A. Ciptadi, M. S. Goodwin, and J. M. Rehg. "Movement pattern histogram for action recognition and retrieval". In: *European conference on computer vision*. Springer. 2014, pp. 695–710.

[12] R. Vemulapalli, F. Arrate, and R. Chellappa. "Human action recognition by representing 3d skeletons as points in a lie group". In: *Proceedings of the*

IEEE conference on computer vision and pattern recognition. 2014, pp. 588–595.

[13] W. Yang et al. “Improved deep convolutional neural network for online handwritten Chinese character recognition using domain-specific knowledge”. In: *2015 13th international conference on document analysis and recognition (ICDAR)*. IEEE. 2015, pp. 551–555.

[14] W. Yang et al. “DropSample: A new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition”. In: *Pattern Recognition* 58 (2016), pp. 190–203.

[15] Z. Xie et al. “Fully convolutional recurrent network for handwritten chinese text recognition”. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE. 2016, pp. 4011–4016.

[16] Z. Xie et al. “Learning spatial-semantic context with fully convolutional recurrent network for online handwritten chinese text recognition”. In: *IEEE transactions on pattern analysis and machine intelligence* 40.8 (2017), pp. 1903–1917.

[17] W. Yang, L. Jin, and M. Liu. “Chinese character-level writer identification using path signature feature, DropStroke and deep CNN”. In: *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*. IEEE. 2015, pp. 546–550.

[18] W. Yang, L. Jin, and M. Liu. “Deepwriterid: An end-to-end online text-independent writer identification system”. In: *IEEE Intelligent Systems* 31.2 (2016), pp. 45–53.

[19] B. Graham. “Sparse arrays of signatures for online character recognition”. In: *arXiv preprint arXiv:1308.0371* (2013).

[20] L. G. Gyurkó et al. “Extracting information from the signature of a financial data stream”. In: *arXiv preprint arXiv:1307.7244* (2013).

[21] J. Diehl. “Rotation invariants of two dimensional curves based on iterated integrals”. In: *arXiv preprint arXiv:1305.6883* (2013).

[22] W. Yang et al. “Rotation-free online handwritten character recognition using dyadic path signature features, hanging normalization, and deep neural network”. In: *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE. 2016, pp. 4083–4088.

[23] I. P. Arribas et al. “A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder”. In: *Translational psychiatry* 8.1 (2018), pp. 1–7.

[24] Y. Du, W. Wang, and L. Wang. “Hierarchical recurrent neural network for skeleton based action recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1110–1118.

[25] J. Liu et al. “Skeleton-based action recognition using spatio-temporal LSTM network with trust gates”. In: *IEEE transactions on pattern analysis and machine intelligence* 40.12 (2017), pp. 3007–3021.

[26] T. Kerola, N. Inoue, and K. Shinoda. “Spectral graph skeletons for 3D action recognition”. In: *Asian conference on computer vision*. Springer. 2014, pp. 417–432.

[27] M. Müller, T. Röder, and M. Clausen. “Efficient content-based retrieval of motion capture data”. In: *ACM SIGGRAPH 2005 Papers*. 2005, pp. 677–685.
[28] G. Evangelidis, G. Singh, and R. Horaud. “Skeletal quads: Human action recognition using joint quadruples”. In: *2014 22nd International Conference on Pattern Recognition*. IEEE. 2014, pp. 4513–4518.
[29] M. Li and H. Leung. “Multiview skeletal interaction recognition using active joint interaction graph”. In: *IEEE Transactions on Multimedia* 18.11 (2016), pp. 2293–2302.
[30] S. Yan, Y. Xiong, and D. Lin. “Spatial temporal graph convolutional networks for skeleton-based action recognition”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 32. 1. 2018.
[31] H. Jhuang et al. “Towards understanding action recognition”. In: *Proceedings of the IEEE international conference on computer vision*. 2013, pp. 3192–3199.
[32] J. Wang et al. “Mining actionlet ensemble for action recognition with depth cameras”. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2012, pp. 1290–1297.
[33] X. Yang and Y. L. Tian. “Effective 3D action recognition using EigenJoints”. In: *Journal of Visual Communication & Image Representation* 25.1 (2014), pp. 2–11.
[34] J. Fan, Z. Zha, and X. Tian. “Action recognition with novel high-level pose features”. In: *IEEE International Conference on Multimedia & Expo Workshops*. 2016.
[35] R. J. Williams and D. Zipser. “A learning algorithm for continually running fully recurrent neural networks”. In: *Neural computation* 1.2 (1989), pp. 270–280.
[36] S. Hochreiter and J. Schmidhuber. “Long short-term memory”. In: *Neural computation* 9.8 (1997), pp. 1735–1780.
[37] W. Zhu et al. “Co-occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. 1. 2016.
[38] V. Veeriah, N. Zhuang, and G.-J. Qi. “Differential recurrent neural networks for action recognition”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 4041–4049.
[39] A. Shahroudy et al. “Ntu rgb+ d: A large scale dataset for 3d human activity analysis”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 1010–1019.
[40] J. Liu et al. “Spatio-temporal lstm with trust gates for 3d human action recognition”. In: *European conference on computer vision*. Springer. 2016, pp. 816–833.
[41] K. Yun et al. “Two-person interaction detection using body-pose features and multiple instance learning”. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE. 2012, pp. 28–35.

