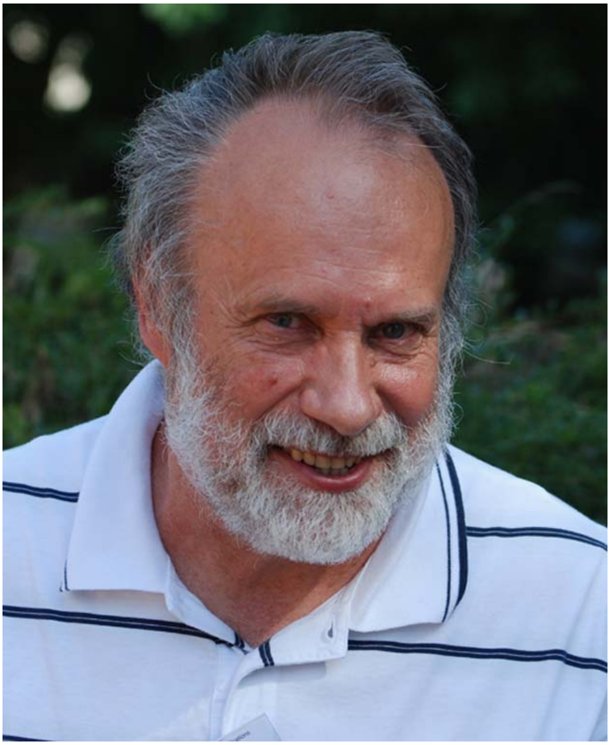
Freddy Delbaen
Miklós Rásonyi
Christophe Stricker
Editors

Optimality and Risk – Modern Trends in Mathematical Finance

S. Biagini
A. Cherny
F. Delbaen
B. Dupire
Ch. Frei
M. Frittelli
A. Gordon
M. Jeanblanc
C. Klüppelberg
V. Yu. Krasin
Y. Le Cam
R. Liptser
A.V. Melnikov
Yu. Mishura
S. Pergamenchtchikov
M. Rásonyi
M. Schweizer
G. Shevchenko
A.N. Shiryaev
I. M. Sonin
L. Stettner
E. Valkeila
P. Y. Zryumov

Springer

Optimality and Risk—
Modern Trends in Mathematical Finance



Freddy Delbaen • Miklós Rásonyi
Christophe Stricker
Editors

Optimality and Risk—Modern Trends in Mathematical Finance

The Kabanov Festschrift

Springer

Freddy Delbaen
Department of Mathematics
ETH Zürich
Rämistr. 101
8092 Zürich
Switzerland
delbaen@math.ethz.ch

Christophe Stricker
Laboratoire de Mathématiques
Université de Franche-Comté
16 route de Gray
25030 Besançon cedex
France
stricker@math.univ-fcomte.fr

Miklós Rásonyi
Computer & Automation
Institute of the Hungarian
Academy of Sciences
Kende utca 13-17
1111 Budapest
Hungary
rasonyi@sztaki.hu

ISBN 978-3-642-02607-2 e-ISBN 978-3-642-02608-9
DOI 10.1007/978-3-642-02608-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009935004

Mathematics Subject Classification (2000): 91B28, 60-06, 91-06

Cover design: WMXDesign GmbH

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This volume is dedicated to Yuri Kabanov,
to celebrate his 60th birthday

Preface

Youri (Yuri) Kabanov was born in a city Cherkassy, now Ukraine, formerly the USSR, in 1948. Even in high school he became interested in mathematics and actively and successfully participated in local and nationwide mathematical competitions for high school students. As a result, without exams he was admitted to a specialized mathematical boarding school at Kiev State University and he graduated from this school with the gold medal.

In the summer of 1966 Yuri came to Moscow with the hopes of studying theoretical physics and to apply to the Moscow Institute of Physics and Technology. This institute, known worldwide as Phystech (sometimes called the Russian version of MIT), was initially created in 1946 to produce high quality specialists in different fields of Physics. But it was located in a distant suburb of Moscow, and for a student who came from a small provincial city, living in a dormitory there meant being cut off from the life of the capital. One of his friends, who had a brother studying at Moscow University, decided to apply to the Department of Mechanics and Mathematics of Moscow State University (the internationally renowned MEKHMAT). He convinced Yuri that MEKHMAT would be a better choice. After passing the difficult entrance exams, Yuri became a student at Moscow University. Who knows, maybe this random event triggered his later interest in random processes.

At that time, known as the Golden Years of Moscow Mathematics (the title of the book, published by the American Math. Society), MEKHMAT had a reputation as one of the leading centers of mathematical research and math education. Students were taught by such great mathematicians as A.N. Kolmogorov, P.S. Aleksandrov, I.M. Gelfand, B.V. Gnedenko as well as then young but well known people like S. Novikov, Yu. Manin, D. Anosov, and Ya. Sinai.

In the beginning Yuri was interested in many fields and attended a variety of courses, but at the end of the second year he had chosen the subfield of probability theory. His scientific advisor was Albert N. Shiryaev, then a young professor, a former student of A. Kolmogorov. Shiryaev later became one of the most prominent leaders of the theory of stochastic processes.

Yuri's interests in his student years were not limited to mathematics only. He was actively involved in sports, becoming a member of the prestigious all-university volleyball team and excelling in track and field and the shot put. Having a phenomenal memory he knew many verses, and delighted his friends and girls by reciting from heart verses from such poets as Baudelaire, Rimbaud, and Gumilyov, who were rarely published during the Soviet years. The city where Yuri was born was a small one but rather old, founded in the late 13th century, and Yuri loves history, in which he is well-versed: today he enjoys presenting a short history of the Roman Empire using his collection of ancient coins as an illustration.

After graduating with distinction from Moscow University in 1971 Yuri became a postgraduate student at the Steklov Mathematical Institute. His first papers were devoted to the integral representation of functionals of Wiener, Poisson, and other processes with independent increments, and were published in the early 1970s in Theory of Probability and its Applications. His next paper "A generalized Itô formula for an extended stochastic integral with respect to Poisson random measure" was published in the prestigious Russian Mathematical Surveys in 1974. This extension was obtained by means of a generalized Cameron-Martin expansion in series of multiple Poisson integrals. Yuri also proved a generalized Itô formula for such extended stochastic integrals. As a result of his work in this new and rapidly developing field of stochastic processes, he obtained his Ph.D. degree with the dissertation topic Point Processes and Extended Stochastic Integrals. His supervisor was his teacher A.N. Shiryaev.

As a graduate student he helped Shiryaev to organize the very popular annual probability conferences in Bakuriani (Georgia). He recalls how he served as a "bodyguard" for a foreign visitor (at the time rare at such conferences), F. Spitzer. A striking feature of these conferences was a matching of intensive scientific spirit during the talks and everyday snow-skiing in the surrounding mountains. The leading figure at these conferences was Shiryaev, who instilled his love of skiing not only in Yuri but in many other participants.

After his postgraduate studies at the Steklov Institute, Yuri started working at one of the best research institutes of the Russian Academy of Sciences—in the Central Mathematical Economical Institute (CEMI) in Moscow. His ties to this Institute continued even after he moved to Besancon, France. At this Institute Yuri continued his studies of stochastic processes and published a few fundamental papers with R. Liptser and A. Shiryaev on absolute continuity and singularity of probability distributions, and on the convergence of distributions of counting processes. He also started working on some problems of stochastic control and problems of mathematical economics. He had strong ties with statisticians, initially working in a group headed by S. Aivasian and later becoming a member of the Probability Laboratory headed by V. Arkin, who assembled a very strong group of probabilists, many of whom became co-authors and friends of Yuri.

In the 1980s, under the influence of his teacher, A. Shiryaev, Yuri started working in the field of functional limit theorems. The summary of his results, on the contiguity of distributions of multivariate point processes and many others, was collected in his D.Sc. thesis, which he defended in 1984. At the end of the 1980s Yuri, together with his student S. Pergamenshchikov, started working in the area of singularly perturbed stochastic equations and partial differential equations. They published more than ten papers on the subject.

From the middle of the 1970s Yuri actively participated in teaching at a variety of Moscow and foreign institutions such as the Moscow Institute of Advanced Studies for Chemistry Managers and Engineers, the Moscow Institute of Aviation Technology (MATI), the Friedrich Schiller University in Jena, Germany, and the Moscow Institute of Electronics and Mathematics (MIEM). He also held a visiting position at Bilkent University in Ankara, Turkey.

In the early 1990s A. Shiryaev, together with his students, organized a seminar on financial mathematics, and this area defined the main directions of his studies thereafter. In 1994 a renowned issue of Theory of Probability and its Applications was almost entirely devoted to financial mathematics. This issue contained four papers which Yuri had participated in.

Since his 1995 nomination as a professor at the University of Besancon, Yuri has revived a forgotten local tradition by devoting himself mainly to mathematical finance: Louis Bachelier, the founder of mathematical finance, was a professor in Besancon from 1927 until his retirement in 1937.

An initial series of his works is related to large financial markets; to the optional decomposition theorem which is an essential tool for solving the superreplication problem; and to a general model of bond markets, which necessitated the introduction of a new type of stochastic integral with respect to processes with values in the space of continuous functions.

He soon became interested in financial markets with transaction costs, a field not too actively studied at the time because of its difficulty. In 1997 he and M. Safarian showed that Leland's approximation strategy did not asymptotically hedge call options, contrary to what was asserted in the literature. Then he launched himself into modeling markets with several (finitely many) basis assets. In fact, all models studied at that time comprised two assets only: a riskless and a risky one. Only the model of Jouini and Kallal allowed for multiple assets, but under very restrictive conditions: direct transfers being forbidden, all exchanges were required to pass by the bank account. In his paper "Hedging and liquidation under transaction costs in currency markets" (Finance and Stochastics, 1999) he proposed a geometric model for markets with transaction costs and showed that the object analogous to martingale measures is vector measures whose density processes take values in the dual of the solvency cone. This kind of model, adopted by all researchers in this field, has greatly contributed to progress, allowed for a satisfactory theory of arbitrage and for attacking important problems of finance (in particular, the superhedging of contingent claims) in the presence of proportional transaction costs. Yuri, working with various co-authors, has obtained criteria for the absence of arbitrage and superreplication theorems: in discrete time with complete or incomplete information; and in continuous time in increasingly general settings. He has just completed his latest book, with M. Safarian: "Markets with transaction costs, mathematical theory".

Yuri has also continued to be interested in stochastic differential equations and in problems of stochastic control. In his book, written together with S. Pergamenshchikov, "Two-scale stochastic systems, asymptotic analysis and control", he established a stochastic version of Tikhonov's theorem, obtained an asymptotic expansion for this system and showed a large deviation limit theorem.

Apart from his activities as a first-class researcher, Yuri has done a great deal for the dissemination of mathematics. In 1997 he actively participated in the founding of Finance and Stochastics, which in only a few years has become one of the leading journals in mathematical finance. The three Bachelier Colloquia he has organized (in 2000 in Besancon, in 2005 and in 2008 in Métabief) were extremely successful, collecting researchers from every continent of the world.

Yuri is now in the prime of his research and career. He is not only one of the most notable figures in financial mathematics and stochastic processes but also a very friendly person with new and fresh ideas, ready to share them with others. Almost 30 people are counted among his co-authors. He is a father of a now adult son and young daughter, and a loving and caring husband. His wife Margarita, or Rita as her friends call her, is a gifted photographer, and those visiting his or her website can see this for themselves.

There is no doubt that his friends, his co-authors, and many others who have had the chance to meet him personally or just to read his papers, will join us in wishing Yuri good health, happiness and many years to come of creative activity.

E. Presman, I. Sonin, and Ch. Stricker

Contents

On the Extension of the Namioka-Klee Theorem and on the Fatou Property for Risk Measures . . . 1
Sara Biagini and Marco Frittelli
1 Introduction . . . 1
2 The Extended Namioka Theorem . . . 7
2.1 The Current Literature . . . 9
3 On Order Lower Semicontinuity in Riesz Spaces . . . 10
3.1 Equivalent Formulations of Order l.s.c. . . . 11
3.2 The Order Continuous Dual $\mathbf{1}_n^{\sim}$ . . . 12
4 On the C-Property . . . 13
4.1 The C-Property in the Representation of Convex and Monotone Functionals . . . 14
5 Orlicz Spaces and Applications to Risk Measures . . . 16
5.1 Orlicz Spaces Have the C-Property . . . 16
5.2 New Insights on the Downside Risk and Risk Measures Associated to a Utility Function u . . . 19
5.3 Quadratic-Flat Utility . . . 26
5.4 Exponential Utility . . . 27
References . . . 28

On Certain Distributions Associated with the Range of Martingales . . . 29
Alexander Cherny and Bruno Dupire
1 Introduction . . . 29
2 Proofs . . . 33
3 Conclusion . . . 37
References . . . 38

Differentiability Properties of Utility Functions . . . 39
Freddy Delbaen
1 Notation and Preliminaries . . . 39
2 The Jouini-Schachermayer-Touzi Theorem . . . 42

3 A Consequence of Ekeland's Variational Principle and Other Family Members of Bishop-Phelps . . . 43
4 A Consequence of Automatic Continuity . . . 44
5 The One-Sided Derivative . . . 44
6 An Example . . . 45
7 The Example of an Incomplete Financial Market . . . 48
References . . . 48

Exponential Utility Indifference Valuation in a General Semimartingale Model . . . 49
Christoph Frei and Martin Schweizer
1 Introduction . . . 49
2 Motivation and Definition of $FER(H)$ . . . 51
3 No-arbitrage and existence of $FER(H)$ . . . 56
4 Relating $FER^\star(H)$ and $FER^\star(0)$ to the Indifference Value . . . 64
5 A BSDE Characterization of the Indifference Value Process . . . 73
6 Application to a Brownian Setting . . . 82
References . . . 85

The Expected Number of Intersections of a Four Valued Bounded Martingale with any Level May be Infinite . . . 87
Alexander Gordon and Isaac M. Sonin
1 Introduction . . . 87
2 Proof of Theorem 2. Cases $N = 2$ and $N = 3$ . . . 90
3 Proof of Theorem 2. Case $N > 3$. An Example . . . 93
References . . . 98

Immersion Property and Credit Risk Modelling . . . 99
Monique Jeanblanc and Yann Le Cam
1 Introduction . . . 99
2 Credit Modelling Framework . . . 101
2.1 The Two Information Flows . . . 102
2.2 Financial Interpretation of This Decomposition . . . 103
2.3 Absence of Arbitrage . . . 104
3 Representation Theorem in the Enlarged Filtration . . . 107
3.1 Representation of the $\mathbb{G}$-Martingales . . . 107
3.2 Change of Probability . . . 112
4 Complete Reference Market . . . 113
4.1 Description of the $\mathbb{G}$-Martingale Probabilities . . . 113
4.2 Completeness of the Full Market . . . 115
4.3 Immersion Property . . . 118
5 Incomplete Markets . . . 121
5.1 The Risk-Neutral Probabilities of the Full Market . . . 122
5.2 Default-Free Pricing Invariance . . . 124
5.3 Immersion Property . . . 129
6 Conclusion . . . 129
References . . . 129

Optimal Consumption and Investment with Bounded Downside Risk for Power Utility Functions . . . 133
Claudia Klüppelberg and Serguei Pergamenchtchikov
1 Introduction . . . 133
2 Formulating the Problem . . . 135
2.1 The Model . . . 135
2.2 The Control Processes . . . 137
2.3 The Cost Functions . . . 138
2.4 The Downside Risk Measures . . . 139
3 Problems and Solutions . . . 141
3.1 The Unconstrained Problem . . . 141
3.2 Value-at-Risk as Risk Measure . . . 143
3.3 Expected Shortfall as Risk Measure . . . 148
4 Proofs . . . 150
4.1 Proof of Theorem 1 . . . 150
4.2 Proof of Theorem 2 . . . 150
4.3 Proof of Theorem 3 . . . 153
4.4 Proof of Theorem 4 . . . 155
4.5 Proof of Theorem 5 . . . 155
4.6 Proof of Lemma 1 . . . 158
4.7 Proof of Theorem 7 . . . 158
4.8 Proof of Theorem 8 . . . 159
4.9 Proof of Theorem 9 . . . 159
Appendix . . . 161
5.1 A Technical Lemma . . . 161
5.2 The Verification Theorem . . . 161
5.3 A Special Version of Itô's Formula . . . 165
References . . . 169

On Comparison Theorem and its Applications to Finance . . . 171
Vladislav Y. Krasin and Alexander V. Melnikov
1 Introduction . . . 171
2 Comparison Theorem . . . 172
3 Applications to Mathematical Finance . . . 177
References . . . 180

Examples of FCLT in Random Environment . . . 183
R. Liptser
1 Introduction . . . 183
2 Assumptions, Notations and Main Result . . . 186
2.1 Notations . . . 186
2.2 Assumptions . . . 186
3 The Proof of Theorem 1 . . . 187
3.1 Auxiliary Lemma . . . 187
3.2 The Proof of (2) . . . 188
4 Diffusion in Random Environment . . . 191

4.1 $b(\omega,u) \equiv 0$ . . . 191
4.2 $b(\omega,u) \not\equiv 0$ . . . 192
5 Markov Chain as Random Environment . . . 192
6 Langevin Random Environment . . . 193
References . . . 194

The Optimal Time to Exchange one Asset for Another on Finite Interval . 197
Yuliya Mishura and Georgiy Shevchenko
1 Introduction . . . 197
2 Basic Properties of Premium Function and Stopping Domain . . . 198
3 Integral Equations for the Premium Function and the Threshold Curve . . . 201
4 Approaching Solution of Integral Equation for Threshold Curve . . 204
References . . . 210

Arbitrage Under Transaction Costs Revisited . . . 211
Miklós Rásonyi
1 Introduction . . . 211
2 Arbitrage and Price Systems . . . 212
3 Markets with One Risky Asset . . . 215
4 Proofs . . . 217
5 Conclusion . . . 220
Appendix . . . 220
References . . . 224

On the Linear and Nonlinear Generalized Bayesian Disorder Problem (Discrete Time Case) . . . 227
Albert N. Shiryaev and Pavel Y. Zryumov
1 Linear Penalty Case . . . 227
2 Nonlinear Penalty Case . . . 230
References . . . 235

Long Time Growth Optimal Portfolio with Transaction Costs . . . 237
Lukasz Stettner
1 Introduction . . . 237
2 Discrete Time Case . . . 241
3 Continuous Time Case . . . 246
References . . . 250

On the Approximation of Geometric Fractional Brownian Motion . . . 251
Esko Valkeila
1 Introduction . . . 251
1.1 Geometric Fractional Brownian Motion . . . 251
1.2 Motivation . . . 252
1.3 The Structure of the Note . . . 253
2 Approximation of fBm . . . 253

2.1 Construction of the Approximation 253
2.2 Further Properties of the Approximation 254
2.3 Approximation to Geometric fBm 257
3 Some Properties of the Approximation 257
3.1 Set-Up 257
3.2 Prelimit Market Models are Arbitrage-Free 258
3.3 Prelimit Market Models are Complete 260
4 Discussion and Conclusion 263
References 265

Contributors

Sara Biagini Dipartimento di Statistica e Mathematica Applicata all'Economia, Università di Pisa, Via Ridolfi, 10, 56124 Pisa, Italy, sara.biagini@ec.unipi.it

Alexander Cherny Department of Probability Theory, Faculty of Mechanics and Mathematics, Moscow State University, 119992 Moscow, Russia, alexander.cherny@gmail.com

Freddy Delbaen Department of Mathematics, ETH Zurich, Rämistr. 101, 8092 Zurich, Switzerland, delbaen@math.ethz.ch

Bruno Dupire Bloomberg L.P., 731 Lexington Avenue, New York, NY 10022, USA, bdupire@bloomberg.net

Christoph Frei Department of Mathematics, ETH Zurich, Rämistr. 101, 8092 Zurich, Switzerland, christoph.frei@math.ethz.ch

Marco Frittelli Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini, 50, 20122 Milano, Italy, marco.frittelli@mat.unimi.it

Alexander Gordon Department of Mathematics and Statistics, UNC Charlotte, 9201 University City Blvd, Charlotte, NC 28223-0001, USA, aygordon@uncc.edu

Monique Jeanblanc Département de Mathématiques, Université d'Évry – Val d'Essonne, boulevard François Mitterrand, 91025 Évry Cedex, France, monique.jeanblanc@univ-evry.fr

Claudia Klüppelberg Center for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, 85747 Garching, Germany, cklu@ma.tum.de

Vladislav Y. Krasin Department of Mathematical and Statistical Sciences, University of Alberta, 632 Central Academic Building, Edmonton, AB T6G 2G1, Canada, vkrasin@math.ualberta.ca

Yann Le Cam Département de Mathématiques, Université d'Évry – Val d'Essonne, boulevard François Mitterrand, 91025 Évry Cedex, France, yann_le_cam@yahoo.fr

R. Liptser Department of Electrical Engineering Systems, Tel Aviv University, 69978 Tel Aviv, Israel, liptser@eng.tau.ac.il

Alexander V. Melnikov Department of Mathematical and Statistical Sciences, University of Alberta, 632 Central Academic Building, Edmonton, AB T6G 2G1, Canada, melnikov@ualberta.ca

Yuliya Mishura Kyiv National Taras Shevchenko University, 64 Volodymyrska, 01033 Kyiv, Ukraine, myus@univ.kiev.ua

Serguei Pergamenchtchikov Laboratoire de Mathématiques, Raphaël Salem Université de Rouen, BP. 12, 76801 Saint Etienne du Rouvray, France, Serge.Pergamenchtchikov@univ-rouen.fr

Miklós Rásonyi Computer and Automation Institute of the Hungarian Academy of Sciences, Kende utca 13-17, 1111 Budapest, Hungary, rasonyi@sztaki.hu

Martin Schweizer Department of Mathematics, ETH Zurich, Rämistr. 101, 8092 Zurich, Switzerland, martin.schweizer@math.ethz.ch

Georgiy Shevchenko Kyiv National Taras Shevchenko University, 64 Volodymyrska, 01033 Kyiv, Ukraine, zhora@univ.kiev.ua

Albert N. Shiryaev Steklov Mathematical Institute, Russian Academy of Sciences, Gubkina str. 8, Moscow 119991, Russia, albertsh@mi.ras.ru

Isaac M. Sonin Department of Mathematics and Statistics, UNC Charlotte, 9201 University City Blvd, Charlotte, NC 28223, USA, imsonin@uncc.edu

Lukasz Stettner Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-956 Warsaw, Poland, stettner@impan.gov.pl

Esko Valkeila Institute of Mathematics, Helsinki University of Technology, P.O. Box 1100, 02015 Helsinki, Finland, esko.valkeila@tkk.fi

Pavel Y. Zryumov Department of Mechanics and Mathematics, Moscow State University, Leninskie gory 1, Moscow 119992, Russia, pavel.zryumov@gmail.com

On the Extension of the Namioka-Klee Theorem and on the Fatou Property for Risk Measures

Sara Biagini and Marco Frittelli

Abstract This paper has been motivated by general considerations on the topic of Risk Measures, which essentially are convex monotone maps defined on spaces of random variables, possibly with the so-called Fatou property.

We show first that the celebrated Namioka-Klee theorem for linear, positive functionals holds also for convex monotone maps π on Frechet lattices.

It is well-known among the specialists that the Fatou property for risk measures on L^∞ enables a simplified dual representation, via probability measures only. The Fatou property in a general framework of lattices is nothing but the lower order semicontinuity property for π. Our second goal is thus to prove that a similar simplified dual representation holds also for *order lower semicontinuous*, convex and monotone functionals π defined on more general spaces $\mathscr{X}$ (locally convex Frechet lattices). To this end, we identify a link between the topology and the order structure in $\mathscr{X}$—*the C-property*—that enables the simplified representation. One main application of these results leads to the study of convex risk measures defined on Orlicz spaces and of their dual representation.

Keywords Convex monotone map · Locally convex Frèchet lattice · Order (lower semi-)continuity · Fatou property · Dual representation

Mathematics Subject Classification (2000) 91B30 · 91B28 · 46A04 · 06B99 · 46N10

1 Introduction

The analysis in this paper was triggered by recent developments in the theory of Risk Measures in Mathematical Finance. Convex risk measures were independently introduced by [13] and [14] as generalization of the concept of a coherent risk measure developed in [3]. Consider a space of financial positions $\mathscr{X}$ (real-valued, measurable functions on a fixed measurable space $(\Omega, \mathscr{F})$) containing the constants.

S. Biagini (✉)
Dipartimento di Statistica e Mathematica Applicata all'Economia, Università di Pisa,
Via Ridolfi, 10, 56124 Pisa, Italy
e-mail: sara.biagini@ec.unipi.it

M. Frittelli
Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini, 50, 20122 Milano, Italy
e-mail: marco.frittelli@mat.unimi.it

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_1,

A convex risk measure on $\mathscr{X}$ is a map $\rho : \mathscr{X} \to (-\infty, +\infty]$ with the following properties:

1. $\rho(0) = 0$ (so ρ is proper, i.e. it does not coincide with $+\infty$)
2. monotonicity: if $X, Y \in \mathscr{X}$, $X \le Y$, then $\rho(X) \ge \rho(Y)$
3. convexity: if $\lambda \in [0, 1]$, then $\rho(\lambda X + (1-\lambda)Y) \le \lambda\rho(X) + (1-\lambda)\rho(Y)$ for any $X, Y \in \mathscr{X}$
4. cash additivity: if $m \in \mathbb{R}$ then $\rho(X + m) = \rho(X) - m$ for any $X \in \mathscr{X}$

   When ρ is also positive homogeneous, i.e.

5. $\rho(\lambda X) = \lambda\rho(X)$ for all $\lambda \ge 0$ the risk measure is called coherent.

If $\mathscr{X}$ is also topological space (as it is always the case in the applications), it is of course useful to have a result on the degree of smoothness of the risk measure ρ. Strangely enough, when this paper was first written to our knowledge there was yet no general result. This is exactly the message of the extended Namioka-Klee Theorem, stated below in Theorem 1. The (topological vector) space of positions $\mathscr{X}$ however must have some other properties, i.e. it must be a *Frechet lattice*.

Recall that a topological vector space $(\mathscr{X}, \tau)$ is a Frechet lattice if:

- its topology τ is induced by a complete distance d
- $\mathscr{X}$ is a lattice, that is it has an order structure $(\mathscr{X}, \le)$ and each pair $X_1, X_2 \in \mathscr{X}$ has a supremum $X_1 \vee X_2$ in $\mathscr{X}$
- $\mathscr{X}$ is locally solid, that is the origin 0 has a fundamental system of solid neighborhoods (a neighborhood U of 0 is solid if for any $X \in U$, $Y \in \mathscr{X}$, $|Y| \le |X| \Rightarrow Y \in U$ where $|X| = X \vee (-X)$).

Note that a Frechet lattice is not necessarily locally convex. Examples of common Frechet lattices are the spaces L^p on a probability space $(\Omega, \mathscr{F}, P)$, for $p \in [0, 1)$ (with the natural, a.s. pointwise order). When $p \ge 1$, L^p is also Frechet lattice, but with an extra property. The topology is induced by the L^p-norm (and thus the space becomes locally convex). Moreover, the norm has a monotonicity property: $|X| \le |Y| \Rightarrow \|X\|_p \le \|Y\|_p$. So L^p, $p \in [1, +\infty]$ is in fact a *Banach lattice*. Other Banach lattices important for our applications belong to the family of Orlicz spaces—denoted with L^{Ψ} for a Young function Ψ—which are described in details in Sect. 5.

Finally, we present the abstract statement of the extended Namioka-Klee Theorem, proved in Sect. 2 (where there is also an extensive comparison with the existing literature, as we discovered that there are a couple of recent, very similar results). The Theorem is stated for convex, monotone *increasing* maps π, not necessarily cash additive. But a similar result clearly holds for monotone decreasing maps. $\mathrm{Dom}(\pi)$ indicates here and in the rest of the paper the subset of $\mathscr{X}$ where π is finite. The topological dual space is denoted by $\mathscr{X}'$ and the set $\mathscr{X}'_+$ indicates the convex cone of those functionals Y in $\mathscr{X}'$ that are positive, i.e. $\langle Y, X\rangle \ge 0$ for all $X \in \mathscr{X}$, $X \ge 0$. The symbol $\langle\, ,\rangle$ indicates the bilinear form for the duality $(\mathscr{X}, \mathscr{X}')$. The map $\pi^* : \mathscr{X}' \to (-\infty, +\infty]$ is the convex conjugate of π, also known as Fenchel transform, and it is defined as

$$\pi^*(Y) = \sup_{X \in \mathscr{X}} \{\langle Y, X\rangle - \pi(X)\}.$$

Theorem 1 (Extended Namioka-Klee) *Any proper convex and monotone increasing functional* $\pi : \mathscr{X} \to (-\infty, +\infty]$ *on a Frechet lattice* $(\mathscr{X}, \tau)$ *is continuous and subdifferentiable on* $\text{int}(\text{Dom}(\pi))$ *(the interior of* $\text{Dom}(\pi)$*). Moreover, it admits a dual representation as*

$$\pi(X) = \max_{Y \in \mathscr{X}'_+} \{\langle Y, X\rangle - \pi^*(Y)\} \quad \forall X \in \text{int}(\text{Dom}(\pi)). \tag{1}$$

To give an idea about the genesis of the second and most innovative part of the paper, let us go back to the financial setup and let us focus first on the case L^∞. A risk measure ρ on L^∞ has the pleasant property of being always finite-valued, thanks to the boundedness of its elements and to the monotonicity and cash additivity property. The theorem above ensures that ρ is continuous and subdifferentiable on the entire L^∞. This implies the existence of a well-known dual representation for ρ over L^∞, namely

$$\rho(X) = \max_{Q \in \mathscr{M}_{1,f}(P)} \{E_Q[-X] - \rho^*(Q)\} \tag{2}$$

where:

(a) $\mathscr{M}_{1,f}(P)$ indicates the set of positive, finitely additive measures Q on $(\Omega, \mathscr{F})$ that are absolutely continuous w.r.t. to P and are normalized $(Q(I_\Omega) = 1)$;
(b) ρ^* is the convex conjugate of ρ and should be interpreted as a penalty functional.

These results are known and proved e.g. [12, Theorem 4.12]), where subdifferentiability is proved by hand for the specific case study L^∞.

Now, let us recall the definition of the Fatou property for risk measures (see e.g. [10] or Sect. 4, [12], in the case L^∞):

Definition 1 (Fatou property) A risk measure $\rho : L^p \to \mathbb{R} \cup \{\infty\}$, $p \in [1, \infty]$, has the Fatou property (F.P.) if given any sequence $\{X_n\}_n$ *dominated* in L^p and converging P-a.s. to X we have:

$$\rho(X) \le \liminf_{n\to\infty} \rho(X_n).$$

This property enables a simplified dual representation of ρ. Instead of the finitely additive measures $\mathscr{M}_{1,f}(P)$, if a convex risk measure $\rho : L^\infty \to \mathbb{R}$ has the F.P. one can write

$$\rho(X) = \sup_{Q \text{ probability } Q \ll P} \{E_Q[-X] - \rho^*(Q)\}, \tag{3}$$

so the supremum can be taken only over *probabilities*, the σ*-additive* elements of $\mathscr{M}_{1,f}$, see [12, Theorem 4.26]. There is a price to pay: the above supremum may not be attained over probability measures, but only on $\mathscr{M}_{1,f}$.

What can be said about the representation problem of a convex risk measures defined on subspaces of L^0?

The spaces L^p are typical examples of spaces of financial positions in the applications (see, for example, [11, 14, 19]). Moreover, in [6] it is shown that Orlicz spaces that can be associated with a utility function are the right framework for the utility maximization problems which commonly arises in financial problems. Motivated by this idea, in the first version of this paper,[1] we initiated the study of risk measures defined on Orlicz spaces L^Ψ and more generally on Frechet lattices. Independently, in [9] convex risk measures were defined on the Morse subspace M^Ψ of the Orlicz space L^Ψ. As we shall see in Sect. 5 some of the findings in [9] are special case our results, while other properties do not hold in the general case, essentially because the topology on the (whole) Orlicz space L^Ψ is not order continuous.

Our generalization of the representation in (3) and its various implications will be stated for maps π, defined of general Frechet lattice, that are convex, monotone and *increasing*, not necessarily translation invariant. This latter implies that the set of dual variables over which the supremum—or the maximum—is taken will not be normalized in general.

To begin with, let us recall a few notions about Riesz spaces, i.e. linear spaces that are lattices (see also Sect. 3). The first is that of order convergence. A generalized sequence, or net, $(X_\alpha)_\alpha$ in a Riesz space $\mathscr{R}$ is *order convergent* to some $X \in \mathscr{R}$, notation $X_\alpha \overset{o}{\to} X$, if there is a net $(Z_\alpha)_\alpha$ in $\mathscr{R}$ satisfying

$$Z_\alpha \downarrow 0 \quad \text{and} \quad |X_\alpha - X| \le Z_\alpha \quad \text{for each } \alpha \tag{4}$$

($Z_\alpha \downarrow 0$ means that $(Z_\alpha)_\alpha$ is monotone decreasing and its infimum is 0).

A functional $f : \mathscr{R} \to \mathbb{R}$ defined on $\mathscr{R}$ is *order continuous* if

$$X_\alpha \overset{o}{\to} X \Rightarrow f(X_\alpha) \to f(X),$$

and a *topology* τ on $\mathscr{R}$ *is order continuous if*

$$X_\alpha \overset{o}{\to} 0 \Rightarrow X_\alpha \overset{\tau}{\to} 0.$$

These definitions readily imply that if the topology τ is order continuous then

$$f \text{ is (topologically) continuous} \Rightarrow f \text{ is order continuous.}$$

We denote with $(\mathscr{X}_n^\sim)$ the cone of *order continuous linear functionals* on $\mathscr{X}$. By the classic Namioka Theorem, see Sect. 2, $(\mathscr{X}_n^\sim)_+$coincides with $\mathscr{X}'_+$, the positive elements of topological dual space $\mathscr{X}'$. We recall that in a Banach lattice $\mathscr{X}$ the norm topology is order continuous if and only if $\mathscr{X}_n^\sim = \mathscr{X}'$ and that the following three classes of spaces all have order continuous norm: (a) L^p when $p \in [1, +\infty)$; (b) M^Ψ when Ψ is a finite valued Young function; (c) L^Ψ when Ψ is a Young function satisfying the Δ_2 condition (in this case $L^\Psi = M^\Psi$).

The above implication, together with Theorem 1, readily imply the following

[1] Presented by the second author at the Workshop on Risk Measures, University of Evry, France, 6–7 July 2006 and at the Conference on Risk Measures and Robust Control in Finance, The Bendheim Center, Princeton University, 6–7 October, 2006.

Corollary 1 *If Frechet lattice* $(\mathscr{X}, \tau)$ *has an order continuous topology and* $\pi : \mathscr{X} \to \mathbb{R}$ *is convex and monotone* (*increasing*), π *is already* order continuous *on* $\mathscr{X}$. *Thus it admits a dual representation as*

$$\pi(X) = \max_{Y \in (\mathscr{X}_n^{\sim})_+} \{\langle Y, X \rangle - \pi^*(Y)\}, \quad X \in \mathscr{X}, \tag{5}$$

where $(\mathscr{X}_n^{\sim})_+ = \mathscr{X}'_+$.

(a) *In the specific case* $\mathscr{X} = L^p$, $p \in [1, +\infty)$ *the representation above becomes*

$$\pi(X) = \max_{Y \in (L^q)_+} \{E[YX] - \pi^*(Y)\}, \quad X \in L^p,$$

(b) *In the specific case* $\mathscr{X} = M^\Psi$ *and* Ψ *is a finite valued Young function, the representation above becomes*

$$\pi(X) = \max_{Y \in (L^{\Psi^*})_+} \{E[YX] - \pi^*(Y)\}, \quad X \in M^\Psi, \tag{6}$$

where Ψ^* *is the conjugate function of* Ψ.

However, the order continuity of a topology is a strong assumption, which is not satisfied by e.g. L^∞, or by L^Ψ for general Young functions Ψ or by other Frechet lattices, as shown in Sect. 5. Moreover, in general $(\mathscr{X}_n^{\sim})_+$ is only a subspace of $\mathscr{X}'_+$. This is exactly what happens with L^∞:

$$((L^\infty)_n^{\sim})_+ = L^1_+ \quad \text{and} \quad (L^\infty)'_+ = L^1_+ \oplus S_+,$$

where S are the purely finitely additive measures.

It is then natural to investigate whether π admits a representation on $(\mathscr{X}_n^{\sim})_+$ under conditions, linking topology and order structure, less restrictive than the order continuity of the topology τ.

As we shall see in Remark 3, when $\mathscr{X} = L^p$, $p \in [0, \infty]$, the Fatou property coincides with order lower semicontinuity, which is the appropriate concept in the present general setting.

Definition 2 A functional $\pi : \mathscr{R} \to (-\infty, +\infty]$ defined on a Riesz space $\mathscr{R}$ is *order lower semicontinuous* if $X_\alpha \xrightarrow{o} X$ implies $\pi(X) \leq \liminf \pi(X_\alpha)$.

From now on, local convexity is needed and in what follows the Frechet lattice $\mathscr{X}$ *is also supposed locally convex.*

As a consequence of the Hahn-Banach theorem in *any* locally convex Frechet lattice if the proper, increasing convex map $\pi : \mathscr{X} \to (-\infty, +\infty]$ is also $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$ lower semicontinuous then

$$\pi(X) = \sup_{Y \in (\mathscr{X}_n^{\sim})_+} \{\langle Y, X \rangle - \pi^*(Y)\}, \quad X \in \mathscr{X}, \tag{7}$$

where π^* is the convex conjugate of π. Therefore π, as the pointwise supremum of a family of order continuous functionals, is also *order lower semicontinuous*.

One could then conjecture that the converse always holds true i.e. any *order lower semicontinuous* π on *any* locally convex Frechet lattice admits a representation in terms of $(\mathscr{X}_n^\sim)_+$, as in (7) or, in lucky cases, (5).

The conjecture is not true in general, see Example 1 at the end of this Introduction. Proposition 1 contains the main result of the paper, that is there exists an *additional* assumption "à la Komloś", linking the topology τ and the order structure, that enables the representation over $(\mathscr{X}_n^\sim)_+$.

Definition 3 (C-property) A linear topology τ on a Riesz space has the C-property[2] if $X_\alpha \xrightarrow{\tau} X$ implies the existence of a subsequence $(X_{\alpha_n})_n$ and convex combinations $Z_n \in \mathrm{conv}(X_{\alpha_n}, \dots)$ such that $Z_n \xrightarrow{o} X$.

This property is quite reasonable, all the details are in Sect. 4. In particular, whenever a locally convex Frechet lattice $(\mathscr{X}, \tau)$ can be embedded in L^1 with a linear lattice embedding, then all the topologies: τ, $\sigma(\mathscr{X}, \mathscr{X}')$ and $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$ have the C-property. A relevant example of spaces with an associated collection of topologies (norm, weak and $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$) satisfying the C-property is the family of Orlicz spaces (Sect. 5.1).

In the case $\mathscr{X} = L^\infty$, it is well known that the sup in (7) in general is not a max, but the sup is attained under some stronger continuity condition. In the general case, in Lemma 7 we show that for a finite valued convex increasing map which is order upper semicontinuous the sup in (7) is indeed a max.

Finally, in Sect. 5.2 we analyze convex risk measures defined on Orlicz spaces and with values in $\mathbb{R} \cup \{+\infty\}$. This new setup allows for an extension of the known dual representation on L^∞. We further provide some new results on the convex risk measures associated to utility functions, as in the case of the entropic risk measure.

Example 1 (When the C-property fails) When the $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$-topology does not satisfy the C-property, there may be order l.s.c. convex functionals (even τ-continuous!) that are not $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$-l.s.c.

Take $\mathscr{X} = C([0, 1])$, the Banach lattice of the continuous functions on $[0, 1]$ with the supremum norm and the pointwise order. The dual $\mathscr{X}'$ consists of the Borel signed measures on $[0, 1]$ and it is known (see e.g. [21, Example 87.5]) that there is no non zero order-continuous functional in $\mathscr{X}'$. The topology $\sigma(\mathscr{X}, \mathscr{X}_n^\sim) = \sigma(\mathscr{X}, \{0\})$ is therefore the indiscrete one and clearly it doesn't have the C-property.

Consider then the convex, increasing 'best case' functional

$$\pi(X) = \max_{t \in [0,1]} X(t)$$

[2]The "C" stands for "convex combinations"...

which is finite valued, so that the extended Namioka-Klee Theorem implies that π is norm continuous and subdifferentiable and it admits a representation on $\mathscr{X}'_+$ as

$$\pi(X) = \max_{Y \in \mathscr{X}'_+} \{\langle Y, X\rangle - \pi^*(Y)\}.$$

But it evidently does not admit a representation on $(\mathscr{X}_n^\sim)_+ = \{0\}$ because it is not constant.

To show that π is order-l.s.c. let $X_\alpha \xrightarrow{o} X$ and suppose by contradiction that there exists a subnet $(X_{\alpha_\beta})_\beta$ (as a subnet, still order convergent to X) such that $\pi(X) > \lim_\beta \pi(X_{\alpha_\beta})$. Let $t^* \in \operatorname{argmax}(X)$. Then

$$\pi(X_{\alpha_\beta}) \geq X_{\alpha_\beta}(t^*)$$

and evidently

$$X(t^*) = \pi(X) > \lim_\beta \pi(X_{\alpha_\beta}) \geq \limsup_\beta X_{\alpha_\beta}(t^*),$$

which contradicts pointwise convergence.

2 The Extended Namioka Theorem

The following is the statement of the well-known Namioka-Klee Theorem in the case of linear functionals φ.

Theorem 2 (Namioka-Klee) *Any linear and positive functional $\varphi : \mathscr{X} \to \mathbb{R}$ on a Frechet lattice $\mathscr{X}$ is continuous* (*see* [17]).

In order to provide a technically straightforward, but quite relevant, extension of Namioka-Klee Theorem to *convex* functionals π, the *positivity* assumption

$$0 \leq Y \Rightarrow 0 \leq \varphi(Y)$$

has to be replaced with the requirement that π is monotone *increasing*

$$X \leq Y \Rightarrow \pi(X) \leq \pi(Y).$$

Monotonicity and positivity are equivalent for linear functionals, but it is straightforward to see that for convex π monotonicity implies positivity (assuming $\pi(0) = 0$). And it is easy to produce a positive but non-monotone convex map by taking $\pi(X) = |X|$ on $\mathscr{X} = \mathbb{R}$.

So, while on one hand one relaxes the linearity assumption, on the other hand a stronger link with the order structure is required.

The properties in the next lemma are straightforward consequences of the definitions.

Lemma 1 *Let $\mathscr{R}$ be a Riesz space and let $\pi : \mathscr{R} \to (-\infty, +\infty]$ be convex, increasing and $\pi(0) = 0$. Then*:

(i) $\pi(\alpha X) \leq \alpha\pi(X)$, $\forall \alpha \in [0, 1]$, $\forall X \in \mathscr{R}$;
(ii) $\alpha\pi(X) \leq \pi(\alpha X)$, $\forall \alpha \in (-\infty, 0] \cup [1 + \infty)$, $\forall X \in \mathscr{R}$;
(iii) $|\pi(X)| \leq \pi(|X|)$, $\forall X \in \mathscr{R}$.

Proof of Theorem 1 Step 1: Continuity. The proof of Namioka-Klee Theorem (see e.g. [1, Theorem 9.6]) can be adapted, in a straightforward manner, to deal with the current weaker assumptions. We repeat the argument so that the paper is self-contained.

W.l.o.g. it can be assumed that the interior of Dom(π), int(Dom(π)), is not empty that $0 \in$ int(Dom(π)) and $\pi(0) = 0$. Let B_r be the centered open ball of radius $r > 0$ in a metric that generates τ. Take any sequence $\{X_n\}_n$ such that $X_n \xrightarrow{\tau} 0$. Fix r small enough, so that $B_{2r} \subseteq$ int(Dom(π)). Then pick a countable base $\{V_n\}_n$ of solid neighborhoods of zero satisfying $V_1 + V_1 \subseteq B_r$ and $V_{n+1} + V_{n+1} \subset V_n$ for each n. Then $V_{n+1} \subseteq V_n \subseteq B_r$ for each n. By passing to a subsequence of X_n, one can suppose $X_n \in \frac{1}{n}V_n$ for each n. Set $Y_n = \sum_{i=1}^n i|X_i|$ and note that $Y_n \leq Y_{n+1}$ and $n|X_n| \leq Y_n$. In addition

$$Y_{n+p} - Y_n = \sum_{i=n+1}^{n+p} i|X_i| \in V_{n+1} + V_{n+2} + \cdots + V_{n+p} \subset V_n.$$

Therefore $Y_n \in B_r$ for each n and $\{Y_n\}_n$ is a Cauchy sequence, so $Y_n \xrightarrow{\tau} Y$ for some Y in $\mathscr{X}$. Since $Y_n \in B_r$, $Y \in \overline{B_r} \subset B_{2r} \subseteq$ int(Dom(π)), $\pi(Y)$ is finite. This Y is an upper bound for the sequence (actually, $\sup_n Y_n = Y$). In fact, fix any n. Since $Y_m - Y_n \in \mathscr{X}_+$ for each $m \geq n$, the sequence $\{Y_m - Y_n : m \geq n\}$ in $\mathscr{X}_+$ satisfies

$$Y_m - Y_n \xrightarrow{\tau} Y - Y_n, \quad \text{as } m \to \infty.$$

Since $\mathscr{X}_+$ is τ-closed [1, Theorem 8.43-1], $Y - Y_n \in \mathscr{X}_+$ for each n. Hence $Y_n \leq Y$ for each n. From Lemma 1, $|\pi(X_n)| \leq \pi(|X_n|) \leq \frac{1}{n}\pi(n|X_n|)$. By monotonicity of π we derive

$$|\pi(X_n)| \leq \frac{1}{n}\pi(n|X_n|) \leq \frac{1}{n}\pi(Y_n) \leq \frac{1}{n}\pi(Y) \to 0$$

which shows that π is continuous at zero and therefore π is continuous on the whole int(Dom(π)) [1, Theorem 5.43].

Step 2: Subdifferentiability. For all $X^* \in$ int(Dom(π)) we must exhibit a subgradient $Y^* \in \mathscr{X}'$, i.e. a Y^* such that

$$\pi(X) - \pi(X^*) \geq \langle Y^*, X - X^* \rangle \quad \text{for all } X \in \mathscr{X}. \tag{8}$$

To this end, again w.l.o.g. we can suppose $X^* = 0 \in$ int(Dom(π)) and $\pi(0) = 0$. Then, the directional derivative functional D in 0

$$D(X) := \lim_{t \downarrow 0} \frac{\pi(tX)}{t}$$

satisfies $D \leq \pi$ thanks to Lemma 1. It is finite valued and convex and thus the first part of this proof implies that it is continuous. By the Hahn-Banach Theorem (see e.g. [1, Theorem 5.53]) there exists a linear functional Y^* which satisfies $\langle Y^*, X\rangle \leq D(X)$ on $\mathscr{X}$ whence Y^* is a continuous subgradient for π at 0.

Step 3: Representation. Fix any $X^* \in \text{int}(\text{Dom}(\pi))$. It is an exercise to show that π increasing implies π^* is finite at most over $\mathscr{X}'_+$. Fix any subgradient Y^* (which is then positive) of π at X^*. Reshuffling (8), this means

$$\langle Y^*, X^*\rangle - \pi(X^*) = \max_{X\in\mathscr{X}}\{\langle Y^*, X\rangle - \pi(X)\} = \pi^*(Y^*)$$

where the last equality follows from the definition of π^*. This chain of equalities in turn implies that $\pi(X^*) = \langle Y^*, X^*\rangle - \pi^*(Y^*) = \max_{Y\in\mathscr{X}'}\{\langle Y, X^*\rangle - \pi^*(Y)\}$ as the inequality $\pi(X^*) \geq \langle Y, X^*\rangle - \pi^*(Y)$ automatically holds for any $Y \in \mathscr{X}'$. □

Remark 1 In [9] there is a formula identical to (1) for π defined on Banach lattices.

Corollary 2 *Every finite-valued convex and monotone functional on a Banach lattice is norm-continuous and subdifferentiable.*

Corollary 3 *If a Frechet lattice $\mathscr{X}$ supports a non-constant convex monotone map π, then necessarily $\mathscr{X}' \neq \{0\}$.*

As a generic Frechet lattice $\mathscr{X}$ is not necessarily locally convex, it may happen that the topological dual $\mathscr{X}'$ is very poor or even $\{0\}$. This is the case, for example, of the spaces $L^p(\Omega, \mathscr{F}, \mu)$, $p \in (0, 1)$, when μ is a nonatomic measure (see [1, Theorem 13.31]) and of the space $L^0(\Omega, \mathscr{F}, \mu)$, when μ is a nonatomic finite measure [1, Theorem 13.41]. Therefore, the only convex monotone πs on these spaces are the constants.

2.1 The Current Literature

Surprisingly enough given their importance in the applications, it seems that results on continuity and subdifferentiability for convex monotone maps have appeared only very recently in the literature.

After finishing the first version of the paper, which did not contain the subdifferentiability additional result, we came to know that in the recent articles [16] and [20] there are statements very close to those of Theorem 1.

To start, in [16] it is shown that:

> *if L is an ordered Banach space, with L_+ closed and such that $L = L_+ - L_+$ then any convex monotone $\pi : U \to \mathbb{R}$ defined on an open set U of L is continuous.*

These hypotheses are stronger than ours on the topological part as L must be a Banach space, but milder on the order part. In fact, their conditions: the positive cone L_+ is closed and generating $L = L_+ - L_+$ are always satisfied in a Frechet lattice. Note that nothing is said about subdifferentiability.

On the contrary, in [20] the authors were the first to prove subdifferentiability of convex monotone maps π, but with the stronger assumption that π is defined on a *Banach* lattice L:

If L is a Banach lattice, $\pi : L \to \overline{\mathbb{R}}$ is proper, convex and monotone, then it is continuous and subdifferentiable on the interior of the proper domain.

The line of their proof is the following. For any fixed $X^* \in \mathrm{int}(\mathrm{Dom}(\pi))$, first one exhibits a positive subgradient, which is then continuous by classic Namioka-Klee theorem. This implies lower semicontinuity of π at X^*, which in turn implies continuity.

Inspired by this work we also prove subdifferentiability of π, in the case of Frechet lattices. However, we reverse the order, since first one proves continuity of π on $\mathrm{int}(\mathrm{Dom}(\pi))$ and then subdifferentiability (this latter in the same way as done by [20]). This is only a matter of taste and it would not be difficult to extend the results in [20] with the same line of reasoning from Banach to Frechet lattices. The only interesting aspect in proving first continuity is that one realizes that the same proof of "classic" Namioka for linear positive functionals still holds, basically unchanged, for convex monotone maps.

The interested reader is also referred to [9] for further developments.

3 On Order Lower Semicontinuity in Riesz Spaces

Let us recall some basic facts about Riesz spaces. The same notation $\leq$ is used for the order relations in $\mathscr{R}$, in $(-\infty, +\infty]$ and for the direction of index sets of nets, as the meaning will be clear from the context.