[42] F. Ofli et al. “Berkeley mhad: A comprehensive multimodal human action database”. In: *2013 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE. 2013, pp. 53–60.
[43] T. Lyons. “Rough paths, signatures and the modelling of functions on streams”. In: *arXiv preprint arXiv:1405.4537* (2014).
[44] K.-T. Chen. “Integration of paths–A faithful representation of paths by noncommutative formal power series”. In: *Transactions of the American Mathematical Society* 89.2 (1958), pp. 395–407.
[45] B. Hambly and T. Lyons. “Uniqueness for the signature of a path of bounded variation and the reduced path group”. In: *Annals of Mathematics* (2010), pp. 109–167.
[46] H. Boedihardjo et al. “The signature of a rough path: uniqueness”. In: *Advances in Mathematics* 293 (2016), pp. 720–737.
[47] F. Yin et al. “ICDAR 2013 Chinese handwriting recognition competition”. In: *2013 12th international conference on document analysis and recognition*. IEEE. 2013, pp. 1464–1470.
[48] I. P. Arribas. “Derivatives pricing using signature payoffs”. In: *arXiv preprint arXiv:1809.09466* (2018).
[49] F. J. Király and H. Oberhauser. “Kernels for sequentially ordered data”. In: *Journal of Machine Learning Research* 20.31 (2019), pp. 1–45.
[50] C. Salvi et al. “The Signature Kernel is the solution of a Goursat PDE”. In: *arXiv e-prints* (2020), arXiv–2006.
[51] P. Kidger et al. “Neural controlled differential equations for irregular time series”. In: *Conference on Neural Information Processing Systems (NeurIPS)*. 2020.
[52] I. Chevyrev, V. Nanda, and H. Oberhauser. “Persistence paths and signature features in topological data analysis”. In: *IEEE transactions on pattern analysis and machine intelligence* 42.1 (2018), pp. 192–202.
[53] W. Yang et al. “Leveraging the path signature for skeleton-based human action recognition”. In: *arXiv preprint arXiv:1707.03993* 1 (2017).
[54] P. Bonnier et al. “Deep Signature Transforms”. In: *Conference on Neural Information Processing Systems (NeurIPS)*. 2019.
[55] S. Liao et al. “Learning stochastic differential equations using RNN with log signature features”. In: *arXiv preprint arXiv:1908.08286* (2019).
[56] J. P. Biyong et al. “Information Extraction from Swedish Medical Prescriptions with Sig-Transformer Encoder”. In: *Proceedings of the 3rd Clinical Natural Language Processing Workshop*. 2020, pp. 41–54.
[57] D. Wilson-Nunn et al. “A path signature approach to online arabic handwriting recognition”. In: *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*. IEEE. 2018, pp. 135–139.
[58] S. Lai, Y. Zhu, and L. Jin. “Encoding Pathlet and SIFT Features With Bagged VLAD for Historical Writer Identification”. In: *IEEE Transactions on Information Forensics and Security* 15 (2020), pp. 3553–3566.

[59] S. Lai and L. Jin. “Recurrent adaptation networks for online signature verification”. In: *IEEE Transactions on Information Forensics and Security* 14.6 (2018), pp. 1624–1637.

[60] G. Biau and A. Fermanian. “Learning with Signatures”. In: *International Workshop on Functional and Operatorial Statistics*. Springer. 2020, pp. 19–26.

[61] T. Ahmad et al. “Human Action Recognition in Unconstrained Trimmed Videos Using Residual Attention Network and Joints Path Signature”. In: *IEEE Access* 7 (2019), pp. 121212–121222.

[62] C. Li et al. “Skeleton-based gesture recognition using several fully connected layers with path signature features and temporal transformer module”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 01. 2019, pp. 8585–8593.

[63] B. Wang et al. “A path signature approach for speech emotion recognition”. In: *Interspeech 2019*. ISCA. 2019, pp. 1661–1665.

[64] Y. Wu et al. “Signature features with the visibility transformation”. In: *arXiv preprint arXiv:2004.04006* (2020).