A subset A of a Riesz space $\mathscr{R}$ is *order bounded* if there exists $X_1 \in \mathscr{R}$ and $X_2 \in \mathscr{R}$ such that $X_2 \leq X \leq X_1$ for all $X \in A$. A net $(X_\alpha)_\alpha$ in $\mathscr{R}$ is *increasing*, written $X_\alpha\uparrow$, if $\alpha \leq \beta$ implies $X_\alpha \leq X_\beta$. A net $(X_\alpha)_\alpha$ in $\mathscr{R}$ is *increasing* to some $X \in \mathscr{R}$, written $X_\alpha \uparrow X$, if $X_\alpha\uparrow$ and $\sup_\alpha X_\alpha = X$. A subset A of $\mathscr{R}$ is *order closed* if $X_\alpha \in A$ and $X_\alpha \overset{o}{\to} X$ implies $X \in A$. The space $\mathscr{R}$ is *order complete* when each order bounded subset A has a supremum (least upper bound) and an infimum (largest lower bound).

Recall [1, Theorem 8.15] that the lattice operations are order continuous. In addition [1, Theorem 8.16], if a net $(X_\alpha)_\alpha$ is order bounded and $\mathscr{R}$ is order complete, then $\liminf_\alpha X_\alpha \triangleq \sup_\alpha \inf_{\beta\geq\alpha} X_\beta$ and $\limsup_\alpha X_\alpha \triangleq \inf_\alpha \sup_{\beta\geq\alpha} X_\beta$ are well defined, and

$$X_\alpha \overset{o}{\to} X \quad \text{iff} \quad X = \liminf_\alpha X_\alpha = \limsup_\alpha X_\alpha.$$

The next lemma is an immediate consequence of the facts and definitions above and of (4).

Lemma 2 *Let $\mathscr{R}$ be a Riesz space.*

(i) *Let $X_\alpha \xrightarrow{o} X$. Then there exists α^* such that $(X_\alpha)_{\alpha \geq \alpha^*}$ is order bounded, i.e. the net is definitely order bounded. In case the index set of the net has a minimum then $(X_\alpha)_\alpha$ is order bounded.*

(ii) *Let $\mathscr{R}$ be order-complete and let $X_\alpha \xrightarrow{o} X$. If $Y_\alpha \triangleq (\inf_{\beta \geq \alpha} X_\beta) \wedge X$, then $Y_\alpha \uparrow X$.*

Example 2 (Order convergence in L^p) In L^p spaces, $p \in [0, \infty]$, the notion of order is the very familiar pointwise one, i.e. $Y \geq X$ iff $Y(\omega) \geq X(\omega)$ P-a.e. As L^p is order separable, see the next section, sequences can be used instead of nets to characterize order convergence. A sequence $(X_n)_n$ in L^p is order bounded iff it is dominated in L^p (i.e. there exists a $Y \in L^p_+$ such that $|X_n| \leq Y$). The order convergence in the L^p case is just dominated pointwise convergence:

$$X_n \xrightarrow{o} X \Leftrightarrow X_n \xrightarrow{P\text{-a.e.}} X \text{ and } (X_n)_n \text{ is } \textit{dominated} \text{ in } L^p. \tag{9}$$

Therefore, the L^p-norm topologies are order continuous for all $p < +\infty$, as the above equivalence implies that Lebesgue dominated convergence theorem can be applied to conclude $X_n \xrightarrow{o} X \Rightarrow X_n \xrightarrow{L^p} X$.

3.1 Equivalent Formulations of Order l.s.c.

Definition 4 A functional $\pi : \mathscr{R} \to (-\infty, +\infty]$ defined on a Riesz space $\mathscr{R}$

(a) is *continuous from below* if $X_\alpha \uparrow X \Rightarrow \pi(X_\alpha) \uparrow \pi(X)$

(a_σ) is *σ-continuous from below* if $X_n \uparrow X \Rightarrow \pi(X_n) \uparrow \pi(X)$

(b) is *order lower semicontinuous* if $X_\alpha \xrightarrow{o} X \Rightarrow \pi(X) \leq \liminf \pi(X_\alpha)$

(b_σ) is *σ-order lower semicontinuous* if

$$X_n \xrightarrow{o} X \Rightarrow \pi(X) \leq \liminf \pi(X_n). \tag{10}$$

Note that the pointwise supremum of a family of order l.s.c. functionals is order l.s.c.

As shown in the next lemma, if π is *increasing* and $\mathscr{R}$ possesses more structure then the conditions (a), (a_σ), (b), (b_σ) are all equivalent. The *order separability of $\mathscr{R}$* (any subset A which admits a supremum in $\mathscr{R}$ contains a countable subset with the same supremum) allows to formulate the order-l.s.c. property with sequences instead of nets (i.e. $(b) \Leftrightarrow (b_\sigma)$).

Lemma 3 *Let $\mathscr{R}$ be an order complete Riesz space and $\pi : \mathscr{R} \to (-\infty, +\infty]$ be increasing. Then*: $(a) \Leftrightarrow (b)$, $(a_\sigma) \Leftrightarrow (b_\sigma)$, $(a) \Rightarrow (a_\sigma)$, $(b) \Rightarrow (b_\sigma)$.

If in addition $\mathscr{R}$ is order separable then (a), (a_σ), (b), (b_σ) are all equivalent.

Proof $(a) \Rightarrow (b)$. Let $X_\alpha \xrightarrow{o} X$ and set $Y_\alpha = (\inf_{\beta \geq \alpha} X_\beta) \wedge X$. By Lemma 2(ii), $Y_\alpha \uparrow X$ and so $\pi(X) \overset{(a)}{=} \lim \pi(Y_\alpha) \overset{(mon)}{\leq} \liminf \pi(X_\alpha)$.

$(b) \Rightarrow (a)$. Since $X_\alpha \uparrow X$ implies $X_\alpha \xrightarrow{o} X$, we get: $\pi(X) \overset{(b)}{\leq} \lim \pi(X_\alpha) \overset{(mon)}{\leq} \pi(X)$.

$(a_\sigma) \Leftrightarrow (b_\sigma)$ follows in the same way as $(a) \Leftrightarrow (b)$, while $(a) \Rightarrow (a_\sigma)$ and $(b) \Rightarrow (b_\sigma)$ are obvious.

To show the last sentence it is sufficient to prove e.g. $(a_\sigma) \Rightarrow (a)$. For any net $X_\alpha \uparrow X$ we can find a countable subnet X_{α_n} such that $X_{\alpha_n} \uparrow X$. Hence

$$\pi(X) \overset{(a_\sigma)}{=} \lim_n \pi(X_{\alpha_n}) \leq \lim \pi(X_\alpha) \overset{(mon)}{\leq} \pi(X). \qquad \square$$

Remark 2 (On order separability) A sufficient condition for $\mathscr{R}$ to be order separable is that, for every principal ideal $\mathscr{R}_X$, there exists a positive linear functional on $\mathscr{R}$ which is strictly positive on $\mathscr{R}_X$ (see [21, Theorem 84.4]).

All Banach lattices with order continuous norm verify this condition, as shown in [2, Theorem 12.14].

Another sufficient condition for order separability is the existence of a linear functional on $\mathscr{R}$ which is strictly positive on the entire $\mathscr{R}$. This implies that all the Orlicz Banach lattices $L^\Psi = L^\Psi(\Omega, \mathscr{F}, P)$ (and henceforth all the L^p spaces, $p \in [1, \infty]$) are order separable (and order complete as well). See Sect. 5.1.

Remark 3 (On the Fatou Property) From (9), Definition 1 and Definition 4, we immediate see that when $\mathscr{R} = L^p$, $p \in [0, \infty]$, order lower semicontinuity coincides with the Fatou Property.

Remark 4 (On decreasing functional) Analogous considerations hold for decreasing functionals: if $\mathscr{R}$ is an order complete and order separable Riesz space and if ρ is *decreasing*, then the conditions: (b), (b_σ), continuity from *above* [i.e.: $X_\alpha \downarrow X \Rightarrow \rho(X_\alpha) \uparrow \rho(X)$] and σ-continuity from *above* [i.e.: $X_n \downarrow X \Rightarrow \rho(X_n) \uparrow \rho(X)$]are all equivalent. These equivalent formulations will be used to study some properties of convex risk measures in Sect. 5.2.

3.2 The Order Continuous Dual $\mathscr{X}_n^\sim$

Given a Frechet lattice $\mathscr{X}$, the space of order bounded linear functionals $\mathscr{X}^\sim$ (those which carry order bounded subset of $\mathscr{X}$ to order bounded sets of $\mathbb{R}$) coincides with the topological dual $\mathscr{X}'$. This is a consequence of Namioka-Klee Theorem 2. From the general theory (see [21]) on the decomposition of $\mathscr{X}^\sim$

$$\mathscr{X}' = \mathscr{X}^\sim = \mathscr{X}_n^\sim \oplus \mathscr{X}_s^\sim$$

where $\mathscr{X}_n^\sim$ is the order closed ideal (*band*) of $\mathscr{X}^\sim$ of all the order continuous linear functionals on $\mathscr{X}$ and it is called the *order continuous dual* of $\mathscr{X}$. The space of *singular* functionals $\mathscr{X}_s^{\sim}$ is defined as the band disjoint complement of $\mathscr{X}_n^{\sim}$ in $\mathscr{X}^{\sim}$. Examples of this decomposition are given in Sect. 5. The main goal of the next section is to provide some criteria that guarantee the C-property of the topology $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$.

4 On the C-Property

The C-property is verified by the strong topology of all Frechet lattices without passing to convex combinations, as shown below.

Lemma 4 *Let $(\mathscr{X}, \tau)$ be a Frechet lattice. If $(X_n)_n$ τ-converges to X, then there exists a subsequence which is order convergent.*

Proof Call d a complete distance that induces τ, which is also absolute, i.e. $d(X, 0) = d(|X|, 0)$. Suppose $d(X_n, X) \to 0$ and select a subsequence such that $\sum_{k\geq 0} d(X_{n_k}, X) = \sum_{k\geq 0} d(X_{n_k} - X, 0) < +\infty$.

Set $Y = \sum_{k\geq 0} |X_{n_k} - X|$. By completeness of d, $Y \in \mathscr{X}$. Now, if $Y_k := \sum_{h\geq k} |X_{n_h} - X|$ then clearly $Y_k\downarrow$ and $Y_k \overset{\tau}{\to} 0$ so by [1, Theorem 8.43] $Y_k \downarrow 0$. As

$$|X_{n_k} - X| \leq Y_k$$

one deduces that X_{n_k} order converges to X. □

We will be mainly concerned with the C-property of weak topologies in locally convex Frechet lattice. This is the reason why the most to hope for is to extract an order convergent subsequence of *convex combinations* from a topologically convergent net, e.g. exactly the C-property.

Lemma 5 *Let $(\mathscr{X}, \tau)$ be a locally convex Frechet lattice. Then the $\sigma(\mathscr{X}, \mathscr{X}')$ topology verifies the C-property.*

Proof Let $W_\alpha \to W$ in the weak topology. By Hahn-Banach Theorem, W belongs to the τ-closure of $\text{conv}(W_\alpha, \ldots)$ for all α and as the topology τ is first countable there exists a subsequence $(\alpha_n)_n$ and a sequence $Y_n \in \text{conv}(W_{\alpha_n}, \ldots)$ which converges to W in the τ topology. Lemma 4 ensures that we can extract a subsequence $(Y_{n_k})_k$ that order converges to W. □

Remark 5 The local convexity assumption cannot be dropped in the statement of the previous lemma. An immediate counterexample is given by the Frechet lattice L^0, since when P has no atoms $(L^0)' = \{0\}$. So, the weak topology $\sigma(L^0, (L^0)')$ is the indiscrete one and doesn't satisfy the C-property.

However, even under the local convexity assumption, the C-property is not preserved if one keeps weakening the topology, from $\sigma(\mathscr{X}, \mathscr{X}')$ to $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$. An extreme situation is the one already encountered in Example 1 where the Banach lattice $\mathscr{X} = C([0,1])$ has dual $\mathscr{X}'$ consisting of the signed measures μ of finite variation on [0, 1], but no μ is order continuous apart from the null measure.

So the following lemma may be helpful.

Lemma 6 *Let $(\mathscr{L}, \tau_{\mathscr{L}}), (\mathscr{X}, \tau)$ be locally convex Frechet lattices and suppose there exists a linear, injective lattice morphism*

$$(\mathscr{X}, \tau) \stackrel{i}{\rightarrow} (\mathscr{L}, \tau_{\mathscr{L}})$$

such that

$$\{Y \circ i \mid Y \in \mathscr{L}'\} \subseteq \mathscr{X}_n^{\sim}. \tag{11}$$

Then $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$ verifies the C-property.

Proof We specify that by "linear, injective lattice morphism" i we mean that i is linear, injective, topologically continuous and preserves the lattice structure. Note that $\mathscr{X}$ needs not to be homeomorphic to $i(\mathscr{X})$.

Let $(X_\alpha)_\alpha$ be a net such that $X_\alpha \stackrel{\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})}{\longrightarrow} X$. The condition (11) implies that $W_\alpha := i(X_\alpha)$ converges to $W := i(X)$ in the $\sigma(\mathscr{L}, \mathscr{L}')$-topology. Applying the same argument and using the same notations of the proof of Lemma 5, there exists $(Y_{n_k})_k$ converging in order to W in $\mathscr{L}$ and so the inverse image $Z_k = i^{-1}(Y_{n_k})$ verifies $Z_k \in \text{conv}(X_{\alpha_{n_k}}, \ldots)$ and $Z_k \stackrel{o}{\rightarrow} X$. □

Condition (11) is evidently satisfied in case $\mathscr{L}' = \mathscr{L}_n^{\sim}$, which is equivalent to the assumption that $\tau_{\mathscr{L}}$ is order continuous. If this holds, essentially the above lemma applies to any locally convex Frechet lattice $\mathscr{X}$ that can be identified with a sublattice of $\mathscr{L}$, so the order structure is identical of that of $\mathscr{L}$, but with possibly finer topology than the one inherited from $\mathscr{L}$. This is the content of the next Corollary, that will be applied for the Orlicz Banach lattice L^Ψ.

Corollary 4 *Any locally convex Frechet lattice $\mathscr{X}$ of random variables that can be injected into L^1 by a linear lattice morphism has τ, $\sigma(\mathscr{X}, \mathscr{X}')$ and $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$ topologies with the C-property.*

4.1 The C-Property in the Representation of Convex and Monotone Functionals

We present the result on the equivalence between the $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$-l.s.c. property for convex functionals on locally convex Frechet lattices and the order-l.s.c. property, under the assumption that the topology $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$ has the C-property.

Proposition 1 *Let $(\mathscr{X}, \tau)$ be a locally convex Frechet lattice and consider the following conditions for a proper, convex functional $\pi : \mathscr{X} \to (-\infty, +\infty]$:*

1. *π is $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$-l.s.c.*
2. *π admits the representation*

$$\pi(X) = \sup_{Y \in \mathscr{X}_n^\sim} \{\langle Y, X\rangle - \pi^*(Y)\}, \quad X \in \mathscr{X}, \tag{12}$$

3. *π is order l.s.c.*

Then (1) $\Rightarrow$ (2) $\Rightarrow$ (3). *If $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$ has the C-property, the three conditions are equivalent.*

If π is in addition monotone increasing, the conclusions are identical and in the representation (12) *$\mathscr{X}_n^\sim$ can be replaced by $(\mathscr{X}_n^\sim)_+$.*

Proof (1) $\Rightarrow$ (2) follows from $(\mathscr{X}, \sigma(\mathscr{X}, \mathscr{X}_n^\sim))' = \mathscr{X}_n^\sim$ and from Fenchel-Moreau Theorem (see e.g. [8, Chap. I]); (2) $\Rightarrow$ (3) Since π is the pointwise supremum of a family of order continuous functionals, it is also order l.s.c. Suppose now that $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$ has the C-property and that (3) holds. To prove (1) we show that for any real k the sublevel

$$A_k = \{X \in \mathscr{X} \mid \pi(X) \le k\}$$

is $\sigma(\mathscr{X}, \mathscr{X}_n^\sim)$-closed. Suppose that $X_\alpha \in A_k$ and $X_\alpha \overset{\sigma(\mathscr{X}, \mathscr{X}_n^\sim)}{\to} X$. By the C-property, there exists $Y_n \in \text{conv}(X_{\alpha_n}, \dots)$ such that $Y_n \overset{o}{\to} X$. The convexity of π implies that $\pi(Y_n) \le k$ for each n. From order l.s.c. of π

$$\pi(X) \le \liminf \pi(Y_n) \le k$$

so that $X \in A_k$. □

Remark 6 Note that the C-property could have been stated with order converging subnets of convex combinations instead of subsequences (as in fact it was in the first version of the present paper). However the current presentation is given with subsequences as the applications rely only on Corollary 4, which in turn is based on Lemma 5.

A natural question is whether the sup in formula (12) is attained when π is finite valued. In general, the answer is no, as shown in the example below, where the max is attained over $\mathscr{X}'_+$ thanks to (1) but not over $(\mathscr{X}_n^\sim)_+$.

Example 3 Consider the classic counterexample [12, Example 4.36] translated in the language of monotone increasing maps, that is take $\pi : L^\infty \to \mathbb{R}$, $\pi(X) = \operatorname{ess\,sup} X$. This map is convex, increasing, positively homogeneous and order l.s.c.

on L^∞. For later use, observe also that it is not order u.s.c. From (12), taking into account the "cash additivity" property $\pi(X+c)=\pi(X)+c$, $c\in\mathbb{R}$,

$$\pi(X)=\sup_{\{Q \text{ probab.}, \, Q\ll P\}} E_Q[X]. \tag{13}$$

Similarly, from (1) and the good properties of π,

$$\pi(X)=\max_{Q\in\mathscr{M}_{1,f}(P)} E_Q[X]$$

which is exactly the representation in (2) with zero penalty function. If X is selected so that its ess-sup is not attained, the sup in (13) cannot be a maximum. We will consider a similar case in the example of Sect. 5.3.

If π is finite valued and order u.s.c., then interestingly enough π admits a representation as in (12) with the supremum replaced by a maximum, *without* the C-property requirement.

Lemma 7 *Let $(\mathscr{X},\tau)$ be a locally convex Frechet lattice and $\pi:\mathscr{X}\to\mathbb{R}$ be a convex increasing map. If π is order u.s.c. then*

$$\pi(X)=\max_{Y\in(\mathscr{X}_n^\sim)_+}\{\langle Y,X\rangle-\pi^*(Y)\},\quad X\in\mathscr{X},$$

and thus a fortiori π is order continuous.

Proof From (1),

$$\pi(X)=\max_{Y\in\mathscr{X}'_+}\{\langle Y,X\rangle-\pi^*(Y)\}\geq\sup_{Y\in(\mathscr{X}_n^\sim)_+}\{\langle Y,X\rangle-\pi^*(Y)\}.$$

We now prove that any Y attaining the max on $\mathscr{X}'_+$ is order continuous. In fact, suppose by contradiction that the max is attained on a positive, non order continuous Y_0. Then, there exists $Z_\alpha\overset{o}{\to}0$ such that $\limsup_\alpha\langle Y_0,Z_\alpha\rangle>0$ and

$$\pi(X)=\{\langle Y_0,X\rangle-\pi^*(Y_0)\}<\limsup_\alpha\{\langle Y_0,X+Z_\alpha\rangle-\pi^*(Y_0)\}\leq\limsup_\alpha\pi(X+Z_\alpha)$$

which is a contradiction with order u.s.c. of π. □

5 Orlicz Spaces and Applications to Risk Measures

5.1 Orlicz Spaces Have the C-Property

The following Orlicz spaces and the L^p spaces, $p\in[0,+\infty]$, are defined on the same probability space $(\Omega,\mathscr{F},P)$.

A Young function Ψ is an even, convex function $\Psi:\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ with the properties:

1. $\Psi(0) = 0$;
2. $\Psi(\infty) = +\infty$;
3. $\Psi < +\infty$ in a neighborhood of 0.

Note that Ψ may jump to $+\infty$ outside of a bounded neighborhood of 0. In case Ψ is finite valued however, it is also continuous by convexity.

The Orlicz space L^{Ψ} is then defined as

$$L^{\Psi} = \{X \in L^0 \mid \exists \alpha > 0 E[\Psi(\alpha X)] < +\infty\}.$$

It is a Banach space with the Luxemburg (or gauge) norm

$$N_{\Psi}(X) = \inf\left\{c > 0 \,\middle|\, E\left[\Psi\left(\frac{X}{c}\right)\right] \leq 1\right\}.$$

With the usual pointwise lattice operations, L^{Ψ} is also a Banach lattice, as the norm satisfies the monotonicity condition

$$|Y| \leq |X| \Rightarrow N_{\Psi}(Y) \leq N_{\Psi}(X).$$

Since Ψ is bounded in a neighborhood of 0 and it is convex and goes to $+\infty$ when $|x| \to \infty$, it is rather easy to prove that

$$L^{\infty} \stackrel{i}{\to} L^{\Psi} \stackrel{i}{\to} L^1 \tag{14}$$

with linear, injective lattice morphisms (the inclusions i). The dual $(L^{\Psi})'$ admits the general decomposition in order continuous band and singular band

$$(L^{\Psi})' = (L^{\Psi})_n^{\sim} \oplus (L^{\Psi})_s^{\sim} \tag{15}$$

and $(L^{\Psi})_n^{\sim}$ can be identified with the Orlicz space L^{Ψ^*} where

$$\Psi^*(y) = \sup_{x \in \mathbb{R}}\{yx - \Psi(x)\}$$

is the Young function conjugate of Ψ. The examples below illustrate different cases and show that the L^p are in fact particular Orlicz spaces.

1. Suppose $p \in (1, +\infty)$ and $\Psi = \Psi_p$ where

   $$\Psi_p(x) = \frac{|x|^p}{p}$$

   then $L^{\Psi_p} = L^p$. Since this space has an order continuous topology, the dual consists only of order continuous functionals. As $(\Psi_p)^* = \Psi_q$ with $q = \frac{p}{p-1}$, one recovers the classic

   $$(L^p)' = L^q.$$

2. $\Psi = \Psi_\infty$, where
$$\Psi_\infty(x) = \begin{cases} 0 & \text{if } |x| \le 1 \\ +\infty & \text{otherwise.} \end{cases}$$
Then the associated Orlicz space L^{Ψ_∞} is exactly L^∞ and as $(\Psi_\infty)^*(y) = |y|$, $L^{(\Psi_\infty)^*} = L^1$.
The decomposition of the dual provided in (15) is nothing but the Yosida-Hewitt decomposition $(L^\infty)' = L^1 \oplus (L^\infty)^\sim_s$ and the singular band $(L^\infty)^\sim_s$ consists of the purely finitely additive measures.
3. $\Psi_e(x) = e^{|x|} - 1$ is a genuine example of Young function which induces an Orlicz space different from the L^p. L^{Ψ_e} is the space of random variables with some finite exponential moment, i.e.
$$L^{\Psi_e} = \{X \in L^0 \mid \exists \alpha > 0 \text{ s.t. } E[e^{\alpha|X|}] < +\infty\}.$$
Analogously to what happens for L^∞, this space has a topology which is *not* order continuous. Thus the dual has the full general decomposition (15), with non-null singular band, as
$$(L^{\Psi_e})' = L^{(\Psi_e)^*} \oplus (L^{\Psi_e})^\sim_s$$
where the conjugate $(\Psi_e)^*$ is given by the function
$$\begin{cases} |y|(\ln|y| - 1) + 1 & \text{if } |y| > 1 \\ 0 & \text{otherwise} \end{cases}$$
which will be indicated with $\widehat{\Phi}$. As better explained below, since $\widehat{\Phi}$ doesn't grow too fast the Orlicz $L^{\widehat{\Phi}}$ displays a behavior similar to that of the L^p, $1 \le p < +\infty$, in the sense that its topology is order continuous. Then, its dual $(L^{\widehat{\Phi}})'$ coincides with $L^{\widehat{\Phi}^*} = L^{\Psi_e}$. The consequence is that the topology induced on L^{Ψ_e} by the order continuous functionals, $\sigma(L^{\Psi_e}, L^{\widehat{\Phi}})$, is nothing but *the weak* topology* on L^{Ψ_e}.

As anticipated in the examples above, when Ψ verifies a slow- growth condition, known in the literature as Δ_2 condition (see e.g. [18]):
$$\exists t^* > 0, \exists K > 0 \quad \text{s.t.} \quad \Psi(2t) \le K\Psi(t) \quad \text{for all } t > t^*$$
then $(L^\Psi)' = (L^\Psi)^\sim_n = L^{\Psi^*}$, that is the norm-topology on L^Ψ is order continuous. So by Lemmata 4 and 5 the norm topology and the weak topology $\sigma(L^\Psi, (L^\Psi)') = \sigma(L^\Psi, L^{\Psi^*})$ have the C-property.

In general, by (14) *and Corollary* 4*, the following topologies on* L^Ψ *all have the C-property*:

(a) *the norm topology*,
(b) *the weak topology*,
(c) *the* $\sigma(L^\Psi, L^{\Psi^*})$*-topology*.

We remark that when it is Ψ^* that satisfies the Δ_2 condition, as in Example 3 above, then the dual space of L^{Ψ^*} coincides with L^{Ψ}. Therefore in this case the topology $\sigma(L^{\Psi}, L^{\Psi^*})$ is nothing *but the weak* topology* on L^{Ψ} and it has the C-property.

One may also consider the Morse subspace M^{Ψ} of the Orlicz space L^{Ψ}:

$$M^{\Psi} = \{X \in L^{\Psi} \mid E[\Psi(kX)] < +\infty \ \forall k > 0\}. \tag{16}$$

When Ψ is finite-valued, M^{Ψ} is a norm closed band of L^{Ψ} and its dual $(M^{\Psi})' = L^{\Psi^*}$, so $\sigma(M^{\Psi}, L^{\Psi^*})$ has the C-property too.

In the context of expected utility maximization, the spaces M^{Ψ} were first used in [5]. They are the object of study in [9] and applied to risk measures. In [9] it has also been shown that a risk measure defined on M^{Ψ} has non empty topological interior if and only if it is finite valued. As the dual of the Morse space M^{Ψ} can be identified with a space of functions, the Orlicz L^{Ψ^*}, these spaces are easier to handle than the whole L^{Ψ}. In particular, since M^{Ψ} has order continuous norm, the dual representation (6) follows immediately from the Extended Namioka Klee Theorem 1.

In [6] and in [7] it has been shown that the full duality $(L^{\widehat{u}}, (L^{\widehat{u}})')$ can also be successfully employed to cover new cases in the applications to expected utility maximization and indifference pricing. In fact the Orlicz space $L^{\widehat{u}}$, defined by the Young function $\widehat{u}(x) = -u(-|x|) + u(0)$ associated to the utility function u, is the natural environment for such investigation. And the results on the indifference price for claims in the general $L^{\widehat{u}}$ obtained in [7] show that in the general setup the result by [9] fails: *a convex risk measure on $L^{\widehat{u}}$ can have non empty topological interior without being finite valued everywhere.* For other examples of this situation self-contained in the present paper, see the next Sect. 5.2 where there are some other interesting applications of the full duality to risk measures.

5.2 *New Insights on the Downside Risk and Risk Measures Associated to a Utility Function u*

We assume that the investment possibilities at a certain date in the future are modeled by elements X of L^0. As straightforward consequences of Proposition 1 we have the following representations of *decreasing* functionals defined on subspaces of L^0.

Corollary 5 *Let $(\mathscr{X}, \tau)$ be a locally convex Frechet lattice contained in L^0. If $\rho : \mathscr{X} \to (-\infty, +\infty]$ is a proper convex order l.s.c.* decreasing *functional and $\sigma(\mathscr{X}, \mathscr{X}_n^{\sim})$ has the C-property, then ρ admits the representation*

$$\rho(X) = \sup_{Y \in (\mathscr{X}_n^{\sim})_+} \{\langle Y, -X\rangle - \rho^*(-Y)\}, \quad X \in \mathscr{X}. \tag{17}$$

If in addition ρ satisfies the cash additivity property

$$\rho(X+c)=\rho(X)-c, \quad \forall c\in\mathbb{R}\ \forall X\in\mathscr{X}, \tag{18}$$

then

$$\rho(X)=\sup_{Y\in(\mathscr{X}_n^{\sim})_+,\langle Y,1\rangle=1}\{\langle Y,-X\rangle-\rho^*(-Y)\}, \quad X\in\mathscr{X}. \tag{19}$$

If in addition ρ is positively homogeneous, then there exists a convex subset $\mathscr{C}\subseteq\{Y\in(\mathscr{X}_n^{\sim})_+\mid\langle Y,1\rangle=1\}$ such that

$$\rho(X)=\sup_{Y\in\mathscr{C}}\langle Y,-X\rangle.$$

Let us consider an agent, whose preferences on the investments X can be represented via expected utility. We assume that the utility function

- $u:\mathbb{R}\to\mathbb{R}$ is increasing and concave (though not necessarily strictly concave) and satisfies $\lim_{x\to-\infty}u(x)=-\infty$.

Without loss of generality, suppose

$$u(0)=0.$$

The goal is that of describing a natural framework associated to the expected utility of the agent, i.e. to the functional

$$E[u(X)],$$

to the related downside risk

$$\Theta(X):=E[-u(X)]$$

and to some associated convex risk measures. As it is not required that u is strictly concave, u can be identically 0 on $\mathbb{R}_+$ and in this case Θ is nothing but the so-called shortfall risk [12].

It turns out that a good setup is that of an Orlicz spaces duality induced by the functional itself. As shown in [5] and [6] the function

$$\widehat{u}(x)=-u(-|x|)$$

is a Young function and defines the Orlicz space $L^{\widehat{u}}$ associated to u. Call

$$\Phi(y)=\sup_{x\in\mathbb{R}}\{u(x)-xy\}$$

the convex conjugate of u. Since $\widehat{u}$ is finite on $\mathbb{R}$, then, as observed right after the definition (16), $M^{\widehat{u}}$ is a norm-closed band of $L^{\widehat{u}}$ and its dual is $L^{\widehat{\Phi}}$.

It is clear that there must be a link between Φ and $\widehat{\Phi}$, the Young function conjugate to $\widehat{u}$. In fact

$$\widehat{\Phi}(y) = \begin{cases} 0 & \text{if } |y| \le \beta \\ \Phi(|y|) & \text{if } |y| > \beta \end{cases}$$

where $\beta \ge 0$ is the right derivative of $\widehat{u}$ at 0, namely $\beta = D^+\widehat{u}(0) = D^- u(0)$. If u is differentiable, note that $\beta = u'(0)$ and it is the unique solution of the equation $\Phi'(y) = 0$. To fix the ideas, consider the following examples.

1. Fix $\gamma > 0$ and take

$$u_\gamma(x) = -e^{-\gamma x} + 1 \tag{20}$$

whence $\widehat{u_\gamma}(x) = e^{\gamma|x|} - 1$ and

$$\Phi_\gamma(y) = \frac{y}{\gamma}\ln\frac{y}{\gamma} - \frac{y}{\gamma} + 1$$

and $\widehat{\Phi}(y) = (|\frac{y}{\gamma}|\ln|\frac{y}{\gamma}| - |\frac{y}{\gamma}| + 1)I_{\{|\frac{y}{\gamma}|\ge 1\}}$. It is not difficult to see that the associated Orlicz spaces do not depend on γ (in the sense that they are physically the same and changing γ amounts to a dilation of the Luxemburg norm) and therefore, as pointed out in Sect. 4.1, Example 3,

$$L^{\widehat{u_\gamma}} = \{X \in L^0 \mid \exists \alpha > 0 \text{ s.t. } E[e^{\alpha|X|}] < +\infty\},$$

$$M^{\widehat{u_\gamma}} = \{X \in L^0 \mid \forall \alpha > 0 E[e^{\alpha|X|}] < +\infty\},$$

$$L^{\widehat{\Phi_\gamma}} = \{Y \in L^0 \mid E[(|Y|\ln|Y|)I_{\{|Y|>1\}}] < +\infty\} \quad \text{and}$$

$$(L^{\widehat{u}})' = L^{\widehat{\Phi}} \oplus (L^{\widehat{u}})^{\sim}_s.$$

2. Let u be the quadratic-flat utility, i.e.

$$u(x) = \begin{cases} -\frac{x^2}{2} & \text{if } x \le 0 \\ 0 & \text{if } x \ge 0. \end{cases} \tag{21}$$

In this case, $\widehat{u}(x) = \frac{x^2}{2} = \widehat{\Phi}(x)$, and all the spaces $L^{\widehat{u}}, M^{\widehat{u}}, L^{\widehat{\Phi}}$ are equal and coincide (modulo an isomorphism) with L^2.

Let us recall that the Orlicz class of $L^{\widehat{u}}$ is defined as

$$\mathscr{L}^{\widehat{u}} = \{X \in L^0 \mid E[\widehat{u}(X)] < +\infty\}$$

and it is a convex subset (not necessarily closed) of $L^{\widehat{u}}$.

The following lemma is a nice consequence of the right choice of the spaces.

Lemma 8 *The downside risk* $\Theta : L^{\widehat{u}} \to (-\infty, +\infty]$, $\Theta(X) = E[-u(X)]$, *is a well-defined, proper, convex and monotone decreasing functional which is order l.s.c. In addition,*

$$\mathrm{Dom}(\Theta) = \{X \in L^{\widehat{u}} \mid X^- \in \mathscr{L}^{\widehat{u}}\}$$

and

$$\mathrm{int}(\mathrm{Dom}(\Theta)) = \{X \in L^{\widehat{u}} \mid \exists \varepsilon > 0 (1+\varepsilon)X^- \in \mathscr{L}^{\widehat{u}}\} \supseteq M^{\widehat{u}}. \tag{22}$$

Moreover, Θ *admits the representation:*

$$\Theta(X) = \sup_{Y \in L_+^{\widehat{\Phi}}} \{E[-XY] - E[\Phi(Y)]\}. \tag{23}$$

Proof If $X \in L^{\widehat{u}}$, then by Jensen's inequality

$$E[-u(X)] \geq -u(E[X]) > -\infty \tag{24}$$

since $E[X] \in \mathbb{R}$ from (14) and $u < +\infty$ on $\mathbb{R}$. So the definition is well-posed and Θ is clearly convex and monotone decreasing. To prove the characterization of $\mathrm{Dom}(\Theta)$, simply note that

$$\begin{aligned} X \in \mathrm{Dom}(\Theta) \quad &\text{iff} \quad E[u(X)] > -\infty \quad \text{iff} \quad E[u(-X^-)] > -\infty \\ &\text{iff} \quad E[\widehat{u}(X^-)] < +\infty \end{aligned}$$

where the second equivalence above is due to the fact that $E[u(X^+)]$ is always finite as

$$u(0) \leq u(x^+) \leq ax^+ + b$$

for some $a, b \in \mathbb{R}$, so that $u(0) \leq E[u(X^+)] \leq aE[X^+] + b < +\infty$.

To prove (22), if $X \in \mathrm{int}(\mathrm{Dom}(\Theta))$ then clearly for some $\varepsilon > 0$ $E[-u(X - \varepsilon X^-)]$ is finite, that is $E[\widehat{u}((1+\varepsilon)X^-)]$ is finite.

Conversely, suppose $(1+\varepsilon)X^- \in \mathscr{L}^{\widehat{u}}$. Then, $(1+\varepsilon)X \in \mathrm{Dom}(\Theta)$ and for any Z with Luxemburg norm $N_{\widehat{u}}(Z) < \frac{\varepsilon}{1+\varepsilon}$, $X + Z \in \mathrm{Dom}(\Theta)$. In fact:

$$\begin{aligned} E[-u(X+Z)] &= E\left[-u\left(\frac{1}{1+\varepsilon}((1+\varepsilon)X) + Z\right)\right] \\ &\leq \frac{1}{1+\varepsilon} E[-u((1+\varepsilon)X)] + \frac{\varepsilon}{1+\varepsilon} E\left[-u\left(\frac{1+\varepsilon}{\varepsilon} Z\right)\right] < +\infty \end{aligned}$$

since $\frac{1+\varepsilon}{\varepsilon} Z$ has Luxemburg norm less than 1 and thus

$$E\left[-u\left(\frac{1+\varepsilon}{\varepsilon} Z\right)\right] \leq E\left[-u\left(-\frac{1+\varepsilon}{\varepsilon} Z^-\right)\right] = E\left[\widehat{u}\left(\frac{1+\varepsilon}{\varepsilon} Z^-\right)\right] \leq 1.$$

Thanks to Remark 4, in order to show that Θ is order l.s.c. one just needs to check whether Θ is σ-continuous from above. But this is an immediate consequence of the monotone convergence theorem and (24). Finally, the $\sigma(L^{\widehat{u}}, L^{\widehat{\Phi}})$ topology has the C-property so the representation (17) on the order continuous dual $L^{\widehat{\Phi}}$ applies

$$\Theta(X) = \sup_{Y \in L_+^{\widehat{\Phi}}} \{E[-YX] - \Theta^*(-Y)\}.$$

By Kozek's results [15] (or directly by hand), Θ^*, the convex conjugate of Θ,

$$\Theta^*(Y) = \sup_{X \in L^{\widehat{u}}} \{E[YX] - \Theta(X)\}$$

verifies

$$\Theta^*(-Y) = E[\Phi(Y)], \quad \text{if } Y \in L^{\widehat{\Phi}}. \tag{25}$$

□

Clearly Θ satisfies all the requirements of a convex risk measure *but* cash additivity.

As shown in [4] in the L^∞ case, the greatest convex risk measure smaller than a convex functional $\theta : L^\infty \to \mathbb{R}$ can be constructed by taking the inf-convolution $\theta \Box \rho_{worst}$ of θ with $\rho_{worst} = \rho_{L_+^\infty}$, which is the risk measure associated to the acceptance set L_+^∞. Then the penalty function of ρ_{worst} is equal to 0 on $\mathscr{M}_{1,f}(P)$, and is equal to ∞ outside $\mathscr{M}_{1,f}(P)$. Since the penalty function of $\theta \Box \rho_{worst}$ is the sum of the penalty function of θ and of ρ_{worst}, the representation of $\theta \Box \rho_{worst}$ will have the same penalty function of θ, but the supremum in such representation is restricted to the set $\mathscr{M}_{1,f}(P)$, i.e. to those positive elements in the dual space that are also *normalized*. The same conclusion holds in our setting, as shown in the following result.

Proposition 2 *The map $\zeta_u : L^{\widehat{u}} \to (-\infty, +\infty]$ defined by*

$$\zeta_u(X) = \sup_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \left\{ E_Q[-X] - E\left[\Phi\left(\frac{dQ}{dP}\right)\right]\right\} \tag{26}$$

is a well-defined order l.s.c. convex risk measure and it is the greatest order l.s.c. convex risk measure smaller than Θ and hence $\zeta_u = \Theta \Box \rho_{L_+^{\widehat{u}}}$. Moreover, the sup *in* (26) *can equivalently be computed on the set*

$$\left\{ Q \text{ probab.}, \ Q \ll P \;\middle|\; E\left[\Phi\left(\frac{dQ}{dP}\right)\right] < +\infty \right\}.$$

Proof It is clear that ζ_u is an order l.s.c convex risk measure. From (23) we also have: $\zeta_u \le \Theta$. We need only to prove that if $\widetilde{\rho} : L^{\widehat{u}} \to (-\infty, +\infty]$ is an order l.s.c.

convex risk measure such that $\widetilde{\rho} \le \Theta$, then $\widetilde{\rho} \le \zeta_u$. Let $\widetilde{\alpha}(Y) = \widetilde{\rho}^*(-Y)$ be the penalty function associated with $\widetilde{\rho}$ in the representation (19)

$$\widetilde{\rho}(X) = \sup_{Y \in L_+^{\widehat{\Phi}}, E[Y]=1} \{E[-XY] - \widetilde{\alpha}(Y)\}.$$

By cash additivity, $\widetilde{\rho}(X + \Theta(X)) = \widetilde{\rho}(X) - \Theta(X) \le 0$, for all $X \in L^{\widehat{u}}$, so that

$$\widetilde{\rho}(X + \Theta(X)) = \sup_{Y \in L_+^{\widehat{\Phi}}, E[Y]=1} \{E[-YX] - \Theta(X) - \widetilde{\alpha}(Y)\} \le 0.$$

This implies that, if $Y \in L_+^{\widehat{\Phi}}$, $E[Y] = 1$,

$$\widetilde{\alpha}(Y) \ge E[-YX] - \Theta(X) \quad \text{for all } X \in L^{\widehat{u}}$$

and, by (25),

$$\widetilde{\alpha}(Y) \ge \sup_{X \in L^{\widehat{u}}} \{E[-YX] - \Theta(X)]\} = \Theta^*(-Y) = E[\Phi(Y)].$$

Therefore,

$$\begin{aligned}\widetilde{\rho}(X) &= \sup_{Y \in L_+^{\widehat{\Phi}}, E[Y]=1} \{E[-YX] - \widetilde{\alpha}(Y)\} \\ &\le \sup_{Y \in L_+^{\widehat{\Phi}}, E[Y]=1} \{E[-YX] - E[\Phi(Y)] = \zeta_u(X).\end{aligned}$$

Since the integrability condition $E[\Phi(Y)] < +\infty$ on $Y \ge 0$ is more severe than the requirement $Y \in L_+^{\widehat{\Phi}}$, the last sentence is obvious. □

To any utility function satisfying our assumptions, one can also associate the map $\rho_u : L^{\widehat{u}} \to (-\infty, +\infty]$ defined by:

$$\rho_u(X) = \inf\{c \in \mathbb{R} \mid X + c \in \mathscr{A}_u\}, \tag{27}$$

where the set $\mathscr{A}_u$ is defined as

$$\mathscr{A}_u := \{X \in L^{\widehat{u}} \mid E[u(X)] \ge u(0) = 0\} = \{X \in L^{\widehat{u}} \mid \Theta(X) \le 0\}.$$

Lemma 9 *$\mathscr{A}_u$ has the properties*:

1. *it is convex*;
2. *if $X \in \mathscr{A}_u$ and $Z \in L^{\widehat{u}}$, $Z \ge X$, then $Z \in \mathscr{A}_u$*;
3. $\inf\{c \in \mathbb{R} \mid c \in \mathscr{A}_u\} > -\infty$;
4. *for any $X \in \mathscr{A}_u$ and $Z \in L^{\widehat{u}}$, the set $\{t \in [0, 1] \mid (1-t)X + tZ \in \mathscr{A}_u\}$ is closed in* $[0, 1]$.

Proof We only prove item 4, as the others are simple consequences of the properties of u. Fix any $X \in \mathscr{A}_u$ and call $\Lambda = \{t \in [0,1] \mid (1-t)X + tZ \in \mathscr{A}_u\}$. For any cluster point t^* of Λ, there exists a sequence $(t_n)_n \in \Lambda, t_n \to t^*$. But then, $(1-t_n)X + t_n Z$ order converges to $(1-t^*)X + t^*Z$. From Lemma 8, Θ is order l.s.c., so $\Theta((1-t^*)X + t^*Z) \le \liminf_n \Theta((1-t_n)X + t_n Z) \le 0$ which means $t^* \in \Lambda$. □

Proposition 3 $\rho_u : L^{\widehat{u}} \to (-\infty, +\infty]$ *is a well-defined order l.s.c. convex risk measure that admits the representation*:

$$\rho_u(X) = \sup_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \{E_Q[-X] - \alpha(Q)\}, \tag{28}$$

where

$$\alpha(Q) = \sup_{X \in \mathscr{A}_u} \{E_Q[-X]\}. \tag{29}$$

Moreover, $\mathscr{A}_{\rho_u} := \{X \in L^{\widehat{u}} \mid \rho_u(X) \le 0\}$, *the acceptance set of* ρ_u, *satisfies*

$$\mathscr{A}_{\rho_u} = \mathscr{A}_u,$$

and, as a consequence, if $\widetilde{\rho} : L^{\widehat{u}} \to \mathbb{R} \cup \{+\infty\}$ *is an order l.s.c. convex risk measure such that* $\widetilde{\rho} \le \Theta$, *then* $\widetilde{\rho} \le \rho_u$.

Proof The facts that ρ_u is a convex risk measure and that its acceptance set $\mathscr{A}_{\rho_u}$ coincides with $\mathscr{A}_u$ are consequences of the above lemma and Propositions 2, 4 in [13]. Now, since Θ is order l.s.c., $\mathscr{A}_u$ is order closed, so that the acceptance set of ρ_u is order-closed. And since $\sigma(L^{\widehat{u}}, L^{\widehat{\Phi}})$ has the C-property the acceptance set $\mathscr{A}_{\rho_u} = \mathscr{A}_u$ is $\sigma(L^{\widehat{u}}, L^{\widehat{\Phi}})$-closed. Hence, by a classic result, as its sublevels are $\sigma(L^{\widehat{u}}, L^{\widehat{\Phi}})$-closed, ρ_u is $\sigma(L^{\widehat{u}}, L^{\widehat{\Phi}})$-l.s.c. But this implies that it is also ρ_u order l.s.c. by the first part of the statement in Proposition 1. Then, the representation (19) on the order continuous dual $L^{\widehat{\Phi}}$ applies:

$$\rho_u(X) = \sup_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \{E_Q[-X] - \alpha(Q)\}, \quad X \in L^{\widehat{u}},$$

$$\alpha(Q) \triangleq \rho_u^*(-Q) = \sup_{X \in L^{\widehat{u}}} \{E_Q[-X] - \rho_u(X)\}.$$

It is straightforward to see that the penalty functional α admits the representation

$$\alpha(Q) = \sup_{X \in \mathscr{A}_{\rho_u}} E_Q[-X].$$

If $\widetilde{\rho} : L^{\widehat{u}} \to (-\infty, +\infty]$ is an order l.s.c. convex risk measure such that $\widetilde{\rho} \le \Theta$ then $\mathscr{A}_{\widetilde{\rho}} = \{X \in L^{\widehat{u}} \mid \widetilde{\rho}(X) \le 0\} \supseteq \mathscr{A}_u = \mathscr{A}_{\rho_u}$, and this implies $\widetilde{\rho} \le \rho_u$. □

Remark 7 Obviously, $\zeta_u \le \rho_u$, but $\zeta_u = \rho_u$ if and only if $\rho_u \le \Theta$ which in general is *not* true. Note that the inequality $\rho_u \le \Theta$ would imply: $\Theta(X + \Theta(X)) \le \Theta(X + \rho_u(X)) \le 0$ (this latter inequality follows from $X + \rho_u(X) \in \mathscr{A}_{\rho_u} = \mathscr{A}_u$)), but in general $\Theta(X + \Theta(X)) \le 0$ does not hold. So, ρ_u and ζ_u may be different (see the Sect. 5.3 below). In Sect. 5.4 there is a case where $\rho_u = \zeta_u$.

5.3 Quadratic-Flat Utility

If u is the quadratic-flat utility function (21), then ζ_u and ρ_u are different. Indeed, $L^{\widehat{u}} = L^2$ and

$$\Theta(X) = \frac{1}{2}E[(X^-)^2] = \sup_{Y \in L^2_+} \left\{ E[-YX] - \frac{1}{2}E[Y^2] \right\}, \quad X \in L^2,$$

$$\zeta_u(X) = \sup_{\frac{dQ}{dP} \in L^2_+} \left\{ E_Q[-X] - \frac{1}{2}E\left[\left(\frac{dQ}{dP}\right)^2\right] \right\}, \quad X \in L^2.$$

Since $\mathscr{A}_u = \{X \in L^2 \mid \Theta(X) \le 0\} = L^2_+$, we have:

$$\rho_u(X) = \inf\{c \in \mathbb{R} \mid X + c \ge 0\} = \rho_{worst}(X) := -\operatorname{ess\,inf}(X).$$

The dual representation in (28) becomes

$$\rho_u(X) = \sup_{Q \ll P, \frac{dQ}{dP} \in L^2_+} E_Q[-X]$$

since, from (29), the penalty term is given by

$$\alpha(Q) = \sup_{X \in L^2, \Theta(X) \le 0} E_Q[-X] = 0, \quad \text{if } \frac{dQ}{dP} \in L^2_+.$$

Note also that

$$\rho_u(X) = \sup_{Q \ll P, \frac{dQ}{dP} \in L^2_+} E_Q[-X] = \sup_{Q \ll P, \frac{dQ}{dP} \in L^2_+} E_Q[X^-] \ge \frac{E[(X^-)^2]}{E[X^-]}$$

and if $0 < E[X^-] < 2$, $\rho_u(X) > \Theta(X)$. Moreover, ρ_u is not even finite-valued. Therefore, while Θ and ζ_u are finite valued and thus continuous and subdifferentiable on L^2, $\mathrm{Dom}(\rho_u)$ has empty interior thanks to the cited result of [9] for risk measures on Morse subspaces (here, $L^{\widehat{u}} = M^{\widehat{u}} = L^2$).

5.4 Exponential Utility

Let $u(x) = -e^{-x} + 1$ be the exponential utility function considered in (20). W.l.o.g. we set $\gamma = 1$. Then,

$$\Theta(X) = E[e^{-X}] - 1, \quad X \in L^{\widehat{u}},$$

and $\Phi(y) = y \ln y - y + 1$. From the definition (27) we have

$$\rho_u(X) = \inf\{c \in \mathbb{R} \mid E[e^{-X-c} - 1] \le 0]\} = \ln E[e^{-X}],$$

with the convention $\ln E[e^{-X}] = +\infty$ if $E[e^{-X}] = +\infty$. Clearly $\rho_u(X) = \ln E[e^{-X}] \le E[e^{-X}] - 1 = \Theta(X)$ and therefore, in this case, $\rho_u = \zeta_u$. So, from (26) we recover the entropic risk measure together with its dual variational identity

$$\begin{aligned}\ln E[e^{-X}] &= \sup_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \left\{ E_Q[-X] - E\left[\Phi\left(\frac{dQ}{dP}\right)\right]\right\} \\ &= \sup_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \left\{ E_Q[-X] - E_Q\left[\ln\left(\frac{dQ}{dP}\right)\right]\right\}, \quad X \in L^{\widehat{u}}.\end{aligned}$$

The novelty here is that the space where this representation holds is $L^{\widehat{u}}$, naturally induced by u and not an arbitrarily selected subspace of L^0 (traditionally, the entropic risk measure is defined on L^∞ and the formula above is provided for $X \in L^\infty \subset L^{\widehat{u}}$, see [12] and the remarks below). And ρ_u is a genuine example of a risk measure on the general Orlicz space $L^{\widehat{u}}$ which is not finite valued everywhere and still has domain with no empty interior: as $\mathrm{Dom}(\rho_u) = \mathrm{Dom}(\Theta)$, the interior of the domain has been computed in (22).

To conclude, let us focus on the restriction of $\rho_u = \zeta_u$ to the subspace $M^{\widehat{u}}$.

Corollary 6 *The restriction υ_u of ζ_u to the subspace $M^{\widehat{u}}$ is a well-defined norm continuous (hence order continuous) convex risk measure $\upsilon_u : M^{\widehat{u}} \to \mathbb{R}$ that admits the representation*

$$\upsilon_u(X) = \max_{Q \ll P, \frac{dQ}{dP} \in L_+^{\widehat{\Phi}}} \left\{ E_Q[-X] - E\left[\Phi\left(\frac{dQ}{dP}\right)\right]\right\}, \quad X \in M^{\widehat{u}}.$$

We thus recover the representation formulae provided by [9] on Morse subspaces and the formula with the max for the entropic risk measure on $L^\infty \subset M^{\widehat{u}}$.