[65] V. M. Zatsiorsky and V. M. Zaciorskij. *Kinetics of human motion*. Human kinetics, 2002.

[66] J. K. Aggarwal and M. S. Ryoo. “Human activity analysis: A review”. In: *ACM Computing Surveys (CSUR)* 43.3 (2011), pp. 1–43.

[67] M. Ye et al. “A survey on human motion analysis from depth data”. In: *Time-of-flight and depth imaging. sensors, algorithms, and applications*. Springer, 2013, pp. 149–187.

[68] L. Wang, D. Q. Huynh, and P. Koniusz. “A comparative review of recent kinect-based action recognition algorithms”. In: *IEEE Transactions on Image Processing* 29 (2019), pp. 15–28.

[69] M. E. Hussein et al. “Human action recognition using a temporal hierarchy of covariance descriptors on 3d joint locations”. In: *Twenty-third international joint conference on artificial intelligence*. 2013.

[70] C. Ellis et al. “Exploring the trade-off between accuracy and observational latency in action recognition”. In: *International Journal of Computer Vision* 101.3 (2013), pp. 420–436.

[71] L. Xia, C.-C. Chen, and J. K. Aggarwal. “View invariant human action recognition using histograms of 3d joints”. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE. 2012, pp. 20–27.

[72] F. Ofli et al. “Sequence of the most informative joints (smij): A new representation for human skeletal action recognition”. In: *Journal of Visual Communication and Image Representation* 25.1 (2014), pp. 24–38.

[73] C. Wang, Y. Wang, and A. L. Yuille. “An approach to pose-based action recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013, pp. 915–922.

[74] P. Wei et al. “Concurrent action detection with structural prediction”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013, pp. 3136–3143.

[75] A. Shahroudy et al. “Multimodal multipart learning for action recognition in depth videos”. In: *IEEE transactions on pattern analysis and machine intelligence* 38.10 (2015), pp. 2123–2129.

[76] Y. Hou et al. “Skeleton optical spectra-based action recognition using convolutional neural networks”. In: *IEEE Transactions on Circuits and Systems for Video Technology* 28.3 (2016), pp. 807–811.

[77] C. Li et al. “Co-occurrence feature learning from skeleton data for action recognition and detection with hierarchical aggregation”. In: *arXiv preprint arXiv:1804.06055* (2018).

[78] M. Rhif, H. Wannous, and I. R. Farah. “Action recognition from 3D skeleton sequences using deep networks on lie group features”. In: *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE. 2018, pp. 3427–3432.

[79] G. Chéron, I. Laptev, and C. Schmid. “P-cnn: Pose-based cnn features for action recognition”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 3218–3226.

[80] J. Liu et al. “Global context-aware attention lstm networks for 3d action recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 1647–1656.

[81] W. Yang et al. “Developing the path signature methodology and its application to landmark-based human action recognition”. In: *arXiv preprint arXiv:1707.03993* (2017).

[82] C. Li et al. “Spatio-Temporal Graph Convolution for Skeleton Based Action Recognition”. In: *Proceedings of the AAAI conference on artificial intelligence*. 2018.

[83] L. Shi et al. “Skeleton-based action recognition with directed graph neural networks”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 7912–7921.

[84] M. Wang, B. Ni, and X. Yang. “Learning Multi-View Interactional Skeleton Graph for Action Recognition”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020).

[85] H. Xia and X. Gao. “Multi-Scale Mixed Dense Graph Convolution Network for Skeleton-Based Action Recognition”. In: *IEEE Access* 9 (2021), pp. 36475–36484.

[86] T. J. Lyons. “Differential equations driven by rough signals”. In: *Revista Matemática Iberoamericana* 14.2 (1998), pp. 215–310.

[87] H. Ni. “A multi-dimensional stream and its signature representation”. In: *arXiv preprint arXiv:1509.03346* (2015).

[88] I. Chevyrev and A. Kormilitzin. “A primer on the signature method in machine learning”. In: *arXiv preprint arXiv:1603.03788* (2016).

[89] H. Boedihardjo, H. Ni, and Z. Qian. “Uniqueness of signature for simple curves”. In: *Journal of Functional Analysis* 267.6 (2014), pp. 1778–1806.