Acknowledgements The first author would like to thank B. Rudloff, P. Cheridito and A. Hamel for some discussions while she was visiting the ORFE Department at Princeton University. The second author would like to thank Marco Maggis, PhD student at Milano University for helpful discussion on this subject.

References

1. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis, 3rd edn. Springer, Berlin (2005)
2. Aliprantis, C.D., Burkinshaw, O.: Positive Operators. Academic Press, San Diego (1985)
3. Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Math. Financ. **4**, 203–228 (1999)
4. Barrieu, P., El Karoui, N.: Pricing, hedging and optimally designing derivatives via minimization of risk measures. In: Carmona, R. (ed.) Indifference Pricing: Theory and Applications. Princeton University Press, Princeton (2008)
5. Biagini, S.: An Orlicz spaces duality for utility maximization in incomplete markets. In: Proceedings of Ascona 2005. Progress in Probability. Birkhäuser, Basel (2007)
6. Biagini, S., Frittelli, M.: A unified framework for utility maximization problems: an Orlicz space approach. Ann. Appl. Prob. **18**(3), 929–966 (2008)
7. Biagini, S., Frittelli, M., Grasselli, M.: Indifference price for general semimartingales. Submitted, 2007
8. Brezis, H.: Analyse fonctionnelle. Masson, Paris (1983)
9. Cheridito, P., Li, T.: Risk measures on Orlicz hearts. Math. Financ. **19**(2), 189–214 (2009)
10. Delbaen, F.: Coherent risk measures on general probability spaces. In: Essays in Honour of Dieter Sondermann. Springer, Berlin (2000)
11. Filipovic, D., Svindland, G.: The canonical model space for law-invariant convex risk measures is L^1. Preprint (2008)
12. Föllmer, H., Schied, A.: Stochastic Finance. An Introduction in Discrete Time, 2nd edn. De Gruyter, Berlin (2004)
13. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Finance Stoch. **6**(4), 429–447 (2002)
14. Frittelli, M., Rosazza Gianin, E.: Putting order in risk measures. J. Bank. Financ. **26**(7), 1473–1486 (2002)
15. Kozek, A.: Convex integral functionals on Orlicz spaces. Ann. Soc. Math. Pol. Ser. 1, Comment. Math. **XXI**, 109–134 (1979)
16. Maccheroni, F., Marinacci, M., Rustichini, A.: A variational formula for the relative Gini concentration index. In press
17. Namioka, I.: Partially Ordered Linear Topological Spaces. Mem. Am. Math. Soc., vol. 24. Princeton University Press, Princeton (1957)
18. Rao, M.M., Ren, Z.D.: Theory of Orlicz Spaces. Marcel Dekker, New York (1991)
19. Rüschendorf, L., Kaina, M.: On convex risk measures on Lp-spaces. Preprint (2007)
20. Ruszczynski, A., Shapiro, A.: Optimization of convex risk measures. Math. Oper. Res. **31**(3), 433–452 (2006)
21. Zaanen, A.C.: Riesz Spaces II. North-Holland Math. Library. North-Holland, Amsterdam (1983)

On Certain Distributions Associated with the Range of Martingales

Alexander Cherny and Bruno Dupire

Abstract We study some properties of a continuous local martingale stopped at the first time when its range (the difference between the running maximum and minimum) reaches a certain threshold. The laws and the conditional laws of its value, maximum, and minimum at this time are simple and do not depend on the local martingale in question. As a consequence, the price and hedge of options which mature when the range reaches a given level are both model-free within the class of arbitrage-free models with continuous paths, which makes these products very convenient for hedging.

Keywords Continuous martingales · Option hedging · Option pricing · Options on the price range

Mathematics Subject Classification (2000) 91B62 · 91B70

1 Introduction

Let $(S_t)_{t\geq 0}$ be a continuous local martingale on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathsf{P})$. We assume that $\mathscr{F}_0$ is trivial. Consider its running minimum and maximum

$$m_t = \inf_{u\leq t} S_u, \qquad M_t = \sup_{u\leq t} S_u, \quad t\geq 0$$

and define the *range process* as

$$R_t = M_t - m_t, \quad t \geq 0.$$

Define the stopping time

$$\tau = \inf\{t \geq 0 : R_t \geq L\},$$

A. Cherny (✉)
Department of Probability Theory, Faculty of Mechanics and Mathematics, Moscow State University, 119992 Moscow, Russia
e-mail: alexander.cherny@gmail.com

B. Dupire
Bloomberg L.P., 731 Lexington Avenue, New York, NY 10022, USA
e-mail: bdupire@bloomberg.net

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_2,

where $L > 0$ is a given threshold. We will be interested in the unconditional and conditional laws of the random variables S_τ and M_τ. Below we will denote by δ_a the Dirac delta-mass concentrated at a point a; by I_A we will denote the indicator of a set A. We have the following result; statement (ii) is actually known and can be found in [2, 5.0.4].

Theorem 1 *Assume that τ is finite almost surely.*

(i) *The distribution* Law S_τ *is given by*

$$(\text{Law}\, S_\tau)(dx) = L^{-2}|x - S_0| I_{(S_0-L,S_0+L)}(x)dx.$$

(ii) *The distribution* Law M_τ *is given by*

$$(\text{Law}\, M_\tau)(dx) = L^{-1} I_{(S_0,S_0+L)} dx.$$

(iii) *The conditional distribution* $\text{Law}(S_\tau \mid \mathscr{F}_t)$ *on* $\{t < \tau\}$ *is given by*

$$\begin{aligned}(\text{Law}(S_\tau \mid \mathscr{F}_t))(dx) = L^{-1}(M_t - S_t)\delta_{M_t-L}(dx) + L^{-1}(S_t - m_t)\delta_{m_t+L}(dx)\\ + L^{-2}(S_t - x) I_{(M_t-L,m_t)}(x)dx\\ + L^{-2}(x - S_t) I_{(M_t,m_t+L)}(x)dx.\end{aligned}$$

(iv) *The conditional distribution* $\text{Law}(M_\tau \mid \mathscr{F}_t)$ *on* $\{t < \tau\}$ *is given by*

$$\begin{aligned}(\text{Law}(S_\tau \mid \mathscr{F}_t))(dx) = L^{-1}(S_t - m_t)\delta_{m_t+L}(dx) + L^{-1}(M_t - S_t)\delta_{M_t}(dx)\\ + L^{-1} I_{(M_t,m_t+L)}(x)dx.\end{aligned}$$

Remark 1 Point (ii) easily follows from (i) if one notes that

$$M_\tau = \begin{cases} S_\tau + L & \text{if } S_\tau \le S_0,\\ S_\tau & \text{if } S_\tau > S_0.\end{cases}$$

Similarly, (iv) is an easy consequence of (iii) since on $\{t < \tau\}$, we have

$$M_\tau = \begin{cases} S_\tau + L & \text{if } S_\tau \le S_t,\\ S_\tau & \text{if } S_\tau > S_t.\end{cases}$$

As $m_\tau = M_\tau - L$, we see that from (i) one can recover the whole $\text{Law}(S_\tau, m_\tau, M_\tau)$, while from (iii) one can recover $\text{Law}(S_\tau, m_\tau, M_\tau \mid \mathscr{F}_t)$.

The above theorem might be given an interesting financial application. Let $(S_t)_{t\ge 0}$ describe the price process of an asset. Imagine an option that pays out the amount $f(S_\tau)$, where τ is the same as above and f is a given function (e.g., $f(x) = (x - K)^+$). Then, as a corollary of the above result, we get the price of this option in any arbitrage-free model with continuous paths. It is given by the theorem below, where we are providing the hedge as well.

Theorem 2 *Assume that the risk-free rate is zero and f is integrable on any compact interval. Assume also that $\tau < \infty$ a.s. and the model is arbitrage-free in the sense that there exists an equivalent measure, under which S is a martingale.*[1]

(i) *For any equivalent martingale measure* Q, *the price* $V_0 = \mathsf{E}_{\mathsf{Q}} f(S_\tau)$ *is given by*

$$V_0 = L^{-2} \int_{S_0-L}^{S_0+L} |x - S_0| f(x) dx.$$

(ii) *For any equivalent martingale measure* Q, *the corresponding price process* $V_t = \mathsf{E}_{\mathsf{Q}}[f(S_\tau) \mid \mathscr{F}_t]$ *is given by*

$$\begin{aligned} V_t = {} & L^{-1}(M_t - S_t) f(M_t - L) + L^{-1}(S_t - m_t) f(m_t + L) \\ & + L^{-2} \int_{M_t-L}^{m_t} (S_t - x) f(x) dx \\ & + L^{-2} \int_{M_t}^{m_t+L} (x - S_t) f(x) dx \quad \text{on } \{t < \tau\} \end{aligned}$$

and $V_t = f(S_\tau)$ *on* $\{t \geq \tau\}$.

(iii) *The hedge* H *is given by*

$$\begin{aligned} H_t = I(t < \tau) \Big[& -L^{-1} f(M_t - L) + L^{-1} f(m_t + L) \\ & + L^{-2} \int_{M_t-L}^{m_t} f(x) dx - L^{-2} \int_{M_t}^{m_t+L} f(x) dx \Big], \quad t \geq 0, \end{aligned}$$

i.e.

$$V_t = V_0 + \int_0^t H_u dS_u, \quad t \geq 0.$$

We see that the price and the hedge do not depend on the equivalent martingale measure and are thus model-independent within the class of arbitrage-free models with continuous paths. These include, in particular, the Bachelier model, the Black-Scholes-Merton model, the local volatility models as well as the stochastic volatility models. What is more important, the price and the hedge admit a simple analytic form. In those respects, options on the range have similarities with options having the payoff $f(S_\sigma)$, where $\sigma = \inf\{t \geq 0 : \langle S \rangle_t \geq L\}$ (for more information on such options, see [1]).

Let us remark that on the set $\{t \leq \tau\}$, V_t has the form $v(S_t, m_t, M_t)$, while H_t has the form $h(m_t, M_t)$, where

$$h(m, M) = \frac{\partial v(S, m, M)}{\partial S},$$

[1] All the results remain the same if the word "martingale" is replaced by "local martingale".

i.e. H is the delta-hedge. The function v is linear in S, so that h does not depend on S, which means that the hedge remains constant until the price S_t reaches its running maximum or running minimum. This implies that the gamma of the option is zero. However, each time when S breaks through its running maximum or minimum, the hedge is updated (as M_t or m_t changes). Loosely speaking, one can say that the option has a non-zero "right-hand gamma"

$$\frac{\partial^2 v(M+\varepsilon, m, M+\varepsilon)}{\partial \varepsilon_+^2} = L^{-2} f(M) - L^{-2} f(M-L) - L^{-1} f'(M-L)$$

at the time when the price breaks through its running maximum and a non-zero "left-hand gamma"

$$\frac{\partial^2 v(m+\varepsilon, m+\varepsilon, M)}{\partial \varepsilon_-^2} = -L^{-2} f(m+L) + L^{-2} f(m) + L^{-1} f'(m+L)$$

at the time when the price breaks through its running minimum. Here by $\partial^2/\partial\varepsilon_+^2$ (resp., $\partial^2/\partial\varepsilon_-^2$) we denote the right-hand (resp., left-hand) second derivative.

Thus, the hedge should be updated only at the times when the price breaks through its running maximum or minimum. As time grows, these "break points" appear more and more rarely (for example, in a random walk model, the number of such points in the interval $[0, N]$ is of order $N^{1/2}$). This makes the product "quite hedgeable".

Let us finally remark that items (ii) and (iv) of Theorem 1 can also be given a financial interpretation. Consider the same setting as before and imagine an option that pays out the amount $f(M_\tau)$. For these options, under the same assumptions as in Theorem 3, the following result holds.

Theorem 3 (i) *For any equivalent martingale measure* Q, *the price* $V_0 = \mathsf{E}_\mathsf{Q} f(M_\tau)$ *is given by*

$$V_0 = L^{-1} \int_{S_0}^{S_0+L} f(x)dx.$$

(ii) *For any equivalent martingale measure* Q, *the corresponding price process* $V_t = \mathsf{E}_\mathsf{Q}[f(M_\tau) \mid \mathscr{F}_t]$ *is given by*

$$V_t = L^{-1}(S_t - m_t) f(m_t + L) + L^{-1}(M_t - S_t) f(M_t) + L^{-1} \int_{M_t}^{m_t+L} f(x)dx \quad \text{on } \{t < \tau\}$$

and $V_t = f(M_\tau)$ *on* $\{t \geq \tau\}$.

(iii) *The hedge* H *is given by*

$$H_t = I(t < \tau)[L^{-1} f(m_t + L) - L^{-1} f(M_t)], \quad t \geq 0.$$

2 Proofs

We will prove Theorems 1 and 2. Theorem 3 is proved in the same way as Theorem 2. We are starting with an auxiliary lemma, which is a generalization of the "plate problem" mentioned at the beginning of the paper.

Lemma 1 *Let $(X_n)_{n=0,1,\dots}$ be a standard symmetric random walk on $\mathbb{Z}$. Denote*

$$R_n = \max_{i \le n} X_i - \min_{i \le n} X_i, \quad n = 0, 1, \dots$$

and let $\tau = \inf\{n : R_n = N\}$, where N is a fixed natural number. Then

$$\mathsf{P}(X_\tau = k) = \frac{|k|}{N(N+1)}, \quad k = -N, \dots, N.$$

Proof Denote $m_n = \min_{i \le n} X_i$. For $k \in \mathbb{Z}$, denote $T_k = \inf\{n : X_n = k\}$. Due to the martingale property of X,

$$\mathsf{P}(m_{T_k} \le l) = \mathsf{P}(T_l < T_k) = \frac{k}{k-l}, \quad l \le 0 < k.$$

Consequently,

$$\mathsf{P}(m_{T_k} = l) = \mathsf{P}(m_{T_k} \le l) - \mathsf{P}(m_{T_k} \le l-1) = \frac{k}{(k-l)(k-l+1)}, \quad l \le 0 < k.$$

As a result,

$$\mathsf{P}(X_\tau = k) = \mathsf{P}(m_{T_k} = k - N) = \frac{k}{N(N+1)}, \quad k = 1, \dots, N.$$

By the symmetry, we get the desired statement for $k = -N, \dots, -1$. □

Proof of Theorem 1 In view of Remark 1, we have to prove only (i) and (iii).

(i) Step 1. Let $(\Omega', \mathscr{F}', (\mathscr{F}'_t)_{t \ge 0}, \mathsf{P}')$ be another probability space with an $(\mathscr{F}'_t)$-Brownian motion B. Consider the enlarged filtered probability space defined by

$$\widetilde{\Omega} = \Omega \times \Omega', \qquad \widetilde{\mathscr{F}} = \mathscr{F} \times \mathscr{F}', \qquad \widetilde{\mathscr{F}}_t = \mathscr{F}_t \times \mathscr{F}'_t, \qquad \widetilde{\mathsf{P}} = \mathsf{P} \times \mathsf{P}'.$$

Then the process

$$\widetilde{S}_t = S_{t \wedge \tau} + \int_0^t I(s \ge \tau) dB_s, \quad t \ge 0$$

is an $(\widetilde{\mathscr{F}}_t, \mathsf{P})$-continuous local martingale that coincides with S up to time τ and satisfies $\langle S \rangle_\infty = \infty$. It is sufficient to prove the desired statement for $\widetilde{S}$ instead of S. Hence, we can assume from the outset that $\langle S \rangle_\infty = \infty$.

Step 2. For $N \in \mathbb{N}$, define the stopping times $\eta_0^N = 0$,

$$\eta_{n+1}^N = \inf\{t \geq \eta_n^N : |S_t - S_{\eta_n^N}| = N^{-1}L\}, \quad n \in \mathbb{N}.$$

Due to the assumption $\langle S \rangle_\infty = \infty$, all the stopping times η_n^N are finite a.s. It follows from the optional stopping theorem that the sequence

$$Y_n^N = S_{\eta_n^N}, \quad n = 0, 1, \ldots$$

is a symmetric random walk (multiplied by $N^{-1}L$). Consider

$$R_n^N = \max_{i \leq n} Y_i^N - \min_{i \leq n} Y_i^N, \quad n = 0, 1, \ldots,$$

$$\sigma^N = \inf\left\{n : R_n^N = L\frac{N-1}{N}\right\}, \qquad \tau^N = \eta_{\sigma^N}^N.$$

It is easy to see that

$$L\frac{N-1}{N} \leq R_{\tau^N} \leq L.$$

This yields the convergence $S_{\tau^N} \to S_\tau$ a.s. Hence, the convergence in law also holds. Employing Lemma 1, we complete the proof.

(iii) Using the same argument as in (i), Step 1, we can assume that $\langle S \rangle_\infty = \infty$. Fix $t \geq 0$ and denote

$$\widetilde{R}_s = \sup_{t \leq u \leq s} S_u - \inf_{t \leq u \leq s} S_u, \quad s \geq t,$$

$$\widetilde{\tau} = \inf\{s \geq t : \widetilde{R}_s = L\}.$$

The assumption $\langle S \rangle_\infty = \infty$ ensures that $\widetilde{\tau} < \infty$ a.s. For a.e. $\omega \in \{\tau > t\}$, the conditional distribution $\mathrm{Law}_{\mathsf{Q}}(S_u; u \geq t \mid \mathscr{F}_t)(\omega)$ is the distribution of a continuous martingale. Applying now (i), we see that, for a.e. $\omega \in \{\tau > t\}$, the conditional distribution $\mathsf{Q}_\omega = \mathrm{Law}_{\mathsf{Q}}(S_{\widetilde{\tau}} \mid \mathscr{F}_t)(\omega)$ has the form

$$\mathsf{Q}_\omega(dx) = L^{-2}|x - S_t(\omega)| I_{(S_t(\omega)-L, S_t(\omega)+L)}(x)dx.$$

A direct analysis of the path behavior shows that

$$S_\tau = \begin{cases} m_t + L & \text{if } m_t + L < S_{\widetilde{\tau}} < S_t + L, \\ S_{\widetilde{\tau}} & \text{if } M_t < S_{\widetilde{\tau}} < m_t + L, \\ M_t - L & \text{if } S_t < S_{\widetilde{\tau}} < M_t, \\ m_t + L & \text{if } m_t < S_{\widetilde{\tau}} < S_t, \\ S_{\widetilde{\tau}} & \text{if } M_t - L < S_{\widetilde{\tau}} < m_t, \\ M_t - L & \text{if } S_t - L < S_{\widetilde{\tau}} < M_t - L. \end{cases}$$

This yields the result. □

Proof of Theorem 2 Statements (i) and (ii) follow directly from Theorem 1, so we only have to prove (iii).

Step 1. Suppose that S is a Brownian motion. The process V remains the same if the filtration $(\mathscr{F}_t)$ is replaced by the natural filtration of S, so we can assume from the outset that $(\mathscr{F}_t)$ is the natural filtration of S. According to the representation theorem for Brownian martingales (see [3, Chap. V, Th. 3.4]), there exists a predictable process $(G_t)_{t\geq 0}$ such that

$$V_t = V_0 + \int_0^t G_u dS_u, \quad t \geq 0.$$

Fix $v \geq 0$ and define the stopping time

$$\sigma_v = \inf\{u \geq v : S_u = m_v \text{ or } S_u = M_v\}.$$

Consider the processes

$$\widetilde{G}_t = G_t I(t \leq \sigma_v), \quad t \geq 0,$$

$$\widetilde{H}_t = H_t I(t \leq \sigma_v), \quad t \geq 0,$$

$$\widetilde{V}_t = V_0 + \int_0^t \widetilde{G}_u dS_u, \quad t \geq 0,$$

$$\widetilde{X}_t = V_0 + \int_0^t \widetilde{H}_u dS_u, \quad t \geq 0.$$

On $\{v \leq t \leq \tau \wedge \sigma_v\}$, the value V_t is an affine function of S_t, and the slope of this function is exactly H_v. Therefore,

$$\widetilde{V}_t - \widetilde{V}_v = H_v(S_t - S_v), \quad v \leq t \leq \tau \wedge \sigma_v.$$

Obviously, H_t is constant on $\{v \leq t \leq \tau \wedge \sigma_v\}$, so that

$$\widetilde{X}_t - \widetilde{X}_v = H_v(S_t - S_v), \quad v \leq t \leq \tau \wedge \sigma_v.$$

As both processes $\widetilde{V}$ and $\widetilde{X}$ are stopped at time $\tau \wedge \sigma_v$, we see that

$$\widetilde{V}_t - \widetilde{V}_v = \widetilde{X}_t - \widetilde{X}_v, \quad t \geq v.$$

From this we deduce that $\widetilde{G} = \widetilde{H}$ $\mu \times \mathsf{P}$-a.e. on $[v, \infty) \times \Omega$, where μ denotes the Lebesgue measure.

Thus, we have proved that $G = H$ $\mu \times \mathsf{P}$-a.e. on every stochastic interval of the form $\{(\omega, t) : v \leq t \leq \tau(\omega) \wedge \sigma_v(\omega)\}$. By the definition of H, $H_t = 0$ on $\{t \geq \tau\}$. As V is stopped at time τ, we can assume from the outset that $G_t = 0$ on $\{t \geq \tau\}$. Thus, $G = H$ $\mu \times \mathsf{P}$-a.e. on every stochastic interval $\{(\omega, t) : v \leq t \leq \sigma_v(\omega)\}$. Obviously,

$$\{(\omega, t) : m_t(\omega) < S_t(\omega) < M_t(\omega)\} \subseteq \bigcup_{v \in \mathbb{Q}_+} \{(\omega, t) : v \leq t \leq \sigma_v(\omega)\}. \quad (1)$$

It is seen from the form of the joint law of (S_t, M_t) (see [3, Chap. III, Ex. 3.14]) that, for any $t \geq 0$, $\mathsf{P}(S_t = M_t) = 0$; similarly, $\mathsf{P}(S_t = m_t) = 0$. By the Fubini theorem,

$$\mu \times \mathsf{P}\big((\omega, t) : S_t(\omega) = m_t(\omega) \text{ or } S_t(\omega) = M_t(\omega)\big)$$
$$= \int_0^\infty \mathsf{P}(S_t = m_t \text{ or } S_t = M_t)dt = 0.$$

Thus, the set on the left-hand side of (1), and hence, the set on its right-hand side, have a complete $\mu \times \mathsf{P}$-measure. We conclude that $G = H$ $\mu \times \mathsf{P}$-a.e. This yields the desired result.

Step 2. Let us now consider the general case. Without loss of generality, $S_0 = 0$. Define the functions

$$g(S, m, M) = L^{-1}(S - m) f(m + L) - L^{-1}(M - S) f(M - L) + L^{-2} \int_{M-L}^{m} (S - x) f(x)dx + L^{-2} \int_{M}^{m+L} (x - S) f(x)dx,$$
$$h(m, M) = -L^{-1} f(M - L) + L^{-1} f(m + L) + L^{-2} \int_{M-L}^{m} f(x)dx - L^{-2} \int_{M}^{m+L} f(x)dx.$$

We can assume from the outset that S is stopped at time τ.

Consider the time change

$$T_t = \inf\{u \geq 0 : \langle S\rangle_u > t\}, \quad t \geq 0,$$

where $\inf \emptyset = \infty$, and define the process $X_t = S_{T_t}$. It follows from [3, Chap. IV, Prop. 1.13] and [3, Chap. V, Prop. 1.5] that the process X is a continuous $(\mathscr{F}_{T_t})$-local martingale with

$$\langle X\rangle_t = \langle S\rangle_{T_t} = t \wedge \langle S\rangle_\tau, \quad t \geq 0.$$

Let R^X denote the range process of X. On $\{t < \langle S\rangle_\tau\}$, we have $T_t < \tau$, and hence, $R_t^X = R_{T_t} < L$; on $\{t \geq \langle S\rangle_\tau\}$, we have $T_t = \infty$, and hence, $R_t^X = R_\infty = L$. This shows that the stopping time $\widetilde{\tau} = \inf\{t \geq 0 : R_t^X \geq L\}$ coincides with $\langle S\rangle_\tau$. Thus, $\langle X\rangle_t = t \wedge \widetilde{\tau}$. In particular, we see that X is stopped at time $\widetilde{\tau}$.

Using the same method as in the proof of Theorem 1(i), we construct (possibly, on an enlarged probability space) a Brownian motion $\widetilde{X}$ that coincides with X up to time $\widetilde{\tau}$. Then $\widetilde{\tau} = \inf\{t \geq 0 : R_t^{\widetilde{X}} \geq L\}$, where $R^{\widetilde{X}}$ is the range process of $\widetilde{X}$. Thus, X appears as the Brownian motion stopped at the first time its range exceeds L.

Now, it follows from the result of Step 1 that on the set $\{t \leq \widetilde{\tau}\}$ we have

$$g(X_t, m_t^X, M_t^X) = g(0, 0, 0) + \int_0^t h(m_u^X, M_u^X)dX_u,$$

where m^X and M^X are the running minimum and maximum of X. As X is stopped at time $\widetilde{\tau}$, we conclude that the above equality is valid for all $t \geq 0$. Thus,

$$\begin{aligned} g(S_{T_t}, m_{T_t}, M_{T_t}) &= g(X_t, m_t^X, M_t^X) \\ &= g(0,0,0) + \int_0^t h(m_u^X, M_u^X) dX_u \\ &= g(0,0,0) + \int_0^t h(m_{T_u}, M_{T_u}) dS_{T_u} \\ &= g(0,0,0) + \int_0^{T_t} h(m_u, M_u) dS_u, \quad t \geq 0. \end{aligned} \tag{2}$$

The last equality is the time change for stochastic integrals (the combination of [3, Chap. V, Prop. 1.5] with [3, Chap. IV, Prop. 1.13]).

It follows from [3, Chap. IV, Prop. 1.13] that, for a.e. ω, the path $S(\omega)$ is constant on all the intervals of constancy of $\langle S\rangle(\omega)$. Hence, the same is true for m and M. Moreover,

$$\left\langle \int_0^{\cdot} h(m_u, M_u) dS_u \right\rangle_t = \int_0^t h^2(m_u, M_u) d\langle S\rangle_u, \quad t \geq 0.$$

Hence, for a.e. ω, the path of the process $\int_0^{\cdot} h(m_u, M_u) dS_u$ is constant on all the intervals of constancy of $\langle S\rangle(\omega)$. The set of points $\{T_t(\omega); t \geq 0\}$ occupies the whole $\mathbb{R}_+$ except for the intervals of constancy of $\langle S\rangle(\omega)$, but contains all the right endpoints of those intervals. Now, we get from (2) the desired result. □

3 Conclusion

We have established the laws of a continuous local martingale and its maximum at a stopping time which is the first time its range reaches a given level. We apply this result to compute the price of options which mature at that stopping time, only under the assumptions of no interest rate, frictionless market, no arbitrage, path continuity, and finiteness of the stopping time. The price is model-free in the sense that it does not depend on the price process. The option is perfectly hedgeable, the hedge is model-free, and it needs rebalancing only when the current minimum or maximum is changed.

Acknowledgements We are thankful to the volume Editor and to the Referee for the careful reading of the manuscript and useful suggestions. We thank David Hobson for having proposed a problem that actually was the starting point for this investigation. The problem (which the Reader is invited to solve before reading the paper) is as follows. There are N people sitting around a round table. The people are numbered $1, \ldots, N$ clockwise. A plate arrives and the first person gets it. After that, the plate performs a symmetric random walk, i.e. each person receiving the plate transfers it to one of his neighbors with probability $1/2$. The random walk stops at the first time everyone has touched the plate. Consider the function

$$p(n) = \mathsf{P}\,(\text{the } n\text{-th person is the last to touch the plate}), \quad n = 2, \ldots, N.$$

The question is at which point/points p attains its maximum.

References

1. Bick, A.: Quadratic-variation-based dynamic strategies. Manag. Sci. **41**, 722–732 (1995)
2. Borodin, A., Salminen, P.: Handbook of Brownian Motion—Facts and Formulae. Birkhäuser, Basel (2002)
3. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)

Differentiability Properties of Utility Functions

Freddy Delbaen

Abstract We investigate differentiability properties of monetary utility functions. At the same time we give a counter-example—important in finance—to automatic continuity for concave functions.

Keywords Risk measures · Monetary utility functions · Differentiability · Automatic continuity

Mathematics Subject Classification (2000) 91B30 · 52A07

1 Notation and Preliminaries

We use standard notation. The triple $(\Omega, \mathscr{F}, \mathbb{P})$ denotes an atomless probability space. In practice this is not a restriction since the property of being atomless is equivalent to the fact that on $(\Omega, \mathscr{F}, \mathbb{P})$, we can define a random variable that has a continuous distribution function. We use the usual—unacceptable—convention to identify a random variable with its class (defined by a.s. equality). By L^p we denote the standard spaces, i.e. for $0 < p < \infty$, $X \in L^p$ if and only if $\mathbb{E}[|X|^p] < \infty$. L^0 stands for the space of all random variables and L^∞ is the space of bounded random variables. The topological dual of L^∞ is denoted by **ba**, the space of bounded finitely additive measures μ defined on $\mathscr{F}$ with the property that $\mathbb{P}[A] = 0$ implies $\mu(A) = 0$. The subset of normalised non-negative finitely additive measures—the so called finitely additive probability measures—is denoted by $\mathscr{P}^{\mathbf{ba}}$. The subset of countably additive elements of $\mathscr{P}^{\mathbf{ba}}$ (a subset of L^1) is denoted by $\mathscr{P}$.

Definition 1 A function $u\colon L^\infty \to \mathbb{R}$ is called a monetary utility function if

(1) $\xi \in L^\infty$ and $\xi \geq 0$ implies $u(\xi) \geq 0$,
(2) u is concave,
(3) for $a \in \mathbb{R}$ and $\xi \in L^\infty$ we have $u(\xi + a) = u(\xi) + a$.

If u moreover satisfies $u(\lambda\xi) = \lambda u(\xi)$ for $\lambda \geq 0$ (positive homogeneity), we say that u is coherent.

The author thanks Credit Suisse for support of his research. Also the support of the NCCR programme FinRisk is appreciated. The paper only reflects the personal opinion of the author.

F. Delbaen (✉)
Department of Mathematics, ETH Zurich, Rämistr. 101, 8092 Zurich, Switzerland
e-mail: delbaen@math.ethz.ch

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_3,

Remark 1 It is not difficult to prove that monetary concave utility functions satisfy a monotonicity property: $\xi \le \eta$ implies $u(\xi) \le u(\eta)$. Consequently we get that $|u(\xi) - u(\eta)| \le \|\xi - \eta\|_\infty$ and $|u(\xi)| \le \|\xi\|_\infty$. With a concave monetary utility function u we associate the convex set $\mathcal{A} = \{\xi \mid u(\xi) \ge 0\}$. This set is necessarily closed for the norm topology of L^∞.

Following Föllmer-Schied [8], we can characterise a concave monetary utility function by its Fenchel-Legendre transform. This transform, denoted by c, is defined on **ba** but outside $\mathscr{P}^{\mathbf{ba}}$ it takes the value $+\infty$. The function $c : \mathscr{P}^{\mathbf{ba}} \to \mathbb{R}_+ \cup \{+\infty\}$ is defined as $c(\mu) = \sup\{\mu(-\xi) \mid \xi \in \mathcal{A}\} = \sup\{u(\xi) - \mu(\xi) \mid \xi \in L^\infty\}$. The function c satisfies

(1) c is convex
(2) it is lower semi-continuous for the weak* topology on $\mathscr{P}^{\mathbf{ba}}$
(3) $\inf_\mu c(\mu) = \min_\mu c(\mu) = 0$.

In case u is coherent we have that c is the indicator of a weak*-closed convex set $\mathcal{S}^{\mathbf{ba}} \subset \mathscr{P}^{\mathbf{ba}}$, this means $c(\mu) = 0$ for $\mu \in \mathcal{S}^{\mathbf{ba}}$ and is $+\infty$ elsewhere. In most applications the utility function will satisfy the following continuity property:

Definition 2 The concave monetary utility function u satisfies the Fatou property if for uniformly bounded sequences ξ_n of L^∞ the convergence in probability of ξ_n to ξ implies $u(\xi) \ge \limsup u(\xi_n)$.

Remark 2 Using the Krein-Smulian theorem one can see that the Fatou property is equivalent to $\mathscr{A}$ being weak*-closed. In this case we get that (see [2] and [4] and see [11] for details on convex analysis):

(1) $u(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) \mid \mathbb{Q} \in \mathscr{P}\}$,
(2) the set $\{(\mu, t) \mid t \ge c(\mu); \mu \in \mathscr{P}^{\mathbf{ba}}\}, t \in \mathbb{R}$ is the weak* closure (in $\mathbf{ba} \times \mathbb{R}$) of the set $\{(\mathbb{Q}, t) \mid t \ge c(\mathbb{Q}); \mathbb{Q} \in \mathscr{P}, t \in \mathbb{R}\}$,
(3) the previous property can be written as: for each $\mu \in \mathscr{P}^{\mathbf{ba}}$ there is a generalised sequence or net $(\mathbb{Q}_\alpha)_\alpha$ in $\mathscr{P}$ so that $\mathbb{Q}_\alpha$ converges weak* to μ and $c(\mu) = \lim_\alpha c(\mathbb{Q}_\alpha)$. This property is even equivalent to the Fatou property.
(4) If u is coherent and satisfies the Fatou property, then the set $\mathscr{S} = \mathscr{S}^{\mathbf{ba}} \cap L^1$ satisfies $u(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] \mid \mathbb{Q} \in \mathscr{S}\}$ and $\mathscr{S}$ is weak* dense in $\mathscr{S}^{\mathbf{ba}}$. This property is even equivalent to the Fatou property.

The aim of this paper is to clarify some issues on the differentiability of monetary concave utility functions u. The differentiability of utility functions is related to equilibrium prices and plays a big role in economic theory. The Gateaux differentiability of utility functions was used in the paper of Gerber and Deprez [6]. They pointed out—without referring to any topology on the underlying space—that premium calculation principles can be derived from such utility functions and they also gave examples. This is in line with general micro-economic principles. Although Gateaux differentiability can be defined without any reference to a topology, the topological properties of the underlying space cannot be avoided. This will become clear when we give some examples. For the moment let us remark that the points where a continuous concave function defined on a *separable* Banach space, is differentiable, form a dense G_δ set. This famous theorem (due to Mazur [10]) does not generalise to general spaces. Especially for L^∞ we will give natural examples of coherent utility functions (with the Fatou property) that are nowhere differentiable.

The study of the differentiability of a monetary utility function can be restricted to the point 0. Indeed if g_0 is another point, then the differentiability at g_0 is the same problem as for the point $g_0 - u(g_0)$. This already means that we may suppose that $u(g_0) = 0$. So let us introduce the new monetary utility function defined as $u_0(\xi) = u(\xi + g_0)$, the corresponding penalty function is then given by $c_0(\mu) = c(\mu) + \mu(g_0)$. From convex analysis, see [11] and [10], we learn that a concave monetary utility function is Gateaux differentiable at ξ if and only if there is exactly one element $\mu \in \mathscr{P}^{\mathbf{ba}}$ such that $u(\xi) = \mu(\xi) + c(\mu)$. This unique element is then the Gateaux derivative. If u is supposed to be coherent, we must have that μ is an extreme (even exposed) point in the set $\mathscr{S}^{\mathbf{ba}}$. For $\xi = 0$ this means that there is exactly one element μ such that $c(\mu) = 0$. For coherent utility functions we immediately get that u can only be Gateaux differentiable at 0 if $\mathscr{S}^{\mathbf{ba}}$ is reduced to one point, i.e. u is linear.

Some of the proofs below use a trick, called homogenisation. This allows the concave utility function to be replaced by a coherent one, on the cost of enlarging the space Ω. We will sketch how this works, leaving most of the elementary details to the reader. The trick is probably not new, it is certainly not very deep but it has some "didactical" values.

We replace the set Ω by the set $\Omega_1 = \Omega \cup \{p\}$ where p is an element not in Ω, e.g. $p = \{\Omega\}$. The sigma algebra, $\mathscr{F}_1$, on Ω_1 is generated by the sets $A \in \mathscr{F}$ and the set $\Omega \subset \Omega_1$. The probability $\mathbb{P}$ is replaced by the probability $\mathbb{P}_1$ defined as $\mathbb{P}_1[A_1] = \frac{1}{2}\mathbb{P}[A_1 \cap \Omega] + \frac{1}{2}\mathbf{1}_{A_1}(p)$. Probabilities on Ω_1 are convex combinations of probabilities of Ω and the Dirac measure Δ_p concentrated in the extra point p. The utility function defined on $L^\infty(\Omega_1) = L^\infty(\Omega) \times \mathbb{R}$, is defined via the acceptance cone $\mathscr{A}_1$. The latter is defined as the norm closed cone generated by the elements of the form $(f, 1)$ where $f \in \mathscr{A}$. An element of the cone $\mathscr{A}_1$ is either of the form (tf, t) with $t > 0$ and $f \in \mathscr{A}$ or of the form $(f, 0)$ where f is in the recession cone (asymptotic cone) of $\mathscr{A}$. This cone is defined as $\bigcap_{t>0} t\mathscr{A}$. The utility function u_1 defined on $L^\infty(\Omega_1)$ does not have an easy expression as a function of u. But we have $u(f) \geq 0$ if and only if $u_1(f, 1) \geq 0$ and $u(f) = 0$ if and only if $u_1(f, 1) = 0$. Important is to note that when $\mathscr{A}$ is weak* closed then also $\mathscr{A}_1$ is weak* closed. This follows from the expression for the recession cone. So u_1 is Fatou when u is. Since we will only apply this trick for Fatou utility functions, we will suppose that u and hence u_1, has the Fatou property. The scenario set $\mathscr{S}_1$ for u_1, a subset of $L^1(\Omega_1)$, is represented as $(\alpha\mathbb{Q}, (1-\alpha)\Delta_p)$ where

(1) $0 \leq \alpha \leq 1$
(2) $\mathbb{Q} \in \mathscr{P}$
(3) for all $\xi \in \mathscr{A}$ and all $t > 0$, we have $\alpha\mathbb{E}_{\mathbb{Q}}[\xi]t + (1-\alpha)t \geq 0$. This is equivalent to $\alpha \leq \frac{1}{1+c(\mathbb{Q})}$. If $c(\mathbb{Q}) = \infty$ this means that the measure becomes Δ_p.

Another representation of the set $\mathscr{S}_1$ is through the Radon-Nikodym derivative with respect to $\mathbb{P}_1$. We find that

$$\mathscr{S}_1 = \left\{ (f, \beta) \mid 0 \le f \in L^1(\Omega),\ \frac{2c(\mathbb{Q})}{c(\mathbb{Q})+1} \ge \beta \ge 0,\ \mathbb{E}[f] + \beta = 2 \right\},$$

where the measure $\mathbb{Q}$ is defined as $d\mathbb{Q} = \frac{f}{\mathbb{E}[f]} d\mathbb{P}$ for $\mathbb{E}[f] > 0$. If $f = 0$ we simply ask that $\beta = 2$ since in this case we get the Dirac measure Δ_p.

2 The Jouini-Schachermayer-Touzi Theorem

This section is devoted to the characterisation of a weak compactness theorem. The theorem is a generalisation of the beautiful result of James' on weakly compact sets, see [7]. The original proof of [9] followed the rather complicated proof of James. The proof below uses the homogenisation trick and allows to apply the original version of James' theorem. Let us recall this theorem

Theorem 1 *Let E be a Banach space and let C be a convex closed subset of E. A necessary and sufficient condition for C to be weakly compact in E is that every continuous linear function $e^* \in E^*$ attains its maximum on C.*

In one direction the statement is trivial: if C is weakly compact then every continuous linear functional attains its maximum on C. From a topological viewpoint the converse is surprising. First a topological space for which every real-valued continuous function attains its maximum need not be compact, second we only need linear functions, a class that is not dense in the space of continuous functions defined on C.

In the following theorem we use the same notation as in the homogenisation.

Theorem 2 (Jouini-Schachermayer-Touzi [9]) *If u is a concave monetary utility function satisfying the Fatou property then are equivalent*:

(1) *For each $\xi \in L^\infty$ there is a $\mathbb{Q} \in \mathscr{P}$ such that $u(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q})$.*
(2) *If $(\xi, t) \in L^\infty(\Omega_1)$ there is a $\mathbb{Q}_1 \in \mathscr{S}_1$ such that $u_1(\xi, t) = \int_{\Omega_1} (\xi, t) d\mathbb{Q}_1$.*
(3) *The set $\mathscr{S}_1$ is weakly compact in $L^1(\Omega_1)$.*
(4) *The homogenisation u_1 satisfies the Lebesgue property. This means that for uniformly bounded sequences $(\theta_n)_n$ in $L^\infty(\Omega_1)$, converging in probability to say θ, we have $u_1(\theta_n) \to u_1(\theta)$.*
(5) *If ξ_n is a uniformly bounded sequence in L^∞ converging in probability to a function ξ, then $u(\xi_n) \to u(\xi)$, i.e. u has the Lebesgue property.*
(6) *For each $k \in \mathbb{R}$ the set $\{\mathbb{Q} \mid c(\mathbb{Q}) \le k\}$ is weakly compact (or uniformly integrable and closed) in L^1, in particular $c(\mu) = +\infty$ for non countably additive elements of $\mathscr{P}^{\mathbf{ba}}$.*

In the sequel we will say that such utility functions satisfy the weak compactness property. For coherent utility functions with the Fatou property, the property means $\mathscr{S}^{\mathbf{ba}} = \mathscr{S}$ is a weakly compact convex set in L^1.

Remark 3 A simple reasoning shows that item (6) is implied by: "There is $k > 0$ such that $\{\mathbb{Q} \mid c(\mathbb{Q}) \le k\}$ is weakly compact".

Proof The proof is divided into different steps (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (1). We start with (1) $\Rightarrow$ (2). Let us consider an element (ξ, t) in $L^\infty(\Omega_1)$. We have to find an element of $\mathscr{S}_1$ so that the minimum is attained. It does not matter if we replace (ξ, t) by an element of the form $(\xi + \eta, t + \eta)$ where $\eta \in \mathbb{R}$. So we may suppose that $u_1(\xi, t) = 0$. There are two possibilities. If $t = 0$ this means that ξ is in the recession cone and hence $\mathbb{E}_\mathbb{Q}[\xi] \ge 0$ whenever $c(\mathbb{Q}) < \infty$. In this case we may take the measure Δ_p to realise the minimum. The other case is when $t > 0$. In this case we may multiply by $1/t$ to get an element $(\xi, 1)$ with $u_1(\xi, 1) = 0$. This implies $u(\xi) = 0$. By hypothesis there is an element $\mathbb{Q}$ such that $\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) = 0$. The element $\mathbb{Q}_1 = \frac{1}{1+c(\mathbb{Q})}\mathbb{Q} + \frac{c(\mathbb{Q})}{1+c(\mathbb{Q})}\Delta_p$ then gives $\mathbb{E}_{\mathbb{Q}_1}[(\xi, 1)] = 0$. This is then a minimum since all other elements of $\mathscr{S}_1$ will give a larger expected value. The implication (2) $\Rightarrow$ (3) is the famous James' theorem. The implication (3) $\Rightarrow$ (4) is standard since weakly compact sets are uniformly integrable. The implication (4) $\Rightarrow$ (5) is easy. Indeed, we can take a bounded sequence ξ_n converging in probability to ξ. We may suppose (substract $u(\xi)$ if necessary), that $u(\xi) = 0$. Then we use that $(\xi_n, 1)$ converges in probability ($\mathbb{P}_1$) to $(\xi, 1)$. This gives $u_1(\xi_n, 1)$ will tend to zero. But as seen above this means that $u(\xi_n)$ tends to zero as well. For the implication (5) $\Rightarrow$ (6) we only need to show that if A_n is a decreasing sequence of sets with $\mathbb{P}[A_n] \to 0$, then $\sup\{\mathbb{Q}[A_n] \mid c(\mathbb{Q}) \le k\}$ tends to zero. Let us put $\xi_n = -\alpha \mathbf{1}_{A_n}$ where $\alpha > 0$. Let us apply item (5). We get that $u(\xi_n)$ tends to zero, or $\inf_\mathbb{Q}(\mathbb{E}_\mathbb{Q}[-\alpha\mathbf{1}_{A_n}] + c(\mathbb{Q}))$ tends to zero. This implies that $\liminf_{n\to\infty} \inf_{c(\mathbb{Q})\le k}(\mathbb{E}_\mathbb{Q}[-\alpha\mathbf{1}_{A_n}] + c(\mathbb{Q})) \ge 0$. Hence we have $\liminf_{n\to\infty} \inf_{c(\mathbb{Q})\le k}(\mathbb{E}_\mathbb{Q}[-\alpha\mathbf{1}_{A_n}] + k) \ge 0$. This is the same as $\limsup_n \sup_{c(\mathbb{Q})\le k} \mathbb{Q}[A_n] \le k/\alpha$. Since α can be taken arbitrary large, we get that $\lim_n \sup_{c(\mathbb{Q})\le k} \mathbb{Q}[A_n] = 0$. The implication (6) $\Rightarrow$ (1) is standard. Let us take $\xi \in L^\infty$. In the calculation of $u(\xi)$ we only need to take $\mathbb{Q}$ with $c(\mathbb{Q}) \le 2\|\xi\|_\infty$. Since this set is weakly compact and since c is lower semi-continuous, we will find an element realising the infimum. □

3 A Consequence of Ekeland's Variational Principle and Other Family Members of Bishop-Phelps

Differentiability properties of convex functions are well studied. We will use [10] as a basic reference. In our context there are two functions that are important: the utility function u and the penalty function c. The function u is defined on the whole space whereas the function c is only defined on a (subset of) $\mathscr{P}$ or $\mathscr{P}^{\mathbf{ba}}$. The subgradient of u at a point ξ is defined as the set of elements $\nu \in \mathbf{ba}$ such that $u(\eta) \le u(\xi) + \nu(\eta - \xi)$. It is also equal to the set

$$\partial u(\xi) = \{\mu \in \mathscr{P}^{\mathbf{ba}} \mid u(\xi) = \mu(\xi) + c(\mu)\}.$$

At the same time we find that the subgradient of c at a point $\mathbb{Q} \in \mathscr{P}$ is the set

$$\{\xi \in L^\infty \mid u(\xi) = \mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q})\}.$$

The Borwein-Rockafellar theorem (see [10]) now shows that the set

$$\{\xi \mid \text{ there is } \mathbb{Q} \in \mathscr{P} \text{ with } u(\xi) = \mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q})\}$$

is norm dense in L^∞. The function u is Gateaux differentiable at ξ if the set $\partial u(\xi)$ is a singleton.

4 A Consequence of Automatic Continuity

Theorem 3 *Suppose that u is a monetary concave utility function satisfying the Fatou property. If u is Gateaux differentiable at g, then its derivative is an element of $\mathscr{P}$, i.e. it is countably additive.*

Proof Since u is Fatou, it is a Borel function for the weak* topology on L^∞. Suppose that $\mu \in \mathscr{P}^{\mathbf{ba}}$ is the derivative of u at the point g, then we have that μ is the pointwise limit of the *sequence*

$$\mu(f) = \lim_n n\left(u\left(g + \frac{1}{n} f\right) - u(g)\right).$$

The linear function μ is then a Borel measurable function for the weak* topology and hence, by the automatic continuity theorem [1], it must be weak* continuous, i.e. induced by an element of L^1. □

Corollary 1 *Under the extra hypothesis that u is also coherent, we have that the derivative $\mathbb{Q}$ is an extreme (even exposed) point of $\mathscr{S}^{\mathbf{ba}}$, but already lying in $\mathscr{S}$.*

5 The One-Sided Derivative

Because of concavity, monetary concave utility functions have a one-sided derivative at a point $g \in L^\infty$, defined as

$$\varphi_g(f) = \lim_{\varepsilon \downarrow 0} \frac{u(g + \varepsilon f) - u(g)}{\varepsilon}.$$

If $g = 0$ we get

$$\varphi(f) = \lim_{\varepsilon \downarrow 0} \frac{u(\varepsilon f)}{\varepsilon}.$$

Proposition 1 *The function φ is the smallest coherent utility function ψ such that $\psi \geq u$.*

Proof This is easy since for each $\varepsilon > 0$ we have $\psi(f) = \psi(\varepsilon f)/\varepsilon \geq u(\varepsilon f)/\varepsilon$. Taking limits gives $\psi(f) \geq \varphi(f)$. □

The acceptance cone for φ is easily obtained.

Proposition 2 *The acceptance cone of φ, $\mathscr{A}_\varphi$, is given by the $\|\cdot\|_\infty$ closure of the union $\bigcup_n n\mathscr{A}$ (where as before $\mathscr{A} = \{f \mid u(f) \geq 0\}$).*

Proof Suppose first that $f \in n\mathscr{A}$, then for $\varepsilon \leq 1/n$ we have by convexity of $\mathscr{A}$ that $u(\varepsilon f) \geq 0$. This shows that $\varphi(f) \geq 0$. It follows that $\bigcup_n n\mathscr{A} \subset \mathscr{A}_\varphi$. Since the latter set is norm closed we have that it also must contain the norm-closure of this union. If $\varphi(f) > 0$, we have that for ε small enough $u(\varepsilon f) > 0$ and hence for n big enough we have $f \in n\mathscr{A}$. This shows the opposite inclusion. □

The scenario set that defines the coherent utility function φ is given by the following theorem

Theorem 4 *With the notation introduced above we have*

$$\varphi(f) = \inf_{\mu \in \mathscr{S}^{\mathbf{ba}}} \mu(f),$$

where the set $\mathscr{S}^{\mathbf{ba}}$ is defined as $\mathscr{S}^{\mathbf{ba}} = \{\mu \in \mathscr{P}^{\mathbf{ba}} \mid c(\mu) = 0\}$.

Proof Because of the previous proposition we have that $\mu \in \mathscr{S}^{\mathbf{ba}}$ if and only if $\mu(f) \geq 0$ for all $f \in \mathscr{A}$. This is equivalent to saying that $c(\mu) = 0$. □

Corollary 2 *The one-sided derivative φ of u at 0 is Fatou if and only if $\{\mathbb{Q} \in \mathscr{P} \mid c(\mathbb{Q}) = 0\}$ is weak* dense in $\{\mu \mid c(\mu) = 0\}$.*

The previous corollary allows for easy constructions of non-Fatou coherent utility functions. For the derivative at a point g we use the transformation $u_g(f) = u(g + f) - u(g)$. It follows that the derivative at a point g is given by

$$\varphi_g(f) = \inf\{\mu(f) \mid c(\mu) + \mu(g) = u(g)\}.$$

6 An Example

Let us fix a countable partition of Ω into a sequence of measurable sets A_n with $\mathbb{P}[A_n] > 0$. For $\mu \in \mathscr{P}^{\mathbf{ba}}$ we define

$$c(\mu) = \sum_n \mu[A_n]^2.$$

Proposition 3 *The function c is convex, takes values in* $[0,1]$, $\min_{\mu\in\mathscr{P}^{\mathbf{ba}}} c(\mu)=0$ *and is lower semi-continuous for the weak* topology on* $\mathscr{P}^{\mathbf{ba}}$. *The utility function u defined by c is Fatou.*

Proof The first four statements are obvious since the mapping $\mu\to\mu[A_n]^2$ is convex and weak* continuous. c is therefore the increasing limit of a sequence of continuous convex functions and hence is lower semi-continuous and convex. The existence of elements in $\mathscr{P}^{\mathbf{ba}}$ so that for all n, $\mu(A_n)=0$ is well known and can be proved using the Hahn-Banach theorem. Of course we have for $\mu\in\mathscr{P}^{\mathbf{ba}}$: $\sum_n\mu(A_n)^2\le\sum_n\mu(A_n)\le 1$. The Fatou property is less trivial. As seen before we must show that for $\mu\in\mathscr{P}^{\mathbf{ba}}$ we can find a generalised sequence or net $\mathbb{Q}_\alpha$ in $\mathscr{P}$ so that $c(\mathbb{Q}_\alpha)$ tends to $c(\mu)$. For this it is sufficient to show the following. Given μ, given $\varepsilon>0$ and given a finite partition of Ω in non-zero sets $B_1,\dots,B_N$ we must find $\mathbb{Q}\in\mathscr{P}$ so that $c(\mathbb{Q})\le c(\mu)+\varepsilon$ and $\mathbb{Q}(B_j)=\mu(B_j)$ for $j=1,\dots,N$. For a set B_j there are two possibilities: either there is s with $B_j\subset\bigcup_{n=1}^s A_n$ or there are infinitely many indices n with $\mathbb{P}[B_j\cap A_n]>0$. Since all the sets A_n have a non-zero measure and since the family $(B_j)_j$ forms a partition of Ω the last alternative must occur for at least one index j. So let us renumber the sets B_j and let us select s so that

(1) for $j\le N'\le N$ there are infinitely many indices with $\mathbb{P}[A_n\cap B_j]>0$,
(2) for $N'<j\le N$ (if any) we have that $B_j\subset\bigcup_{n=1}^s A_n$.

Fix now an integer $L\ge 1$ so that $1/L\le\varepsilon$. We will define the measure $\mathbb{Q}$ by its Radon-Nikodym density. For $j\le N'$ we find indices as follows, we take L indices $s<n_1^1<n_2^1<\cdots<n_L^1$ so that $\mathbb{P}[A_{n_k^1}\cap B_1]>0$. We then take indices $n_L^1<n_1^2<n_2^2<\cdots<n_L^2$ with $\mathbb{P}[A_{n_k^2}\cap B_2]>0$ and so on. We can now define the density of $\mathbb{Q}$ as

$$\frac{d\mathbb{Q}}{d\mathbb{P}}=\sum_{j=1}^N\sum_{k=1}^s\frac{\mu(B_j\cap A_k)}{\mathbb{P}[B_j\cap A_k]}\mathbf{1}_{B_j\cap A_k}+\sum_{j=1}^N\sum_{p=1}^L\frac{\mu(B_j\cap(\bigcup_{n>s}A_n))}{L\mathbb{P}[B_j\cap A_{n_p^j}]}\mathbf{1}_{B_j\cap A_{n_p^j}}.$$

The reader can convince himself that there is no reason to drop the terms with denominator zero. For all $j\le N$ we have that $\mathbb{Q}[B_j]=\mu(B_j)$. Furthermore we have that for $n\le s:\mathbb{Q}[A_n]=\mu(A_n)$. For indices $n>s$ there is at most one of the N sets $B_j\cap A_n$ that is chosen. So we get for $n>s$:

$$\mathbb{Q}[A_{n_p^j}]=\frac{1}{L}\mu\left(B_i\cap\left(\bigcup_{n>s}A_n\right)\right)\quad\text{and for other indices } n \text{ we get } =0.$$

Finally we find

$$\begin{aligned}c(\mathbb{Q})&=\sum_n\mathbb{Q}[A_n]^2=\sum_{n\le s}\mathbb{Q}[A_n]^2+\sum_{n>s}\mathbb{Q}[A_n]^2\\&=\sum_{n\le s}\mu(A_n)^2+\sum_{n>s}\mathbb{Q}[A_n]^2\end{aligned}$$

$$\leq c(\mu) + \sum_{j=1}^{N} \sum_{p=1}^{L} \frac{1}{L^2} \mu\left(B_j \cap \left(\bigcup_{n>s} A_n\right)\right)^2$$

$$\leq c(\mu) + \frac{1}{L} \sum_{j=1}^{N} \mu(B_j)^2$$

$$\leq c(\mu) + \varepsilon. \qquad \square$$

Corollary 3 *The scenario set* $\mathcal{S}^{\mathbf{ba}}$ *for* φ *consists of purely finitely additive measures* $\mathcal{S}^{\mathbf{ba}} = \{\mu \in \mathcal{P}^{\mathbf{ba}} \mid$ *for all* $n : \mu(A_n) = 0\}$. *Consequently the coherent utility function* φ *is not Fatou.*

Remark 4 The coherent utility function φ is weak* Borel measurable since it is the limit of a sequence of Borel measurable functions. The acceptance cone $\mathcal{A}_\varphi$ is norm closed, is Borel measurable but is not weak* closed. With some little effort one can show that $\mathcal{A}_\varphi$ is a weak* $F_{\sigma\delta}$. This shows—the already known fact—that the automatic continuity theorems do not apply to concave or convex functions.

Proposition 4 *For all* s *we have* $u(\mathbf{1}_{\bigcup_{n=1}^{s} A_n}) = 0$ *also for each* $2 \geq \varepsilon \geq 0$ *we have* $u(-\mathbf{1}_{A_n}) = -\frac{\varepsilon^2}{4}$. *For* $\varepsilon \geq 2$ *we have* $u(-\varepsilon \mathbf{1}_{A_n}) = -\varepsilon + 1$.

Proof The first statement is easy since it is sufficient to take an element $\mu \in \mathcal{P}^{\mathbf{ba}}$ with $\mu(A_n) = 0$ for all n. For the other equality, we may without loss of generality suppose that $n = 1$. Let us write

$$u(-\varepsilon \mathbf{1}_{A_1}) \leq \mu(-\varepsilon \mathbf{1}_{A_1}) + \sum_n \mu(A_n)^2 = \mu(-\varepsilon \mathbf{1}_{A_1}) + \mu(A_1)^2 + \sum_{n\geq 2} \mu(A_n)^2,$$

and observe that the minimum is taken for elements μ satisfying $\mu(A_1) = \varepsilon/2$ and $\mu(A_k) = 0$ for $k \geq 2$. For $\varepsilon \geq 2$ we find that $u(-\varepsilon \mathbf{1}_{A_1}) = -\varepsilon + 1$ and the minimum is taken for elements μ with $\mu(A_1) = 1$. $\square$

Remark 5 The last line of the proof shows that for some points g, the set $\{\mathbb{Q} \mid \mathbb{Q} \in \mathcal{P}, \mathbb{E}_{\mathbb{Q}}[g] + c(\mathbb{Q}) = u(g)\}$ is weak* dense in $\{\mu \mid \mu \in \mathcal{P}^{\mathbf{ba}}, \mu(g) + c(\mu) = u(g)\}$. Indeed this is the case for $\alpha \mathbf{1}_{A_n}$ with $\alpha \leq -2$. This means that the one-sided derivative at these points g is Fatou.

Proposition 5 *For* $f \in L^\infty$ *we have*

$$\varphi(f) = \liminf_n \operatorname{ess\,inf}(\mathbf{1}_{A_n} f) = \liminf_n \operatorname{ess\,inf}(\mathbf{1}_{\bigcup_{m\geq n} A_m} f).$$

Proof The proof follows from the description of $\mathcal{S}^{\mathbf{ba}}$. We leave the details to the reader. $\square$

Remark 6 The previous corollary shows that the function φ is related to the function lim inf on the space l^∞. This function is also Borel measurable for the weak* topology. Its set of representing measures is the convex weak* closed hull of the Banach limits. None of these representing measures is Borel measurable.

7 The Example of an Incomplete Financial Market

In this section we take a filtered probability space $(\Omega, (\mathscr{F}_t)_{t\in[0,1]}, \mathbb{P})$. We suppose that the filtration is continuous meaning that all martingales are continuous, equivalently that all stopping times are predictable. On this space we consider a d-dimensional *continuous* price process $S : [0, 1] \times \Omega \to \mathbb{R}^d$. We assume that the price process S satisfies the *NFLVR* property (see [5]). More precisely and to make the notation easier we will suppose that the process S is a bounded martingale for $\mathbb{P}$. The market generated by S is supposed to be incomplete which in this case means that the set $\mathbb{M}^a = \{\mathbb{Q} \ll \mathbb{P} \mid S \text{ ia a } \mathbb{Q}\text{-martingale}\}$ is strictly bigger than $\{\mathbb{P}\}$. From [3] we know that the set $\mathbb{M}^a$ is then a set without extreme points. The utility function u is defined as the bid price

$$u(\xi) = \inf\{\mathbb{E}_{\mathbb{Q}}[\xi] \mid \mathbb{Q} \in \mathbb{M}^a\}.$$

This is a Fatou coherent utility function. If u were Gateaux differentiable at a point ξ, then by the automatic continuity theorem we have that its derivative would be an extreme point of $\mathbb{M}^a$. Since this set has no extreme points we get

Theorem 5 *The function u is nowhere Gateaux differentiable.*

References

1. Christensen, J.P.R.: Topology and Borel Structure: Descriptive Topology and Set Theory with Applications to Functional Analysis and Measure Theory. North-Holland, Amsterdam (1974)
2. Delbaen, F.: Coherent risk measures on general probability spaces. In: Sandmann, K., Schönbucher, P.J. (eds.) Advances in Finance and Stochastics, pp. 1–37 (2002)
3. Delbaen, F.: Representing martingale measures when asset prices are continuous and bounded. Math. Financ. **2**, 107–130 (1992)
4. Delbaen, F.: Coherent Risk Measures. Scuola Normale Superiore di Pisa (2000)
5. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
6. Deprez, O., Gerber, H.U.: On convex principles of premium calculation. Insur. Math. Econ. 179–189 (1985)
7. Diestel, J.: Geometry of Banach Spaces—Selected Topics. Springer, Berlin (1975)
8. Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time. de Gruyter Studies in Mathematics, vol. 27. Walter de Gruyter, Berlin (2004)
9. Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. Adv. Math. Econ. **9**, 49–71 (2006)
10. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability, 2nd edn. Springer, Berlin (1993)
11. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)

Exponential Utility Indifference Valuation in a General Semimartingale Model

Christoph Frei and Martin Schweizer

Abstract We study the exponential utility indifference valuation of a contingent claim H when asset prices are given by a general semimartingale S. Under mild assumptions on H and S, we prove that a no-arbitrage type condition is fulfilled if and only if H has a certain representation. In this case, the indifference value can be written in terms of processes from that representation, which is useful in two ways. Firstly, it yields an interpolation expression for the indifference value which generalizes the explicit formulas known for Brownian models. Secondly, we show that the indifference value process is the first component of the unique solution (in a suitable class of processes) of a backward stochastic differential equation. Under additional assumptions, the other components of this solution are *BMO*-martingales for the minimal entropy martingale measure. This generalizes recent results by Becherer (Ann. Appl. Probab. 16:2027–2054, 2006) and Mania and Schweizer (Ann. Appl. Probab. 15:2113–2143, 2005).

Keywords Exponential utility · Indifference valuation · Minimal entropy martingale measure · BSDE · *BMO*-martingales · Fundamental entropy representation (*FER*)

Mathematics Subject Classification (2000) 91B28 · 60G48

1 Introduction

One general approach to the problem of valuing contingent claims in incomplete markets is utility indifference valuation. Its basic idea is that the investor valuing a contingent claim H should achieve the same expected utility in the two cases where (1) he does not have H, or (2) he owns H but has his initial capital reduced by the amount of the *indifference value* of H. Exponential utility indifference valuation means that the utility function one uses is exponential.

This paper is dedicated to Yuri Kabanov on the occasion of his 60th birthday. We hope he likes it even if it is not short...

C. Frei · M. Schweizer (✉)
Department of Mathematics, ETH Zurich, Rämistr. 101, 8092 Zurich, Switzerland
e-mail: martin.schweizer@math.ethz.ch

C. Frei
e-mail: christoph.frei@math.ethz.ch

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_4,

Even in a concrete model, it is difficult to obtain a closed-form formula for the indifference value. The majority of existing explicit results are for Brownian settings; see for instance Frei and Schweizer [10] and the references therein. In more general situations, Becherer [2] and Mania and Schweizer [19] derive a backward stochastic differential equation (BSDE) for the indifference value process. While [19] assumes a continuous filtration, the framework in [2] has a continuous price process driven by Brownian motions and a filtration generated by these and a random measure allowing the modeling of non-predictable events.

The main contribution of this paper is to extend the above results to a setting where asset prices are given by a *general semimartingale*. We show that the exponential utility indifference value can still be written in a closed-form expression similar to that known for Brownian models, although the structure of this formula is here much less explicit. Independently from that, we establish a BSDE formulation for the dynamic indifference value process. Both of these results are based on a representation of the claim H and on the relationship between a notion of no-arbitrage, the form of the so-called minimal entropy martingale measure, and the indifference value.

As our starting point, we take the work of Biagini and Frittelli [3, 4]. Their results yield a representation of the minimal entropy martingale measure which we can use to derive a decomposition of a fixed payoff H in a similar way as in Becherer [1]. We call this decomposition, which is closely related to the minimal entropy martingale measure, the *fundamental entropy representation* of H $(FER(H))$. It is central to all our results here, because we can express the indifference value for H as a difference of terms from $FER(H)$ and $FER(0)$. We derive from this a fairly explicit formula for the indifference value by an interpolation argument, and we also establish a BSDE representation for the indifference value process. Its proof is based on the idea that the two representations $FER(H)$ and $FER(0)$ can be merged to yield a single BSDE. This direct procedure allows us to work with a general semimartingale, whereas Becherer [2] as well as Mania and Schweizer [19] use more specific models because they first prove some results for more general classes of BSDEs and then apply these to derive the particular BSDE for the indifference value. The price to pay for working in our general setting is that we must restrict the class of solutions of the BSDE to get uniqueness. Under additional assumptions, the components of the solution to the BSDE for the indifference value are again BMO-martingales for the minimal entropy martingale measure; this applies in particular to the value process of the indifference hedging strategy.

The paper is organized as follows. Section 2 lays out the model, motivates, and introduces the important notion of $FER(H)$. In Sect. 3, we prove that the existence of $FER(H)$ is essentially equivalent to an absence-of-arbitrage condition. Moreover, we develop a uniqueness result for $FER(H)$ and its relationship to the minimal entropy martingale measure. Section 4 establishes the link between the exponential indifference value of H and the two decompositions $FER(H)$ and $FER(0)$. By an interpolation argument, we derive a fairly explicit formula for the indifference value. In Sect. 5, we extend to a general filtration the BSDE representation of the indifference value by Becherer [2] and Mania and Schweizer [19]. We further provide conditions under which the components of the solution to the BSDE are *BMO*-martingales for the minimal entropy martingale measure. Section 6 rounds off with an application to a Brownian setting.

2 Motivation and Definition of *FER*(*H*)

We start with a probability space $(\Omega, \mathscr{F}, P)$, a finite time interval $[0, T]$ for a fixed $T > 0$ and a filtration $\mathbb{F} = (\mathscr{F}_t)_{0\le t\le T}$ satisfying the usual conditions of right-continuity and completeness. For simplicity, we assume that $\mathscr{F}_0$ is trivial and $\mathscr{F}_T = \mathscr{F}$. For a positive process Z, we use the abbreviation $Z_{t,s} := Z_s/Z_t$, $0 \le t \le s \le T$.

In our financial market, there are d risky assets whose price process $S = (S_t)_{0\le t\le T}$ is an $\mathbb{R}^d$-valued semimartingale. In addition, there is a riskless asset, chosen as numeraire, whose price is constant at 1. Our investor's risk preferences are given by an exponential utility function $U(x) = -\exp(-\gamma x)$, $x \in \mathbb{R}$, for a fixed $\gamma > 0$. We always consider a fixed contingent claim H which is a real-valued $\mathscr{F}$-measurable random variable satisfying $E_P[\exp(\gamma H)] < \infty$. Expressions depending on H are introduced with an index H so we can later use them also in the absence of the claim by setting $H = 0$. However, the dependence on γ is not explicitly mentioned. We define by $\frac{\mathrm{d}P_H}{\mathrm{d}P} := \exp(\gamma H)/E_P[\exp(\gamma H)]$ a probability measure P_H on $(\Omega, \mathscr{F})$ equivalent to P. Note that $P_0 = P$. We denote by $L(S)$ the set of all $\mathbb{R}^d$-valued predictable S-integrable processes, so that $\int \vartheta \,\mathrm{d}S$ is a well-defined semimartingale for each ϑ in $L(S)$.

We always impose without further mention the following *standing assumption*, introduced by Biagini and Frittelli [3, 4] for $H = 0$. We assume that

$$\mathscr{W}_H \neq \emptyset \quad \text{and} \quad \mathscr{W}_0 \neq \emptyset, \tag{1}$$

where $\mathscr{W}_H$ is the set of *loss variables* W which satisfy the following two conditions:

(1) $W \ge 1$ P-a.s., and for every $i = 1, \dots, d$, there exists some $\beta^i \in L(S^i)$ such that $P[\exists t \in [0, T] \text{ s.t. } \beta^i_t = 0] = 0$ and $|\int_0^t \beta^i_s \mathrm{d}S^i_s| \le W$ for all $t \in [0, T]$;
(2) $E_{P_H}[\exp(cW)] < \infty$ for all $c > 0$.

Clearly, $\mathscr{W}_H = \mathscr{W}_0$ if H is bounded. Lemma 1 at the beginning of Sect. 3 gives a less restrictive condition on H for $\mathscr{W}_H = \mathscr{W}_0$. The standing assumption (1) is automatically fulfilled if S is locally bounded since then $1 \in \mathscr{W}_H \cap \mathscr{W}_0$ by Proposition 1 of Biagini and Frittelli [3], using $P_H \approx P$. But (1) is for example also satisfied if H is bounded and $S = S^1$ is a scalar compound Poisson process with Gaussian jumps. This follows from Sect. 3.2 in Biagini and Frittelli [3]. So the model with condition (1) is a genuine generalization of the case of a locally bounded S.

To assign to H at time $t \in [0, T]$ a value based on our exponential utility function, we first fix an $\mathscr{F}_t$-measurable random variable x_t, interpreted as the investor's starting capital at time t. Then we define

$$V_t^H(x_t) := \operatorname*{ess\,sup}_{\vartheta \in \mathscr{A}_t^H} E_P\left[-\exp\left(-\gamma x_t - \gamma \int_t^T \vartheta_s \mathrm{d}S_s + \gamma H\right)\middle|\mathscr{F}_t\right], \tag{2}$$

where the set $\mathscr{A}_t^H$ of *H-admissible strategies* on $(t, T]$ consists of all processes $\vartheta I_{]\!]t,T]\!]}$ with $\vartheta \in L(S)$ and such that $\int \vartheta \mathrm{d}S$ is a Q-supermartingale for every $Q \in \mathbb{P}_H^{e,f}$; the set $\mathbb{P}_H^{e,f}$ is defined in the paragraph after the next. We recall that $x_t + \int_t^T \vartheta_s \mathrm{d}S_s$ is the investor's final wealth when starting with x_t and investing according to the self-financing strategy ϑ over $(t, T]$. Therefore, $V_t^H(x_t)$ is the maximal conditional expected utility the investor can achieve from the time-t initial capital x_t by trading during $(t, T]$ and paying out H (or receiving $-H$) at the maturity T.

The *indifference (seller) value* $h_t(x_t)$ at time t for H is implicitly defined by

$$V_t^0(x_t) = V_t^H(x_t + h_t(x_t)). \tag{3}$$

This says that the investor is indifferent between solely trading with initial capital x_t, versus trading with initial capital $x_t + h_t(x_t)$ but paying an additional cash-flow H at maturity T.

To define our strategies, we need the sets

$$\mathbb{P}_H^f := \big\{Q \ll P_H | I(Q|P_H) < \infty \text{ and } S \text{ is a } Q\text{-sigma-martingale}\big\},$$
$$\mathbb{P}_H^{e,f} := \big\{Q \approx P_H | I(Q|P_H) < \infty \text{ and } S \text{ is a } Q\text{-sigma-martingale}\big\},$$

where

$$I(Q|P_H) := \begin{cases} E_Q[\log \frac{\mathrm{d}Q}{\mathrm{d}P_H}] & \text{if } Q \ll P_H \\ +\infty & \text{otherwise} \end{cases}$$

denotes the relative entropy of Q with respect to P_H. Since P_H is equivalent to P, the sets $\mathbb{P}_H^f$ and $\mathbb{P}_H^{e,f}$ depend on H only through the condition $I(Q|P_H) < \infty$. By Proposition 3 and Remark 3 of Biagini and Frittelli [3], applied to P_H instead of P, there exists a unique $Q_H^E \in \mathbb{P}_H^f$ that minimizes $I(Q|P_H)$ over all $Q \in \mathbb{P}_H^f$, provided of course that $\mathbb{P}_H^f \neq \emptyset$. We call Q_H^E the *minimal H-entropy measure*, or H-MEM for short. If $\mathbb{P}_H^{e,f} \neq \emptyset$, then Q_H^E is even equivalent to P_H, i.e., $Q_H^E \in \mathbb{P}_H^{e,f}$; see Remark 2 of Biagini and Frittelli [3]. Note that the proper terminology would be "minimal H-entropy sigma-martingale measure" or H-MEσMM, but this is too long.

We briefly recall the relation between Q_H^E, Q_0^E and the indifference value $h_0(x_0)$ at time 0 to motivate the definition of *FER*(H), which we introduce later in this section. Assume $\mathbb{P}_H^{e,f} \neq \emptyset$ and $\mathbb{P}_0^{e,f} \neq \emptyset$. The P_H-density of Q_H^E and the P-density of Q_0^E have the form

$$\frac{\mathrm{d}Q_H^E}{\mathrm{d}P_H} = c^H \exp\bigg(\int_0^T \zeta_s^H \mathrm{d}S_s\bigg) \quad \text{and} \quad \frac{\mathrm{d}Q_0^E}{\mathrm{d}P_0} = c^0 \exp\bigg(\int_0^T \zeta_s^0 \mathrm{d}S_s\bigg) \tag{4}$$

for some positive constants c^H, c^0 and processes ζ^H, ζ^0 in $L(S)$ such that $\int \zeta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^f$ and $\int \zeta^0 \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_0^f$, whence $\zeta^H \in \mathscr{A}_0^H$ and $\zeta^0 \in \mathscr{A}_0^0$. This was first shown by Kabanov and Stricker [16] in their Theorem 2.1 for a locally bounded S (and $H = 0$), and extended by Biagini and Frittelli [4] in their Theorem 1.4 to a general S for $H = 0$ (under the assumption $\mathscr{W}_0 \neq \emptyset$). By using this result also under P_H instead of P, we immediately obtain (4). It is now straightforward to calculate (and also well known—at least for locally bounded S) that for $x_0 \in \mathbb{R}$, we can write

$$\begin{aligned}
V_0^H(x_0) &= \sup_{\vartheta \in \mathscr{A}_0^H} E_P\left[-\exp\left(-\gamma x_0 - \gamma \int_0^T \vartheta_s \mathrm{d}S_s + \gamma H\right)\right] \\
&= -\exp(-\gamma x_0) E_P[\exp(\gamma H)] \inf_{\vartheta \in \mathscr{A}_0^H} E_{P_H}\left[\exp\left(-\gamma \int_0^T \vartheta_s \mathrm{d}S_s\right)\right] \\
&= -\exp(-\gamma x_0) E_P[\exp(\gamma H)] \\
&\quad \times \inf_{\vartheta \in \mathscr{A}_0^H} E_{Q_H^E}\left[\frac{1}{c^H} \exp\left(\int_0^T (-\gamma \vartheta_s - \zeta_s^H) \mathrm{d}S_s\right)\right] \\
&= -\frac{\exp(-\gamma x_0) E_P[\exp(\gamma H)]}{c^H}
\end{aligned} \tag{5}$$

and therefore

$$h_0(x_0) = h_0 = \frac{1}{\gamma} \log \frac{c^0 E_P[\exp(\gamma H)]}{c^H}. \tag{6}$$

In Sect. 4, we study the relation between Q_H^E, Q_0^E and $V_t^H(x_t)$, h_t for arbitrary $t \in [0, T]$. From this we can derive, on the one hand, an interpolation formula for each h_t in Sect. 4 and, on the other hand, a BSDE characterization of the process h in Sect. 5. To generalize the static relations (5), (6) to dynamic ones, we introduce a certain representation of H that we call *fundamental entropy representation of H* (*FER*(H)). Its link to the minimal H-entropy measure is elaborated in the next section. We give two different versions of this representation. The idea is that the first definition only requires some minimal conditions, whereas the second strengthens the conditions to guarantee uniqueness of the representation and ensure the identification of the H-MEM; see Proposition 2.

Definition 1 We say that *FER*(H) *exists* if there is a decomposition

$$H = \frac{1}{\gamma} \log \mathscr{E}(N^H)_T + \int_0^T \eta_s^H \mathrm{d}S_s + k_0^H, \tag{7}$$

where

(i) N^H is a local P-martingale null at 0 such that $\mathscr{E}(N^H)$ is a positive P-martingale and S is a $P(N^H)$-sigma-martingale, where $P(N^H)$ is defined by $\frac{\mathrm{d}P(N^H)}{\mathrm{d}P} := \mathscr{E}(N^H)_T$;
(ii) η^H is in $L(S)$ and such that $\int_0^T \eta_s^H \mathrm{d}S_s \in L^1(P(N^H))$;
(iii) $k_0^H \in \mathbb{R}$ is constant.

In this case, we say that (N^H, η^H, k_0^H) is an *FER*(H). If moreover

$$\int_0^T \eta_s^H \mathrm{d}S_s \in L^1(Q) \quad \text{and} \quad E_Q\bigg[\int_0^T \eta_s^H \mathrm{d}S_s\bigg] \le 0 \quad \text{for all } Q \in \mathbb{P}_H^f$$
$$\text{and} \quad \int \eta^H \mathrm{d}S \quad \text{is a } P(N^H)\text{-martingale,} \tag{8}$$

we say that (N^H, η^H, k_0^H) is an *FER*$^\star(H)$. For any *FER*(H) (N^H, η^H, k_0^H), we set

$$k_t^H := k_0^H + \frac{1}{\gamma}\log \mathscr{E}(N^H)_t + \int_0^t \eta_s^H \mathrm{d}S_s \quad \text{for } t \in [0,T] \tag{9}$$

and call $P(N^H)$ the probability measure *associated with* (N^H, η^H, k_0^H).

Because $\mathscr{E}(N^H)$ is by (i) a positive P-martingale, the local P-martingale N^H has no negative jumps whose absolute value is 1 or more, and $P(N^H)$ is a probability measure equivalent to P. We consider two *FER*(H) (N^H, η^H, k_0^H) and $(\widetilde{N}^H, \widetilde{\eta}^H, \widetilde{k}_0^H)$ as equal if $\widetilde{N}^H$ and N^H are versions of each other (hence indistinguishable, since both are RCLL), $\int \widetilde{\eta}^H \mathrm{d}S$ is a version of $\int \eta^H \mathrm{d}S$, and $\widetilde{k}_0^H = k_0^H$. For future use, we note that (7) and (9) combine to give

$$H = k_t^H + \frac{1}{\gamma}\log \mathscr{E}(N^H)_{t,T} + \int_t^T \eta_s^H \mathrm{d}S_s \quad \text{for } t \in [0,T]. \tag{10}$$

The next result shows that for continuous asset prices, we can write *FER*(H) in a different (and perhaps more familiar) form. For its formulation, we need the following definition. We say that S satisfies the *structure condition* (*SC*) if

$$S^i = S_0^i + M^i + \sum_{j=1}^d \int \lambda^j \mathrm{d}\langle M^i, M^j\rangle, \quad i = 1, \dots, d,$$

where M is a locally square-integrable local P-martingale null at 0 and λ is a predictable process such that the (final value of the) mean-variance tradeoff, $K_T = \sum_{i,j=1}^d \int_0^T \lambda_s^i \lambda_s^j \mathrm{d}\langle M^i, M^j\rangle_s = \langle \int \lambda \mathrm{d}M\rangle_T$, is almost surely finite.

Proposition 1 *Assume that S is continuous. Then a triple (N^H, η^H, k_0^H) is an FER(H) if and only if S satisfies (SC) and $\widetilde{N}^H = N^H + \int \lambda \mathrm{d}M$, $\widetilde{\eta}^H = \eta^H - \frac{1}{\gamma}\lambda$, $\widetilde{k}_0^H = k_0^H$ satisfy*

$$H = \frac{1}{\gamma}\log \mathscr{E}(\widetilde{N}^H)_T + \int_0^T \widetilde{\eta}_s^H \mathrm{d}S_s + \frac{1}{2\gamma}\Big\langle \int \lambda \mathrm{d}M\Big\rangle_T + \widetilde{k}_0^H \tag{11}$$

and

(i′) *$\widetilde{N}^H$ is a local P-martingale null at 0 and strongly P-orthogonal to each component of M, and $\mathscr{E}(\widetilde{N}^H)\mathscr{E}(-\int \lambda \mathrm{d}M)$ is a positive P-martingale*;

(ii′) $\widetilde{\eta}^H$ *is in* $L(S)$ *and such that* $\int_0^T (\widetilde{\eta}^H_s + \frac{1}{\gamma}\lambda_s)\mathrm{d}S_s$ *is* $P(N^H)$*-integrable, where* $\frac{\mathrm{d}P(N^H)}{\mathrm{d}P} := \mathscr{E}(\widetilde{N}^H)_T \mathscr{E}(-\int \lambda \mathrm{d}M)_T$;

(iii′) $\widetilde{k}^H_0 \in \mathbb{R}$ *is constant.*

Proof Let first (N^H, η^H, k^H_0) be an *FER*(H). Its associated measure $P(N^H)$ is equivalent to P and S is a local $P(N^H)$-martingale since S is continuous. By Theorem 1 of Schweizer [23], S satisfies (SC) and we can write $N^H = \widetilde{N}^H - \int \lambda \mathrm{d}M$, where $\widetilde{N}^H$ is a local P-martingale null at 0 and strongly P-orthogonal to each component of M, and $\mathscr{E}(N^H) = \mathscr{E}(\widetilde{N}^H)\mathscr{E}(-\int \lambda \mathrm{d}M)$. The last equality uses that $[\widetilde{N}^H, \int \lambda \mathrm{d}M] = 0$ due to the continuity of M. Hence conditions (i)–(iii) of *FER*(H) imply (i′)–(iii′), and (7) is equivalent to (11) by (SC) and the continuity of S.

Conversely, let $(\widetilde{N}^H, \widetilde{\eta}^H, \widetilde{k}^H_0)$ be as in the proposition. We claim that the triple $(\widetilde{N}^H - \int \lambda \mathrm{d}M, \widetilde{\eta}^H + \frac{1}{\gamma}\lambda, \widetilde{k}^H_0)$ is an *FER*(H). Because M is a local P-martingale and $\mathscr{E}(N^H) = \mathscr{E}(\widetilde{N}^H)\mathscr{E}(-\int \lambda \mathrm{d}M)$ is the P-density process of $P(N^H)$, the process L defined by

$$L_t := M_t - \langle N^H, M\rangle_t, \quad t \in [0, T]$$

is a local $P(N^H)$-martingale by Girsanov's theorem; see for instance Theorem III.40 of Protter [21] and observe that $\langle \mathscr{E}(N^H), M\rangle = \int \mathscr{E}(N^H)_- \mathrm{d}\langle N^H, M\rangle$ exists since M is continuous like S. Because $\widetilde{N}^H$ is strongly P-orthogonal to each component of M and M is continuous, we have

$$\langle N^H, M^i\rangle = \left\langle \widetilde{N}^H - \int \lambda \mathrm{d}M, M^i \right\rangle = -\sum_{j=1}^d \int \lambda^j \mathrm{d}\langle M^j, M^i\rangle, \quad i = 1, \ldots, d,$$

and so (SC) shows that $S = L + S_0$ is also a local $P(N^H)$-martingale. The other conditions of *FER*(H) are easy to check. □

Remark 1 (1) Suppose that S is continuous and satisfies (SC). If the stochastic exponential $\mathscr{E}(-\int \lambda \mathrm{d}M)$ is a P-martingale, conditions (i′) and (ii′) in Proposition 1 can be written under the probability measure $\widehat{P}$ defined by $\frac{\mathrm{d}\widehat{P}}{\mathrm{d}P} := \mathscr{E}(-\int \lambda \mathrm{d}M)_T$, which is called the *minimal local martingale measure* in the terminology of Föllmer and Schweizer [9]. This means that condition (i′) in Proposition 1 is equivalent to

(i″) $\widetilde{N}^H$ is a local $\widehat{P}$-martingale null at 0 and strongly $\widehat{P}$-orthogonal to each component of S, and $\mathscr{E}(\widetilde{N}^H)$ is a positive $\widehat{P}$-martingale,

and $P(N^H)$ can be defined by $\frac{\mathrm{d}P(N^H)}{\mathrm{d}\widehat{P}} := \mathscr{E}(\widetilde{N}^H)_T$. To prove the equivalence of (i′) and (i″), first assume that $\widetilde{N}^H$ is a local P-martingale null at 0 and strongly P-orthogonal to each M^i. Then

$$\left[\widetilde{N}^H, \int \lambda \mathrm{d}M\right] = \left\langle \widetilde{N}^H, \int \lambda \mathrm{d}M\right\rangle = 0$$

by the continuity of M, and hence $\widetilde{N}^H$ is also a local $\widehat{P}$-martingale by Girsanov's theorem; see, for instance, Theorem III.40 of Protter [21]. The continuity of S, (SC) and the strong P-orthogonality of $\widetilde{N}^H$ to M entail

$$[\widetilde{N}^H, S^i] = \langle \widetilde{N}^H, M^i \rangle = 0, \quad i = 1, \dots, d,$$

implying that $\widetilde{N}^H$ is strongly $\widehat{P}$-orthogonal to each component of S. The proof of "(i″) $\Longrightarrow$ (i′)" goes analogously.

(2) Assume that S is not necessarily continuous but locally bounded and satisfies (SC) with $\lambda^i \in L^2_{loc}(M^i)$, $i = 1, \dots, d$, and let (N^H, η^H, k_0^H) be an $FER(H)$. Then we can still write $N^H = \widetilde{N}^H - \int \lambda \mathrm{d}M$ for a local P-martingale $\widetilde{N}^H$ null at 0 and strongly P-orthogonal to each component of M, by using Girsanov's theorem, (SC) and the fact that $\mathscr{E}(N^H)$ defines an equivalent local martingale measure. However, we cannot separate $\mathscr{E}(\widetilde{N}^H - \int \lambda \mathrm{d}M)$ into two factors.

3 No-arbitrage and existence of $FER(H)$

Theorem 1 below says that a certain notion of no-arbitrage is equivalent to the existence of $FER(H)$. It can be considered as an exponential analogue to the L^2-result of Theorem 3 in Bobrovnytska and Schweizer [5]. For a locally bounded S, the implication "$\Longrightarrow$" roughly corresponds to Proposition 2.2 of Becherer [1], who makes use of the idea to consider known results under P_H instead of P. This technique, which already appears in Delbaen et al. [6], will also be central for the proofs of our Theorem 1 and Proposition 2.

We start with a result that gives sufficient conditions for $\mathscr{W}_H \subseteq \mathscr{W}_0$ and $\mathbb{P}_0^{e,f} \subseteq \mathbb{P}_H^{e,f}$ as well as for $\mathscr{W}_0 = \mathscr{W}_H$ and $\mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$. The relation between $\mathbb{P}_0^{e,f}$ and $\mathbb{P}_H^{e,f}$ will be used later, while $\mathscr{W}_0 = \mathscr{W}_H$ is helpful in applications to verify the condition (1).

Lemma 1 *If H satisfies*

$$E_P[\exp(-\varepsilon H)] < \infty \quad \textit{for some } \varepsilon > 0, \tag{12}$$

then $\mathscr{W}_H \subseteq \mathscr{W}_0$, $\mathbb{P}_0^f \subseteq \mathbb{P}_H^f$ and $\mathbb{P}_0^{e,f} \subseteq \mathbb{P}_H^{e,f}$. If H satisfies

$$E_P[\exp((\gamma + \varepsilon) H)] < \infty \quad \textit{and} \quad E_P[\exp(-\varepsilon H)] < \infty$$
$$\textit{for some } \varepsilon > 0, \tag{13}$$

then $\mathscr{W}_0 = \mathscr{W}_H$, $\mathbb{P}_0^f = \mathbb{P}_H^f$ and $\mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$.

Proof We first show $\mathscr{W}_H \subseteq \mathscr{W}_0$ under (12). For $c > 0$, Hölder's inequality yields

$$E_P[\exp(cW)] = E_P\left[\exp\left(cW + \frac{\varepsilon\gamma}{\varepsilon + \gamma} H\right) \exp\left(-\frac{\varepsilon\gamma}{\varepsilon + \gamma} H\right)\right]$$

$$\leq \left(E_P\left[\exp\left(\frac{\varepsilon+\gamma}{\varepsilon}cW+\gamma H\right)\right]\right)^{\frac{\varepsilon}{\varepsilon+\gamma}}(E_P[\exp(-\varepsilon H)])^{\frac{\gamma}{\varepsilon+\gamma}}$$
$$= \left(E_{P_H}\left[\exp\left(\frac{\varepsilon+\gamma}{\varepsilon}cW\right)\right]E_P[\exp(\gamma H)]\right)^{\frac{\varepsilon}{\varepsilon+\gamma}}$$
$$\times (E_P[\exp(-\varepsilon H)])^{\frac{\gamma}{\varepsilon+\gamma}}. \tag{14}$$

Because of $E_P[\exp(\gamma H)] < \infty$ and (12), this is finite if $W \in \mathscr{W}_H$, and then $W \in \mathscr{W}_0$.

To prove $\mathscr{W}_0 = \mathscr{W}_H$ under (13), we only need to show $\mathscr{W}_0 \subseteq \mathscr{W}_H$. For $c > 0$ and $W \in \mathscr{W}_0$, we obtain similarly to (14) that

$$E_{P_H}[\exp(cW)] \leq \frac{(E_P[\exp((\varepsilon+\gamma)H)])^{\frac{\gamma}{\varepsilon+\gamma}}}{E_P[\exp(\gamma H)]}\left(E_P\left[\exp\left(\frac{\varepsilon+\gamma}{\varepsilon}cW\right)\right]\right)^{\frac{\varepsilon}{\varepsilon+\gamma}} < \infty$$

by (13), and hence $W \in \mathscr{W}_H$.

The remainder of the second part follows from Lemma A.1 in Becherer [1]. The proof of the rest of the first part is very similar. Indeed, (12) and the standing assumption that $E_P[\exp(\gamma H)] < \infty$ imply $E_P[\exp(\widetilde{\varepsilon}|H|)] < \infty$, where $\widetilde{\varepsilon} := \min(\varepsilon, \gamma)$. Lemma 3.5 of Delbaen et al. [6] yields

$$E_Q[\widetilde{\varepsilon}|H|] \leq I(Q|P) + \frac{1}{e}E_P[\exp(\widetilde{\varepsilon}|H|)] \quad \text{for } Q \ll P. \tag{15}$$

If $Q \in \mathbb{P}_0^f$, the right-hand side is finite, thus $E_Q[|H|] < \infty$, and we have

$$I(Q|P_H) = E_Q\left[\log\frac{\mathrm{d}Q}{\mathrm{d}P} - \log\frac{\mathrm{d}P_H}{\mathrm{d}P}\right] = I(Q|P) + \log E_P[\exp(\gamma H)] - \gamma E_Q[H],$$

which is finite. This shows $Q \in \mathbb{P}_H^f$, and $\mathbb{P}_0^{e,f} \subseteq \mathbb{P}_H^{e,f}$ follows analogously. □

Theorem 1 *We have that*

$$\mathbb{P}_H^{e,f} \neq \emptyset \iff \mathit{FER}^\star(H) \text{ exists} \iff \mathit{FER}(H) \text{ exists.}$$

In particular, if $\mathbb{P}_0^{e,f} \neq \emptyset$ *and* H *satisfies* (12), *then* $\mathit{FER}^\star(H)$ *exists.*

Proof We first show that $\mathbb{P}_H^{e,f} \neq \emptyset$ yields the existence of $\mathit{FER}^\star(H)$. As already mentioned, $\mathbb{P}_H^{e,f} \neq \emptyset$ (and the standing assumption $\mathscr{W}_H \neq \emptyset$) imply by Proposition 3 and Remarks 2, 3 of Biagini and Frittelli [3], applied to P_H instead of P, existence and uniqueness of the H-MEM $Q_H^E \in \mathbb{P}_H^{e,f}$. Using $Q_H^E \approx P_H \approx P$, we can write

$$\frac{\mathrm{d}Q_H^E}{\mathrm{d}P} = \mathscr{E}(N^H)_T \tag{16}$$

for some local P-martingale N^H null at 0 such that $\mathscr{E}(N^H)$ is a positive P-martingale and S is a Q_H^E-sigma-martingale. Moreover, by Theorem 1.4 of Biagini and Frittelli [4], applied to P_H instead of P, we have as in (4)

$$\frac{\mathrm{d}Q_H^E}{\mathrm{d}P_H} = c^H \exp\left(\int_0^T \zeta_s^H \mathrm{d}S_s\right) \tag{17}$$

for a constant $c^H > 0$ and some ζ^H in $L(S)$ such that $\int \zeta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^f$. Since $\frac{\mathrm{d}P_H}{\mathrm{d}P} = \exp(\gamma H)/E_P[\exp(\gamma H)]$, comparing (17) with (16) gives

$$\mathscr{E}(N^H)_T = c_1^H \exp\left(\int_0^T \zeta_s^H \mathrm{d}S_s + \gamma H\right),$$

where $c_1^H := c^H / E_P[\exp(\gamma H)]$ is a positive constant. We thus obtain

$$H = \frac{1}{\gamma}\log\mathscr{E}(N^H)_T - \frac{1}{\gamma}\int_0^T \zeta_s^H \mathrm{d}S_s + c_2^H \quad \text{with } c_2^H := -\frac{1}{\gamma}\log c_1^H,$$

and hence $(N^H, -\frac{1}{\gamma}\zeta^H, c_2^H)$ is an $FER^\star(H)$. Note that $\int \zeta^H \mathrm{d}S$ is a $P(N^H)$-martingale because the H-MEM Q_H^E equals the probability measure $P(N^H)$ associated with $(N^H, -\frac{1}{\gamma}\zeta^H, c_2^H)$ by construction; compare (16).

To establish the equivalences of Theorem 1, it remains to show that the existence of $FER(H)$ implies $\mathbb{P}_H^{e,f} \neq \emptyset$, because every $FER^\star(H)$ is obviously an $FER(H)$. So let (N^H, η^H, k_0^H) be an $FER(H)$ and recall that its associated measure $P(N^H)$ is defined by $\frac{\mathrm{d}P(N^H)}{\mathrm{d}P} := \mathscr{E}(N^H)_T$. We prove that $P(N^H) \in \mathbb{P}_H^{e,f}$. By condition (i) on $FER(H)$, $P(N^H)$ is a probability measure equivalent to P and S is a $P(N^H)$-sigma-martingale. To show that $P(N^H)$ has finite relative entropy with respect to P_H, we write

$$\begin{aligned}\frac{\mathrm{d}P(N^H)}{\mathrm{d}P_H} &= \frac{\mathrm{d}P(N^H)}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}P_H} = \mathscr{E}(N^H)_T \exp(-\gamma H) E_P[\exp(\gamma H)] \\ &= \exp(-\gamma k_0^H) E_P[\exp(\gamma H)] \exp\left(-\gamma\int_0^T \eta_s^H \mathrm{d}S_s\right),\end{aligned} \tag{18}$$

where the last equality is due to the decomposition (7) in $FER(H)$. This yields by (ii) of $FER(H)$ that

$$\begin{aligned} I(P(N^H)|P_H) &= E_{P(N^H)}\left[\log\frac{\mathrm{d}P(N^H)}{\mathrm{d}P_H}\right] \\ &= -\gamma k_0^H + \log E_P[\exp(\gamma H)] - \gamma E_{P(N^H)}\left[\int_0^T \eta_s^H \mathrm{d}S_s\right] \\ &< \infty. \end{aligned}$$

Finally, the last assertion follows directly from the first part of Lemma 1. □

While the *existence* of $FER(H)$ and of $FER^{\star}(H)$ is equivalent by Theorem 1, the two representations are obviously different since $FER^{\star}(H)$ imposes more stringent conditions. The next result serves to clarify this difference.

Proposition 2 *Assume* $\mathbb{P}_H^{e,f} \neq \emptyset$ *and let* (N^H, η^H, k_0^H) *be an* $FER(H)$ *with associated measure* $P(N^H)$. *Then the following are equivalent*:

(a) (N^H, η^H, k_0^H) *is an* $FER^{\star}(H)$, *i.e.*, (N^H, η^H, k_0^H) *satisfies* (8);
(b) $P(N^H)$ *equals the* H-*MEM* Q_H^E, *and* $\int \eta^H \mathrm{d}S$ *is a* $P(N^H)$-*martingale*;
(c) $\int \eta^H \mathrm{d}S$ *is a* Q_H^E-*martingale and* $E_{P(N^H)}[\int_0^T \eta_s^H \mathrm{d}S_s] = 0$;
(d) $\int \eta^H \mathrm{d}S$ *is a* Q-*martingale for every* $Q \in \mathbb{P}_H^f$.

Moreover, the class of $FER^{\star}(H)$ *consists of a singleton.*

Proof Clearly, (d) implies (a), and also (c) since Q_H^E exists by Proposition 3 of Biagini and Frittelli [3], using $\mathbb{P}_H^{e,f} \neq \emptyset$ and the standing assumption $\mathscr{W}_H \neq \emptyset$. We prove "(a) $\Longrightarrow$ (b)", "(c) $\Longrightarrow$ (b)" and finally "(b) $\Longrightarrow$ (d)". The first implication goes as in the proof of Theorem 2.3 of Frittelli [11], because we have by (18) that

$$\frac{\mathrm{d}P(N^H)}{\mathrm{d}P_H} = c_3^H \exp\left(-\gamma \int_0^T \eta_s^H \mathrm{d}S_s\right)$$
$$\text{with } c_3^H := \exp(-\gamma k_0^H) E_P[\exp(\gamma H)]. \tag{19}$$

The implication "(c) $\Longrightarrow$ (b)" follows from the first part of the proof of Proposition 3.2 of Grandits and Rheinländer [12], which does not use the assumption that S is locally bounded. To show "(b) $\Longrightarrow$ (d)", note that (b), (17) and (19) yield

$$c_3^H \exp\left(-\gamma \int_0^T \eta_s^H \mathrm{d}S_s\right) = c^H \exp\left(\int_0^T \zeta_s^H \mathrm{d}S_s\right) \quad P\text{-a.s.}, \tag{20}$$

where ζ^H in $L(S)$ is such that $\int \zeta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^f$. Taking logarithms and $P(N^H)$-expectations in (20), we obtain $c_3^H = c^H$ by using that $P(N^H) \in \mathbb{P}_H^{e,f}$ by the proof of Theorem 1. Thus $\int_0^T \eta_s^H \mathrm{d}S_s = -\frac{1}{\gamma}\int_0^T \zeta_s^H \mathrm{d}S_s$ P-a.s. and hence $\int \eta^H \mathrm{d}S = -\frac{1}{\gamma}\int \zeta^H \mathrm{d}S$ since both $\int \eta^H \mathrm{d}S$ and $\int \zeta^H \mathrm{d}S$ are $P(N^H)$-martingales. Therefore, $\int \eta^H \mathrm{d}S = -\frac{1}{\gamma}\int \zeta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^f$.

Theorem 1 implies the existence of $FER^{\star}(H)$ because $\mathbb{P}_H^{e,f} \neq \emptyset$. To show uniqueness, let (N^H, η^H, k_0^H) and $(\widetilde{N}^H, \widetilde{\eta}^H, \widetilde{k}_0^H)$ be two $FER^{\star}(H)$. Since the minimal H-entropy measure is unique by Proposition 3 of Biagini and Frittelli [3], we have from "(a) $\Longrightarrow$ (b)" that

$$\mathscr{E}(N^H)_T = \frac{\mathrm{d}Q_H^E}{\mathrm{d}P} = \mathscr{E}(\widetilde{N}^H)_T.$$

So $\mathscr{E}(\widetilde{N}^H)$ is a version of $\mathscr{E}(N^H)$ since both are P-martingales, and taking stochastic logarithms implies that $\widetilde{N}^H$ is a version of N^H. Similarly, (19) and (c) yield

$$-\gamma k_0^H + \log(E_P[\exp(\gamma H)]) = E_{Q_H^E}\left[\log \frac{\mathrm{d}Q_H^E}{\mathrm{d}P_H}\right] \\ = -\gamma \widetilde{k}_0^H + \log(E_P[\exp(\gamma H)]),$$

thus $\widetilde{k}_0^H = k_0^H$, and therefore again from (19) that

$$\int_0^T \eta_s^H \mathrm{d}S_s = -\frac{1}{\gamma}\log\left(\frac{1}{c_3^H}\frac{\mathrm{d}Q_H^E}{\mathrm{d}P_H}\right) = \int_0^T \widetilde{\eta}_s^H \mathrm{d}S_s.$$

But both $\int \eta^H \mathrm{d}S$ and $\int \widetilde{\eta}^H \mathrm{d}S$ are Q_H^E-martingales due to (d), and so $\int \widetilde{\eta}^H \mathrm{d}S$ is a version of $\int \eta^H \mathrm{d}S$. □

Remark 2 Exploiting Proposition 3.4 of Grandits and Rheinländer [12], applied to P_H instead of P, gives a sufficient condition for $FER^\star(H)$ by using our Proposition 2. Indeed, assume that S is locally bounded and $\mathbb{P}_H^{e,f} \neq \emptyset$. If for an $FER(H)$ (N^H, η^H, k_0^H), $\int \eta^H \mathrm{d}S$ is a $BMO(P(N^H))$-martingale and $E_{P_H}[|\frac{\mathrm{d}P(N^H)}{\mathrm{d}P_H}|^{-\varepsilon}] < \infty$ for some $\varepsilon > 0$, then (N^H, η^H, k_0^H) is the $FER^\star(H)$.

Another sufficient criterion is obtained from Proposition 3.2 of Rheinländer [22] in view of our Proposition 2. Namely, if S is locally bounded and for an $FER(H)$ (N^H, η^H, k_0^H) there exists $\varepsilon > 0$ such that $E_{P_H}[\exp(\varepsilon[\int \eta^H \mathrm{d}S]_T)] < \infty$, then (N^H, η^H, k_0^H) is the $FER^\star(H)$.

While there is always at most one $FER^\star(H)$ by Proposition 2, the next example shows that there may be several $FER(H)$. This also illustrates that the uniqueness for $FER^\star(H)$ is closely related to integrability properties.

Example 1 Take two independent P-Brownian motions W and $W^\perp$, denote by $\mathbb{F}$ their P-augmented filtration and choose $d = 1$, $S = W$ and $H \equiv 0$. The MEM Q_0^E then equals P since S is a P-martingale, and $(0, 0, 0)$ is the unique $FER^\star(0)$.