[90] G. Flint, B. Hambly, and T. Lyons. "Discretely sampled signals and the rough Hoff process". In: *Stochastic Processes and their Applications* 126.9 (2016), pp. 2593–2614.
[91] T. Lyons et al. *Esig on PyPi derived from CoRoPa: Computational Rough Paths software library.* https://github.com/datasig-ac-uk/esig,https://github.com/terrylyons/libalgebra, http://coropa.sourceforge.net/. 2007–2021.
[92] J. Reizenstein and B. Graham. "The iisignature library: efficient calculation of iterated-integral signatures and log signatures". In: *arXiv preprint arXiv:1802.08252* (2018).
[93] P. Kidger and T. Lyons. "Signatory: differentiable computations of the signature and logsignature transforms, on both CPU and GPU". In: *International Conference on Learning Representations*. 2020.
[94] M. Pfeffer, A. Seigal, and B. Sturmfels. "Learning paths from signature tensors". In: *SIAM Journal on Matrix Analysis and Applications* 40.2 (2019), pp. 394–416.
[95] C. Améndola, P. Friz, and B. Sturmfels. "Varieties of signature tensors". In: *Forum of Mathematics, Sigma*. Vol. 7. Cambridge University Press. 2019.
[96] *Implementation of the Path-signature-feature methodology for creating datasets for human action recognition from landmark data.* https://github.com/kschlegel/PSFDataset/. 2020.
[97] *CMU graphics lab motion capture database.* http://mocap.cs.cmu.edu/. 2003.
[98] *Microsoft Kinect.* http://www.xbox.com/en-US/Kinect/.
[99] J. Shotton et al. "Real-time human pose recognition in parts from single depth images". In: *Communications of the ACM* 56.1 (2013), pp. 116–124.
[100] L. Wan et al. "Regularization of neural networks using dropconnect". In: *International conference on machine learning*. PMLR. 2013, pp. 1058–1066.
[101] G. E. Hinton et al. "Improving neural networks by preventing co-adaptation of feature detectors". In: *arXiv preprint arXiv:1207.0580* (2012).
[102] H. Wang and C. Schmid. "Action recognition with improved trajectories". In: *Proceedings of the IEEE international conference on computer vision*. 2013, pp. 3551–3558.
[103] D. Oneata, J. Verbeek, and C. Schmid. "Action and event recognition with fisher vectors on a compact feature set". In: *Proceedings of the IEEE international conference on computer vision*. 2013, pp. 1817–1824.
[104] H.-S. Fang et al. "Rmpe: Regional multi-person pose estimation". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017, pp. 2334–2343.
[105] U. Iqbal, M. Garbade, and J. Gall. "Pose for action-action for pose". In: *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. IEEE. 2017, pp. 438–445.

[106] B. Xiaohan Nie, C. Xiong, and S.-C. Zhu. “Joint action recognition and pose estimation from video”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 1293–1301.
[107] Y. Ji, G. Ye, and H. Cheng. “Interactive body part contrast mining for human interaction recognition”. In: *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE. 2014, pp. 1–6.
[108] W. Li et al. “Category-blind human action recognition: A practical recognition system”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 4444–4452.
[109] S. Song et al. “An end-to-end spatio-temporal attention model for human action recognition from skeleton data”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017.
[110] Q. Ke et al. “Skeletonnet: Mining deep part features for 3-d action recognition”. In: *IEEE signal processing letters* 24.6 (2017), pp. 731–735.
[111] S. Vantigodi and R. V. Babu. “Real-time human action recognition from motion capture data”. In: *2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*. IEEE. 2013, pp. 1–4.
[112] S. Vantigodi and V. B. Radhakrishnan. “Action recognition from motion capture data using meta-cognitive rbf network classifier”. In: *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. IEEE. 2014, pp. 1–6.
[113] I. Kapsouras and N. Nikolaidis. “Action recognition on motion capture data using a dynemes and forward differences representation”. In: *Journal of Visual Communication and Image Representation* 25.6 (2014), pp. 1432–1445.
[114] C. De Boor and C. De Boor. *A practical guide to splines*. Vol. 27. springer-verlag New York, 1978.
[115] J.-F. Hu et al. “Jointly learning heterogeneous features for RGB-D activity recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 5344–5352.
[116] C. Li et al. “Joint distance maps based action recognition with convolutional neural networks”. In: *IEEE Signal Processing Letters* 24.5 (2017), pp. 624–628.

Correction to: Stochastic Analysis, Filtering, and Stochastic Optimization

George Yin and Thaleia Zariphopoulou

Correction to:
G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6

The original version of the book was inadvertently published with incorrect abstracts in the chapters. This has now been amended.

In addition to this, the affiliation of author Dr. Bertram Tschiderer has been changed to *Faculty of Mathematics, University of Vienna* in the online version of Chapter 10.

The updated version of the book can be found at
https://doi.org/10.1007/978-3-030-98519-6
https://doi.org/10.1007/978-3-030-98519-6_10

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6_19

Index

G. Yin, T. Zariphopoulou (eds.), *Stochastic Analysis, Filtering, and Stochastic Optimization*,
https://doi.org/10.1007/978-3-030-98519-6