To construct another $FER(0)$, choose $N^0 := W^\perp$. Then $\mathscr{E}(N^0) = \mathscr{E}(W^\perp)$ is clearly a positive P-martingale strongly P-orthogonal to $S = W$ so that condition (i) in $FER(0)$ holds. Define $P(N^0)$ as usual by $\frac{\mathrm{d}P(N^0)}{\mathrm{d}P} := \mathscr{E}(N^0)_T = \mathscr{E}(W^\perp)_T$. By Girsanov's theorem, W and $\widetilde{W}_t^\perp := W_t^\perp - t$, $0 \le t \le T$, are then $P(N^0)$-Brownian motions and we can explicitly compute

$$E_P[\log \mathscr{E}(N^0)_T] = E_P[W_T^\perp - T/2] = -T/2, \\ I(P(N^0)|P) = E_{P(N^0)}[\log \mathscr{E}(N^0)_T] = E_{P(N^0)}[\widetilde{W}_T^\perp + T/2] = T/2. \tag{21}$$

This shows that $P(N^0) \in \mathbb{P}_0^{e,f}$. Since $S = W$ is a P-Brownian motion, Proposition 1 of Emery et al. [8] now yields for every $c \in \mathbb{R}$ a process $\eta^0(c)$ in $L(S)$ such that

$$-\frac{1}{\gamma}\log \mathscr{E}\big(W^{\perp}\big)_T - c = \int_0^T \eta_s^0(c)\mathrm{d}S_s \quad P\text{-a.s.} \tag{22}$$

Because $I(P(N^0)|P) < \infty$, using the inequality $x|\log x| \leq x\log x + 2e^{-1}$ shows that $\int_0^T \eta_s^0(c)\mathrm{d}S_s$ is in $L^1(P(N^0))$ so that (ii) of $FER(0)$ is also satisfied. Hence $(N^0, \eta^0(c), c)$ is an $FER(0)$, but does not coincide with $(0,0,0)$ which is the $FER^\star(0)$. To check that property (8) indeed fails, we can easily see from (21) and (22) that $\int \eta^0(c)\mathrm{d}S$ cannot be a $P(N^0)$-martingale if $c \neq -\frac{1}{2\gamma}T$. If $c = -\frac{1}{2\gamma}T$, we can simply compute, for $P \in \mathbb{P}_0^f$, that

$$E_P\left[\int_0^T \eta_s^0(c)\mathrm{d}S_s\right] = -\frac{1}{\gamma}E_P[\log \mathscr{E}(N^0)_T] + \frac{1}{2\gamma}T = \frac{1}{\gamma}T > 0.$$

We have just constructed an $FER(0)$ different from $FER^\star(0)$. Yet another $FER(0)$ can be obtained by choosing for $k \in \mathbb{R}\backslash\{0\}$ a process $\beta^0(k)$ in $L(S)$ such that

$$\int_0^{T/2} \beta_s^0(k)\mathrm{d}S_s = k \quad \text{and} \quad \int_{T/2}^T \beta_s^0(k)\mathrm{d}S_s = -k \quad P\text{-a.s.,}$$

which is possible by Proposition 1 of Emery et al. [8]. Clearly, $\int_0^T \beta_s^0(k)\mathrm{d}S_s = 0$ P-a.s. and $(0, \beta^0(k), 0)$ is an $FER(0)$ (with associated measure P), which even satisfies $E_Q[\int_0^T \beta_s^0(k)\mathrm{d}S_s] = 0$ for all $Q \in \mathbb{P}_0^f$; but $\int \beta^0(k)\mathrm{d}S$ is not a P-martingale. This ends the example.

Example 1 shows that we should focus on $FER^\star(H)$ if we want to obtain good results. If S is continuous and we impose additional assumptions, the next result gives *BMO*-properties for the components of $FER^\star(H)$. This will be used later when we give a BSDE description for the exponential utility indifference value process. We first recall some definitions.

Let Q be a probability measure on $(\Omega, \mathscr{F})$ equivalent to P and $p > 1$. An adapted positive RCLL stochastic process Z is said to satisfy the *reverse Hölder inequality* $R_p(Q)$ if there exists a positive constant C such that

$$\operatorname*{ess\,sup}_{\substack{\tau \text{ stopping} \\ \text{time}}} E_Q\left[\left(\frac{Z_T}{Z_\tau}\right)^p \middle| \mathscr{F}_\tau\right] = \operatorname*{ess\,sup}_{\substack{\tau \text{ stopping} \\ \text{time}}} E_Q[(Z_{\tau,T})^p|\mathscr{F}_\tau] \leq C.$$

Recall that $Z_{\tau,T} = Z_T/Z_\tau$ for a positive process Z. We say that Z satisfies the *reverse Hölder inequality* $R_{L\log L}(Q)$ if there exists a positive constant C such that

$$\operatorname*{ess\,sup}_{\substack{\tau \text{ stopping} \\ \text{time}}} E_Q[Z_{\tau,T}\log^+ Z_{\tau,T}|\mathscr{F}_\tau] \leq C.$$

Z satisfies *condition* (*J*) if there exists a positive constant C such that

$$\frac{1}{C}Z_- \le Z \le CZ_-.$$

Theorem 2 *Assume that S is continuous, H is bounded and there exists $Q \in \mathbb{P}_0^{e,f}$ whose P-density process satisfies $R_{L\log L}(P)$. Let (N^H, η^H, k_0^H) be an FER(H). Then the following are equivalent*:

(a) (N^H, η^H, k_0^H) *is the FER*$^\star(H)$;
(b) N^H *is a BMO*(P)*-martingale*, $\mathscr{E}(N^H)$ *satisfies condition* (*J*), *and* $\int \eta^H \mathrm{d}S$ *is a* $P(N^H)$*-martingale*;
(c) N^H *is a BMO*(P)*-martingale*, $\mathscr{E}(N^H)$ *satisfies condition* (*J*), *and* $\int \eta^H \mathrm{d}S$ *is a BMO*$(P(N^H))$*-martingale*;
(d) $\int \eta^H \mathrm{d}M$ *is a BMO*(P)*-martingale*, *where M is the P-local martingale part of S*;
(e) *there exists $\varepsilon > 0$ such that* $E_P[\exp(\varepsilon[\int \eta^H \mathrm{d}S]_T)] < \infty$.

The hypotheses of Theorem 2 are for instance fulfilled if H is bounded, S is continuous and satisfies (SC), and $\int \lambda \mathrm{d}M$ is a $BMO(P)$-martingale. To see this, note that $\mathscr{E}(-\int \lambda \mathrm{d}M)$ then satisfies the reverse Hölder inequality $R_p(P)$ for some $p > 1$ by Theorem 3.4 of Kazamaki [18]. The fact that there exists $k < \infty$ such that $x\log x \le k + x^p$ for all $x > 0$ now implies that $\mathscr{E}(-\int \lambda \mathrm{d}M)$ also satisfies $R_{L\log L}(P)$. Hence the minimal local martingale measure $\widehat{P}$ given by $\frac{\mathrm{d}\widehat{P}}{\mathrm{d}P} := \mathscr{E}(-\int \lambda \mathrm{d}M)_T$ is in $\mathbb{P}_0^{e,f}$ and its P-density process satisfies $R_{L\log L}(P)$.

Proof of Theorem 2 By Lemma 1, $\mathbb{P}_H^{e,f} = \mathbb{P}_0^{e,f} \neq \emptyset$ so that there exists an $FER(H)$ (N^H, η^H, k_0^H) by Theorem 1. Before we show that (a)–(e) are equivalent, we need some preparation. Let $\widetilde{Q}$ be a probability measure equivalent to P. Denoting by Z the P-density process of $\widetilde{Q}$ and by Y the P_H-density process of $\widetilde{Q}$, we prove that

$$Z \text{ satisfies } R_{L\log L}(P) \quad \text{if and only if} \quad Y \text{ satisfies } R_{L\log L}(P_H), \tag{23}$$

$$Z \text{ satisfies condition (J)} \quad \text{if and only if} \quad Y \text{ satisfies condition (J)}. \tag{24}$$

To that end, observe first that because H is bounded, there exists a positive constant k with $\frac{1}{k} \le \frac{\mathrm{d}P_H}{\mathrm{d}P} \le k$, which yields

$$\frac{1}{k}Z \le Y \le kZ. \tag{25}$$

For any stopping time τ, (25) implies

$$E_{P_H}[Y_{\tau,T}\log^+ Y_{\tau,T}|\mathscr{F}_\tau] \le E_P[Z_{\tau,T}\log^+(Z_{\tau,T}k^2)|\mathscr{F}_\tau],$$

and so the inequality $\log^+(ab) \le \log^+ a + \log b$ for $a > 0$ and $b \ge 1$ yields

$$E_P[Z_{\tau,T}\log^+(Z_{\tau,T}k^2)|\mathscr{F}_\tau] \le E_P[Z_{\tau,T}\log^+ Z_{\tau,T}|\mathscr{F}_\tau] + 2\log k,$$

which is bounded independently of τ if Z satisfies $R_{L\log L}(P)$. If Z satisfies condition (J) with constant C, then (25) gives

$$Y \le kZ \le kCZ_- \le k^2CY_- \quad \text{and} \quad Y \ge \frac{1}{k}Z \ge \frac{1}{kC}Z_- \ge \frac{1}{k^2C}Y_-.$$

So the "only if" part of both (23) and (24) is clear, and the "if" part is proved symmetrically.

By assumption, there exists $Q \in \mathbb{P}_0^{e,f}$ whose P-density process satisfies $R_{L\log L}(P)$, and so the P_H-density process of Q satisfies $R_{L\log L}(P_H)$ by (23). Because $\mathbb{P}_H^{e,f} = \mathbb{P}_0^{e,f}$ is nonempty, the unique minimal H-entropy measure Q_H^E exists, and its P_H-density process also satisfies $R_{L\log L}(P_H)$ by Lemma 3.1 of Delbaen et al. [6], used for P_H instead of P. Since S is continuous, the P_H-density process of Q_H^E also satisfies condition (J) by Lemma 4.6 of Grandits and Rheinländer [12]. It follows from (23), (24) and Lemma 2.2 of Grandits and Rheinländer [12] that

$$\text{the } P\text{-density process } Z^{Q_H^E,P} \text{ of } Q_H^E \text{ satisfies } R_{L\log L}(P) \text{ and condition (J),}$$
$$\text{and the stochastic logarithm of } Z^{Q_H^E,P} \text{ is a } BMO(P)\text{-martingale.} \tag{26}$$

"(a) $\Longrightarrow$ (b)". Since (N^H, η^H, k_0^H) is the $FER^\star(H)$, Proposition 2 implies that the P-density process $Z^{Q_H^E,P}$ of Q_H^E is given by $\mathscr{E}(N^H)$ and that $\int \eta^H \mathrm{d}S$ is a $P(N^H)$-martingale. We deduce (b) from (26).

"(b) $\Longrightarrow$ (c)". We have to show that $\int \eta^H \mathrm{d}S$ is in $BMO(P(N^H))$. By conditioning (7) under $P(N^H)$ on $\mathscr{F}_\tau$ for a stopping time τ, we obtain by (b)

$$\int_0^\tau \eta_s^H \mathrm{d}S_s = -\frac{1}{\gamma} E_{P(N^H)}[\log \mathscr{E}(N^H)_T | \mathscr{F}_\tau] + E_{P(N^H)}[H|\mathscr{F}_\tau] - k_0^H,$$

and hence

$$\begin{aligned}\int_\tau^T \eta_s^H \mathrm{d}S_s = &-\frac{1}{\gamma}\log \mathscr{E}(N^H)_T + \frac{1}{\gamma}E_{P(N^H)}[\log \mathscr{E}(N^H)_T|\mathscr{F}_\tau] \\ &+ H - E_{P(N^H)}[H|\mathscr{F}_\tau].\end{aligned}$$

By Proposition 6 of Doléans-Dade and Meyer [7], there is a $BMO(P(N^H))$-martingale $\widehat{N}^H$ with $\mathscr{E}(N^H)^{-1} = \mathscr{E}(\widehat{N}^H)$. This uses that $Z^{Q_H^E,P} = \mathscr{E}(N^H)$ satisfies condition (J) and N^H is a $BMO(P)$-martingale by (26). Since H is bounded, we get

$$\begin{aligned} &E_{P(N^H)}\left[\left|\int_\tau^T \eta_s^H \mathrm{d}S_s\right| \middle| \mathscr{F}_\tau\right] \\ &\le 2\|H\|_{L^\infty(P)} + \frac{1}{\gamma}E_{P(N^H)}\Big[\big|\log \mathscr{E}(N^H)_T - E_{P(N^H)}[\log \mathscr{E}(N^H)_T|\mathscr{F}_\tau]\big|\Big|\mathscr{F}_\tau\Big] \\ &= 2\|H\|_{L^\infty(P)} + \frac{1}{\gamma}E_{P(N^H)}\Big[\big|\log \mathscr{E}(\widehat{N}^H)_T - E_{P(N^H)}[\log \mathscr{E}(\widehat{N}^H)_T|\mathscr{F}_\tau]\big|\Big|\mathscr{F}_\tau\Big], \end{aligned} \tag{27}$$

and now we proceed like on page 1031 in Grandits and Rheinländer [12] to show that (27) is bounded uniformly in τ. This proves the assertion since S is continuous.

"(c) $\Longrightarrow$ (d)". Due to (26), Proposition 7 of Doléans-Dade and Meyer [7] implies that $\int \eta^H \mathrm{d}S + [\int \eta^H \mathrm{d}S, N^H]$ is a $BMO(P)$-martingale. By Proposition 1, S satisfies (SC) and $N^H = \widetilde{N}^H - \int \lambda \mathrm{d}M$ for a local P-martingale $\widetilde{N}^H$ null at 0 and strongly P-orthogonal to each component of M. Since S is continuous and satisfies (SC),

$$\Big[\int \eta^H \mathrm{d}S, N^H\Big] = \Big[\int \eta^H \mathrm{d}M, N^H\Big] = -\Big[\int \eta^H \mathrm{d}M, \int \lambda \mathrm{d}M\Big]$$
$$= -\sum_{i,j=1}^{d} \int (\eta^H)^i \lambda^j \mathrm{d}\langle M^i, M^j\rangle.$$

Hence $\int \eta^H \mathrm{d}S + [\int \eta^H \mathrm{d}S, N^H] = \int \eta^H \mathrm{d}M$ is a $BMO(P)$-martingale.

"(d) $\Longrightarrow$ (e)". We set

$$\varepsilon := \frac{1}{2\|\int \eta^H \mathrm{d}M\|^2_{BMO_2(P)}} \quad \text{and} \quad L := \sqrt{\varepsilon}\int \eta^H \mathrm{d}M.$$

Clearly, L is like $\int \eta^H \mathrm{d}M$ a continuous $BMO(P)$-martingale and we have that $\|L\|_{BMO_2(P)} = 1/\sqrt{2} < 1$. Since S is continuous, the John-Nirenberg inequality (see Theorem 2.2 of Kazamaki [18]) yields

$$E_P\Big[\exp\Big(\varepsilon\Big[\int \eta^H \mathrm{d}S\Big]_T\Big)\Big] = E_P[\exp([L]_T)] \le \frac{1}{1 - \|L\|^2_{BMO_2(P)}} < \infty.$$

"(e) $\Longrightarrow$ (a)". This is based on the same idea as the proof of Proposition 3.2 of Rheinländer [22]. Lemma 3.5 of Delbaen et al. [6] yields

$$E_Q\Big[\varepsilon\Big[\int \eta^H \mathrm{d}S\Big]_T\Big] \le I(Q|P_H) + \frac{1}{e}E_{P_H}\Big[\exp\Big(\varepsilon\Big[\int \eta^H \mathrm{d}S\Big]_T\Big)\Big] < \infty$$

for any $Q \in \mathbb{P}^f_H$ because H is bounded and (e) holds. So $[\int \eta^H \mathrm{d}S]_T$ is Q-integrable and thus the local Q-martingale $\int \eta^H \mathrm{d}S$ is a square-integrable Q-martingale for any $Q \in \mathbb{P}^f_H$. This concludes the proof in view of Proposition 2. □

4 Relating $FER^\star(H)$ and $FER^\star(0)$ to the Indifference Value

In this section, we establish the connection between $FER^\star(H)$, $FER^\star(0)$ and the indifference value process h. We then derive and study an interpolation formula for h. Throughout this section, we assume that

$$\mathbb{P}^{e,f}_H \neq \emptyset \quad \text{and} \quad \mathbb{P}^{e,f}_0 \neq \emptyset,$$

and we denote by (N^H, η^H, k_0^H) and (N^0, η^0, k_0^0) the unique $FER^\star(H)$ and $FER^\star(0)$ with associated measures $P(N^H) = Q_H^E$ and $P(N^0) = Q_0^E$, respectively.

Our first result expresses the maximal expected utility and the indifference value in terms of the given $FER^\star(H)$ and $FER^\star(0)$. For a locally bounded S, this is very similar to Becherer [1]; see in particular there Propositions 2.2 and 3.5 and the discussion on page 12 at the end of Sect. 3. Indeed, the main differences are that the representation in [1] is given in terms of certainty equivalents instead of maximal conditional expected utilities and S is locally bounded; but the results are the same.

Theorem 3 *V^H, V^0 and h are well defined and, for any $t \in [0,T]$ and any $\mathscr{F}_t$-measurable random variable x_t, we have*

$$V_t^H(x_t) = -\exp(-\gamma x_t + \gamma k_t^H) \tag{28}$$

and

$$h_t(x_t) = h_t = k_t^H - k_t^0, \tag{29}$$

where k_t^H (and k_t^0, with the obvious adaptations) are defined in (9).

Proof Let us first write (2) as

$$V_t^H(x_t) = -\exp(-\gamma x_t) \operatorname*{ess\,inf}_{\vartheta \in \mathscr{A}_t^H} \varphi_t^H(\vartheta) \tag{30}$$

with the abbreviation

$$\varphi_t^H(\vartheta) := E_P\left[\exp\left(-\gamma \int_t^T \vartheta_s \mathrm{d}S_s + \gamma H\right)\middle|\mathscr{F}_t\right].$$

Because (N^H, η^H, k_0^H) is the $FER^\star(H)$, $\varphi_t^H(\vartheta)$ can be written by (10) as

$$\begin{aligned}\varphi_t^H(\vartheta) &= \exp(\gamma k_t^H) E_P\left[\mathscr{E}(N^H)_{t,T} \exp\left(\gamma \int_t^T (\eta_s^H - \vartheta_s)\mathrm{d}S_s\right)\middle|\mathscr{F}_t\right] \\ &= \exp(\gamma k_t^H) E_{P(N^H)}\left[\exp\left(\gamma \int_t^T (\eta_s^H - \vartheta_s)\mathrm{d}S_s\right)\middle|\mathscr{F}_t\right],\end{aligned} \tag{31}$$

using Bayes' formula. Since $P(N^H) = Q_H^E \in \mathbb{P}_H^{e,f}$ and $\int \vartheta \mathrm{d}S$ is a Q-supermartingale and $\int \eta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^{e,f}$, we have

$$E_{P(N^H)}\left[\int_t^T (\eta_s^H - \vartheta_s)\mathrm{d}S_s\middle|\mathscr{F}_t\right] \geq 0$$

which implies $\varphi_t^H(\vartheta) \geq \exp(\gamma k_t^H)$ by Jensen's inequality and (31). On the other hand, the choice

$$\vartheta_s^\star := \eta_s^H, \quad s \in (t,T], \tag{32}$$

gives $\varphi_t^H(\vartheta^\star) = \exp(\gamma k_t^H)$ by (31). Because $\int \vartheta^\star \mathrm{d}S = \int \eta^H \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_H^{e,f}$, $\vartheta^\star$ is in $\mathscr{A}_t^H$, and (28) now follows from (30).

By the same reasoning as for (28), we obtain

$$V_t^0(x_t) = -\exp(-\gamma x_t + \gamma k_t^0).$$

Solving the implicit equation (3) for $h_t(x_t)$ then immediately leads to (29). □

The proof of Theorem 3, especially (32), gives an interpretation for the $FER^\star(H)$. An investor who must pay out the claim H at time T uses, under exponential utility preferences, the decomposition (7). The portion of H that he hedges by trading in S is $\int_0^T \eta_s^H \mathrm{d}S_s$, whereas $\frac{1}{\gamma}\log \mathscr{E}(N^H)_T$ remains unhedged. Moreover, the proof of Theorem 3 shows that for $t \in [0,T]$ and an $\mathscr{F}_t$-measurable x_t, the value of $V_t^H(x_t)$ is not affected if we restrict the set $\mathscr{A}_t^H$ to those $\vartheta \in \mathscr{A}_t^H$ such that $\int \vartheta \mathrm{d}S$ is not only a Q-supermartingale, but a Q-martingale for every $Q \in \mathbb{P}_H^{e,f}$.

Proposition 3 *Assume that H satisfies (12). Then for any $Q \in \mathbb{P}_0^f$ and $t \in [0,T]$,*

$$h_t = E_Q[H|\mathscr{F}_t] - \frac{1}{\gamma} E_Q\left[\log \frac{\mathscr{E}(N^H)_{t,T}}{\mathscr{E}(N^0)_{t,T}}\middle|\mathscr{F}_t\right]. \tag{33}$$

In particular,

$$h_0 = E_Q[H] + \frac{1}{\gamma}\big(I(Q|Q_H^E) - I(Q|Q_0^E)\big). \tag{34}$$

The decomposition (34) of the indifference value h_0 can be described as follows. The first term, $E_Q[H]$, is the expected payoff under a measure $Q \in \mathbb{P}_0^f$. This is linear in the number of claims. The second term is a nonlinear correction term or safety loading. It can be interpreted as the difference of the distances from Q_H^E and Q_0^E to Q (although $I(\cdot|\cdot)$ is not a metric). This correction term is not based on all of H, but only on the processes N^H and N^0 from the $FER^\star(H)$ and $FER^\star(0)$, i.e., on the unhedged parts of H and 0, respectively. A similar decomposition also appears for indifference pricing under quadratic preferences; see Schweizer [24].

If H satisfies (12), then the indifference value process h is a Q_0^E-supermartingale. In fact, Jensen's inequality and (33) with $Q = Q_0^E$ yield $h_t \geq E_{Q_0^E}[H|\mathscr{F}_t]$ P-a.s. for $t \in [0,T]$ and so $h_t^- \in L^1(Q_0^E)$ since H is Q_0^E-integrable due to (12); compare (15). Moreover, $Z := \mathscr{E}(N^H)/\mathscr{E}(N^0)$ is a Q_0^E-martingale as it is the Q_0^E-density process of Q_H^E. Thus $\log Z$ has the Q_0^E-supermartingale property by Jensen's inequality, and so has h since $h_t = E_{Q_0^E}[H|\mathscr{F}_t] - \frac{1}{\gamma}E_{Q_0^E}[\log Z_T|\mathscr{F}_t] + \frac{1}{\gamma}\log Z_t$ for $t \in [0,T]$ by (33). Now $E_{Q_0^E}[h_t] \leq h_0 < \infty$ shows that h_t is Q_0^E-integrable for every $t \in [0,T]$.

Proof of Proposition 3 Since $Q \in \mathbb{P}_0^f \subseteq \mathbb{P}_H^f$ by Lemma 1, $\int \eta^H \mathrm{d}S$ is a Q-martingale by Proposition 2. Moreover, H is Q-integrable due to (12); compare (15). From

(10), we thus obtain for $t \in [0,T]$ that

$$k_t^H = E_Q\Big[H - \frac{1}{\gamma}\log\mathscr{E}(N^H)_{t,T}\Big|\mathscr{F}_t\Big]. \tag{35}$$

Plugging (35) and the analogous expression for k_t^0 into (29) leads to (33).

To prove (34), we first show that $I(Q|Q_0^E)$ is finite. We can write

$$\begin{aligned} I(Q|Q_0^E) &= E_Q\Big[\log\frac{\mathrm{d}Q}{\mathrm{d}P} + \log\frac{\mathrm{d}P}{\mathrm{d}Q_0^E}\Big] \\ &= I(Q|P) - E_Q[\log\mathscr{E}(N^0)_T] < \infty \end{aligned} \tag{36}$$

because $Q \in \mathbb{P}_0^f$ and $-E_Q[\log\mathscr{E}(N^0)_T] = \gamma k_0^0$ by (35) for $H=0$ and $t=0$. Moreover, $Q \ll P \approx Q_H^E$ gives $\frac{\mathrm{d}Q}{\mathrm{d}P} > 0$ Q-a.s. and thus from

$$\frac{\mathrm{d}Q}{\mathrm{d}Q_H^E} = \frac{\mathrm{d}Q}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}Q_H^E} = \frac{\mathrm{d}Q}{\mathrm{d}P}\frac{1}{\mathscr{E}(N^H)_T} \quad Q\text{-a.s.}$$

that

$$-\log\mathscr{E}(N^H)_T = \log\frac{\mathrm{d}Q}{\mathrm{d}Q_H^E} - \log\frac{\mathrm{d}Q}{\mathrm{d}P} \quad Q\text{-a.s.},$$

and analogously for 0 instead of H. Hence

$$\begin{aligned} E_Q\Big[-\log\frac{\mathscr{E}(N^H)_T}{\mathscr{E}(N^0)_T}\Big] &= E_Q\Big[\log\frac{\mathrm{d}Q}{\mathrm{d}Q_H^E} - \log\frac{\mathrm{d}Q}{\mathrm{d}Q_0^E}\Big] \\ &= I(Q|Q_H^E) - I(Q|Q_0^E), \end{aligned}$$

where we have used (36) for the last equality. Now (34) follows from (33). □

We next come to the announced interpolation formula for the indifference value.

Theorem 4 *Let $Q \in \mathbb{P}_H^{e,f}$ and φ in $L(S)$ be such that $\int\varphi\mathrm{d}S$ is a Q- and Q_H^E-martingale. Fix $t \in [0,T]$, denote by Z the P-density process of Q, set*

$$\Psi_t^H := \frac{\exp(\gamma H + \int_t^T \varphi_s\mathrm{d}S_s)}{Z_{t,T}} \tag{37}$$

and assume that Ψ_t^H and $\log\Psi_t^H$ are Q-integrable. Then there exists an $\mathscr{F}_t$-measurable random variable $\delta_t^H : \Omega \to [1,\infty]$ such that for almost all $\omega \in \Omega$,

$$k_t^H(\omega) = \frac{1}{\gamma}\log\big(E_Q[|\Psi_t^H|^{1/\delta}|\mathscr{F}_t](\omega)\big)^{\delta}\big|_{\delta=\delta_t^H(\omega)}, \tag{38}$$

where

$$\log(E_Q[|\Psi_t^H|^{1/\delta}|\mathscr{F}_t](\omega))^\delta|_{\delta=\infty} := \lim_{\delta\to\infty} \log(E_Q[|\Psi_t^H|^{1/\delta}|\mathscr{F}_t](\omega))^\delta$$
$$= E_Q[\log \Psi_t^H|\mathscr{F}_t](\omega) \tag{39}$$

for almost all $\omega \in \Omega$.

In view of $h_t = k_t^H - k_t^0$ by Theorem 3, (38) gives us a quasi-explicit formula for the exponential utility indifference value if H is bounded and if we can find a measure $Q \in \mathbb{P}_0^{e,f}$ such that the corresponding Ψ_t^0 given in (37) and $\log \Psi_t^0$ are Q-integrable for some predictable φ such that $\int \varphi \mathrm{d}S$ is a Q-, Q_0^E- and Q_H^E-martingale. For $t = 0$, one possible choice is the minimal 0-entropy measure Q_0^E which is by (19) and Proposition 2 of the form $\frac{\mathrm{d}Q_0^E}{\mathrm{d}P} = c_3^0 \exp(\int_0^T \zeta_s^0 \mathrm{d}S_s)$ for a constant c_3^0 and a predictable process ζ^0 such that $\int \zeta^0 \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_0^f$. One disadvantage of this choice is that Q_0^E is in general unknown; a second is that we still need to find some φ, and we know almost nothing about the potential candidate ζ^0. In Corollary 1, we give conditions under which the explicitly known minimal local martingale measure $\widehat{P}$ satisfies the assumptions of Theorem 4.

Proof of Theorem 4 From (10) and (37), we obtain via $\frac{\mathrm{d}Q_H^E}{\mathrm{d}P} = \mathscr{E}(N^H)_T$ and Bayes' formula that

$$\begin{aligned}
\exp(-\gamma k_t^H) E_Q[\Psi_t^H|\mathscr{F}_t] &= E_Q\left[\frac{\mathscr{E}(N^H)_{t,T}}{Z_{t,T}} \exp\left(\int_t^T (\varphi_s + \gamma \eta_s^H) \mathrm{d}S_s\right)\middle|\mathscr{F}_t\right] \\
&= E_{Q_H^E}\left[\exp\left(\int_t^T (\varphi_s + \gamma \eta_s^H) \mathrm{d}S_s\right)\middle|\mathscr{F}_t\right] \\
&\ge \exp\left(E_{Q_H^E}\left[\int_t^T (\varphi_s + \gamma \eta_s^H) \mathrm{d}S_s\middle|\mathscr{F}_t\right]\right) \\
&= 1
\end{aligned} \tag{40}$$

by Jensen's inequality and because $\int \varphi \mathrm{d}S$ and $\int \eta^H \mathrm{d}S$ are Q_H^E-martingales. Hence

$$k_t^H \le \frac{1}{\gamma} \log E_Q[\Psi_t^H|\mathscr{F}_t]. \tag{41}$$

On the other hand, (35), (37) and Jensen's inequality yield

$$\begin{aligned}
\gamma k_t^H &= E_Q[\gamma H - \log \mathscr{E}(N^H)_{t,T}|\mathscr{F}_t] \\
&= E_Q\left[\log \Psi_t^H - \log \frac{\mathscr{E}(N^H)_{t,T}}{Z_{t,T}}\middle|\mathscr{F}_t\right] \\
&\ge E_Q[\log \Psi_t^H|\mathscr{F}_t].
\end{aligned} \tag{42}$$

Consider the stochastic process $f(\cdot,\cdot):[1,\infty)\times\Omega\to\mathbb{R}$ defined by

$$f(\delta,\omega) := \log(E_Q[|\Psi_t^H|^{\frac{1}{\delta}}|\mathscr{F}_t](\omega))^{\delta}, \quad (\delta,\omega)\in[1,\infty)\times\Omega.$$

Because $|\Psi_t^H|^{1/\delta} \le 1+\Psi_t^H \in L^1(Q)$ for all $\delta\in[1,\infty)$, Lebesgue's dominated convergence theorem and Jensen's inequality for conditional expectations allow us to choose a version of f which is continuous and nonincreasing in δ for all fixed $\omega\in\Omega$, so that by monotonicity, the limit $f(\infty,\omega) := \lim_{\delta\to\infty} f(\delta,\omega)$ exists for all $\omega\in\Omega$. We next show that

$$f(\infty,\omega) = E_Q[\log\Psi_t^H|\mathscr{F}_t](\omega) \quad \text{for almost all } \omega\in\Omega. \tag{43}$$

To ease the notation, we define $g(\cdot,\cdot):[1,\infty)\times\Omega\to\mathbb{R}$ by

$$g(\delta,\omega) := (\exp(f(\delta,\omega)))^{\frac{1}{\delta}} = E_Q[|\Psi_t^H|^{\frac{1}{\delta}}|\mathscr{F}_t](\omega), \quad (\delta,\omega)\in[1,\infty)\times\Omega$$

so that $f(\delta,\omega) = \delta\log g(\delta,\omega)$. Again since $|\Psi_t^H|^{1/\delta} \le 1+\Psi_t^H \in L^1(Q)$ for all $\delta\in[1,\infty)$, dominated convergence gives

$$\lim_{n\to\infty} g(n,\omega) = 1 \quad \text{for almost all } \omega\in\Omega. \tag{44}$$

For $x > 1/2$ we have $x-1 \ge \log x \ge x-1-|x-1|^2$, from which we obtain by (44) that for almost all $\omega\in\Omega$, there exists $n_0(\omega)\in\mathbb{N}$ such that

$$n(g(n,\omega)-1) \ge f(n,\omega) \ge n(g(n,\omega)-1) - n|g(n,\omega)-1|^2, \quad n\ge n_0(\omega). \tag{45}$$

In view of (44) and (45), we get (43) if we show that

$$\lim_{n\to\infty} n(g(n,\omega)-1) = E_Q[\log\Psi_t^H|\mathscr{F}_t](\omega) \quad \text{for almost all } \omega\in\Omega. \tag{46}$$

But (46) follows from Lebesgue's convergence theorem and

$$\lim_{n\to\infty} n(|\Psi_t^H|^{\frac{1}{n}}-1) = \lim_{n\to\infty} n\left(\exp\left(\frac{1}{n}\log\Psi_t^H\right)-1\right) = \log\Psi_t^H \quad P\text{-a.s.}$$

if we show that $n||\Psi_t^H|^{1/n}-1|$, $n\in\mathbb{N}$, is dominated by a Q-integrable random variable. Due to $e^x-1\ge x$ for $x\in\mathbb{R}$ and

$$\frac{\mathrm{d}}{\mathrm{d}x}x(a^{\frac{1}{x}}-1) = a^{\frac{1}{x}}\left(1-\frac{1}{x}\log a\right)-1 \le a^{\frac{1}{x}}\exp\left(-\frac{1}{x}\log a\right)-1 = 0$$

for $a>0$ and $x>0$, it follows for $a=\Psi_t^H$ that

$$\log\Psi_t^H \le n\left(\exp\left(\frac{1}{n}\log\Psi_t^H\right)-1\right) \le \Psi_t^H-1, \quad n\in\mathbb{N}.$$

This gives $n||\Psi_t^H|^{1/n}-1| \le |\log\Psi_t^H| + \Psi_t^H \in L^1(Q)$, $n\in\mathbb{N}$, and proves (43).

Combining (41), (42) and (43) yields $f(\infty,\omega) \le \gamma k_t^H(\omega) \le f(1,\omega)$ for almost all $\omega \in \Omega$. By the intermediate value theorem, the set

$$\Delta(\omega) := \{\delta \in [1,\infty] \mid f(\delta,\omega) = \gamma k_t^H(\omega)\}$$

is thus nonempty for almost all $\omega \in \Omega$. Define $\delta_t^H : \Omega \to [1,\infty]$ by

$$\delta_t^H(\omega) := \sup \Delta(\omega), \quad \omega \in \Omega, \tag{47}$$

setting $\delta_t^H := 1$ on the P-null set $\{\omega \in \Omega | \Delta(\omega) = \emptyset\}$. By continuity of f in δ, $\Delta(\omega)$ is closed in $\mathbb{R} \cup \{+\infty\}$ for all $\omega \in \Omega$, and we get for almost all $\omega \in \Omega$ that

$$f(\delta_t^H(\omega),\omega) = \gamma k_t^H(\omega). \tag{48}$$

It remains to prove that the mapping $\omega \mapsto \delta_t^H(\omega)$ is $\mathscr{F}_t$-measurable. Because f is nonincreasing and due to (47) and (48), we have for any $a \in [1,\infty]$ that

$$\begin{aligned}\{\omega \in \Omega | \delta_t^H(\omega) < a\} &= \{\omega \in \Omega | f(\delta_t^H(\omega),\omega) > f(a,\omega)\} \\ &= \{\omega \in \Omega | \gamma k_t^H(\omega) > f(a,\omega)\} \\ &= \bigcup_{q \in \mathbb{Q}} \big(\{\omega \in \Omega | \gamma k_t^H(\omega) > q\} \cap \{\omega \in \Omega | q > f(a,\omega)\}\big)\end{aligned}$$

up to a P-null set. The last set is in $\mathscr{F}_t$ because k_t^H and $f(a,\cdot)$ for fixed $a \in [1,\infty]$ are $\mathscr{F}_t$-measurable random variables. Since $\mathscr{F}_t$ is complete, $\{\omega \in \Omega | \delta_t^H(\omega) < a\}$ is in $\mathscr{F}_t$ for every $a \in \mathbb{R} \cup \{+\infty\}$, and so δ_t^H is $\mathscr{F}_t$-measurable. □

The next result provides a simplified version of Theorem 4 based on the use of the minimal local martingale measure $\widehat{P}$.

Corollary 1 *Fix $t \in [0,T]$ and assume that H is bounded and S satisfies (SC). Suppose further that $\widehat{P}$ given by $\frac{\mathrm{d}\widehat{P}}{\mathrm{d}P} := \mathscr{E}(-\int \lambda \mathrm{d}M)_T$ is in $\mathbb{P}_0^{e,f}$, that $\int \lambda \mathrm{d}S$ is a $\widehat{P}$-, Q_0^E- and Q_H^E-martingale, and that the random variable*

$$\exp\Big(-\Big\langle \int \lambda \mathrm{d}M \Big\rangle + \frac{1}{2}\Big[\int \lambda \mathrm{d}M\Big]^c\Big)_{t,T} \prod_{t<s\le T} \frac{e^{-\lambda_s \cdot \Delta M_s}}{1-\lambda_s \cdot \Delta M_s}$$

and its logarithm are $\widehat{P}$-integrable. Then there exist $\mathscr{F}_t$-measurable random variables $\delta_t^0, \delta_t^H : \Omega \to [1,\infty]$ such that for almost all $\omega \in \Omega$,

$$\begin{aligned}h_t(\omega) &= \frac{1}{\gamma}\log(E_{\widehat{P}}[|\Psi_t^H|^{1/\delta}|\mathscr{F}_t](\omega))^{\delta}\big|_{\delta=\delta_t^H(\omega)} \\ &\quad - \frac{1}{\gamma}\log(E_{\widehat{P}}[|\Psi_t^0|^{1/\delta'}|\mathscr{F}_t](\omega))^{\delta'}\big|_{\delta'=\delta_t^0(\omega)},\end{aligned}$$

where we use the convention (39) *and the definition*

$$\Psi_t^H := \frac{\exp(\gamma H - \int_t^T \lambda_s \mathrm{d}S_s)}{\mathscr{E}(-\int \lambda \mathrm{d}M)_{t,T}} = \frac{e^{\gamma H} \exp(-\int \lambda \mathrm{d}S)_{t,T}}{\mathscr{E}(-\int \lambda \mathrm{d}M)_{t,T}}. \tag{49}$$

Proof We only need to check that Ψ_t^0, Ψ_t^H given by (49) and $\log \Psi_t^0$, $\log \Psi_t^H$ are $\widehat{P}$-integrable as the result then follows from Theorems 3 and 4 with the choice $Q := \widehat{P}$ and $\varphi := -\lambda$. Using the formula for the stochastic exponential and (SC), we get

$$\Psi_t^0 = \exp\Big(-\Big\langle \int \lambda \mathrm{d}M \Big\rangle + \frac{1}{2}\Big[\int \lambda \mathrm{d}M\Big]^c\Big)_{t,T} \prod_{t<s\leq T} \frac{e^{-\lambda_s \cdot \Delta M_s}}{1 - \lambda_s \cdot \Delta M_s},$$

and thus $\Psi_t^0, \log \Psi_t^0 \in L^1(\widehat{P})$ by assumption. The same is true for Ψ_t^H because H is bounded by assumption. □

To the best of our knowledge, results like Theorem 4 and Corollary 1 have not been available in the literature so far. A closed-form expression for the exponential utility indifference value has been known only in specific cases when the asset prices are modeled by continuous semimartingales; see for example [10] for explicit expressions of the indifference value in two Brownian settings. There the adapted process δ^H, called the *distortion power*, is closely related to the instantaneous correlation between the driving Brownian motions. The model in [10] consists of a risk-free bank account and a stock $S = S^1$ driven by a Brownian motion W. The claim H depends on another Brownian motion Y which has a time-dependent and fairly general instantaneous stochastic correlation ϱ with W, with $|\varrho|$ uniformly bounded away from 1. Theorem 2 of [10] proves that the indifference value is of the form of Corollary 1 above, with δ_t^H and δ_t^0 taking values between

$$\underline{\delta}_t := \inf_{s\in[t,T]} \frac{1}{\|1 - |\varrho_s|^2\|_{L^\infty(P)}} \quad \text{and} \quad \overline{\delta}_t := \sup_{s\in[t,T]} \left\| \frac{1}{1 - |\varrho_s|^2} \right\|_{L^\infty(P)}.$$

For small $|\varrho|$ (uniformly in s, in the L^∞-norm), the claim H is almost unhedgeable and $1/\delta^H$ is nearly 1, whereas for $|\varrho|$ close to 1, the claim H is well hedgeable and $1/\delta^H$ is nearly 0. So in that Brownian model, $1/\delta^H$ is closely related to some kind of distance of H from being attainable or hedgeable. In the subsequent discussion, we extend this idea to a more general setting, while we come back to the Brownian model in Sect. 6.

Consider the setting of Corollary 1 where S is (in addition) continuous and satisfies (SC), and H is bounded. Then the P-martingale part M of S is also continuous and the mean-variance tradeoff process $K = \langle \int \lambda \mathrm{d}M \rangle = \langle \int \lambda \mathrm{d}S \rangle$ is P-a.s. finite by (SC). The quantity Ψ_t^H from (49) then reduces to $\Psi_t^H = \exp(\gamma H - \frac{1}{2}(K_T - K_t))$, and the assumptions of Corollary 1 are satisfied if K_T is bounded, because $\int \lambda \mathrm{d}M$ is then a $BMO(P)$-martingale. If we now even suppose that K_T is deterministic, the indifference value at time 0 simplifies to

$$h_0 = \frac{1}{\gamma}\log(E_{\widehat{P}}[\exp(\gamma H/\delta)])^{\delta}|_{\delta=\delta_0^H} \tag{50}$$

by Corollary 1. If $\delta_0^H < \infty$, we can write

$$h_0 = -\widetilde{U}_H^{-1}(E_{\widehat{P}}[\widetilde{U}_H(-H)]), \quad \text{where } \widetilde{U}_H(x) := -\exp(-\gamma x/\delta_0^H),\ x \in \mathbb{R},$$

which means that $-h_0$ is a certainty equivalent of $-H$. Note, however, that this is done under $\widehat{P}$, not P, and with respect to the utility function $\widetilde{U}_H$, not U, where $\widetilde{U}_H$ depends itself on the claim H. If $\delta_0^H = 1$, then $\widetilde{U}_H$ and U coincide and H is valued by the U-certainty equivalent under $\widehat{P}$. Moreover, (38) shows that we then must have equality in (40) for $t = 0$, which implies that $\int_0^T(\gamma\eta_s^H - \lambda_s)\mathrm{d}S_s$ is deterministic, hence $\int(\gamma\eta^H - \lambda)\mathrm{d}S = 0$. In other words, the equivalent formulation (11) of $FER(H)$ in Proposition 1 simplifies in this case to

$$H = \frac{1}{\gamma}\log\mathscr{E}(\widetilde{N}^H)_T + \frac{1}{2\gamma}K_T + k_0^H,$$

which means that H consists only of a constant plus an unhedged term. This may be interpreted as saying that H has maximal distance to attainability. On the opposite extreme, the case $\delta_0^H = \infty$ leads by (50) and (39) (and still under the same assumptions) to $h_0 = E_{\widehat{P}}[H]$. Hence for $\delta_0^H = \infty$, we get a familiar no-arbitrage value for H. In this case, (38) and (39) show that we must have equality in (42) for $t = 0$; hence $\mathscr{E}(N^H) = \mathscr{E}(-\int\lambda\mathrm{d}M)$ and thus (11) simplifies to

$$H = \int_0^T \widetilde{\eta}_s^H\mathrm{d}S + \frac{1}{2\gamma}K_T + k_0^H,$$

showing that H is attainable. Summing up, we can interpret $1/\delta^H$ as the distance of H from being attainable; for $1/\delta^H = 0$ (convention: $1/\infty = 0$), the distance is minimal, whereas for $1/\delta^H = 1$, it is maximal. The following remark shows how this idea can be made mathematically more precise.

Remark 3 Assume that S is continuous, satisfies (SC) and that $K_T = \langle\int\lambda\mathrm{d}M\rangle_T$ is bounded, but not necessarily deterministic. By Theorem 4 and Corollary 1, we can attribute to any $H \in L^\infty(P)$ a number $\delta(H) := \delta_0^H$ in $[1,\infty]$ uniquely defined via (47) with $Q = \widehat{P}$ and $\varphi = -\lambda$. Defining for $G, H \in L^\infty(P)$

$$G \sim H \ :\Longleftrightarrow\ \delta\Big(G + \frac{1}{2\gamma}K_T\Big) = \delta\Big(H + \frac{1}{2\gamma}K_T\Big)$$

gives an equivalence relation on $L^\infty(P)$. We denote by $D := L^\infty(P)/\sim$ the set of its equivalence classes and associate to each equivalence class a representative. We further define the mapping $d : D \times D \to [0, 1]$ for $G, H \in D$ by

$$d(G, H) := \left| \frac{1}{\delta(G + \frac{1}{2\gamma} K_T)} - \frac{1}{\delta(H + \frac{1}{2\gamma} K_T)} \right|.$$

Clearly, d is a metric on D. A claim $G \in L^\infty(P)$ is called ($\widehat{P}$-)*attainable* if it can be written as $G = E_{\widehat{P}}[G] + \int_0^T \beta_s \mathrm{d}S_s$ for a predictable process β such that $\int \beta \mathrm{d}S$ is a $\widehat{P}$-martingale, which is then even a $BMO(\widehat{P})$-martingale. If G is attainable, the $FER^\star$ of $G + \frac{1}{2\gamma} K_T$ equals $(-\int \lambda \mathrm{d}M, \beta + \frac{1}{\gamma}\lambda, E_{\widehat{P}}[G])$, and so the term $\log \frac{\mathscr{E}(N^H)_T}{\mathscr{E}(-\int \lambda \mathrm{d}M)_T}$ vanishes identically. This implies $\delta(G + \frac{1}{2\gamma} K_T) = \infty$ by the proof of Theorem 4, hence $G \sim 0$. Therefore,

$$d(0, H) = \frac{1}{\delta(H + \frac{1}{2\gamma} K_T)}$$

is a distance of $H \in L^\infty(P)$ from attainability.

The maximal value of $d(0, \cdot)$ depends on the diversity of the filtration $\mathbb{F}$. If S has the predictable representation property in $\mathbb{F}$ in the sense that any $H \in L^\infty(P)$ is attainable (as above), then $\sim$ has only one equivalence class and $d \equiv 0$. On the other hand, suppose that there exists a nondeterministic local $\widehat{P}$-martingale N null at 0 and strongly $\widehat{P}$-orthogonal to each component of S such that $\mathscr{E}(N)$ is a $\widehat{P}$-martingale bounded away from zero and infinity. The maximal distance to attainability is then attained by $\frac{1}{\gamma} \log \mathscr{E}(N)_T$ since $d(0, \frac{1}{\gamma} \log \mathscr{E}(N)_T) = 1$.

5 A BSDE Characterization of the Indifference Value Process

In this section, we prove that the indifference value process h is (the first component of) the unique solution, in a suitable class of processes, of a backward stochastic differential equation (BSDE). This result is similar to Becherer [2] and Mania and Schweizer [19], but obtained here in a general (not even locally bounded) semimartingale model.

We assume throughout this section that

$$\mathbb{P}_0^{e,f} \neq \emptyset$$

and denote by Q_0^E the minimal 0-entropy measure. Let us consider the BSDE

$$\Gamma_t = \Gamma_0 + \frac{1}{\gamma} \log \mathscr{E}(L)_t + \int_0^t \psi_s \mathrm{d}S_s, \quad t \in [0, T] \tag{51}$$

with the boundary condition

$$\Gamma_T = H. \tag{52}$$

We introduce three different notions of solutions to (51), (52).

Definition 2 We say that the triple (Γ, ψ, L) is a *solution* of (51), (52) if

(Si) Γ is a real-valued semimartingale;
(Sii) ψ is in $L(S)$;
(Siii) L is a local Q_0^E-martingale null at 0 such that $\mathscr{E}(L)$ is a positive Q_0^E-martingale and S is a $Q(L)$-sigma-martingale, where $Q(L)$ is defined by $\frac{\mathrm{d}Q(L)}{\mathrm{d}Q_0^E} := \mathscr{E}(L)_T$.

We call (Γ, ψ, L) a *special solution* of (51), (52) if furthermore

(Siv) $\int \psi \,\mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_0^{e,f}$;
(Sv) $E_P[\mathscr{E}(L)_T \frac{\mathrm{d}Q_0^E}{\mathrm{d}P} \log(\mathscr{E}(L)_T \frac{\mathrm{d}Q_0^E}{\mathrm{d}P})] < \infty$, i.e., the probability measure $Q(L)$ defined by $\frac{\mathrm{d}Q(L)}{\mathrm{d}Q_0^E} := \mathscr{E}(L)_T$ has finite relative entropy with respect to P.

If S is locally bounded, we say that (Γ, ψ, L) is an *orthogonal solution* of (51), (52) if it satisfies (51), (52), *(Si)*, *(Sii)* and

(Siii′) L is a local Q_0^E-martingale null at 0 and strongly Q_0^E-orthogonal to every component of S and such that $\mathscr{E}(L)$ is positive.

Under the assumption that S is locally bounded,

$$\text{a triple } (\Gamma, \psi, L) \text{ is a solution of (51), (52) if and only if}$$
$$\text{it is an orthogonal solution and } \mathscr{E}(L) \text{ is a } Q_0^E\text{-martingale.} \tag{53}$$

To see this, note first that a locally bounded S is a $Q(L)$-sigma-martingale if and only if $\mathscr{E}(L)S$ is a local Q_0^E-martingale, under the assumption that $Q(L)$ is a probability measure. If (Γ, ψ, L) is a solution, then (Siii) holds and all of $\mathscr{E}(L)S$, $\mathscr{E}(L)$ and S are local Q_0^E-martingales. Hence $\mathscr{E}(L)$ is strongly Q_0^E-orthogonal to every component of S, and therefore so is L. Conversely, if (Siii′) holds, then $\mathscr{E}(L)$ is like L strongly Q_0^E-orthogonal to every component of the local Q_0^E-martingale S. Hence $\mathscr{E}(L)S$ is a local Q_0^E-martingale and thus S is a $Q(L)$-sigma-martingale if $\mathscr{E}(L)$ is a Q_0^E-martingale.

Our main result in this section is then

Theorem 5 *Assume that H satisfies* (13). *Then the indifference value process h is the first component of the unique special solution of the BSDE* (51), (52).

Theorem 5 looks at first glance like Theorem 13 of Mania and Schweizer [19]. The important difference, however, is that we do not suppose that the filtration $\mathbb{F}$ is continuous, i.e., that all local P-martingales are continuous. If $\mathbb{F}$ is continuous, then $\frac{1}{\gamma} \log \mathscr{E}(L) = L/\gamma - \frac{\gamma}{2}\langle L/\gamma \rangle$ and Theorem 5 corresponds to Theorem 13 of Mania and Schweizer [19]. (Since H is allowed to be unbounded in Theorem 5, there are some differences in the integrability properties.) However, recovering the latter result in precise form and almost full strength from Theorem 5 requires some additional work which we discuss at the end of this section. The derivation in [19] uses the martingale optimality principle, the existence of an optimal strategy for the indifference value process, and a comparison theorem for BSDEs. Our proof is completely different; it is based on our results for the $FER^\star(H)$ and its relation to the indifference value.

Theorem 4.4 of Becherer [2] is another similar result. Instead of a continuous filtration, the framework in [2] has a continuous price process driven by Brownian motions, and a filtration generated by these and a random measure allowing the modeling of non-predictable events. Again, to regain from Theorem 5 the same statement as in Theorem 4.4 of Becherer [2], some additional work is necessary.

In Corollary 3.6 of the earlier paper [1], Becherer gives a characterization of $\frac{\mathrm{d}Q_H^E}{\mathrm{d}Q_0^E}$ in a locally bounded semimartingale model. Theorem 5 can be viewed as a dynamic extension of that result to a general semimartingale model.

Proof of Theorem 5 By Lemma 1, (13) implies that $\mathbb{P}_H^{e,f} = \mathbb{P}_0^{e,f} \neq \emptyset$, and so Theorem 3 and (9) yield

$$h_t = k_t^H - k_t^0 = h_0 + \frac{1}{\gamma}\log\frac{\mathscr{E}(N^H)_t}{\mathscr{E}(N^0)_t} + \int_0^t (\eta_s^H - \eta_s^0)\mathrm{d}S_s, \quad 0 \le t \le T,$$

where (N^H, η^H, k_0^H) and (N^0, η^0, k_0^0) are the $FER^\star(H)$ and $FER^\star(0)$; see Proposition 2 for their properties. Then $\psi := \eta^H - \eta^0$ is in $L(S)$ and $\int \psi \mathrm{d}S$ is a Q-martingale for every $Q \in \mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$. By Bayes' formula, $\mathscr{E}(N^H)/\mathscr{E}(N^0)$ is the Q_0^E-density process of Q_H^E, and so it is a positive Q_0^E-martingale and its stochastic logarithm L, defined by $\mathscr{E}(L) = \mathscr{E}(N^H)/\mathscr{E}(N^0)$, is a local Q_0^E-martingale null at 0. Moreover, $\frac{\mathrm{d}Q(L)}{\mathrm{d}P} = \mathscr{E}(L)_T\frac{\mathrm{d}Q_0^E}{\mathrm{d}P} = \frac{\mathrm{d}Q_H^E}{\mathrm{d}P}$ shows $Q(L) = Q_H^E$. Hence S is a $Q(L)$-sigma-martingale and (Sv) is satisfied because Q_H^E has finite relative entropy with respect to P. Since $h_T = H$ by definition, we see that h is the first component of a special solution of the BSDE (51), (52).

To prove uniqueness, let (Γ, ψ, L) be any special solution of (51), (52). Denote by (N^0, η^0, k_0^0) the unique $FER^\star(0)$, and define

$$N := N^0 + L + [N^0, L], \qquad \eta := \eta^0 + \psi \quad \text{and} \quad k_0 := k_0^0 + \Gamma_0. \tag{54}$$

We claim that

$$(N, \eta, k_0) \quad \text{is the unique } FER^\star(H). \tag{55}$$

For the proof, we first note that $\mathscr{E}(N^0)\mathscr{E}(L) = \mathscr{E}(N^0 + L + [N^0, L]) = \mathscr{E}(N)$ by Yor's formula. Using (51), (52) and (7) for $H = 0$ thus yields

$$\begin{aligned} H &= \frac{1}{\gamma}\log(\mathscr{E}(N^0)_T\mathscr{E}(L)_T) + \int_0^T (\eta_s^0 + \psi_s)\mathrm{d}S_s + k_0^0 + \Gamma_0 \\ &= \frac{1}{\gamma}\log\mathscr{E}(N)_T + \int_0^T \eta_s \mathrm{d}S_s + k_0. \end{aligned}$$

Therefore (N, η, k_0) satisfies (7) for H, and it is enough to show that the assumptions on N and η for $FER^\star(H)$ are fulfilled. By Bayes' formula, $\mathscr{E}(N) = \mathscr{E}(N^0)\mathscr{E}(L)$ is a positive P-martingale, because $\mathscr{E}(L)$ is a positive Q_0^E-martingale by (Siii) and $\mathscr{E}(N^0)$ is the P-density process of Q_0^E. Writing next

$$\frac{\mathrm{d}P(N)}{\mathrm{d}Q_0^E} = \frac{\mathrm{d}P(N)}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}Q_0^E} = \mathscr{E}(N)_T/\mathscr{E}(N^0)_T = \mathscr{E}(L)_T,$$

we see that $P(N) = Q(L)$ which implies that

$$I(P(N)|P) = E_P\left[\mathscr{E}(L)_T\frac{\mathrm{d}Q_0^E}{\mathrm{d}P}\log\left(\mathscr{E}(L)_T\frac{\mathrm{d}Q_0^E}{\mathrm{d}P}\right)\right] < \infty$$

by (Sv) and that S is a $P(N)$-sigma-martingale by (Siii). Because (N^0, η^0, k_0^0) is the $FER^\star(0)$, $\int \eta \mathrm{d}S = \int \eta^0 \mathrm{d}S + \int \psi \mathrm{d}S$ is by Proposition 2 and (Siv) a Q-martingale for every $Q \in \mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$, hence also for $P(N)$ and Q_H^E, and so (N, η, k_0) is an $FER(H)$ satisfying (c) from Proposition 2. This implies (55). Uniqueness of the $FER^\star(H)$ and (54) now imply that Γ_0, ψ are unique; so is L due to $\mathscr{E}(L) = \mathscr{E}(N)/\mathscr{E}(N^0)$, and finally also Γ by (51). This ends the proof. □

The above argument shows in particular a close link between the $FER^\star(H)$ and the BSDE (51), (52). Provided we have the $FER^\star(0)$, we can construct $FER^\star(H)$ from the special solution of (51), (52), and vice versa. This is familiar from exponential utility indifference valuation; indeed, knowing $FER^\star(0)$ corresponds to knowing the minimal 0-entropy measure Q_0^E.

Remark 4 If S is locally bounded and H is bounded, there is another way to prove uniqueness of the first component of a special solution of the BSDE (51), (52), which we briefly sketch here. If (Γ, ψ, L) is a special solution of (51), (52), the idea is to show that Γ equals the indifference value process h, which then yields the desired uniqueness result. Let $t \in [0, T]$ and replace in the definition of $\mathscr{A}_t^H$ the condition that $\int \vartheta \mathrm{d}S$ is a Q-supermartingale for every $Q \in \mathbb{P}_H^{e,f}$ by assuming that it is a Q-martingale for every $Q \in \mathbb{P}_H^{e,f}$. We do the analogous change for $\mathscr{A}_t^0$ and note that this does not affect the values of V_t^H and V_t^0, as mentioned after the proof of Theorem 3. We now apply Proposition 3 of Mania and Schweizer [19] to obtain

$$h_t = \frac{1}{\gamma}\log \operatorname*{ess\,inf}_{\vartheta \in \mathscr{A}_t^H} E_{Q_0^E}\left[\exp\left(\gamma H - \gamma\int_t^T \vartheta_s \mathrm{d}S_s\right)\middle|\mathscr{F}_t\right]. \tag{56}$$

Using (51), (52) gives

$$\gamma H = \gamma\Gamma_0 + \log\mathscr{E}(L)_T + \gamma\int_0^T \psi_s \mathrm{d}S_s = \gamma\Gamma_t + \log\frac{\mathscr{E}(L)_T}{\mathscr{E}(L)_t} + \gamma\int_t^T \psi_s \mathrm{d}S_s,$$

which we plug into (56) to obtain

$$h_t = \Gamma_t + \frac{1}{\gamma}\log \operatorname*{ess\,inf}_{\vartheta\in\mathscr{A}_t^H} E_{Q(L)}\left[\exp\left(\gamma\int_t^T(\psi_s-\vartheta_s)\mathrm{d}S_s\right)\middle|\mathscr{F}_t\right] =: \Gamma_t + \frac{1}{\gamma}\log\Lambda,$$

where the probability measure $Q(L)$ is defined by $\frac{\mathrm{d}Q(L)}{\mathrm{d}Q_0^E} := \mathscr{E}(L)_T$. To show that $\Lambda = 1$, we first note that $Q(L)\in\mathbb{P}_0^{e,f}$ by (Sv), $\mathbb{P}_H^{e,f} = \mathbb{P}_0^{e,f}$ by Lemma 1, and $\int\psi\mathrm{d}S$ as well as $\int\vartheta\mathrm{d}S$ are Q-martingales for every $Q\in\mathbb{P}_H^{e,f} = \mathbb{P}_0^{e,f}$ by (Siv) and because $\vartheta\in\mathscr{A}_t^H$. Jensen's inequality then yields $\Lambda\geq 1$, and we obtain $\Lambda\leq 1$ by the choice $\vartheta^\star := \psi\in\mathscr{A}_t^H$. Note that also for this uniqueness proof, we have used the assumption that (Γ,ψ,L) is a *special* solution of the BSDE (51), (52), i.e., that it also satisfies (Siv), (Sv).

We have seen in Sect. 3 that the difference between $FER(H)$ and the (unique) $FER^\star(H)$ is an issue of integrability. The same thing happens here: The next example shows that the BSDE (51), (52) may have many solutions if we omit the requirement (Siv) (which corresponds to (d) in Proposition 2).

Example 2 As in Example 1, take independent P-Brownian motions W and $W^\perp$, their P-augmented filtration $\mathbb{F}$ and $d=1$, $S=W$, $H\equiv 0$. Then $Q_0^E = P$ and $(0,0,0)$ is the unique special solution of (51), (52).

As in Example 1, take $N^0 = W^\perp$ and use Proposition 1 of Emery et al. [8] to find for any $c\in\mathbb{R}$ a process $\psi(c)$ in $L(S)$ such that

$$-\frac{1}{\gamma}\log\mathscr{E}\left(N^0\right)_T - c = \int_0^T\psi_s(c)\mathrm{d}S_s \quad P\text{-a.s.}$$

If we then set $\Gamma_t(c) := c + \frac{1}{\gamma}\log\mathscr{E}(N^0)_t + \int_0^t\psi_s(c)\mathrm{d}S_s$ for $t\in[0,T]$, we easily see as in Example 1 that $(\Gamma(c),\psi(c),N^0)$ is a solution to (51), (52) and satisfies (Sv), but not (Siv). So we clearly have multiple solutions.

Theorem 5 allows us to obtain a result similar to Proposition 3.

Corollary 2 *Assume that H satisfies* (13). *Then we have for any probability measure $Q\in\mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$ and $t\in[0,T]$ that*

$$h_t = E_Q[H|\mathscr{F}_t] - \frac{1}{\gamma}E_Q[\log\mathscr{E}(L)_{t,T}|\mathscr{F}_t], \tag{57}$$

where L is the third component of the unique special solution of the BSDE (51), (52). *In particular,*

$$h_0 = E_{Q_0^E}[H] + \frac{1}{\gamma}I(Q_0^E|Q(L)), \tag{58}$$

where $\frac{\mathrm{d}Q(L)}{\mathrm{d}Q_0^E} := \mathscr{E}(L)_T$.

Proof Equation (57) follows from Theorem 5 by taking conditional Q-expectations between t and T in (51), using (52) and (Siv). Equation (58) follows for $Q = Q_0^E$. □

Remark 5 Corollary 2 raises the question if one can find a probability measure $Q \in \mathbb{P}_0^{e,f}$ such that the indifference value is the Q-conditional expectation of H. From (57) we see that $\log \mathscr{E}(L)$ must then be a Q-martingale, and if we write the Q_0^E-density process of Q as $\mathscr{E}(R)$ for some local Q_0^E-martingale R, Bayes' formula tells us that we want $\mathscr{E}(R) \log \mathscr{E}(L)$ to be a Q_0^E-martingale. Itô's formula gives

$$\begin{aligned} \mathrm{d}(\mathscr{E}(R)\log\mathscr{E}(L))_t &= \log\mathscr{E}(L)_{t-}\mathrm{d}\mathscr{E}(R)_t + \frac{\mathscr{E}(R)_{t-}}{\mathscr{E}(L)_{t-}}\mathrm{d}\mathscr{E}(L)_t \\ &\quad + \mathscr{E}(R)_{t-}\mathrm{d}\left[L^c, R^c - \frac{1}{2}L^c\right]_t \\ &\quad + \mathscr{E}(R)_{t-}((\Delta R_t + 1)\log(1+\Delta L_t) - \Delta L_t), \end{aligned}$$

where L^c and R^c denote the continuous local Q_0^E-martingale parts of L and R. For $\mathscr{E}(R)\log\mathscr{E}(L)$ to be a local Q_0^E-martingale, we must have that $R^c = \frac{1}{2}L^c$ on $\{L^c \neq 0\}$ and $\Delta R_t = \frac{\Delta L_t - \log(1+\Delta L_t)}{\log(1+\Delta L_t)}$ on $\{\Delta L_t \neq 0\}$. Therefore, we define $R = R^c + R^d$ by

$$R_t^c := \frac{1}{2}L_t^c \quad \text{and} \quad R_t^d := \sum_{0<s\le t} \frac{\Delta L_s - \log(1+\Delta L_s)}{\log(1+\Delta L_s)} I_{\Delta L_s \neq 0} - A_t, \tag{59}$$

where A is the dual predictable projection under Q_0^E of the sum in (59). Note that R^d is well defined, since $\Delta L_s > -1$, $\Delta L_s \neq 0$ implies that

$$\left|\frac{\Delta L_s - \log(1+\Delta L_s)}{\log(1+\Delta L_s)}\right| \le |\Delta L_s|;$$

in fact, $\log(1+x) \ge \frac{x}{1+x}$ for $x > -1$ implies that $|\frac{x-\log(1+x)}{\log(1+x)}| \le |x|$ for $x > -1$, $x \neq 0$. By this construction, $\mathscr{E}(R)$ and $\mathscr{E}(R)\log\mathscr{E}(L)$ are local Q_0^E-martingales, but it is not clear whether they are true Q_0^E-martingales. If they are and if Q defined by $\frac{\mathrm{d}Q}{\mathrm{d}Q_0^E} := \mathscr{E}(R)_T$ is in $\mathbb{P}_0^{e,f}$, then we obtain indeed $h_t = E_Q[H|\mathscr{F}_t]$ for all $t \in [0,T]$. In general, this representation is not linear in H since the probability measure Q may (via L) depend on H. Mania and Schweizer [19] showed in their Proposition 11 that a representation of this type exists if the filtration is continuous and H is bounded, in which case $R = \frac{1}{2}L$.

Becherer [2] and Mania and Schweizer [19] show *BMO*-estimates for all components of the solution to the BSDE for the indifference value process h. It seems doubtful if one can obtain such results in our general framework here, but under a mild additional assumption, we can still characterize (Siv) via *BMO*-properties without being more specific about the filtration $\mathbb{F}$; see Theorem 6 below.

The indifference hedging strategy β is defined as the difference of the strategies which attain $V_0^H(h_0)$ and $V_0^0(0)$, i.e., as that extra trading we do in the optimization which can be attributed to the presence of a claim. If H satisfies (13), we have $\beta = \eta^H - \eta^0 = \psi$ by (32) and the proof of Theorem 5, where ψ is the second component of the unique special solution of the BSDE (51), (52). Hence it is of particular interest to know when $\int \psi \mathrm{d}S$ is a $BMO(Q_0^E)$-martingale.

Theorem 6 *Assume that S is continuous, H is bounded and there exists $Q \in \mathbb{P}_0^{e,f}$ whose P-density process satisfies $R_{L\log L}(P)$. Let (Γ, ψ, L) be a solution of the BSDE* (51), (52) *which satisfies* (Sv). *Then the following are equivalent*:

(a) (Γ, ψ, L) *is the special solution of* (51), (52), *i.e., it also satisfies* (Siv);
(b) L *is a $BMO(Q_0^E)$-martingale, $\mathscr{E}(L)$ satisfies condition (J), and $\int \psi \mathrm{d}S$ is a Q_0^E-martingale*;
(c) $\int \psi \mathrm{d}S$ *is a $BMO(Q_0^E)$-martingale*;
(d) $\int \psi \mathrm{d}M$ *is a $BMO(P)$-martingale, where M is the P-local martingale part of S*;
(e) *there exists $\varepsilon > 0$ such that $E_P[\exp(\varepsilon[\int \psi \mathrm{d}S]_T)] < \infty$.*

Proof "(a) $\Longrightarrow$ (b)". Denote by (N^H, η^H, k_0^H) and (N^0, η^0, k_0^0) the unique $FER^\star(H)$ and $FER^\star(0)$. Theorem 2 implies that N^H, N^0 are $BMO(P)$-martingales and $\mathscr{E}(N^H)$, $\mathscr{E}(N^0)$ satisfy condition (J), say with constants C^H and C^0. By the proof of Theorem 5, we have $\mathscr{E}(L) = \mathscr{E}(N^H)/\mathscr{E}(N^0)$ and thus $\mathscr{E}(L)$ satisfies condition (J) with constant $C^H C^0$. Since $1/\mathscr{E}(N^0)$ is the Q_0^E-density process of P, $\mathscr{E}(N^0)^{-1} = \mathscr{E}(\widehat{N}^0)$ for a local Q_0^E-martingale $\widehat{N}^0$, and so $\mathscr{E}(L) = \mathscr{E}(N^H + \widehat{N}^0 + [N^H, \widehat{N}^0])$ by Yor's formula. Due to the properties of N^0 and N^H, both $\widehat{N}^0$ and $N^H + [N^H, \widehat{N}^0]$ are $BMO(Q_0^E)$-martingales by Propositions 6 and 7 of Doléans-Dade and Meyer [7], and hence so is $L = \widehat{N}^0 + N^H + [N^H, \widehat{N}^0]$. Finally, $\int \psi \mathrm{d}S$ is a Q_0^E-martingale by Siv).

"(b) $\Longrightarrow$ (c)", "(c) $\Longrightarrow$ (d)" and "(d) $\Longrightarrow$ (e)". These go along the same lines as the proofs of the corresponding implications in Theorem 2. Instead of (7) we take (51), (52), and we replace $P(N^H)$ by Q_0^E.

"(e) $\Longrightarrow$ (a)". Like for the corresponding implication in Theorem 2, we obtain that $\int \psi \mathrm{d}S$ is a square-integrable Q-martingale for any $Q \in \mathbb{P}_0^{e,f} = \mathbb{P}_H^{e,f}$, which implies (Siv). □

Remark 6 Example 2 also shows that even if the assumptions of Theorem 6 are satisfied, none of the equivalent statements (a)–(e) need hold. This is another way of saying that there exist solutions of (51), (52) which are not special solutions.

Corollary 3 *Suppose the assumptions of Theorem 6 hold. Let (Γ, ψ, L) be an orthogonal solution of the BSDE* (51), (52). *Then (Γ, ψ, L) is the special solution of* (51), (52) *if and only if both L and $\int \psi \mathrm{d}S$ are $BMO(Q_0^E)$-martingales and $\mathscr{E}(L)$ is a Q_0^E-martingale which satisfies condition (J).*

Proof The "only if" part follows immediately from Theorem 6. For the "if" part, note first that (Γ, ψ, L) is a solution of (51), (52) by (53). So we need only show that (Γ, ψ, L) satisfies (Sv) in view of Theorem 6. We first prove that $\int \psi \mathrm{d}S$ is a $BMO(Q(L))$-martingale, where $\frac{\mathrm{d}Q(L)}{\mathrm{d}Q_0^E} = \mathscr{E}(L)_T$. Because $1/\mathscr{E}(L)$ is the $Q(L)$-density process of Q_0^E, it can be written as $\mathscr{E}(L)^{-1} = \mathscr{E}(\widehat{L})$ for a local $Q(L)$-martingale $\widehat{L}$ which must satisfy $L + \widehat{L} + [L, \widehat{L}] = 0$ by Yor's formula. The continuity of S and the strong Q_0^E-orthogonality of L to S entail

$$\left[\int \psi \mathrm{d}S, \widehat{L}\right] = -\left[\int \psi \mathrm{d}S, L\right] = 0.$$

This yields by Proposition 7 of Doléans-Dade and Meyer [7] that $\int \psi \mathrm{d}S$ is a $BMO(Q(L))$-martingale. For the second component η^0 of the $FER^\star(0)$, we similarly have that $\int \eta^0 \mathrm{d}S$ is a $BMO(Q(L))$-martingale since $\int \eta^0 \mathrm{d}S$ is a $BMO(Q_0^E)$-martingale by Theorem 2. Because (Γ, ψ, L) is a solution of (51), (52), we can write

$$\log \mathscr{E}(L)_T = -\gamma \int_0^T \psi_s \mathrm{d}S_s + \gamma H - \gamma \Gamma_0,$$

and similarly, we have for the $FER^\star(0)$ (N^0, η^0, k_0^0) that

$$\log \frac{\mathrm{d}Q_0^E}{\mathrm{d}P} = \log \mathscr{E}(N^0)_T = -\gamma \int_0^T \eta_s^0 \mathrm{d}S_s - \gamma k_0^0.$$

Because $\int (\eta^0 + \psi) \mathrm{d}S$ is a $BMO(Q(L))$-martingale, we thus obtain

$$\begin{aligned} E_{Q(L)}\left[\log\left(\mathscr{E}(L)_T \frac{\mathrm{d}Q_0^E}{\mathrm{d}P}\right)\right] &= -\gamma\Gamma_0 - \gamma k_0^0 + \gamma E_{Q(L)}\left[H - \int_0^T (\eta_s^0 + \psi_s)\mathrm{d}S_s\right] \\ &= -\gamma\Gamma_0 - \gamma k_0^0 + \gamma E_{Q(L)}[H] < \infty \end{aligned}$$

since H is bounded. Hence (Γ, ψ, L) satisfies (Sv) and we are done. □

Corollary 3 allows us to recover Theorem 13 of Mania and Schweizer [19] from our Theorem 5. However, this still requires some work which is done in the next two results. A similar approach can be used to recover Theorem 4.4 of Becherer [2] from our Theorem 5, but we do not detail this here. Although the following lemma is a special case of Proposition 7 of Mania and Schweizer [19], we give the proof here as well, both for completeness and because it is quite simple in this case.

Lemma 2 *Assume that the filtration $\mathbb{F}$ is continuous, H is bounded and let (Γ, ψ, L) be an orthogonal solution of the BSDE (51), (52) with bounded first component Γ. Then L and $\int \psi \mathrm{d}S$ are $BMO(Q_0^E)$-martingales.*

Proof If L and $\int \psi \mathrm{d}S$ are true Q_0^E-martingales, (51) yields by continuity of L

$$E_{Q_0^E}[\langle L\rangle_T - \langle L\rangle_\tau | \mathscr{F}_\tau] = 2\gamma E_{Q_0^E}[\Gamma_\tau - \Gamma_T | \mathscr{F}_\tau] \quad \text{for any stopping time } \tau. \quad (60)$$

Because Γ is bounded, the right-hand side of (60) is bounded independently of τ, and thus L is a $BMO(Q_0^E)$-martingale. This implies that $(E_{Q_0^E}[\langle L\rangle_T|\mathscr{F}_s])_{0\le s\le T}$ is also a continuous $BMO(Q_0^E)$-martingale, because

$$E_{Q_0^E}[|\langle L\rangle_T - E_{Q_0^E}[\langle L\rangle_T|\mathscr{F}_\tau]||\mathscr{F}_\tau] \le 2E_{Q_0^E}[\langle L\rangle_T - \langle L\rangle_\tau|\mathscr{F}_\tau] \le 2\|L\|^2_{BMO_2(Q_0^E)}$$

for any stopping time τ. Taking conditional Q_0^E-expectations in (51) with $t = T$ gives

$$\int_0^s \psi_y \mathrm{d}S_y = E_{Q_0^E}[\Gamma_T - \Gamma_0|\mathscr{F}_s] - \frac{1}{\gamma}L_s + \frac{1}{2\gamma}E_{Q_0^E}[\langle L\rangle_T|\mathscr{F}_s], \quad 0 \le s \le T,$$

and so $\int \psi \mathrm{d}S$ is a $BMO(Q_0^E)$-martingale as well. Note that we obtain bounds for the $BMO_2(Q_0^E)$-norms of L and $\int \psi \mathrm{d}S$ that depend on Γ (and γ) alone.

For general L and $\int \psi \mathrm{d}S$, we stop at τ_n and apply the above argument with T replaced by τ_n. Letting $n \to \infty$ then completes the proof. □

A closer look at the proof of Lemma 2 shows that we did not use the property that L is strongly Q_0^E-orthogonal to S. However, this is of course necessary if we want to prove a uniqueness result. By combining Lemma 2 and Corollary 3, we obtain the following sufficient conditions for the uniqueness of an orthogonal solution of (51), (52) with bounded first component.

Proposition 4 *Assume that* $\mathbb{F}$ *is continuous*, H *is bounded*, *and there exists* $Q \in \mathbb{P}_0^{e,f}$ *whose* P*-density process satisfies* $R_{L\log L}(P)$. *Then the indifference value process* h *is the first component of the unique orthogonal solution of* (51), (52) *with bounded first component*. *Moreover*, L *and* $\int \psi \mathrm{d}S$ *are* $BMO(Q_0^E)$*-martingales*.

Proof By Theorem 5 and (53), h is the first component of an orthogonal solution of (51), (52). Using the definition (3) of h and $V_t^H(h_t) = \exp(-\gamma h_t)V_t^H(0)$ easily implies that the indifference value process h is bounded by $\|H\|_{L^\infty(P)}$. If (Γ, ψ, L) is any orthogonal solution of the BSDE (51), (52) with bounded Γ, then L and $\int \psi \mathrm{d}S$ are $BMO(Q_0^E)$-martingales by Lemma 2. By Corollary 3, (Γ, ψ, L) is then a special solution, which is unique by Theorem 5. □

Proposition 4 is almost identical to Theorem 13 in Mania and Schweizer [19]; the only difference is that we have here the additional assumption that there exists $Q \in \mathbb{P}_0^{e,f}$ whose P-density process satisfies $R_{L\log L}(P)$. The explanation for this is that we actually prove more than we really need for Proposition 4. Mania and Schweizer [19] use a comparison result for BSDEs (their Theorem 8) to deduce directly that one has uniqueness of orthogonal solutions to the BSDE within the class of those with bounded first component. In contrast, the proof of Proposition 4 actually shows that under the $R_{L\log L}$-condition, any solution with bounded first component is even a special solution—and then one appeals to Theorem 5 which asserts uniqueness within that class.

6 Application to a Brownian Setting

In this section, we consider as a special case a model with one risky asset driven by a Brownian motion and a claim coming from a second, correlated Brownian motion. All processes are indexed by $0 \le s \le T$. Let W and Y be two Brownian motions with constant instantaneous correlation ϱ satisfying $|\rho| < 1$. Choose as $\mathbb{F}$ the P-augmentation of the filtration generated by the pair (W, Y), and denote by $\mathbb{Y} = (\mathscr{Y}_s)_{0\le s\le T}$ the P-augmentation of the filtration generated by Y alone.

As usual, the risk-free *bank account* has zero interest rate. The single *tradable stock* has a price process given by

$$\mathrm{d}S_s = \mu_s S_s \mathrm{d}s + \sigma_s S_s \mathrm{d}W_s, \quad 0 \le s \le T,\ S_0 > 0, \tag{61}$$

where drift μ and volatility σ are $\mathbb{F}$-predictable processes. We assume for simplicity that μ is bounded and σ is bounded away from zero and infinity. We further assume that

the instantaneous Sharpe ratio $\frac{\mu}{\sigma}$ of the tradable stock is $\mathbb{Y}$-predictable.

In the notation of Sect. 2, $S = S_0 + M + \int \lambda \mathrm{d}\langle M\rangle$, where $M := \int \sigma S \mathrm{d}W$ is a local $(\mathbb{F}, P)$-martingale and $\lambda := \frac{\mu}{\sigma}\frac{1}{\sigma S}$ is $\mathbb{F}$-predictable. Since μ is bounded and σ is bounded away from zero, the Sharpe ratio $\frac{\mu}{\sigma}$ is also bounded, and thus $\int \lambda \mathrm{d}M = \int \frac{\mu}{\sigma}\mathrm{d}W$ is a $BMO(\mathbb{F}, P)$-martingale and $\mathscr{E}(-\int \lambda \mathrm{d}M)$ is an $(\mathbb{F}, P)$-martingale. We suppose that the contingent claim H is a bounded $\mathscr{Y}_T$-measurable random variable. Together with the structure of S in (61), this assumption on H formalizes the idea that the payoff H is driven by Y, whereas hedging can only be done in S which is imperfectly correlated with the factor Y.

In the literature, there are three main approaches to obtain explicit formulas for the resulting optimization problem (2). In a Markovian setting, Henderson [13], Henderson and Hobson [14, 15], and Musiela and Zariphopoulou [20], among others, first derive the Hamilton-Jacobi-Bellman nonlinear PDE for the value function of the underlying stochastic control problem. This PDE is then linearized by a power transformation with a constant exponent, called the *distortion power*, which corresponds to δ_0^H from Theorem 4 and Corollary 1. This method works only if one has a Markovian model. Using general techniques, Tehranchi [25] first proves a Hölder-type inequality, which he then applies to the portfolio optimization problem. The distortion power there arises as an exponent in the Hölder-type inequality. A third approach based on martingale arguments allows us in [10] to consider a more general framework with a fairly general stochastic correlation ϱ. In [10], we prove that the explicit form of the indifference value from Musiela and Zariphopoulou [20] or Tehranchi [25] is preserved, except that the distortion power, which is shown to exist but not explicitly determined, may be random and depend on H like in our general semimartingale model; compare Theorem 4 and Corollary 1.

We give here another proof based on the results of the previous sections. While there are no new results, the arguments in comparison to [10] are easier and shorter, give new insights, and show the advantage of $FER^\star(H)$ compared to the BSDE formulation (51), (52) in Sect. 5. Indeed, $FER^\star(H)$ is a representation under the original probability measure P, whereas in the BSDE formulation (51), (52), one must first determine the minimal 0-entropy measure.

Proposition 5 *For any $t \in [0,T]$ and any $\mathscr{F}_t$-measurable random variable x_t,*

$$V_t^H(x_t) = -\exp(-\gamma x_t) E_{\widehat{P}}[|\Psi_t^H|^{1-|\varrho|^2} | \mathscr{Y}_t]^{\frac{1}{1-|\varrho|^2}},$$

where $\Psi_t^H = \exp(\gamma H - \frac{1}{2}\int_t^T |\frac{\mu_s}{\sigma_s}|^2 \mathrm{d}s)$ and the minimal martingale measure $\widehat{P}$ is given by

$$\frac{\mathrm{d}\widehat{P}}{\mathrm{d}P} = \mathscr{E}\Big(-\int \frac{\mu}{\sigma}\mathrm{d}W\Big)_T. \tag{62}$$

The exponential utility indifference value h_t of H at time t equals

$$h_t = \frac{1}{\gamma(1-|\varrho|^2)} \log \frac{E_{\widehat{P}}[|\Psi_t^H|^{1-|\varrho|^2} | \mathscr{Y}_t]}{E_{\widehat{P}}[|\Psi_t^0|^{1-|\varrho|^2} | \mathscr{Y}_t]}.$$

In Corollary 1, we have shown that

$$h_t(\omega) = \frac{1}{\gamma} \log(E_{\hat{P}}[|\Psi_t^H|^{1/\delta} | \mathscr{F}_t](\omega))^{\delta}|_{\delta=\delta_t^H(\omega)}$$
$$- \frac{1}{\gamma} \log(E_{\widehat{P}}[|\Psi_t^0|^{1/\delta'} | \mathscr{F}_t](\omega))^{\delta'}|_{\delta'=\delta_t^0(\omega)},$$

and have related $1/\delta^H$ to a kind of distance of H from attainability. Here we have $1/\delta^H = 1 - |\varrho|^2$, which confirms our interpretation: The closer $1/\delta^H$ is to one, the greater is the distance of H from being attainable, because a smaller correlation ϱ between W and Y makes hedging more difficult.

Proof of Proposition 5 The idea is to explicitly derive the $FER^\star(H)$ and $FER^\star(0)$, from which the result follows by Theorem 3. In view of Proposition 1 and (10), we thus look for suitable real-valued processes $\widetilde{N}^H$ and $\widetilde{\eta}^H$ and an $\mathscr{F}_t$-measurable random variable k_t^H such that

$$H = \frac{1}{\gamma}\log \mathscr{E}(\widetilde{N}^H)_{t,T} + \int_t^T \widetilde{\eta}_s^H \sigma_s S_s \mathrm{d}\widehat{W}_s + \frac{1}{2\gamma}\int_t^T \left|\frac{\mu_s}{\sigma_s}\right|^2 \mathrm{d}s + k_t^H, \tag{63}$$

where $\widehat{W} := W + \int \frac{\mu}{\sigma}\mathrm{d}s$ is by Girsanov's theorem a Brownian motion under the minimal martingale measure $\hat{P}$ given by (62). Using Itô's representation theorem as in Lemma 1.6.7 of Karatzas and Shreve [17] for $|\Psi_t^H|^{1-|\varrho|^2}$ under $\mathbb{Y}$ and $\widehat{P}$ restricted to $\mathscr{Y}_T$, we can find a $\mathbb{Y}$-predictable process ζ with $E_{\hat{P}}[\int_0^T |\zeta_s|^2 \mathrm{d}s] < \infty$ such that

$$|\Psi_t^H|^{1-|\varrho|^2} = E_{\widehat{P}}[|\Psi_t^H|^{1-|\varrho|^2} | \mathscr{Y}_t] \mathscr{E}\Big(\int \zeta \mathrm{d}\widehat{Y}\Big)_{t,T}, \tag{64}$$

where the $(\mathbb{Y}, \widehat{P})$-Brownian motion $\widehat{Y}$ is defined by

$$\widehat{Y}_s := Y_s + \int_0^s \varrho \frac{\mu_y}{\sigma_y} \mathrm{d}y \quad \text{for } s \in [0, T].$$

Note that this argument uses that Ψ_t^H is $\mathscr{Y}_T$-measurable because $\frac{\mu}{\sigma}$ is $\mathbb{Y}$-predictable and H is $\mathscr{Y}_T$-measurable by assumption. We can write $\widehat{Y} = \varrho \widehat{W} + \sqrt{1 - |\varrho|^2} \widehat{W}^\perp$ for an $(\mathbb{F}, \hat{P})$-Brownian motion $\widehat{W}^\perp$ independent of $\widehat{W}$. Taking the logarithm in (64) results in

$$H = \frac{1}{\gamma} \int_t^T \frac{\zeta_s}{1 - |\varrho|^2} \mathrm{d}\widehat{Y}_s - \frac{1}{2\gamma} \int_t^T \frac{|\zeta_s|^2}{1 - |\varrho|^2} \mathrm{d}s + \frac{1}{2\gamma} \int_t^T \left| \frac{\mu_s}{\sigma_s} \right|^2 \mathrm{d}s + k_t^H,$$

where

$$k_t^H := \frac{1}{\gamma(1 - |\varrho|^2)} \log E_{\widehat{P}}[|\Psi_t^H|^{1 - |\varrho|^2} | \mathscr{Y}_t].$$

But this is (63) with

$$\widetilde{N}^H := \int \frac{\zeta}{\sqrt{1 - |\varrho|^2}} \mathrm{d}\widehat{W}^\perp \quad \text{and} \quad \widetilde{\eta}^H := \frac{\varrho \zeta}{\gamma(1 - |\varrho|^2)} \frac{1}{\sigma S}.$$

Clearly, $\widetilde{N}^H$ is a local $\widehat{P}$-martingale strongly $\widehat{P}$-orthogonal to S, hence also a local P-martingale strongly P-orthogonal to M. Moreover, Ψ_t^H is bounded away from zero and infinity, which implies by (64) that $\mathscr{E}(\int \zeta \mathrm{d}\widehat{Y})$ is uniformly bounded away from zero and infinity. By Theorem 3.4 of Kazamaki [18], $\int \zeta \mathrm{d}\widehat{Y}$ is then a $BMO(\mathbb{F}, \widehat{P})$-martingale and thus so is $\widetilde{N}^H$ because

$$\langle \widetilde{N}^H \rangle = \frac{1}{1 - |\varrho|^2} \int |\zeta|^2 \mathrm{d}s = \frac{1}{1 - |\varrho|^2} \left\langle \int \zeta \mathrm{d}\widehat{Y} \right\rangle.$$

This implies first that $\mathscr{E}(\widetilde{N}^H)$ is an $(\mathbb{F}, \widehat{P})$-martingale so that $\mathscr{E}(\widetilde{N}^H)\mathscr{E}(-\int \lambda \mathrm{d}M)$ is an $(\mathbb{F}, P)$-martingale, and then that also

$$\begin{aligned} \int (\gamma \widetilde{\eta}^H + \lambda) \mathrm{d}S &= \int \gamma \widetilde{\eta}^H \sigma S \mathrm{d}\widehat{W} + \int \frac{\mu}{\sigma} \mathrm{d}\widehat{W} \\ &= \frac{1}{1 - |\varrho|^2} \int \zeta \mathrm{d}\widehat{Y} - \widetilde{N}^H + \int \frac{\mu}{\sigma} \mathrm{d}\widehat{W} \end{aligned}$$

is a $BMO(\mathbb{F}, \widehat{P})$-martingale. So if we set $\frac{\mathrm{d}P(N^H)}{\mathrm{d}\widehat{P}} = \mathscr{E}(\widetilde{N}^H)_T$, then $\int (\widetilde{\eta}^H + \frac{1}{\gamma} \lambda) \mathrm{d}S$ is also a $BMO(\mathbb{F}, P(N^H))$-martingale by Theorem 3.6 of Kazamaki [18]. By Proposition 1, $(\widetilde{N}^H - \int \frac{\mu}{\sigma} \mathrm{d}W, \widetilde{\eta}^H + \frac{\mu}{\gamma\sigma} \frac{1}{\sigma S}, k_t^H)$ is thus an $FER(H)$ on $[t, T]$, and because the P-density process of $\widehat{P}$ satisfies $R_{L \log L}(P)$ since $\frac{\mu}{\sigma}$ is bounded, this $FER(H)$ is even the unique $FER^\star(H)$ on $[t, T]$ by Theorem 2. The unique $FER^\star(0)$ (N^0, η^0, k_t^0) on $[t, T]$ is constructed analogously, with Ψ_t^H replaced by Ψ_t^0. This concludes the proof in view of Theorem 3. □

Remark 7 Proposition 5 can be extended to the more general framework of case (I) in Frei and Schweizer [10] where the correlation ϱ is no longer constant, but $\mathbb{Y}$-predictable with absolute value uniformly bounded away from one. The explicit form of the indifference value is then essentially preserved; see Theorem 2 of [10] for the precise formulation. This can also be proved with our methods here, but we only sketch the main steps for $t = 0$ since the full details are a bit technical. First, one calls a triple (N^H, η^H, k_0^H) an *upper* (or *lower*) *FER*$^\star(H)$ if it has the properties of an *FER*$^\star(H)$, except that the equality sign in (7) is replaced by "$\geq$" (or "$\leq$"). One then shows that for an upper (lower) *FER*$^\star(H)$, (28) is satisfied with "$\leq$" ("$\geq$") instead of equality. In a third step, one defines constants

$$\overline{\delta} := \sup_{s\in[0,T]} \left\| \frac{1}{1-|\varrho_s|^2} \right\|_{L^\infty(P)} \quad \text{and} \quad \underline{\delta} := \inf_{s\in[0,T]} \frac{1}{\|1-|\varrho_s|^2\|_{L^\infty(P)}}$$

and finds, in the spirit of (64), $\mathbb{Y}$-predictable processes $\overline{\zeta}$ and $\underline{\zeta}$ such that

$$|\Psi_0^H|^{1/\overline{\delta}} = E_{\widehat{P}}[|\Psi_0^H|^{1/\overline{\delta}}]\mathscr{E}\Big(\int \overline{\zeta}\,\mathrm{d}\widehat{Y}\Big)_T \quad \text{and} \quad E_{\widehat{P}}\bigg[\int_0^T |\overline{\zeta}_s|^2 \mathrm{d}s\bigg] < \infty,$$

with an analogous construction for $\underline{\zeta}$. For this one uses that $\widehat{Y}$ is $\mathbb{Y}$-adapted because ϱ is $\mathbb{Y}$-predictable. Similarly to the proof of Proposition 5, one shows that $(\overline{N}^H, \overline{\eta}^H, \overline{k}_0^H)$ is an upper *FER*$^\star(H)$, where $\overline{N}^H = \int \overline{\delta}\,\overline{\zeta}\sqrt{1-|\varrho|^2}\mathrm{d}\widehat{W}^\perp - \int \frac{\mu}{\sigma}\mathrm{d}W$, $\overline{\eta}^H = \frac{\overline{\delta}\varrho\overline{\zeta}}{\gamma}\frac{1}{\sigma S} + \frac{\mu}{\gamma\sigma}\frac{1}{\sigma S}$ and $\overline{k}_0^H = \frac{\overline{\delta}}{\gamma}\log E_{\widehat{P}}[|\Psi_0^H|^{1/\overline{\delta}}]$. A completely analogous result holds for $\underline{\delta}$. Therefore, one obtains

$$-\exp(-\gamma x_0 + \gamma \underline{k}_0^H) \leq V_0^H(x_0) \leq -\exp(-\gamma x_0 + \gamma \overline{k}_0^H)$$

by the above versions of (28). Because $\delta \mapsto \delta \log E_{\widehat{P}}[|\Psi_0^H|^{1/\delta}]$ is continuous on $[\underline{\delta}, \overline{\delta}]$, interpolation then yields the existence of $\delta_0^H \in [\underline{\delta}, \overline{\delta}]$ such that

$$V_0^H(x_0) = -\exp(-\gamma x_0) E_{\widehat{P}}[|\Psi_0^H|^{1/\delta_0^H}]^{\delta_0^H}.$$

Solving the implicit equation (3) with respect to h_0 finally gives an explicit expression for h_0.

Acknowledgements Financial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management) is gratefully acknowledged. The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

References

1. Becherer, D.: Rational hedging and valuation of integrated risks under constant absolute risk aversion. Insur. Math. Econ. **33**, 1–28 (2003)

2. Becherer, D.: Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging. Ann. Appl. Probab. **16**, 2027–2054 (2006)
3. Biagini, S., Frittelli, M.: Utility maximization in incomplete markets for unbounded processes. Finance Stoch. **9**, 493–517 (2005)
4. Biagini, S., Frittelli, M.: The supermartingale property of the optimal wealth process for general semimartingales. Finance Stoch. **11**, 253–266 (2007)
5. Bobrovnytska, O., Schweizer, M.: Mean-variance hedging and stochastic control: Beyond the Brownian setting. IEEE Trans. Automat. Contr. **49**, 396–408 (2004)
6. Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M., Stricker, C.: Exponential hedging and entropic penalties. Math. Financ. **12**, 99–123 (2002)
7. Doléans-Dade, C., Meyer, P.A.: Inégalités de normes avec poids. In: Séminaire de Probabilités XIII. Lecture Notes in Mathematics, vol. 721, pp. 313–331. Springer, Berlin (1979)
8. Emery, M., Stricker, C., Yan, J.A.: Valeurs prises par les martingales locales continues à un instant donné. Ann. Probab. **11**, 635–641 (1983)
9. Föllmer, H., Schweizer, M.: Hedging of contingent claims under incomplete information. In: Davis, M., Elliott, R. (eds.) Applied Stochastic Analysis. Stochastics Monographs, vol. 5, pp. 389–414. Gordon and Breach, New York (1991)
10. Frei, C., Schweizer, M.: Exponential utility indifference valuation in two Brownian settings with stochastic correlation. Adv. Appl. Probab. **40**, 401–423 (2008)
11. Frittelli, M.: The minimal entropy martingale measure and the valuation problem in incomplete markets. Math. Financ. **10**, 39–52 (2000)
12. Grandits, P., Rheinländer, T.: On the minimal entropy martingale measure. Ann. Probab. **30**, 1003–1038 (2002)
13. Henderson, V.: Valuation of claims on nontraded assets using utility maximization. Math. Financ. **12**, 351–373 (2002)
14. Henderson, V., Hobson, D.: Real options with constant relative risk aversion. J. Econ. Dyn. Control. **27**, 329–355 (2002)
15. Henderson, V., Hobson, D.: Substitute hedging. RISK **15**, 71–75 (2002)
16. Kabanov, Yu., Stricker, C.: On the optimal portfolio for the exponential utility maximization: Remarks to the six-author paper. Math. Financ. **12**, 125–134 (2002)
17. Karatzas, I., Shreve, S.: Methods of Mathematical Finance. Applications of Mathematics, vol. 39. Springer, New York (1998)
18. Kazamaki, N.: Continuous Exponential Martingales and *BMO*. Lecture Notes in Mathematics, vol. 1579. Springer, New York (1994)
19. Mania, M., Schweizer, M.: Dynamic exponential utility indifference valuation. Ann. Appl. Probab. **15**, 2113–2143 (2005)
20. Musiela, M., Zariphopoulou, T.: An example of indifference prices under exponential preferences. Finance Stoch. **8**, 229–239 (2004)
21. Protter, P.: Stochastic Integration and Differential Equations. 2nd edn., version 2.1. Stochastic Modelling and Applied Probability, vol. 21. Springer, New York (2005)
22. Rheinländer, T.: An entropy approach to the Stein and Stein model with correlation. Finance Stoch. **9**, 399–413 (2005)
23. Schweizer, M.: On the minimal martingale measure and the Föllmer-Schweizer decomposition. Ann. Probab. **24**, 573–599 (1995)
24. Schweizer, M., From actuarial to financial valuation principles. Insur. Math. Econ. **28**, 31–47 (2001)
25. Tehranchi, M.: Explicit solutions of some utility maximization problems in incomplete markets. Stoch. Process. Appl. **114**, 109–125 (2004)

The Expected Number of Intersections of a Four Valued Bounded Martingale with any Level May be Infinite

Alexander Gordon and Isaac M. Sonin

Abstract According to the well-known Doob's lemma, the expected number of crossings of every fixed interval (a, b) by trajectories of a bounded martingale (X_n) is finite on the infinite time interval. For such a random sequence (r.s.) with an extra condition that X_n takes no more than N, $N < \infty$, values at each moment $n \geq 1$, this result was refined in Sonin (Stochastics 21:231–250, 1987) by proving that inside any interval (a, b) there are non-random sequences (barriers) (d_n), such that the expected number of intersections of d_n by (X_n) is finite on the infinite time interval. This result left open the problem of whether for such r.s. any *constant* barriers $d_n \equiv d$, $n \geq 1$, exist. The main result of this paper is an example of a bounded martingale X_n, $0 \leq X_n \leq 1$, with at most four values at each moment n, such that no constant d, $0 < d < 1$, is a barrier for (X_n). We also discuss the relationship of this problem with such problems as the behavior of a general finite nonhomogeneous Markov chain and the behavior of the simplest model of an irreversible process.

Keywords Martingale · Finite nonhomogeneous Markov chain · Irreversible process

Mathematics Subject Classification (2000) 60G42 · 60J10 · 82B35

1 Introduction

In this note we present some results that shed light on particular properties of random sequences in discrete time (X_n) which satisfy two key assumptions. First, (X_n) is a bounded (sub)(super)martingale in forward or reverse time. Second, (X_n), at each moment n, takes no more that N values, where $N < \infty$. In other words, there exists a sequence of finite sets (G_n) such that $P(X_n \in G_n) = 1$ and

---

To Youri M. Kabanov with deep respect and best wishes.

A. Gordon · I.M. Sonin (✉)
Department of Mathematics and Statistics, UNC Charlotte, 9201 University City Blvd, Charlotte, NC 28223-0001, USA
e-mail: imsonin@uncc.edu

A. Gordon
e-mail: aygordon@uncc.edu

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*, DOI 10.1007/978-3-642-02608-9_5,

$|G_n| \le N < \infty,\ n \ge 1$. The class of all random sequences that have the latter property is denoted by $\mathcal{G}^N$; the class of all random sequences that have both properties is denoted by $\mathcal{M}^N$.

The random sequences from $\mathcal{M}^N$ appear very naturally, for example, in the study of *finite nonhomogeneous Markov chains* (MCs). Let S be a countable set, (P_n) be a sequence of stochastic matrices, $Z = (Z_n)$ be a Markov chain from a family of MCs defined by a Markov model $(S, (P_n))$. Let Φ be the corresponding "tail" σ-algebra for Z, i.e. $\Phi = \bigcap_n F_{n\infty}$, where $F_{n\infty}$ a σ-algebra generated by $(Z_n, Z_{n+1}, \ldots)$. It is easy to check that if $A \in \Phi$ and $\beta_n(i) = P(A|Z_n = i)$, then the r.s. (X_n), where $X_n = \beta(Z_n)$, is a *martingale* (in forward time). Another, even more important family of (sub)martingales can be obtained as follows. Let D_1 be a subset of S. Let us set, for $n \ge 1$ and $i \in S$,

$$\alpha_n(i) = \begin{cases} P(Z_1 \in D_1|Z_n = i) & \text{if } P(Z_n = i) > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

It is easy to verify that the r.s. (Y_n) specified by $Y_n = \alpha_n(Z_n)$, $n \ge 1$, is a *martingale in reverse time*. If a subset D_1 is replaced by a sequence of sets $D_n \subseteq S$, and $\alpha_n(i)$ is defined as $\alpha_n(i) = P(Z_s \in D_s, s = 1, 2, \ldots, n|Z_n = i)$, then (Y_n) is a *submartingale* in reverse time. Obviously, if $|S| = N < \infty$ then the martingales and submartingales described above belong to $\mathcal{M}^N$.

The random sequences from $\mathcal{M}^N$ have much stronger properties than implied by the well-known Doob's convergence theorem, i.e. a theorem about the existence of limits of trajectories of a bounded (sub)martingale when time tends to infinity. Theorem 1 below describing these properties played a key role in the proof of the final part of a general theorem describing the behavior of a family of *finite nonhomogeneous* Markov chains defined by a finite Markov model $(S, (P_n))$, where S is a finite state space and (P_n) is a sequence of stochastic matrices. The striking feature of this theorem called a Decomposition-Separation (DS) theorem in [7], is that *no assumptions* on the sequence of stochastic matrices (P_n) are made. The DS theorem was initiated by a small paper of A. Kolmogorov [4, 5] and was proved in steps in a series of papers: D. Blackwell [1], H. Cohn ([2, 3] and other papers) and I. Sonin ([6–8] and other papers). We refer the reader to [7], where the final version of the DS theorem was presented and a brief survey of related results was given, and to a current expository paper [9]. Theorem 1 left an open problem described below and the main goal of our paper is to give an answer to that problem.

Before formulating Theorem 1 and the main result of this paper, Theorem 2, let us recall the well-known Doob's Upcrossing Lemma (see [10]), which lies at the foundation of Doob's convergence theorem. If (X_n) is a r.s., then the number of upcrossings of an interval (a, b) by a trajectory $X_1, X_2, \ldots$ on the infinite time interval is the number of times when a transition, maybe in a few steps, occurs from values less than a to values larger than b.

Doob's Lemma *If $X = (X_n)$ is a (sub)martingale, then the expected number of upcrossings of a fixed interval (a, b) by the trajectories of X on the infinite time interval is bounded by* $\sup_n E(X_n - a)^+/(b - a)$.

Similar statements are true for the number of downcrossings and the number of crossings. The condition $\sup_n E(X_n - a)^+ < \infty$ obviously holds if (X_n) is bounded, so for simplicity we will consider only bounded random sequences with $0 \le X_n \le 1$ for all n.

Note that the width of the interval $(b - a)$ is in the denominator of the above estimate, so Doob's lemma does not imply that inside the interval there exists a *level* d, such that the expected number of intersections of this level is finite, and in general such levels may not exist at all. But for the random sequences from $\mathcal{M}^N$ Doob's lemma can be substantially strengthened.

The following definition was introduced in [6]. A nonrandom sequence (d_n) is called a *barrier* for the r.s. $X = (X_n)$, if the *expected number of intersections* of (d_n) by trajectories of X on the infinite time interval is *finite*, i.e.

$$\sum_{n=1}^{\infty}[P(X_n \le d_n, X_{n+1} > d_{n+1}) + P(X_n > d_n, X_{n+1} \le d_{n+1})] < \infty. \tag{2}$$

Theorem 3 in [6] about the existence of barriers for processes with finite variation and a bounded number of values implies the following

Theorem 1 *Let (X_n) be a bounded r.s. from $\mathcal{M}^N$. Then inside each interval (a, b) there exists a barrier (d_n), $d_n \in (a, b)$, $n \ge 1$.*

An example in Sonin [8] shows that the barriers may not exist inside a given interval if a bounded martingale (X_n) takes a *countable* number of values, but for the random sequences from $\mathcal{M}^N$ Theorem 1 left an open problem.

Problem 1 *Is it true that for any r.s. from $\mathcal{M}^N$ in any interval (a, b) there exists a constant barrier $d_n \equiv d$, $n \ge 1$?*

For a r.s. (X_n) defined on a finite or infinite time interval $\{1, 2, \ldots, T\}$, $T \le \infty$, with values in [0, 1], denote by $N_T(x, X)$ the expected number of intersections of level x by this sequence, i.e. the value of the sum in (2) when n runs from 1 to $T - 1$ and $d_n \equiv x$ for all n. We will omit the indication of X and T, if $T = \infty$ and X is clear from the context. Similarly, by $N^+(x)$ we denote the expected number of up-intersections, i.e. the first sum in (2) when $T = \infty$ and $d_n = x$ for all $n \ge 1$. Obviously, both $N(x)$ and $N^+(x)$ are finite or both are infinite.

In the sequel, the abbreviation MCM will mean a (nonhomogeneous) MC (X_i) defined on a finite or infinite time interval $[1, 2, \ldots, T]$, $T \le \infty$, which is also a *martingale*. We also assume that $0 \le X_i \le 1$ for all i. The main result of this paper is

Theorem 2 1. *For any $X = (X_i) \in \mathcal{M}^2$, in forward time, any value $x \in (0, 1)$ is a level barrier, i.e. $N(x) < \infty$ for all $x \in (0, 1)$. In reverse time, one exception may occur: if the two possible values of a martingale have the same limit x_0, then $N(x_0)$ may equal ∞.*

2. *For any* $X = (X_i) \in \mathcal{M}^3$, $N(x) < \infty$ *for Lebesgue almost every (a.e.)* $x \in (0, 1)$, *and it is possible that* $N(x) = \infty$ *for all* $x \in G$, *where* $G \subset (0, 1)$ *is a countable set.*

3. *There is a MCM* $X = (X_i) \in \mathcal{M}^4$, *such that* $N(x) = \infty$ *for all* $x \in (0, 1)$.

We prove only the most important first statement in part 2, and we leave out the construction of a set G.

Remark 1 A Markov model $(S, (P_n))$ has a transparent deterministic interpretation (see [7]), and the DS theorem mentioned above has such an interpretation as well. According to this interpretation, the states of MC (Z_n) are represented by "cups" containing some solution (liquid), say tea. The entry $p_n(i, j)$ of the stochastic matrix P_n represents the proportion of the solution transferred from cup i to cup j at moment n. Correspondingly, $P(Z_n = i) = m_n(i)$ represents the volume of the solution in cup i at the moment n; $\alpha_n(i)$ introduced in (1) can be interpreted as the "concentration" of tea in the cup i at the moment n, and so on. Such a deterministic "colored" flow is the simplest example of an *irreversible process*. The DS theorem presented in the language of colored flows states that for any sequence (P_n) of $N \times N$ stochastic matrices the set of cups can be decomposed into a number of groups, with the decomposition possibly depending on time n, such that in each group, except possibly one, both the total volume and the concentration of tea have limits. In the "exceptional group", the total volume tends to zero, but the concentration may oscillate. The total volume of tea exchanged between these groups is *finite* on the infinite time interval. The number of groups and the decomposition are unique (up to a certain equivalence) and depend only on the sequence (P_n). Problem 1 described above is equivalent to the question of whether such a decomposition can be provided by constant values of the concentration. Accordingly, our Theorem 2 can be reformulated as follows. If there are only two cups and the concentrations of tea in those cups do not tend to a common limit, then the total amount of liquid exchanged between the cups with the concentration higher than x (before the transfer) and lower than x (after the transfer), or vice versa, is finite for any x. For three cups—such values of x form a subset of (0, 1) of full measure. For four or more cups, such x may not exist at all. We are going to present this and other possible interpretations, as well as some related results, in a separate paper. The authors would like to thank Robert Anderson and Joseph Quinn who read the first version of this paper and made valuable comments, and the anonymous referee for careful reading and thoughtful remarks.

2 Proof of Theorem 2. Cases $N = 2$ and $N = 3$

To simplify the presentation, we will consider only martingales in forward time. We denote by EX and $V(X)$, respectively, the expected value and the variance of a random variable (r.v.) X.

Case $N = 2$. The definition of $X = (X_n) \in \mathcal{M}^2$ implies that at each moment n, $X_n \in \{a_n, b_n\}$, where $0 \le a_n < b_n \le 1$, a_n is decreasing and b_n is increasing. Let $a_\infty = \lim a_n$, $b_\infty = \lim b_n$. For any $x \le a_\infty$ or $x \ge b_\infty$, obviously, $N(x) = 0$. If $x \in (a_\infty, b_\infty)$, select d_1 and d_2 so that $a_\infty < d_1 < x < d_2 < b_\infty$, and let n_0 be a number such that $a_n < d_1$, $b_n > d_2$ for all $n \ge n_0$. Then

$$N^+(x) = c + \sum_{n=n_0}^{\infty} P(X_n \le d_1, X_{n+1} > d_2), \tag{3}$$

where c is the sum of the first n_0 terms and therefore is finite.

Proposition 1 *For any two r.v.'s* X_1, X_2 *and any two numbers* $d_1 < d_2$,

$$P(X_1 \le d_1, X_2 > d_2) \le E(X_1 - X_2)^2/(d_2 - d_1)^2. \tag{4}$$

Proof The assertion of Proposition 1 follows immediately from the implications $(X_1 \le d_1, X_2 \ge d_2) \subset (|X_1 - X_2| \ge h)$, $h = d_2 - d_1$, and the Chebyshev's inequality, $P(|Y| \ge h) \le EY^2/h^2$ for any r.v. Y.

To prove part 1 of Theorem 2, note that for a martingale (X_n) we have $E(X_{n+1}|X_n) = X_n$, and $V(X_{n+1} - X_n) = E(X_{n+1} - X_n)^2 = EX_{n+1}^2 - EX_n{}^2$ and hence for a bounded martingale (X_n), $0 \le X_n \le 1$,

$$\sum_{n=k}^{T-1} E(X_{n+1} - X_n)^2 = EX_T^2 - EX_k{}^2 \le 1. \tag{5}$$

Then formula (5) and Proposition 1 imply that the sum in (3) is finite for any $T \le \infty$. Part 1 of Theorem 2 is proved.

For any r.v. X with $EX = m$, let us denote $M^+(X) = E(X - m)^+$. Then $E(X - m)^+ = E(m - X)^+$, and $M(X)$, the mean absolute deviation of X, is

$$M(X) = E|X - m| = E(X - m)^+ + E(m - X)^+ = 2M^+(X). \tag{6}$$

Let us also put $M(X|c) = E|X - c|$ and $M^+(X|c) = E(X - c)^+$. If (X_n) is a martingale, then the equality $E(X_{n+1}|X_n) = X_n$ and (6) imply that

$$M(X_{n+1}|X_n) = 2M^+(X_{n+1}|X_n). \tag{7}$$

To study the case of $N = 3$, we need some simple properties of r.v.'s and martingales with two and three values. They are described in Propositions 2–4. Let $X \in G^2$, i.e. an r.v. with two values a and b, $a < b$, $b - a = d$ and $P(X = b) = p$, $P(X = a) = q = 1 - p$. Then it is easy to check that the following statement is true. □

Proposition 2 *If* $X \in G^2$, *then* $M(X) = 2pqd$, $V(X) = pqd^2$, *and hence*

$$M^+(X) = V(X)/d. \tag{8}$$

Let X be an r.v. with three values $a \le e \le b$, and $P(X=b)=p$, $P(X=e)=r$, $P(X=a)=q$, where $p+r+q=1$. Let us denote by X^U an r.v. obtained from X by averaging the two upper values, i.e. X^U takes two values a and $b'=(er+bp)/(r+p)$ with probabilities q and $p'=r+p$. Similarly, we define an r.v. X^L as an r.v. obtained from X by averaging the two lower values, i.e. X^L takes two values $a'=(aq+er)/(q+r)$ and b with probabilities $q'=q+r$ and p. It is easy to check that for such r.v.'s the following statement is true

Proposition 3 (a) $a \le a' \le e \le b' \le b$, $EX = EX^A$, $V(X^A) \le V(X)$, $A = U$ *or* L,
(b) $M(X^U) \le M(X)$, $M^+(X^U) \le M^+(X)$, *with equalities when* $EX \le e$,
(c) $M(X^L) \le M(X)$, $M^+(X^L) \le M^+(X)$, *with equalities when* $EX \ge e$.

For a r.s. $(X_n) \in \mathcal{M}^3$, i.e. for a martingale with three values $a_n \le e_n \le b_n$ at moment n, similarly to Proposition 3, it is easy to obtain

Proposition 4 (a) *If* $(X_n) \in \mathcal{M}^3$, *then* $E(X^A_{n+1}|X_n) = X_n$, $E((X^A_{n+1} - X_n)^2|X_n) \le E((X_{n+1} - X_n)^2|X_n)$, $V(X^A_{n+1} - X_n) \le V(X_{n+1} - X_n)$, $A = U$ *or* L,
(b) $M^+(X^U_{n+1}|X_n) \le M^+(X_{n+1}|X_n)$, *with equality when* $X_n \le e_{n+1}$,
(c) $M^+(X^L_{n+1}|X_n) \le M^+(X_{n+1}|X_n)$, *with equality when* $X_n \ge e_{n+1}$.

Now we can prove part 2 of Theorem 2 (case $N=3$). The situation in this case is substantially different from $N=2$ and $N>3$. It is possible to have $N(x)=\infty$, for example, for all rational numbers; nevertheless the Lebesque measure of the set of all such x is always zero. We prove here only the latter statement.

Lemma 1 *For any r.s.* (X_n), $0 \le X_n \le 1$, *and* $T \le \infty$,

$$\int_0^1 N_T^+(x)dx = \sum_{n=1}^{T-1} EM^+(X_{n+1}|X_n). \tag{9}$$

Proof The proof follows immediately from the definition of $N_T^+(x) = \sum_{n=1}^{T-1} P(X_n \le x, X_{n+1} > x)$ and the equalities: (1) $P(X_n \le x, X_{n+1} > x) = EI(X_n, X_{n+1}|x)$, where $I(c,d|x) = 1$ if $c \le x < d$, and 0 otherwise, (2) $\int_0^1 I(c,d|x)dx = (d-c)^+$ and (3) $E(X_{n+1} - X_n)^+ = EM^+(X_{n+1}|X_n)$.

Let $(X_n) \in \mathcal{M}^3$ and $\{a_n, e_n, b_n\}$ be the ordered set of possible values of X_n at the moment n, $0 \le a_n \le e_n \le b_n \le 1$. The definition of a martingale again implies that the sequence (a_n) can only decrease, the sequence (b_n) can only increase but the sequence (e_n) may oscillate between a_n and b_n. WLOG we can assume that $\lim a_n = 0$, $\lim b_n = 1$.

Note that if, given a, b, $0 \le a < b \le 1$, in the left side of (9) we change the limits of integration from 0 and 1, to a and b, i.e. consider $\int_a^b$, then the equality (9) remains true with $M^+(X_{n+1}|X_n)$ replaced by $M^+(X_{n+1}|X_n, a, b) \equiv E(\min(b, X_{n+1}) - \max(a, X_n))^+$. For simplicity we will denote $M^+(X_{n+1}|X_n, 2\varepsilon, 1-2\varepsilon)$ as $M^+(X_{n+1}|X_n, \varepsilon)$.

It is sufficient to prove that the integral similar to that in (9), with integration limits replaced by 2ε and $1-2\varepsilon$, is finite for any $\varepsilon > 0$. Then obviously $N_T^+(x) < \infty$ almost surely. To prove this we will show that for any $(X_n) \in \mathcal{M}^3$, any ε, $0 < \varepsilon < 1/4$, and sufficiently large n,

$$M^+(X_{n+1}|X_n, \varepsilon) \le E((X_{n+1} - X_n)^2|X_n)/\varepsilon. \tag{10}$$

If (X_n) is a martingale, then $E(E((X_{n+1} - X_n)^2|X_n)) = E(X_{n+1} - X_n)^2 = V(X_{n+1} - X_n)$. Since series in (5) is convergent, estimate (10) will prove that the integral in (9) is finite.

Let $\varepsilon > 0$ and n_0 be a number such that $a_n < \varepsilon$, and $1 - \varepsilon < b_n$ for all $n \ge n_0$. In the sequel we will consider only $n \ge n_0$.

If $X_n > 1 - 2\varepsilon$, then, obviously, $M^+(X_{n+1}|X_n, \varepsilon) = 0$. Since $b_n > 1 - \varepsilon$, we need to consider further only the cases $X_n = a_n$ or $X_n = e_n$.

If $X_n = a_n$ or e_n, and $e_{n+1} \le 2\varepsilon$, then $M^+(X_{n+1}|X_n, \varepsilon) = M^+(X^L_{n+1}|X_n, \varepsilon)$. Using formula (8) applied to X^L_{n+1} we have $M^+(X^L_{n+1}|X_n, \varepsilon) \le M^+(X^L_{n+1}|X_n) = E((X^L_{n+1} - X_n)^2|X_n)/(b_{n+1} - a'_{n+1})$ and then using point (a) of Proposition 4, we obtain $M^+(X^L_{n+1}|X_n, \varepsilon) \le E((X_{n+1} - X_n)^2|X_n)/(b_{n+1} - a'_{n+1})$. Since $a'_{n+1} \le e_{n+1} \le 2\varepsilon$ and $b_{n+1} > 1 - \varepsilon$, we have $b_{n+1} - a'_{n+1} \ge 1 - 3\varepsilon \ge \varepsilon$, and thus (10) holds.

If $e_{n+1} \ge 2\varepsilon$ and $X_n \le 2\varepsilon$ then $M^+(X_{n+1}|X_n, \varepsilon) = M^+(X^U_{n+1}|X_n, \varepsilon)$. Using formula (8) applied to X^U_{n+1} we have $M^+(X^U_{n+1}|X_n, \varepsilon) \le M^+(X^U_{n+1}|X_n) = E((X^U_{n+1} - X_n)^2|X_n)/(b'_{n+1} - a_{n+1})$ and then using point (a) of Proposition 4, we obtain $M^+(X^U_{n+1}|X_n, \varepsilon) \le E((X_{n+1} - X_n)^2|X_n)/(b'_{n+1} - a_{n+1})$. Since $b'_{n+1} \ge e_{n+1} \ge 2\varepsilon$ and $a_{n+1} < \varepsilon$, we have $b'_{n+1} - a_{n+1} \ge \varepsilon$, and thus (10) holds. What remains are two cases when $2\varepsilon \le X_n = e_n \le 1 - 2\varepsilon$, $e_{n+1} \ge 2\varepsilon$ and $X_n = e_n \le e_{n+1}$ or $X_n = e_n \ge e_{n+1}$. The proofs are similar to the above. Part 2 of Theorem 2 is also proved. □

3 Proof of Theorem 2. Case $N > 3$. An Example

We prove part 3 of Theorem 2 by a direct construction of the MCM $X = (X_i)$ for $N = 4$. First, we construct an auxiliary MCM $U = (U_i)$. Let (a_k), (b_k), $k = 1, 2, \dots$ be two deterministic sequences such that:

$$\begin{aligned} &1 > a_1 > a_2 > \cdots > 0, \qquad a_1 < b_1 < b_2 < \cdots < 1, \\ &\lim a_k = 0, \qquad \lim b_k = 1. \end{aligned} \tag{11}$$

Given such sequences (a_k) and (b_k), we can define a MC $U = (U_i)$, $i = 1, 2, \dots$, such that $P(U_1 = b_1) = 1$, $U_{2k-1} \in \{a_k, b_k, 1\}$, $U_{2k} \in \{a_k, 1\}$, and the transition probabilities $u_i(x, y)$ are: $u_k(1, 1) = 1$, $u_{2k-1}(a_k, a_k) = 1$, $u_{2k-1}(b_k, 1) + u_{2k-1}(b_k, a_k) = 1$, $u_{2k}(a_k, b_{k+1}) + u_{2k}(a_k, a_{k+1}) = 1$, $k \ge 1$. To obtain not just a MC but a (unique) MCM, it is sufficient to define

$$u_{2k-1}(b_k, 1) = \frac{b_k - a_k}{1 - a_k}, \qquad u_{2k}(a_k, b_{k+1}) = \frac{a_k - a_{k+1}}{b_{k+1} - a_{k+1}}, \quad k \geq 1.$$

Assumptions (11) imply that this is possible and that $m_k = P(U_{2k-1} = b_k) > 0$ for all $k \geq 1$.

The MCM $U = (U_i)$ will serve as a "frame sequence" for MCM $X = (X_i)$, i.e. (X_i) will consist of "blocks" (X_i^k), $k \geq 1$, where each (X_i^k) is a MCM defined on a time interval $[t_k, e_k]$, $t_1 = 1$, $t_{k+1} = e_k + 1$, $k \geq 1$, and each block is "inserted" into a constructed above "frame sequence" $U = (U_i)$ so that the time interval $[2k-1, 2k]$ "stretches" into the time interval $[t_k, e_k]$, $k \geq 1$. More precisely, the values and the transition probabilities $p_i(x, y)$ for MCM (X_i) are defined as follows: $X_i = X_i^k$ for $i \in [t_k, e_k]$, $i = 1, 2, \ldots$; $P(X_1 = X_1^1 = b_1) = 1$, $P(X_{e_1} = X_{e_1}^k \in \{a_1, 1\}) = 1$. Any other k-th block, $k \geq 2$, has three entrance points $\{a_k, b_k, 1\}$ and two exit points $\{a_k, 1\}$, i.e. $P(X_{t_k} = X_{t_k}^k \in \{a_k, b_k, 1\}) = 1$, and $P(X_{e_k} = X_{e_k}^1 \in \{a_k, 1\}) = 1$. The state 1 is an absorbing state, $p_i(1, 1) = 1$ for all $i \geq 1$. The transition probabilities *between* blocks, i.e. at moments e_k, $k \geq 1$, are defined using the transition probabilities from MCM U: if $i = e_k$ then $p_i(a_k, y) = u_{2k}(a_k, y)$, where $y = a_{k+1}$ or b_{k+1}. The transition probabilities of k-th block $p_i(x, y)$, $t_k \leq i < e_k$, are as follows: $p_i(a_k, a_k) = 1$ for all i, the other transition probabilities are the "shifted" probabilities from MCM $Y^k = (Y_i^k)$, $k \geq 1$, where (Y_i^k) is defined on the time interval $[1, T_k]$, $T_k = e_k - t_k + 1$, i.e. $p_{t_k+i-1}(x, y) = q_i^k(x, y)$, where $i = 1, 2, \ldots, T_k$ and $q_i^k(x, y)$ are transition probabilities for (Y_i^k). We say that block X^k is obtained from a block Y^k by a *shift* from interval $[1, T_k]$ to interval $[t_k, e_k]$, $e_k = t_k + T_k - 1$.

The structure of each MCM $Y^k = (Y_i^k)$, $k = 1, 2, \ldots$ is similar and its properties are described in Lemma 2 which is the key element of our construction.

Lemma 2 *For every tuple* $\beta = (a, b, \varepsilon, C)$, $0 \leq a < b < 1$, $0 < \varepsilon < b - a$, $C > 0$, *there is a MCM* $Y = (Y_i)$ *defined on a finite time interval* $[1, 2, \ldots, T]$, $T = T(\beta)$, *and such that*

(1) $P(Y_1 = b) = 1$, $P(Y_T \in \{a, 1\}) = 1$, *and for all other* i, $1 < i < T$, Y_i *takes no more than three values* $r_i, s_i, 1$, $a \leq r_i < s_i \leq 1$.
(2) $N_T(x, Y) \geq C$ *for each* $x \in (a + \varepsilon, b)$.

We will prove Lemma 2 later. Assuming that Lemma 2 holds, we next construct a MCM (X_i) satisfying part (3) of Theorem 2.

Let (ε_k) be a sequence, $\varepsilon_k > 0$, $\lim \varepsilon_k = 0$, and let (a_k), (b_k) be two sequences satisfying conditions (11). Let (U_i) be a corresponding "frame" MCM, $u_i(x, y)$ its transitional probabilities, $i = 1, 2, \ldots$ and $m_k = P(U_{2k-1} = b_k) > 0$, $k \geq 1$. Given $k \geq 1$, let $Y^k = (Y_i^k)$, $i = 1, 2, \ldots, T_k$, be a MCM satisfying the conditions of Lemma 2 with parameters $(a_k, b_k, \varepsilon_k, C_k)$, where $C_k = 1/m_k$. We define sequences (t_k) and (e_k) by: $t_1 = 1$, $e_k = t_k + T_k - 1$, $t_{k+1} = e_k + 1$, $k \geq 1$. Let us denote by (X_i) the combined MC consisting of blocks X^k obtained by the corresponding shift from Y^k and connected by the frame sequence (U_i) as described above.

Let us denote by $N^k(x) \equiv N_{T_k}(x, Y^k)$ the expected number of intersections of level x by a r.s. $Y^k = (Y_i^k)$, $k \geq 1$, and $N(x) \equiv N(x, X)$ the expected number of intersections of level x by a r.s. $X = (X_i)$. By our choice of C_k we have $N^k(x) \geq 1/m_k$ for each $x \in (a_k + \varepsilon_k, b_k)$, where $m_k = P(X_{t_k} = b_k), k \geq 1$. Let $x \in (0, 1)$ and let $k(x)$ be a number such that $x \in (a_k + \varepsilon_k, b_k)$ for all $k \geq k(x)$. Then by our construction

$$N(x) \geq \sum_{k=1}^{\infty} P(X_{t_k} = b_k) N^k(x) \geq \sum_{k \geq k(x)}^{\infty} m_k/m_k = \infty.$$

Note that MCM (X_i) takes no more than four values on each time interval $[t_k, e_k]$, $k \geq 1$. Three values, r_i^k, s_i^k and 1 are from MCM Y^k in Lemma 2, and the fourth value is a_k.

Thus, to prove part 3 of Theorem 2 we only need to prove Lemma 2. From now on, the numbers (indices) i, k and n and such notation as $p_i(x, y)$ have a new meaning.

We prove Lemma 2 for a special case when $a = 0$, $b = \frac{1}{2}$. The general case requires only minor changes in notation.

We will construct (Y_i) combining a finite number of MCMs, having a similar structure. To avoid confusion with the "blocks" used above, we call these MCMs *modules*. Each module $(L_i^{k,r})$ is a MCM characterized by two parameters (k, r), $k \geq 1, 0 \leq r < 1$, and defined on the time interval $[1, 2, \ldots, k]$.

First we describe the *standard module* with parameters $(k, 0)$. This is a MC $(L_i^{k,0}) \equiv (S_i)$ defined on $[1, 2, \ldots, k]$, and taking at each moment i two values 0 and s_i, where (s_i) is a deterministic sequence given by formula

$$s_i = \frac{1}{k + 1 - i}, \quad i = 1, 2, \ldots, k. \tag{12}$$

Obviously $0 < \frac{1}{k} = s_1 < \cdots < s_{k-1} = \frac{1}{2} < s_k = 1$, and s_i satisfy $s_{i+1}(1 - s_i) = s_i$. Point s_1 is an initial point for a r.s. (S_i), i.e. $P(S_1 = s_1) = 1$. The transition probabilities $p_i(x, y)$ are defined as follows. State 0 is absorbing for all i, i.e. $p_i(0, 0) = 1$, $i = 2, 3, \ldots, k - 1$. The other transition probabilities are given by

$$p_i(s_i, 0) = s_i, \qquad p_i(s_i, s_{i+1}) = 1 - s_i = \frac{k - i}{k + 1 - i}, \quad i = 1, \ldots, k - 1. \tag{13}$$

It is easy to see that $E(S_{i+1}|S_i = s_i) = s_{i+1}(1 - s_i) = s_i$. Therefore, the r.s. (S_i) is also a martingale, i.e (S_i) is a MCM.

It is easy to check that

$$P(S_i = s_i) = \prod_{j=1}^{i-1} p_j(s_j, s_{j+1}) = \frac{k + 1 - i}{k} = \frac{1}{k s_i}, \quad i = 1, 2, \ldots, k. \tag{14}$$

Let us denote by $N^k(x)$ the expected number of intersections of level x by the r.s. (S_i). If $x \in (s_i, s_{i+1})$, $i = 1, 2, \ldots, k-2$, then every trajectory can intersect x on the way up and after that on the way down, so $N^k(x) = P(S_{i+1} = s_{i+1}) + P(S_{i+1} = s_{i+1}, S_k = 0) = 2P(S_{i+1} = s_{i+1}) - P(S_k = 1) = \frac{2(k-i)-1}{k} \geq 1/ks_i$. These relations imply that

$$N^k(x) \geq f^k(x), \quad \text{where } f^k(x) = \frac{1}{kx}, \text{ if } \frac{1}{k} < x \leq 1/2. \tag{15}$$

The *module* $(L_i^{k,r})$ with parameters (k,r), $0 \leq r < 1$, $k \geq 1$ is a r.s. defined on the finite time interval $i = 1, 2, \ldots, k$ by equalities

$$L_i^{k,r} = r + (1-r)S_i, \quad i = 1, 2, \ldots, k, \tag{16}$$

where $(S_i) = (L_i^{k,0})$ is a standard module with parameters $(k, 0)$.

Formula (16) implies that the initial point for $(L_i^{k,r})$ is

$$r + \frac{1-r}{k} \tag{17}$$

and that $(L_i^{k,r})$ is also a MCM with the same transitional probabilities as in (13) but with possible values r and $r + (1-r)s_i$ instead of 0 and s_i. The value r is the smallest of possible values for this module, so later we will refer to r as to the "floor" of this module. The intersection function $N^{k,r}(x)$ for $(L_i^{k,r})$, instead of (15), satisfies the inequality

$$N^{k,r}(x) \geq f^k\left(\frac{x-r}{1-r}\right), \quad r + \frac{1-r}{k} \leq x < r + \frac{1-r}{2} = \frac{1+r}{2}. \tag{18}$$

Formula (14) for $i = k$ and formula (16) imply that

$$P(L_k^{k,r} = 1) = \frac{1}{k}, \qquad P(L_k^{k,r} = r) = \frac{k-1}{k}. \tag{19}$$

Now we will construct a sequence of MCMs (Y_i^n), $n = 1, 2, \ldots, i = 1, 2, \ldots, T_n$, and we will show that for any $\varepsilon > 0$ and any C each of these MCMs will satisfy the condition of Lemma 2 for sufficiently large n. Each (Y_i^n) consists of n modules connected subsequently, each with parameters (k_j, r_j), $j = 1, 2, \ldots, n$. The parameters (k_j, r_j), $j = 1, 2, \ldots, n$, given $n = 1, 2, \ldots$, are selected as follows

$$k_j = n + j, \qquad r_j = \frac{n-j}{2n}, \quad j = 1, \ldots, n. \tag{20}$$

It is easy to check that

$$r_{j-1} = r_j + \frac{1-r_j}{k_j} = r_j + \frac{1}{2n}, \quad j = 2, \ldots, n. \tag{21}$$

Thus, for each n, points r_j divide the interval $(0, \frac{1}{2})$ into n equal parts of size $1/2n$ and the interval $(1-r_j, 1)$ contains k_j subintervals of this size. Let us denote, for the sake of brevity, $(L_i^{k_j,r_j})$ by (L_i^j). Formulas (17) and (21) imply that the floor r_j of the module (L_i^{j-1}) serves as the initial point for the next module (L_i^j).

Let $T_n = \sum_{j=1}^n k_j - n + 1$ be the total length of the time interval where these modules are sequentially defined. We define $(Y_i) \equiv (Y_i^n)$, $1 \le i \le T_n$, as follows. State 1 is absorbing for all i. At moment 1 r.s. (Y_i) starts at $r_0 = \frac{1}{2}$ and on the time interval $[1, k_1]$ coincides with the module $(L_i^{k_1,r_1}) = (L_i^1)$. Then, at moment k_1 according to (19), we have

$$\begin{aligned} P(Y_{k_1} = 1) &= P(L_{k_1}^1 = 1) = \frac{1}{k_1}, \\ P(Y_{k_1} = r_1) &= P(L_{k_1}^1 = r_1) = \frac{k_1 - 1}{k_1}. \end{aligned} \tag{22}$$

On the time interval $[k_1, k_1 + k_2 - 1]$ r.s. (Y_i) stays at 1 with probability $\frac{1}{k_1}$ and with probability $m_2 = \frac{k_1 - 1}{k_1}$ coincides with the module (L_i^2). As mentioned above, the floor r_1 of the module (L_i^1) serves as the initial point for the next module (L_i^2). And so on. Obviously, MCM (Y_i) satisfies the condition (1) of Lemma 2 with $b = \frac{1}{2}$ and $a = 0$.

From the above construction, using the last equality in formula (14) for $i = k = k_j$ and denoting $m_0 = 1$, we also obtain that for $j = 1, 2, \dots, n$,

$$\begin{aligned} m_j &= P(Y_{k_1 + \dots + k_j - j + 1} = r_j) \\ &= m_{j-1} P(L_{k_j}^j = r_j) = m_{j-1} \frac{k_j - 1}{k_j} = \frac{n}{n + j}. \end{aligned} \tag{23}$$

Our last step is to estimate $N^{(n)}(x)$, the expected number of intersections of level x by the constructed MCM (Y_i^n), $i = 1, 2, \dots, T_n$, and to show that for any ε, $0 < \varepsilon < 1/2$, $\lim_n N^{(n)}(x) = \infty$ uniformly for all x, $\varepsilon \le x \le 1/2$. Therefore, given any number C, for sufficiently large n, MCMs (Y_i^n) will satisfy the condition (2) of Lemma 2.

By our construction, we have obviously $N^{(n)}(x) = \sum_{j=1}^n m_{j-1} N^j(x)$, where $N^j(x)$ is the expected number of intersections of level x by module (L_i^j). Using for each j the estimate (18) with $k = k_j$ and $r = r_j$ taken from (20), and taking into account that by (23), $\frac{1}{2} \le m_j \le 1$ for all j, we obtain that

$$N^{(n)}(x) \ge \frac{1}{2} \sum_{j=1}^n f^{k_j}\left(\frac{x - r_j}{1 - r_j}\right), \tag{24}$$

where $f^{k_j}(\frac{x - r_j}{1 - r_j})$ is defined by (15) (see also (18)) for $r_{j-1} \le x < (1 + r_j)/2$ and for other x's can be defined to be equal to zero. Hence $f^{k_j}(\frac{x - r_j}{1 - r_j}) \ge \frac{1 - r_j}{k_j(x - r_j)} = \frac{1 - r_j}{2n(x - r_j)}$ for $r_{j-1} \le x \le (1 + r_j)/2$. Using formulas (20) and (21), we obtain that for any x,

$0 < x \le 1/2$,

$$N^{(n)}(x) \ge \frac{1}{2} \sum_{j: r_{j-1} \le x}^{n} \frac{1}{2n(x - r_j)} = \frac{1}{2} \sum_{j: n-j+1 \le 2nx}^{n} \frac{1}{2n(x - \frac{n-j}{2n})}$$
$$= \frac{1}{2} \sum_{k=0}^{[2nx]-1} \frac{1}{2n(x - \frac{k}{2n})}, \tag{25}$$

where $[a]$ is an integer part of a.

The last sum is just a Riemann sum of the integral $\int_0^x \frac{dy}{x-y} = \infty$. Thus, for large n the sum $N^{(n)}(x)$ in (25) can be made arbitrarily large uniformly for all x, $0 < \varepsilon \le x \le 1/2$. This proves Lemma 2 and therefore part 3 of Theorem 2.

Remark A slightly different but similar construction proves the analog of Theorem 2 for the case where (X_i) is a martingale in reverse time.

References

1. Blackwell, D.: Finite nonhomogeneous Markov chains. Ann. Math. **46**, 594–599 (1945)
2. Cohn, H.: Finite nonhomogeneous Markov chains: asymptotic behavior. Adv. Appl. Probab. **8**, 502–516 (1976)
3. Cohn, H.: Products of stochastic matrices and applications. Int. J. Math. Sci. **12**, 209–333 (1989)
4. Kolmogoroff, A.N., Zur Theorie der Markoffschen Ketten. Math. Ann. **112** (1936)
5. Shiryaev, A.N. (ed.): Selected Works of A.N. Kolmogorov, vol. 2. Probability Theory and Mathematical Statistics, Kluwer Academic, Dordrecht (1992)
6. Sonin, I.: Theorem on separation of jets and some properties of random sequences. Stochastics **21**, 231–250 (1987)
7. Sonin, I.: The asymptotic behaviour of a general finite nonhomogeneous Markov chain (the decomposition-separation theorem). In: Fergusson, T.S., Shapley, L.S., MacQueen, J.B. (eds.) Statistics, Probability and Game Theory, papers in Honor of David Blackwell. Lecture Notes-Monograph Series, vol. 30, pp. 337–346. IMS (1996)
8. Sonin, I.: On some asymptotic properties of nonhomogeneous Markov chains and random sequences with countable number of values. In: Kabanov, Y., Rozovskii, B., Shiryaev, A. (eds.) Statistics and Control of Stochastic Processes. The Liptser Festschrift, Proceedings of Steklov Mathematical Institute Seminar, pp. 297–313. World Scientific, Singapore (1997)
9. Sonin, I.: The Decomposition-separation theorem for finite nonhomogeneous Markov chains and related problems. In: Ethier, S., Feng, J., Stockbridge, R.H. (eds.) Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz. pp. 1–15. Institute of Mathematical Statistics, Beachwood (2008)
10. Stroock, D.W.: Probability Theory, an Analytic View. Cambridge University Press, Cambridge (1999)

Immersion Property and Credit Risk Modelling

Monique Jeanblanc and Yann Le Cam

Abstract The goal of this paper is to study the immersion property through its links with credit risk modelling. The construction of a credit model by the enlargement of a reference filtration with the progressive knowledge of a credit event occurrence has become a standard for reduced form modelling. It is known that such a construction rises mathematical difficulties, mainly relied to the properties of the random time. Whereas the invariance of the property of semi-martingale in the enlargement is implied by the absence of arbitrage, we address in this paper the question of the invariance of the martingale property.

Keywords Initial and progressive enlargement of filtration · Credit risk · Risk-neutral probability

Mathematics Subject Classification (2000) 60G46

1 Introduction

Most of the literature on credit risk focuses on pricing problems and postulates the existence of a pricing measure, without questioning its features. The purpose of this paper is to propose a study of the set of equivalent martingale measures (e.m.ms) in the context of credit modelling. Within the reduced form approach and particularly under the filtration enlargement framework, such questions may be precisely studied, and lead to interesting properties. Only finite time horizon problems will be treated in this paper.

Three parts will be developed in the sequel, so that to present the issues relative to the discussion about the completeness of a market potentially exposed to a credit event.

- The first one presents the credit modelling framework and discusses the meaning of the options taken. We adopt a reduced form model, and specify the split

This research benefited from the support of the “Chaire Risque de Crédit”, Fédération Bancaire Française.

M. Jeanblanc (✉) · Y. Le Cam
Département de Mathématiques, Université d'Évry – Val d'Essonne,
boulevard François Mitterrand, 91025 Évry Cedex, France
e-mail: monique.jeanblanc@univ-evry.fr

Y. Le Cam
e-mail: yann_le_cam@yahoo.fr

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_6, © Springer-Verlag Berlin Heidelberg 2009

of the full information $\mathbb{G}$ between a “reference filtration” $\mathbb{F}$ and the credit event (generating a filtration $\mathbb{H}$) so that to benefit the methodology based on the hazard process.[1] A subsection is dedicated to the presentation of each filtration, and another one to the financial interpretation of that splitting.

We model the credit event by a random time belonging to the class of *initial times*.[2] The choice of initial times insures that the semi-martingales in the reference filtration remain semi-martingales in the full filtration $\mathbb{G}$.

- The second one is a mathematical part that aims at proving a representation theorem for martingales of the full filtration, as soon as the martingales of the reference market can be represented on a finite set of martingales, and when the credit event is an initial time.

  This theorem emphasizes the major role played by initial times in credit modelling. A corollary allows to describe the positive $\mathbb{G}$-martingales, hence the set of $\mathscr{G}_T$-equivalent probabilities once an “historical probability” is given.
- The third one is a study of the special—and fundamental—case where the “reference market” is complete and arbitrage free. We describe thanks to the previous martingale representation theorem the set of the probabilities equivalent to the historical one on $\mathscr{G}_T$ under which the reference assets remain martingales. Then it is established under mild assumptions that the full market, where a credit event-sensitive asset is added to the collection of the reference assets, is also complete and without arbitrage opportunities.

  A following section presents the links between the completeness and the immersion property of the filtration enlargement: this property, often referred to as $(\mathscr{H})$-hypothesis, denotes the fact that the $\mathbb{F}$-martingales remain $\mathbb{G}$-martingales. We shall prove that immersion holds under any $\mathbb{G}$-e.m.m., and characterize the change of probability that allows to go from a “reference neutral-risk” probability[3] to a neutral risk probability, under which immersion holds. We derive the important corollary that if the $\mathbb{F}$-martingale part of the survival process defined under a “reference neutral-risk” probability $\mathbb{P}^*$ is not equal to zero, $\mathbb{P}^*$ is not neutral-risk probability of the full market.
- The last part is devoted to the case where the reference market is incomplete and a default sensitive asset is traded. Starting from a reference e.m.m. $\mathbb{P}^*$, we construct a unique e.m.m. in the full filtration that preserves the main properties of the reference market.

  Precisely, we prove that there exists a unique probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}^*$ such that the price process (composed of the reference assets and the default sensitive asset) is a $(\mathbb{G},\mathbb{Q})$-martingale that preserves the “reference pricing”, i.e., such that $\mathbb{E}^{\mathbb{Q}}(X_T)=\mathbb{E}^*(X_T)$, for any $X_T\in L^2(\mathscr{F}_T)$. We establish

---

[1] See Jeanblanc and Le Cam [23] for a survey on reduced form modelling and hazard process.

[2] See the following section, Jiao [21], El Karoui et al. [11] or the paper of Jeanblanc and Le Cam [22] for a study of the properties of initial times and their application to progressive enlargement of filtrations.

[3] A probability under which the reference assets are $\mathbb{F}$-martingales.

that here again, immersion property of the filtration enlargement holds under this probability.

In this paper, all the processes are constructed on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$, where the probability measure $\mathbb{P}$ is referred to as the historical probability.

A financial market is represented in the sequel by a $(n+2)$-dimensional price process $\widetilde{S} = (S_t^0, \dots, S_t^{n+1}; 0 \le t \le T)$, in which S^0 will denote the saving accounts, i.e., the risk-free asset,[4] and its information by $\mathbb{G}$ assumed to be the natural (augmented) filtration generated by $\widetilde{S}$. We do not assume that $\mathscr{G}_T = \mathscr{A}$, and we emphasize that $\mathbb{P}$ is a probability measure defined on $\mathscr{A}$ (even if we shall be interested in the sequel in restrictions of the probabilities on sub-σ-algebras of $\mathscr{A}$).

We denote by $\Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$ the set of $\mathbb{G}$-e.m.ms, i.e., the set of probability measures $\mathbb{Q}$ defined on $\mathscr{A}$, equivalent to $\mathbb{P}$ on $\mathscr{A}$, such that $\widetilde{S}/S^0 \in \mathscr{M}^{loc}(\mathbb{G}, \mathbb{Q})$, i.e., such that the discounted process $(\widetilde{S}_t/S_t^0, t \le T)$ is a $(\mathbb{G}, \mathbb{Q})$-local martingale. In what follows, we assume for the sake of simplicity that $S^0 \equiv 1$.

It is well known that there are strong links between no-arbitrage hypothesis and the existence of an equivalent martingale measure (see Kabanov [25], Delbaen and Schachermayer [10]). In this paper we are interested with the property of $\Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$ being not empty. This condition is equivalent to the No Free Lunch with Vanishing Risk condition (a condition slightly stronger than absence of arbitrage).

Recall that the market where the assets $S^i, i = 0, \dots, n+1$ are traded is complete if any contingent claim is replicable: For any payoff $X_T \in L^2(\mathscr{G}_T)$ there exists a $\mathbb{G}$-predictable self-financed strategy with terminal value X_T. Our aim in this paper is neither to make a fine discussion on the best hypotheses on trading strategies, nor to use the most precise and efficient assumptions in that matter (the interested reader may refer for example to Delbaen and Schachermayer [10], Kabanov [25] or/and Protter [30]).

Given an e.m.m. $\mathbb{Q} \in \Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$, we say that the $\mathbb{Q}$-local martingale $\widetilde{S}$ enjoys the predictable representation property[5] (under $\mathbb{Q}$), if every $(\mathbb{Q}, \mathbb{G})$-local martingale M can be written $M = M_0 + m \star \widetilde{S}$, where M_0 belongs to $\mathscr{G}_0$ and $m \star \widetilde{S}$ is the process $\int_0^t m_s d\widetilde{S}_s$ with m a $\mathbb{G}$-predictable locally bounded process. If $\widetilde{S}$ enjoys the *PRP*, the market is complete. When only considering probabilities equivalent to a given one (in our case the historical probability $\mathbb{P}$), it is straightforward that if $\Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$ restricted to $\mathscr{G}_T$ is a singleton, *PRP* holds under this martingale probability and the market is complete.

2 Credit Modelling Framework

We work in this study within a progressive enlargement of filtration set-up, so that to study the pricing of derivatives written on underlyings sensitive to a credit event τ.

---

[4] S^{n+1} will represent in the sequel an asset on which may be read the occurrence of the credit event, at the opposite of the $n+1$ first ones.

[5] General presentations of *PRP* are available in Revuz and Yor [32], Jacod and Shiryaev [19] or Protter [31].

We refer the reader to Elliott et al. [14] or to Jeanblanc and Rutkowski [24] for a detailed presentation of this approach, and to Jeanblanc and Le Cam [23] for the reasons that lead us to adopt it in this context.

In this framework, we shall split the information induced by the market into two components: a first one generated by what will be called the *default-free* assets, and a second by the knowledge of the occurrence of a credit event (the probability of occurrence of this event will depend on factors adapted to the first filtration). Precisely, two sources of information are introduced, the *reference filtration* and the *default-sensitive asset*.

2.1 The Two Information Flows

The reference filtration. We consider an $(n+1)$-dimensional vector S of assets $S^0, \dots, S^n$ and its natural filtration $\mathbb{F}$, that will be referred to as the reference filtration[6] in the sequel.[7] These assets may be shares, vanilla options, interest rates, change rates, etc.: All listed information that may be used by the market to build its anticipations on the probability of occurrence of the risk—and impact the bid-ask price of instruments written on τ. For example, if τ is the default time of a bond issued by a firm X, it is in general not a stopping time with respect to the filtration generated by the stock of X, or by interest rates (even it is far from being independent of such variables). The information flow $\mathbb{F}$ does not contain the information of the occurrence of the credit event, i.e., τ is not an $\mathbb{F}$-stopping time.

We denote by $\Theta_{\mathbb{P}}^{\mathbb{F}}(S)$ the set of $\mathbb{F}$-e.m.ms.[8] The next hypothesis will be systematically imposed on the model, and implies the absence of arbitrage in the reference market:[9] $\Theta_{\mathbb{P}}^{\mathbb{F}}(S)$ is not empty. This hypothesis will hold until the end of the article.

Default-sensitive asset. We introduce an asset S^{n+1}, that bears direct information on τ, i.e., that satisfies:

$$\mathscr{H}_t \subset \sigma(S_s^{n+1}, 0 \le s \le t) \subset \mathscr{H}_t \vee \mathscr{F}_t \quad \text{for any } t \ge 0, \tag{1}$$

where the notation $\mathbb{H} = (\mathscr{H}_t, t \ge 0)$ stands for the natural augmentation of the filtration generated by the process $H_t = \mathbb{1}_{\tau \le t}$. This filtration models the knowledge of the occurrence of the credit event.

[6] In [3] Bélanger et al. refer to $\mathbb{F}$ as the *non firm specific information*. For us, this information flow must be considered as the "market risk" information, and can bear assets linked to the firm, for example its equity or even its directly its spread risk, see later.

[7] If needed, we set $\mathscr{F}_s = \mathscr{F}_T$ for $s > T$.

[8] The set of probabilities $\mathbb{Q}$ defined on $\mathscr{A}$, equivalent to $\mathbb{P}$ on $\mathscr{A}$, such that $S \in \mathscr{M}^{loc}(\mathbb{F}, \mathbb{Q})$, i.e., such that the process $S = (S_t, t \le T)$ is an $(\mathbb{F}, \mathbb{Q})$-local martingale.

[9] Recall that we have assumed null interest rate to ease the presentation.

This relation explains the fact that the default can be read on the paths of S^{n+1}, and that this asset can be priced in terms of τ and $\mathbb{F}$ (think of a risky bond, a defaultable zero-coupon or a credit default swap).[10]

We denote by $\widetilde{S}$ the vector $(S^0, S^1, \dots, S^{n+1}) = (S, S^{n+1})$, and by $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$, the natural augmentation of the filtration generated by $\widetilde{S}$ (the full information of the market).

2.2 Financial Interpretation of This Decomposition

In terms of risk analysis, S^{n+1} bears in general two types of risks (we consider in this survey only the single default case, hence do not enter in a discussion about the correlation risk):

- a "market risk"—typically a spread risk, i.e., the natural variation of the price of the asset when time goes on without default;
- a jump risk—the specific risk of default, due to the occurrence of τ.

This framework is based on the assumption that *the market risk can be hedged with $\mathbb{F}$-adapted instruments, and that the jump risk relies on $\mathbb{H}$-adapted instruments.* Two points of view can be considered to justify this assumption.

1. The first one—based on economic analysis—lies on the observation that the spread risk is mainly ruled out by the same noise sources as the assets that generate the reference filtration. For example:

- In the context of firm bonds pricing, Bélanger et al. link in [3] the spread risk of the defaultable zero-coupon to the stochastic interest rates. In such a modelling, the credit event is constructed as the hitting time of an independent random barrier by an increasing $\mathbb{F}$-adapted increasing process (the $\mathbb{F}$-intensity), where $\mathbb{F}$ is the filtration bearing the stochastic interest rates movements.

  The parameters of the intensity process may depend on the firm (see also Ehlers and Schönbucher [12], where the authors insist on the rôle of the systemic risk implied by the interest rates on a portfolio of credit risks).
- Moreover in a very close matter, Carr and Wu in [7] or Cremers et al. in [9] show that corporate CDS spreads covary with both the stock option implied volatilities and skewness.

  It confirms that the factors ruling out the movements of the spreads are linked to the variations of the interest rate and of the equity (and its volatility).
- In the context of modelling CDS on debt issued by states (in their example Mexico and Brazil), Carr and Wu study in [8] the correlation between the currency options and the credit spreads. They prove that these quantities are deeply linked and propose a model in which the alea driving the intensity of the default is composed by the sum of a function of the alea of the stochastic volatility of the FX (see Heston [16]), and an independent noise (see also Ehlers and Schönbucher in [13]).

[10] *CDS* will refer to Credit Default Swap in the sequel.

- More generally, this vision is shared by the supporters of structural modelling, in which the default time is triggered by a barrier reached by the equity value (see Merton [29] or Black and Cox [5] for example). In [2], Atlan and Leblanc model the credit time as the reaching time by the Equity of the firm of value zero, the stock following a CEV (see also Albanese and Chen [1] or Linetsky [28]).

2. The second one is based on the introduction of a new noise source, this alea driving the spread risk, considered as having its own evolution (both approaches can be combined). In this construction as well, the "market risk of the defaultable security" does not contain the default occurrence knowledge, and can be sorted in the $\mathbb{F}$-information with the other "market risks sensible" assets.

In practice, it is easy to synthesize an asset that is sensible to this spread risk and not to the jump risk. Take two instruments as S^{n+1} of different maturity for example, denoted by X^1 and X^2, and assume the market risk is modelled by a (risk-neutral) Brownian motion W. Assume that M is the compensated martingale associated to H, and that $dX_t^i = \beta_t^i dM_t + \delta_t^i dW_t$ under the e.m.m. Set up the self-financed portfolio Π that is long at any time of β_t^2 of the asset X^1 and short of β_t^1 of X^2 (and has a position in the savings account to stay self-financed). This portfolio has only sensitivity against the spread risk, and does not jump with τ. Indeed,

$$d\Pi_t = r\Pi_t dt + \beta_t^2 dX_t^1 - \beta_t^1 dX_t^2 = r\Pi_t dt + (\beta_t^2\delta_t^1 - \beta_t^1\delta_t^2) dW_t.$$

Remark that with a δ-combination (instead of the β-combination), we can set up a portfolio only sensible of the jump risk (and that has no spread risk):

$$d\Pi_t' = r\Pi_t' dt + \delta_t^2 dX_t^1 - \delta_t^1 dX_t^2 = r\Pi_t' dt + (\beta_t^1\delta_t^2 - \beta_t^2\delta_t^1) dM_t.$$

The two points of view (that need to be combined to achieve a maximum of precision in calibration procedures) converge on the idea that *splitting the information of the market in two sub-filtrations is finally quite natural.* Another nomenclature may consist in "market risk filtration" for $\mathbb{F}$, and "default risk filtration" for $\mathbb{H}$.

2.3 Absence of Arbitrage

Starting from a reference market with no arbitrage, the absence of arbitrage of the full market is not automatic and deeply depends on the nature of τ. For it to hold, it is necessary to work in a mathematical set up where $\mathbb{F}$-semi-martingales remain $\mathbb{G}$-semi-martingales.[11]

As developed in Jeanblanc and Le Cam [23], this property does not hold for any random time τ, and we choose to work in all the paper, under the following condition on τ.

[11] So that $\Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$ be not empty.

Hypothesis $\mathbf{H}_1$ The credit event is an initial time, that is, there exists a family of processes $(\alpha^u, u \in \mathbb{R}^+)$ such that for any $u \geq 0$, the process $(\alpha^u_t, 0 \leq t \leq T)$ is an $\mathbb{F}$-martingale and that satisfies

$$\mathbb{P}(\tau > \theta | \mathscr{F}_t) = \int_\theta^\infty \alpha^u_t du$$

for any $\theta \geq 0$, and any $t \leq T$.

There exists a $\mathscr{O}(\mathbb{F} \otimes \mathscr{B})$-measurable version[12] of the mapping $(\omega, u, t) \to \alpha^u_t(\omega)$, right-continuous with left limits.[13] We shall consider this version in the sequel. In our setting, the law of τ admits a density w.r.t. Lebesgue measure, equal to α^u_0. In the general definition of initial times, $\alpha_0(u)du$ may be replaced by any probability measure on $\mathbb{R}^+$ $\nu(du)$.

We denote by G the Azéma super-martingale

$$G_t := \mathbb{P}(\tau > t | \mathscr{F}_t).$$

We write $G = Z - A$ the $\mathbb{F}$-Doob-Meyer decomposition of this super-martingale (of class (D)). From hypothesis $\mathbf{H}_1$, every $(\mathbb{F}, \mathbb{P})$-martingale X is a $(\mathbb{G}, \mathbb{P})$-semi-martingale, and if the $\mathbb{F}$-martingales are continuous:

$$X_t - \int_0^{t \wedge \tau} \frac{d\langle X, Z\rangle_u}{G_{u-}} - \int_{t\wedge\tau}^t \frac{d\langle X, \alpha^\theta\rangle_u}{\alpha^\theta_{u-}}\Bigg|_{\theta=\tau} \in \mathscr{M}(\mathbb{G}, \mathbb{P}) \tag{2}$$

(see Jeanblanc and Le Cam [22]).

Proposition 1 *When the law ν of the initial time τ has no atoms, for example under the hypothesis* **H**, *it avoids the* $\mathbb{F}$*-stopping times, i.e.,*

$$\mathbb{P}(\tau = T) = 0, \quad \forall T \textit{ finite } \mathbb{F}\textit{-stopping time}.$$

Proof This result is a consequence of Lemma 2 of [22], that states that if τ is an initial time and if T is a finite $\mathbb{F}$-stopping time,

$$\mathbb{E}(1_{\{\tau=T\}} | \mathscr{F}_T) = \alpha^T_T \nu(\{T\}) \quad \text{a.s.}$$

It follows that if ν has no atoms,

$$\mathbb{P}(\tau = T) = \mathbb{E}(1_{\{\tau=T\}}) = 0,$$

hence τ avoids the $\mathbb{F}$-stopping times. □

As we shall focus on in this paper in change of probabilities, it is necessary to ensure that the initial property does not depend on the historical probability. It follows from:

[12] The σ-field $\mathscr{O}(\mathbb{F} \otimes \mathscr{B})$ is the optional σ-field on $(\Omega \times \mathbb{R}^+) \times [0, T]$.

[13] See Jacod [18] for a presentation of the paths regularity of the martingale density family.

Proposition 2 *If τ is an initial time under $\mathbb{P}$, and if $\mathbb{Q}$ is a probability measure equivalent to $\mathbb{P}$, then τ is a $\mathbb{Q}$-initial time.*

Proof The definition of τ being an initial time can be formulated in the following way: if $Q_t^{\mathbb{P}}$ denotes a regular version of the conditional law of τ: $Q_t(\omega,]\theta, \infty[) := \mathbb{P}(\tau > \theta|\mathscr{F}_t)(\omega)$, there exists a deterministic measure ν such that $Q_t^{\mathbb{P}}(\omega, d\theta) \ll \nu(d\theta)$ $\mathbb{P}$-*a.s.* As two equivalent probabilities have the same null sets, if $\mathbb{Q} \sim \mathbb{P}$, $Q_t^{\mathbb{Q}}(\omega, d\theta) \ll \nu(d\theta)$ $\mathbb{Q}$-*a.s.*, and the proposition is proved. The following alternative proof allows to derive the new martingale density.

Let $(\eta_t, t \geq 0)$ be the $\mathbb{G}$-martingale which is the Radon-Nikodym density of $\mathbb{Q}$ w.r.t. $\mathbb{P}$

$$d\mathbb{Q}|_{\mathscr{G}_t} = \eta_t d\mathbb{P}|_{\mathscr{G}_t}.$$

Let T be fixed and $\theta, t < T$. Then, the Bayes rule implies:

$$\mathbb{Q}(\tau > \theta|\mathscr{F}_t) = \mathbb{E}^{\mathbb{Q}}((1 - H_\theta)|\mathscr{F}_t) = \frac{\mathbb{E}^{\mathbb{P}}((1 - H_\theta)\eta_T|\mathscr{F}_t)}{\mathbb{E}^{\mathbb{P}}(\eta_T|\mathscr{F}_t)}$$

since $(1 - H_\theta)$ is $\mathscr{G}_\theta$ hence $\mathscr{G}_T$-measurable. Assume in a first step that $\eta_T = \widetilde{\eta}_T h(\tau \wedge T)$ where $\widetilde{\eta}_T$ is a (bounded) $\mathscr{F}_T$-measurable random variable and h is a (bounded) deterministic function. We have:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}((1 - H_\theta)\eta_T|\mathscr{F}_t) &= \mathbb{E}^{\mathbb{P}}((1 - H_\theta)\widetilde{\eta}_T h(\tau \wedge T)|\mathscr{F}_t) \\
&= \mathbb{E}^{\mathbb{P}}(\widetilde{\eta}_T \mathbb{E}^{\mathbb{P}}((1 - H_\theta)h(\tau \wedge T)|\mathscr{F}_T)|\mathscr{F}_t) \\
&= \mathbb{E}^{\mathbb{P}}\left(\widetilde{\eta}_T \int_\theta^\infty h(u \wedge T)\alpha_T^u du \,\Big|\, \mathscr{F}_t\right) \\
&= \int_\theta^\infty \mathbb{E}^{\mathbb{P}}(\widetilde{\eta}_T \alpha_T^u|\mathscr{F}_t)h(u \wedge T)du
\end{aligned}$$

It follows that

$$\mathbb{Q}(\tau > \theta|\mathscr{F}_t) = \int_\theta^\infty \frac{\mathbb{E}^{\mathbb{P}}(\widetilde{\eta}_T \alpha_T^u|\mathscr{F}_t)}{\mathbb{E}^{\mathbb{P}}(\eta_T|\mathscr{F}_t)} h(u \wedge T)du.$$

Moreover, if μ_T denotes the $\mathscr{F}_T$-density of $\mathbb{Q}$ w.r.t. $\mathbb{P}$, i.e., $d\mathbb{Q}|_{\mathscr{F}_T} = \mu_T d\mathbb{P}|_{\mathscr{F}_T}$, μ_T writes

$$\mu_T = \mathbb{E}^{\mathbb{P}}(\eta_T|\mathscr{F}_T) = \widetilde{\eta}_T \mathbb{E}^{\mathbb{P}}(h(\tau \wedge T)|\mathscr{F}_T) = \widetilde{\eta}_T \int_0^\infty h(u \wedge T)\alpha_T^u du := \widetilde{\eta}_T \phi_T.$$

We now introduce the family of $(\mathbb{F}, \mathbb{Q})$-martingales $\widehat{\alpha}_{\cdot}^u$, defined for any $u \geq 0$ by

$$\widehat{\alpha}_t^u := h(u \wedge T)\mathbb{E}^{\mathbb{Q}}(\alpha_T^u/\phi_T|\mathscr{F}_t).$$

Bayes rule implies:

$$\mathbb{E}^{\mathbb{Q}}(\alpha_T^u/\phi_T|\mathscr{F}_t) = \frac{\mathbb{E}^{\mathbb{P}}(\alpha_T^u \mu_T/\phi_T|\mathscr{F}_t)}{\mathbb{E}^{\mathbb{P}}(\mu_T|\mathscr{F}_t)} = \frac{\mathbb{E}^{\mathbb{P}}(\alpha_T^u \widetilde{\eta}_T|\mathscr{F}_t)}{\mathbb{E}^{\mathbb{P}}(\eta_T|\mathscr{F}_T|\mathscr{F}_t)} = \frac{\mathbb{E}^{\mathbb{P}}(\widetilde{\eta}_T \alpha_T^u|\mathscr{F}_t)}{\mathbb{E}^{\mathbb{P}}(\eta_T|\mathscr{F}_t)},$$

and that, for any T, for any $t, \theta < T$,

$$\mathbb{Q}(\tau > \theta|\mathscr{F}_t) = \int_\theta^\infty \widehat{\alpha}_t^u du,$$

which means that τ is an $\mathbb{F}$-initial time under $\mathbb{Q}$. The general case follows by application of the monotone class theorem. □

Remark 1 When the credit event avoids the $\mathbb{F}$-stopping times, immersion property under the probability measure $\mathbb{P}$—$\mathscr{M}(\mathbb{F}, \mathbb{P}) \subset \mathscr{M}(\mathbb{G}, \mathbb{P})$—is equivalent to the property that for any $u \geq 0$, the martingale α^u is constant after u (see Jeanblanc and Le Cam [22]), i.e.,

$$\alpha_t^u = \alpha_{t\wedge u}^u, \quad \text{for any } (u, t) \geq 0 \tag{3}$$

That is the case under hypothesis **H**.

We recall that immersion property is not preserved by a change of probability (see for example Kusuoka [27]).

3 Representation Theorem in the Enlarged Filtration

Under hypothesis $\mathbf{H}_1$ the $\mathbb{F}$-martingales are $\mathbb{G}$-semi-martingales and the initial time property is stable when changing the probability. We also assume, to ease the proofs the following condition:

Hypothesis H₂ The process S is continuous.

This hypothesis will hold until the end of the paper. The aim of the following first subsection is to prove that under a progressive enlargement of a filtration by an initial time, if the reference filtration $\mathbb{F}$ enjoys a predictable representation property, the enlarged filtration $\mathbb{G}$ enjoys the same property. The goal of the second one is to apply this result to the description of all the $\mathbb{G}$-martingales and to parameterize the change of equivalent probabilities.

3.1 Representation of the $\mathbb{G}$-Martingales

We starting this section by defining the:

Hypothesis $\mathbf{H}_3$ We assume that the $\mathbb{F}$-market is complete and arbitrage free, i.e., that the $(\mathbb{F}, \mathbb{P}^*)$-local martingale S enjoys the *PRP* (with $\mathbb{P}^* \in \Theta_{\mathbb{P}}^{\mathbb{F}}(S)$).

As a consequence of Hypotheses $\mathbf{H}_2$ and $\mathbf{H}_3$, $\mathbb{F}$-martingales are continuous (in particular the densities). We start with a brief presentation of the two fundamental (local) martingales on which we shall prove that any $\mathbb{G}$-martingale can be represented.

The first one, M, is the martingale part of the Doob-Meyer decomposition of the increasing $\mathbb{G}$-adapted process H. It is well known (see for example, Bielecki and Rutkowski [4]) that with no particular condition on τ, the $\mathbb{G}$-compensator of H writes: $(1-H)dA/G_-$, where A is the $\mathbb{F}$-compensator of the $\mathbb{F}$-super-martingale G, hence

$$M_t = H_t - \int_0^{t\wedge\tau} \frac{dA_u}{G_{u-}}.$$

From hypothesis $\mathbf{H}_1$, the conditional survival probability writes

$$G_t^\theta := \mathbb{P}^*(\tau > \theta|\mathscr{F}_t) = \int_\theta^\infty \alpha_t^u du \tag{4}$$

that allows to explicitly compute A, from

$$G_t = G_t^t = \int_0^\infty \alpha_{t\wedge u}^u du - \int_0^t \alpha_u^u du \equiv Z_t - A_t, \tag{5}$$

and to conclude (using that G is continuous):

$$M_t = H_t - \int_0^t \frac{1-H_u}{G_u}\alpha_u^u du. \tag{6}$$

The second one is the local martingale part of the decomposition of the special $\mathbb{G}$-semi-martingale S ($\mathbf{H}_1$ implies that the $\mathbb{F}$-martingales remain $\mathbb{G}$-semi-martingales from (2), and we will see that the form of τ makes them special). From the *PRP* of $\mathbb{F}$, there exist:

- An n-dimensional $\mathbb{F}$-predictable process $z = (z^1, \ldots, z^n)$ such that Z (defined by (5)) writes
$$Z = Z_0 + z \star S,$$
where $z \star S$ stands for the process $t \longmapsto \sum_{i\le n} \int_0^t z_s^i dS_s^i$,
- A family of n-dimensional $\mathbb{F}$-predictable processes a, such that for any $u \ge 0$ the $(\mathbb{F}, \mathbb{P}^*)$-martingale α^u writes
$$\alpha^u = \alpha_0^u + a^u \star S,$$
where $a^u \star S$ stands for the process $t \longmapsto \sum_{i\le n} \int_0^t a_s^{u,i} dS_s^i$.

The quadratic covariations $\langle S, Z\rangle$ and $\langle S, \alpha^\theta\rangle$ are well defined, and from (2), the process

$$\widehat{S}_t := S_t - \int_0^t \frac{(1-H_s)}{G_S} d\langle S, Z\rangle_S + \frac{H_S}{\alpha_S^\theta} d\langle S, \alpha^\theta\rangle_s|_{\theta=\tau} \tag{7}$$

is a $(\mathbb{G}, \mathbb{P}^*)$-local martingale. It follows that for any $i \le n$

$$\widehat{S}^i_t = S^i_t - \int_0^t \left(\frac{1 - H_S}{G_S} z_S + \frac{H_S}{\alpha^\tau_s} a^\tau_s \right) \cdot d\langle S^i, S\rangle_s := S^i_t - C_t,$$

where $z \cdot d\langle S^i, S\rangle$ (resp. $a^\tau \cdot d\langle S^i, S\rangle$) stands for $d\langle S^i, z \star S\rangle$ (resp. $d\langle S^i, a^\tau \star S\rangle$).

Predictable representation. The next theorem establishes a predictable representation property for $\mathbb{G}$-local martingales under $\mathbb{P}^*$, as soon as the $\mathbb{F}$-market enjoys this predictable representation property. Indeed, any η which belongs to $\mathcal{M}^{loc}(\mathbb{G}, \mathbb{P}^*)$ will write as the sum of an integral[14] with respect to the $(\mathbb{G}, \mathbb{P}^*)$-martingale M and an integral with respect to the $(\mathbb{G}, \mathbb{P}^*)$-local martingale $\widehat{S}$. This result extends the representation theorem by Kusuoka [27] to any complete reference market and to the case where immersion does not hold.

Theorem 1 *Assume that* $\mathbf{H}_1$ *to* $\mathbf{H}_3$ *hold. Denote by* $\widehat{S}$ *the local martingale part of the decomposition of* S *as* $\mathbb{G}$*-semi-martingale, given by* (7)*. For every* $\eta \in \mathcal{M}^{loc}(\mathbb{G}, \mathbb{P}^*)$*, there exist* $n + 1$ $\mathbb{G}$*-predictable processes* β *and* γ *such that*

$$\eta_t = \eta_0 + (\beta \star M)_t + (\gamma \star \widehat{S})_t.$$

Proof Without loss of generality we prove the theorem for $n = 1$ (considering only one component of the vector S), to ease the notations. The vectorial version of the proof is a straightforward generalization. By localization, we only consider martingales. Let $\eta \in \mathcal{M}(\mathbb{G}, \mathbb{P}^*)$, and as we are only interested in finite time horizon, we write $\eta_t = \mathbb{E}^*(\eta_T|\mathcal{G}_t)$ for $t < T$. By a monotone class argument, we reduce ourself to the case where η_T writes $F_T h(\tau \wedge T)$, with $F_T \in \mathcal{F}_T$, assumed to be bounded and h is a (bounded) deterministic function. We split the problem in three parts:

$$\begin{aligned}
\eta_t &= \mathbb{E}^*(F_T h(T) 1_{\tau > T}|\mathcal{G}_t) + \mathbb{E}^*(F_T h(\tau) 1_{\tau \le T}|\mathcal{G}_t) \\
&= a_t + \mathbb{E}^*(F_T h(\tau) 1_{\tau \le T}|\mathcal{G}_t) \\
&= \underbrace{L_t h(T)\mathbb{E}^*(F_T G_T|\mathcal{F}_t)}_{a_t} + \underbrace{L_t \mathbb{E}^*(F_T h(\tau) 1_{t < \tau \le T}|\mathcal{F}_t)}_{b_t} \\
&\quad + \underbrace{H_t \mathbb{E}^*(F_T h(\tau) 1_{\tau \le t}|\mathcal{F}_t \vee \sigma(\tau))}_{c_t}
\end{aligned}$$

with $L_t = (1 - H_t)/G_t = D_t(1 - H_t) \in \mathcal{M}(\mathbb{G}, \mathbb{P}^*)$, where $D_t = G_t^{-1}$. From the decomposition (5), we have $dG_t = -\alpha^t_t dt + z_t dS_t$, and from Itô's formula, $dD_t = D_t^2(\alpha^t_t dt + z_t^2 D_t d\langle S\rangle_t) - D_t^2 z_t dS_t$.

Let us start by developing a: We first remark that the process a defined by $a_t := L_t h(T)\mathbb{E}^*(F_T G_T|\mathcal{F}_t)$ is a $\mathbb{G}$-martingale, so one knows in advance that this particular semi-martingale has a null predictable bounded variation part; nevertheless, we

[14]Recall for any $X \in \mathcal{H}_1$, one has $\mathbb{E}([X, M]_\infty) = \mathbb{E}X_\tau \le \|X\|_{\mathcal{H}_1}$ hence M is a *BMO* (the dual of $\mathcal{H}_1$) martingale.

keep all these terms in our computation. The process N where $N_t := \mathbb{E}^*(F_T G_T|\mathscr{F}_t)$ is an $\mathbb{F}$-martingale, and writes by representation theorem $n + \int_0^t n_s dS_s$, with $(n_s, s \geq 0)$ predictable. Since S, D and N are continuous, $[S, H] = 0$, $[D, N] = \langle D, N\rangle$, and, from $(h(T))^{-1} a_t = (1 - H_t) D_t N_t$, one gets

$$
\begin{aligned}
(h(T))^{-1} da_t &= -D_t N_t dH_t + (1 - H_t) D_t dN_t + (1 - H_t) N_t dD_t \\
&\quad + (1 - H_t) d\langle D, N\rangle_t \\
&= -D_t N_t dH_t + (1 - H_t) D_t n_t dS_t + (1 - H_t) N_t D_t^2 \alpha_t^t dt \\
&\quad + (1 - H_t) N_t z_t^2 D_t^3 d\langle S\rangle_t - (1 - H_t) N_t D_t^2 z_t dS_t \\
&\quad - (1 - H_t) D_t^2 n_t z_t d\langle S\rangle_t \\
&= -D_t N_t dM_t - (1 - H_t) D_t^2 N_t \alpha_t^t dt + (1 - H_t)(D_t n_t - N_t D_t^2 z_t) dS_t \\
&\quad + (1 - H_t) N_t D_t^2 \alpha_t^t dt + (1 - H_t)(N_t z_t D_t - n_t) D_t^2 z_t d\langle S\rangle_t.
\end{aligned}
$$

In the third equality, we have used that the $\mathbb{G}$-Doob-Meyer decomposition of the increasing process H writes $dH_t = dM_t + (1 - H_t) D_t \alpha_t^t dt$ (from $dA_t = \alpha_t^t dt$), with $M \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*)$ (see (6)). Moreover $S_t = \widehat{S}_t + C_t$ with $\widehat{S} \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*)$, and from (7) $(1 - H_t) dC_t = (1 - H_t) z_t D_t d\langle S\rangle_t$. It follows

$$
\begin{aligned}
(h(T))^{-1} da_t &= -D_t N_t dM_t + (1 - H_t)(D_t n_t - N_t D_t^2 z_t) d\widehat{S}_t \\
&\quad + (1 - H_t)((D_t n_t - N_t D_t^2 z_t) z_t D_t + N_t z_t^2 D_t^3 - n_t D_t^2 z_t) d\langle S\rangle_t \\
&= -D_t N_t dM_t + (1 - H_t)(D_t n_t - N_t D_t^2 z_t) d\widehat{S}_t.
\end{aligned}
$$

To explicit the decomposition of the special $\mathbb{G}$*-semi-martingale* b, where $b_t = L_t \mathbb{E}^*(F_T h(\tau) 1_{t<\tau\leq T}|\mathscr{F}_t)$, we introduce for any $u \geq 0$ the $\mathbb{F}$-martingale $N_t^u = \mathbb{E}^*(F_T \alpha_T^u|\mathscr{F}_t)$ and its decomposition on $S : N_t^u = y^u + \int_0^t y_s^u dS_s$ provided by the martingale representation theorem on $\mathbb{F}$. By definition of initial times, it follows:

$$
b_t = L_t \mathbb{E}^*\left(F_T \int_t^T h(u) \alpha_T^u du \,\middle|\, \mathscr{F}_t\right) = L_t \int_t^T h(u) N_t^u du
$$

hence, one can differentiate using Itô Wentcell formula:

$$
\begin{aligned}
db_t = -D_t\left(\int_t^T h(u) N_t^u du\right) dH_t &+ (1 - H_t)\left(\int_t^T h(u) N_t^u du\right) dD_t \\
&- (1 - H_t) D_t h(t) N_t^t dt + (1 - H_t) D_t\left(\int_t^T h(u) y_t^u du\right) dS_t \\
&- (1 - H_t) D_t^2 z_t\left(\int_t^T h(u) y_t^u du\right) d\langle S\rangle_t
\end{aligned}
$$

and, introducing the $\mathbb{G}$-decomposition of the semi-martingale S and the compensator of H, we obtain finally:

$$\begin{aligned}
db_t &= -D_t\left(\int_t^T h(u)N_t^u du\right)dH_t \\
&\quad + (1-H_t)\left(\alpha_t^t D_t^2 \int_t^T h(u)N_t^u du - D_t h(t)N_t^t\right)dt \\
&\quad + (1-H_t)\left(D_t \int_t^T h(u)(y_t^u - D_t N_t^u z_t)du\right)dS_t \\
&\quad - (1-H_t)\left(D_t^2 z_t \int_t^T h(u)(y_t^u - D_t N_t^u z_t)du\right)d\langle S\rangle_t \\
&= -D_t\left(\int_t^T h(u)N_t^u du\right)dM_t \\
&\quad + (1-H_t)\left(D_t \int_t^T h(u)(y_t^u - D_t N_t^u z_t)du\right)d\widehat{S}_t - (1-H_t)D_t h(t)N_t^t dt.
\end{aligned}$$

Decomposition of c. We can write $c_t = H_t\mathbb{E}^*(F_T h(\tau)1_{\tau\le t}|\mathscr{F}_t \vee \sigma(\tau)) = H_t F(t,\tau)$, where for each $u \ge 0$ the random variable $F(t,u)$ is $\mathscr{F}_t$-measurable and for any $t \ge 0$, $u \longmapsto F(t,u)$ is a Borel function. Using the properties of initial times, we compute from last expression $F(t,u) = h(u)N_t^u/\alpha_t^u$. For any $u \ge 0$, the dynamics write (using $dN_t^u = y_t^u dS_t$ and $d\alpha_t^u = a_t^u dS_t$):

$$d_t F(t,u) = h(u)\left(\frac{y_t^u}{\alpha_t^u} - \frac{N_t^u a_t^u}{(\alpha_t^u)^2}\right)dS_t + h(u)\left(N_t^u \frac{(a_t^u)^2}{(\alpha_t^u)^3} - \frac{a_t^u y_t^u}{(\alpha_t^u)^2}\right)d\langle S\rangle_t.$$

It follows that, since

$$\int_0^t F(s,\tau)dH_s = F(\tau,\tau)\mathbb{1}_{\tau\le t} = \int_0^t F(s,s)dH_s,$$

we can write the decomposition of c:

$$\begin{aligned}
dc_t &= F(t,\tau)dH_t + H_t h(\tau)\left(\frac{y_t^\tau}{\alpha_t^\tau} - \frac{N_t^\tau a_t^\tau}{(\alpha_t^\tau)^2}\right)dS_t \\
&\quad + H_t h(\tau)\left(\frac{N_t^\tau (a_t^\tau)^2}{(\alpha_t^\tau)^3} - \frac{a_t^\tau y_t^\tau}{(\alpha_t^\tau)^2}\right)d\langle S\rangle_t \\
&= F(t,t)dM_t + H_t h(\tau)\left(\frac{y_t^\tau}{\alpha_t^\tau} - \frac{N_t^\tau a_t^\tau}{(\alpha_t^\tau)^2}\right)d\widehat{S}_t + (1-H_t)D_t F(t,t)\alpha_t^t dt \\
&\quad + H_t h(\tau)\left(\frac{y_t^\tau}{\alpha_t^\tau} - \frac{N_t^\tau a_t^\tau}{(\alpha_t^\tau)^2}\right)dC_t + H_t h(\tau)\left(\frac{N_t^\tau (a_t^\tau)^2}{(\alpha_t^\tau)^3} - \frac{a_t^\tau y_t^\tau}{(\alpha_t^\tau)^2}\right)d\langle S\rangle_t
\end{aligned}$$

$$= F(t,t)dM_t + H_t h(\tau)\left(\frac{y_t^\tau}{\alpha_t^\tau} - \frac{N_t^\tau a_t^\tau}{(\alpha_t^\tau)^2}\right)d\widehat{S}_t + (1-H_t)D_t F(t,t)\alpha_t^t dt$$

where the last equality comes from the expression (7) of dC on $\{\tau \le t\}$.

Conclusion. Adding the three parts a, b, and c, we conclude, since $F(t,t)\alpha_t^t = h(t)N_t^t$, that the $\mathbb{G}$-martingale can be decomposed on the two martingales $(M, \widehat{S})$ and writes:

$$d\eta_t = \gamma_t d\widehat{S}_t + \beta_t dM_t$$

where

$$\gamma_t = (1-H_t)D_t\left((n_t - N_t D_t z_t)h(T) + \int_t^T h(u)(y_t^u - N_t^u D_t z_t)du\right)$$
$$+ H_t\left(\frac{y_t^\tau}{\alpha_t^\tau} - \frac{N_t^\tau a_t^\tau}{(\alpha_t^\tau)^2}\right)$$
$$\beta_t = F(t,t) - D_t\left(N_t h(T) + \int_t^T h(u)N_t^u du\right)$$

which concludes the proof. □

3.2 Change of Probability

Once a probability $\mathbb{P}$ is given on $\mathscr{A}$, each probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ on $\mathscr{G}_T$ is fully described by its $\mathbb{G}$-martingale density w.r.t. $\mathbb{P}$. The representation theorem established in the last section allows to describe all the $(\mathbb{P}, \mathbb{G})$-martingales, hence to describe all the probabilities equivalent to $\mathbb{P}$ on $\mathscr{G}_T$.

Applying Theorem 1 to the particular case of a strictly positive martingale—in particular the density of a change of probability—we derive the

Proposition 3 *If the filtration $\mathbb{F}$ is generated by the $(\mathbb{F}, \mathbb{P})$-martingale S and enjoys the PRP, and if η is a strictly positive $(\mathbb{G}, \mathbb{P})$-martingale, then, there exists a pair of predictable processes γ, β such that*

$$\frac{d\eta_t}{\eta_{t-}} = \gamma_t d\widehat{S}_t + \beta_t dM_t,$$

with $\beta > -1$, i.e., $\eta = \mathscr{E}(\gamma \star \widehat{S})\mathscr{E}(\beta \star M)$, with

$$\begin{cases} \mathscr{E}(\gamma \star \widehat{S})_t = \exp(\int_0^t \gamma_u d\widehat{S}_u - \frac{1}{2}\int_0^t \gamma_u^2 d\langle \widehat{S}\rangle_u) \\ \mathscr{E}(\beta \star M)_t = \exp(\int_0^t \ln(1+\beta_s)dH_s - \int_0^t \beta_s \frac{1-H_s}{G_s}\alpha_s^s ds). \end{cases}$$

4 Complete Reference Market

In this section, we still make the assumption that the reference market is arbitrage free and complete: For any $X_T \in L^2(\mathscr{F}_T)$, there exist a constant x and n $\mathbb{F}$-predictable processes φ^i such that $X_T = x + \int_0^T \sum_{1\le i\le n} \varphi^i_u dS^i_u$ (we recall that we assume null interest rate). As recalled, this property is equivalent to the fact that the restriction of $\Theta^{\mathbb{F}}_{\mathbb{P}}(S)$ on $\mathscr{F}_T$ is a singleton (Jacod and Yor theorem [20]).

This assertion does not imply that there exists a unique probability measure $\mathbb{Q}$ on $\mathscr{A}$ such that S is an $(\mathbb{F},\mathbb{Q})$-martingale, but that if two probabilities $\mathbb{P}^*$ and $\mathbb{Q}^*$ belong to $\Theta^{\mathbb{F}}_{\mathbb{P}}(S)$, then, their restriction to $\mathscr{F}_T$ are equal: $\mathbb{P}^*|_{\mathscr{F}_T} = \mathbb{Q}^*|_{\mathscr{F}_T}$.

We shall prove in this section that the full market described by $\widetilde{S} = (S^0, \dots, S^{n+1})$ is also complete, and study the links of this property with immersion. Precisely,

- The first subsection describes the set $\Theta^{\mathbb{G}}_{\mathbb{P}}(S)$, i.e., the set of probabilities equivalent to $\mathbb{P}$ on $\mathscr{G}_T$ such that the $(n+1)$-dimensional process S remains a local-martingale[15]);
- The second subsection describes the unique martingale measure on the full market, i.e., the unique element of $\Theta^{\mathbb{G}}_{\mathbb{P}}(\widetilde{S})$;
- The third subsection presents the links of this construction with the immersion property.

4.1 Description of the $\mathbb{G}$-Martingale Probabilities

In this first section, we study the behaviour of the reference assets in the full market, i.e., the properties of the $(\mathbb{F},\mathbb{P})$-martingale S viewed as a $\mathbb{G}$-adapted process. The goal of this part is to describe the set $\Theta^{\mathbb{G}}_{\mathbb{P}}(S)$ of probabilities under which this $\mathbb{G}$-semi-martingale[16] is a martingale

Before stating the proposition, we start by a technical remark. The following proof is based on the martingale property of the process $\eta = \mathscr{E}(\ell) := \mathscr{E}(-\vartheta \star \widehat{S})$ where

$$\vartheta_t = (1 - H_t)\frac{z_t}{G_t} + H_t \frac{a^\tau_t}{\alpha^\tau_t}. \tag{8}$$

As $\widehat{S}$ is a $(\mathbb{G},\mathbb{P}^*)$-martingale, η is a local-martingale, and extra conditions have to be assumed so that it be a true martingale and might be used for changing the probability. However, the conditions on the process $-\vartheta \star \widehat{S}$ may be brought to $\mathbb{F}$-adapted processes, in the following way.

---

[15] We shall prove that the restrictions of $\Theta^{\mathbb{F}}_{\mathbb{P}}(S)$ and $\Theta^{\mathbb{G}}_{\mathbb{P}}(S)$ on $\mathscr{F}_T$ are the same.

[16] Recall that the initial property of τ ($\mathbf{H}_2$ holds) ensures that S remains a $\mathbb{G}$-semi-martingale.

Define by R the process $\int_0^t dZ_u/G_u$, and $\Phi_t = \int_0^t (1-H_u)\frac{z_u}{G_u} d\widehat{S}_u$. If $R \in \mathscr{M}(\mathbb{F}, \mathbb{P})$,

$$\widehat{R}_t = R_t - \int_0^{t\wedge\tau} \frac{d\langle R, Z\rangle_u}{G_u} - \int_{t\wedge\tau}^t \frac{d\langle R, \alpha^\theta\rangle_u}{\alpha_u^\theta}\bigg|_{\theta=\tau}$$

$$= R_t - \int_0^t \vartheta_u \frac{z_u}{G_u} d\langle S\rangle_u \in \mathscr{M}(\mathbb{G}, \mathbb{P})$$

hence[17] $\Phi_t = \int_0^t (1-H_u)d\widehat{R}_u \in \mathscr{M}(\mathbb{G}, \mathbb{P})$. For the same reasons, if we define the family of processes $R_t^x = \int_0^t d\alpha_u^x/\alpha_u^x$, and $\Phi_t^x = \int_0^t H_u a_u^x/\alpha_u^x d\widehat{S}_u$ for any $x \geq 0$, the condition $R^x \in \mathscr{M}(\mathbb{F}, \mathbb{P})$ implies $\Phi^x \in \mathscr{M}(\mathbb{G}, \mathbb{P})$. Indeed,

$$\widehat{R_t^x} = R_t^x - \int_0^{t\wedge\tau} \frac{d\langle R^x, Z\rangle_u}{G_u} - \int_{t\wedge\tau}^t \frac{d\langle R^x, \alpha^\theta\rangle_u}{\alpha_u^\theta}\bigg|_{\theta=\tau}$$

$$= R_t^x - \int_0^t \vartheta_u \frac{a_u^x}{\alpha_u^x} d\langle S\rangle_u \in \mathscr{M}(\mathbb{G}, \mathbb{P})$$

and[18] $\Phi_t^x = \int_0^t H_u d\widehat{R_u^x} \in \mathscr{M}(\mathbb{G}, \mathbb{P})$.

It follows that if R and R^x (whose definitions only depends on the construction of the random time) are $\mathbb{F}$-martingales, ℓ is an $\mathbb{G}$-martingale. In the same way, if these processes satisfy a Novikov-type condition, ℓ also satisfy one and the density η is a true martingale. For example, if $\mathbb{E}(\exp 2\langle R\rangle_\infty) < \infty$ and $\mathbb{E}(\exp 2\langle R^x\rangle_\infty) < K$ for any $x \geq 0$, then $\ell = \Phi + \Phi^\tau$ satisfies Novikov criterion:

$$\mathbb{E}\exp\left(\frac{1}{2}\langle\ell\rangle_\infty\right) \leq \mathbb{E}\exp\langle\Phi\rangle_\infty \exp\langle\Phi^\tau\rangle_\infty$$

$$\leq (\mathbb{E}\exp 2\langle\Phi\rangle_\infty)^{1/2}(\mathbb{E}\exp 2\langle\Phi^\tau\rangle_\infty)^{1/2}$$

and the result follows from

$$\langle\Phi\rangle_\infty = \int_0^t (1-H_u)d\langle\widehat{R}\rangle_u = \int_0^t (1-H_u)d\langle R\rangle_u \leq \langle R\rangle_\infty$$

$$\langle\Phi^\tau\rangle_\infty = \int_0^t H_u d\langle\widehat{R^\tau}\rangle_u = \int_0^t H_u d\langle R^\tau\rangle_u \leq \langle R^\tau\rangle_\infty.$$

Hypothesis H$_4$ We assume that $\mathscr{E}(-\vartheta \star \widehat{S})$ is a true martingale.

As noticed in the last lines, this hypothesis holds for example if R and R^x satisfy a Novikov-type condition. Under this condition, we can prove the

[17] $(1-H_u)d\widehat{S}_u = (1-H_u)(dS_u - z_u D_u d\langle S\rangle_u) \in \mathscr{M}(\mathbb{G}, \mathbb{P})$.

[18] $H_u d\widehat{S}_u = H_u(dS_u - \frac{a_u^\theta}{\alpha_u^\theta}|_{\theta=\tau} d\langle S\rangle_u) \in \mathscr{M}(\mathbb{G}, \mathbb{P})$.

Proposition 4 *Assume that* $\mathbf{H}_1$ *to* $\mathbf{H}_4$ *hold. Then, the set* $\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S)$ *is not empty and we can fully describe it as*

$$\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S) = \left\{ \mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}^*}\bigg|_{\mathscr{G}_t} = \mathscr{E}(-\vartheta \star \widehat{S})_t \mathscr{E}(\beta \star M)_t \right\},$$

with ϑ *given by* (8), *and* $\beta > -1$ *a* $\mathbb{G}$*-predictable process.*

Proof If the $\mathbb{F}$-conditional survival process writes $G_t = \mathbb{P}^*(\tau > t|\mathscr{F}_t) = Z_t - A_t$, the $(\mathbb{G}, \mathbb{P}^*)$-dynamics of S follow the decomposition (7):

$$S_t = \widehat{S}_t + \int_0^t \vartheta_u d\langle S\rangle_u \quad \text{with } \widehat{S} \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*),$$

where ϑ was defined in (8). Hence $\mathbb{P}^*$ is not a $\mathbb{G}$-e.m.m. From Proposition 3, the set of $\mathbb{G}$-e.m.m can be perfectly described as:

$$\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S) = \left\{ \mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}^*}\bigg|_{\mathscr{G}_t} = \mathscr{E}(-\vartheta \star \widehat{S})_t \mathscr{E}(\beta \star M)_t \right\},$$

where β is a $\mathbb{G}$-predictable processes, taking values in $]-1, \infty[$. As a check, under such a probability $\mathbb{Q}$, as $\widehat{S} \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*)$, one has, setting $\eta_t = \mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}^*}|_{\mathscr{G}_t}$

$$\begin{aligned} \widehat{S}_t - \int_0^t \frac{d\langle \widehat{S}, \eta\rangle_u}{\eta_u} &= \widehat{S}_t - \int_0^t d\langle \widehat{S}, -\vartheta \star \widehat{S} + \beta \star M\rangle_u \\ &= \widehat{S}_t + \int_0^t \vartheta_u d\langle S\rangle_u = S_t \end{aligned}$$

where the second equality comes from the fact that $\langle \widehat{S}\rangle = \langle S\rangle$ and $\langle \widehat{S}, M\rangle = 0$, since $\widehat{S}$ is continuous and M purely discontinuous. It follows from Girsanov's theorem that S is a $(\mathbb{G}, \mathbb{Q})$-martingale. □

As a conclusion, there exists at least a probability $\mathbb{Q}$ such that S is a $(\mathbb{G}, \mathbb{Q})$-martingale. We shall see in the immersion part that the drift relative to the change of probability may be interpreted as a risk premium.

4.2 Completeness of the Full Market

When considering also a $(n+2)$th asset S^{n+1} satisfying condition (1), we shall be able to select an e.m.m., in a unique way, and prove that the market defined by the price process $\widetilde{S}$ is complete.

4.2.1 The Default-Sensitive Asset

We start this section by emphasizing that it is necessary to introduce the asset S^{n+1} to the collection $(S^0, \dots, S^n)$ when working on derivatives whose pay-off depends on τ: It is not possible to hedge the jumping risk with $\mathbb{F}$-adapted assets and without such a product.

Let us consider for example a credit default swap[19] (*CDS*). To ease the discussion we take a continuous tenor, with a proportional continuous premium κ, and a constant recovery fee δ (S_t^0 denotes the saving account, i.e., the value at t of one unit invested at 0, and $\mathbb{Q}$ a martingale measure associated to this numeraire—that exists by absence of arbitrage). The price of a *CDS* is the difference between the value of the protection leg and the premium leg: $CDS(t, \delta, \kappa, T) = \mathrm{Prot}_t - \mathrm{Prem}_t$ with

$$\mathrm{Prem}_t = S_t^0 \kappa \mathbb{E}\left(\int_t^T \frac{1 - H_u}{S_u^0} du \middle| \mathcal{G}_t \right) = (1 - H_t) \frac{S_t^0 \kappa}{G_t} \int_t^T \mathbb{E}\left(\frac{1 - H_u}{S_u^0} \middle| \mathcal{F}_t \right) du$$

and

$$\mathrm{Prot}_t = S_t^0 \delta \mathbb{E}\left(\int_t^T \frac{dH_u}{S_u^0} \middle| \mathcal{G}_t \right) = (1 - H_t) \frac{S_t^0 \delta}{G_t} \int_t^T \mathbb{E}\left(\frac{(1 - H_u)\alpha_u^u}{S_u^0 G_u} \middle| \mathcal{F}_t \right) du.$$

Both legs have a value whose variation may be due to two factors: (i) $\mathbb{F}$-events (through the $\mathbb{F}$-conditional expectation), that evolve according to the alea structure underlying the filtration $\mathbb{F}$, and (ii) $\mathbb{H}$ events, mainly the occurrence of default (through the expression $1 - H$). Whereas it is reasonable to think that—under suitable assumptions—the "market/spread" variation of the value of the *CDS* (linked to the $\mathbb{F}$-events) may be hedged with $\mathbb{F}$-adapted instruments, the "jump" risk being not $\mathbb{F}$-adapted will not be hedgeable with assets of the reference market $\mathbb{F}$. It follows that a model containing only $\mathbb{F}$-adapted assets S would not be able to remove the jump risk of defaultable portfolios.

For quoted instruments like *CDS*, a formula like above allows to calibrate the parameters involved in the construction of the default time. A natural class of assets for S^{n+1} would be the risky bonds associated to τ or a *CDS*.

4.2.2 The Unique Martingale Probability

We introduce the asset S^{n+1} that is sensitive to the jump risk, i.e., for any t, the r.v. $\tau \wedge t$ is $\sigma(S_s^{n+1}, s \le t)$-measurable.[20] *Our aim is to prove that if the $\mathbb{F}$-market is complete, the $\mathbb{G}$-market is complete as well, under weak assumption on S^{n+1}.*

[19] Contract in which the holder buys a protection in paying a premium at each date of a tenor to the seller until a predefined credit event occurs, and receives a recovery fee if default occurs.

[20] Precisely w.r.t. its natural augmented càd version.

From Hypothesis $\mathbf{H}_1$, the $\mathbb{F}$-compensator of G writes $dA_t = \alpha_t^t dt$. From

$$dM_t = dH_t - \frac{1 - H_t}{G_t}\alpha_t^t dt, \tag{9}$$

we can derive the quadratic variation of process M:

$$[M]_t = [H]_t = \sum_{s \le t} \Delta H_s^2 = \sum_{s \le t} \Delta H_s = H_t,$$

since the two processes of the right-hand member of the equality (9) have finite variation paths and the second one is continuous. It follows that

$$[M]_t - \int_0^t \frac{1 - H_s}{G_s}\alpha_s^s ds = M_t \in \mathscr{M}(\mathbb{G}, \mathbb{P}),$$

hence that the angle brackets $\langle M \rangle$ exist and satisfy $d\langle M \rangle_t = (1 - H_t)\alpha_t^t / G_t dt$.

According to the last section, we consider the following framework for the full market:

1. *Reference assets.* The "reference" assets are defined under the historical probability $\mathbb{P}$ by:

$$\begin{cases} dS_t = b_t dt + dS_t^* \\ S_0 = x, \end{cases}$$

where $S^* \in \mathscr{M}(\mathbb{F}, \mathbb{P})$ continuous has a quadratic variation assumed to be absolutely continuous w.r.t. Lebesgue measure, $d\langle S^* \rangle_t = s_t dt$. Let $\widehat{S}$ be the $(\mathbb{G}, \mathbb{P})$-martingale part of the decomposition of S^* viewed as a $\mathbb{G}$-semi-martingale, which writes:

$$d\widehat{S}_t = dS_t^* - c_t dt,$$

with

$$c_u = (1 - H_u) z_u s_u / G_u + H_u a_u^\tau s_u / \alpha_u^\tau = s_u \vartheta_u,$$

with previous notations. It follows that the $\mathbb{G}$-decomposition of the semi-martingale S writes under the historical probability $\mathbb{P}$:

$$dS_t = (b_t + c_t)dt + d\widehat{S}_t := \nu_t dt + d\widehat{S}_t.$$

2. *Default sensitive asset.* We postulate for the asset S^{n+1} the general form:

$$dS_t^{n+1} = \mu_t dt + \varepsilon_t d\widehat{S}_t + \zeta_t dM_t,$$

where μ_t is a drift term and the three processes μ, ε and ζ are $\mathbb{G}$-predictable, and where ζ_t does not vanish (such a decomposition is quite general, from Theorem 1, the only assumption being the absolute continuity of the drift w.r.t. Lebesgue measure).

We can state the most important result of this part.

Theorem 2 *Assume that the reference market $\mathbb{F}$ is complete. If* $\mathbf{H}_1$ *to* $\mathbf{H}_5$ *hold and if the quadratic variation of S and the drifts of the assets are absolutely continuous w.r.t. Lebesgue measure the full market composed of $\widetilde{S}$ is complete.*

Proof The set of martingale probabilities that make S a $\mathbb{G}$-martingale writes, by Proposition 3:

$$\Theta_{\mathbb{P}}^{\mathbb{G}}(S) = \Big\{ \mathbb{Q} \sim \mathbb{P}, \exists \gamma, \beta \mathbb{G} \text{ predictable } \beta < 1, \\ \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathscr{G}_t} = \eta_t = \mathscr{E}(-\gamma \star \widehat{S})_t \mathscr{E}(-\beta \star M)_t \Big\}.$$

It follows that the exists a unique $\mathbb{G}$-e.m.m, i.e., a unique probability that makes $\widetilde{S} = (1, S^1, \ldots, S^n, S^{n+1})$ a $\mathbb{G}$-martingale. It is defined by:

$$\gamma_t = \nu_t, \quad \text{and} \quad \beta_t = G_t \frac{\mu_t - \varepsilon_t \nu_t}{\alpha_t^t \zeta_t},$$

by application of Girsanov's theorem. □

Once this probability has been defined, it is possible to price and hedge the τ-sensitive claims with $(S^0, \ldots, S^{n+1})$, like for example *CDS* on τ (of shorter maturities if S^{n+1} is a *CDS*) or derivatives written on S^{n+1}.

4.3 Immersion Property

We shall emphasize in this section the deep links between immersion and completeness. We start with some general results, precise them in the case where the credit event is an initial time and conclude with some considerations about the credit risk premium.

4.3.1 Immersion and Completeness in an Arbitrage Free Set Up

Proposition 5 *Assume that* (i) *the reference market is complete and* (ii) *the full market is arbitrage free. Then*:

1. *The restrictions of all the $\mathbb{G}$-e.m.m. on $\mathscr{F}_T$ are the same and*
2. *Immersion holds under every $\mathbb{G}$-e.m.m.*

Proof 1. *The restrictions of all the $\mathbb{G}$-e.m.m. on $\mathscr{F}_T$ are the same*. For any $\mathbb{Q} \in \Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S})$, it is straightforward[21] that $\mathbb{Q} \in \Theta_{\mathbb{P}}^{\mathbb{F}}(S)$. As the reference market is complete and S is finite dimensional, the restriction of all the $\mathbb{Q}$'s to the σ-algebra $\mathscr{F}_T$ is unique: All the e.m.ms of the full market have the same restriction on $\mathscr{F}_T$.

[21] $\Theta_{\mathbb{P}}^{\mathbb{G}}(\widetilde{S}) \subset \Theta_{\mathbb{P}}^{\mathbb{G}}(S) \subset \Theta_{\mathbb{P}}^{\mathbb{F}}(S)$, because S is $\mathbb{F}$-adapted.

2. *Immersion holds under every* $\mathbb{G}$*-e.m.m.* Indeed if P_1 holds, let $\mathbb{Q} \in \Theta_{\mathbb{P}}^{\mathbb{G}}(S)$ and $X \in \mathscr{M}^2(\mathbb{F}, \mathbb{Q})$, with $X_t = \mathbb{E}^{\mathbb{Q}}(X_T | \mathscr{F}_t)$. As recalled above, the completeness of the market implies the existence of a constant x and $\mathbb{F}$-predictable processes $(\varphi^i, i = 1, \dots, n)$ such that $X_T = x + \int_0^T \sum_{1 \le i \le n} \varphi_u^i dS_u^i$. Therefore,

$$\mathbb{E}^{\mathbb{Q}}(X_T | \mathscr{F}_t) = x + \sum_{1 \le i \le n} \int_0^t \varphi_u^i dS_u^i + \mathbb{E}^{\mathbb{Q}}\left(\int_t^T \varphi_u^i dS_u^i \Big| \mathscr{F}_t \right) = x + \sum_{i \le n} \int_0^t \varphi_u^i dS_u^i,$$

where the first equality comes from the fact that the random variable $\int_0^t \varphi_u^i dS_u^i$ is $\mathscr{F}_t$-measurable (the process φ^i is $\mathbb{F}$-predictable), and the second, from the fact that the process $\int_0^\cdot \varphi_u^i dS_u^i$ is a $(\mathbb{G}, \mathbb{Q})$-martingale (φ is $\mathbb{G}$-predictable and $\mathbb{Q} \in \Theta_{\mathbb{P}}^{\mathbb{G}}(S)$), and an $(\mathbb{F}, \mathbb{Q})$-martingale since it is $\mathbb{F}$-adapted. Hence $\mathbb{E}^{\mathbb{Q}}(\int_t^T \varphi_u^i dS_u^i | \mathscr{F}_t) = 0$ and $\mathbb{E}^{\mathbb{Q}}(\int_t^T \varphi_u^i dS_u^i | \mathscr{G}_t) = 0$. Therefore

$$\mathbb{E}^{\mathbb{Q}}(X_T | \mathscr{F}_t) = x + \sum_{i \le n} \int_0^t \varphi_u^i dS_u^i + \mathbb{E}^{\mathbb{Q}}\left(\int_t^T \varphi_u^i dS_u^i \Big| \mathscr{G}_t \right) = \mathbb{E}^{\mathbb{Q}}(X_T | \mathscr{G}_t),$$

hence $X \in \mathscr{M}^2(\mathbb{G}, \mathbb{Q})$ and immersion holds under $\mathbb{Q}$. Such a result had already been pointed out by Blanchet-Scaillet and Jeanblanc in [6]. □

4.3.2 Case Where the Credit Event is an Initial Time

The last result can be refined in the set up where the credit event is an initial time. Indeed we shall prove the

Proposition 6 *Assume that* $\mathbf{H}_1$ *to* $\mathbf{H}_4$ *hold. Then, immersion holds under any* $\mathbb{Q} \in \Theta_{\mathbb{P}^*}^{\mathbb{G}}(S)$.

This result implies that immersion holds under any $\mathbb{G}$-e.m.m[22] for S (it does not have to be a martingale probability of the full market).

Proof Let β be any predictable process such that $\beta_t > -1$, and $\mathbb{Q}^\beta$ be the corresponding e.m.m.:

- $d\mathbb{Q}^\beta|_{\mathscr{F}_\infty} = d\mathbb{P}^*|_{\mathscr{F}_\infty}$. Indeed, for any bounded $F_t \in \mathscr{F}_t$ with $\mathbb{P}^*$-null expectation, $F_t = \int_0^t f_s dS_s$ by *PRP* and

$$\mathbb{E}^\beta(F_t) = \mathbb{E}^*(F_t \eta_t) = \mathbb{E}^*\left(\int_0^t \eta_s f_s dS_s + \int_0^t F_s \, d\eta_s + \int_0^t f_s d\langle S, \eta \rangle_s \right)$$

[22] The set of $\mathbb{G}$-e.m.m. is infinite, parameterized by the predictable processes β.

$$= \mathbb{E}^*\left(\int_0^t \eta_s f_s dS_s + \int_0^t F_s\, d\eta_s + \int_0^t f_s \vartheta_s \eta_s d\langle \widehat{S}\rangle_s\right)$$

$$= \mathbb{E}^*\left(\int_0^t \eta_s f_s d\widehat{S}_s + \int_0^t F_s\, d\eta_s\right) = 0 = \mathbb{E}^*(F_t),$$

where the first equality is obtained by the integration by parts formula (f is predictable), the second comes from the definition of the dynamics of the density η and the third from the definition of $\widehat{S}$, the expectation being null since $\widehat{S}$ and η belong to $\mathscr{M}(\mathbb{G}, \mathbb{P}^*)$. It follows that $d\mathbb{Q}^\beta|_{\mathscr{F}_\infty} = d\mathbb{P}^*|_{\mathscr{F}_\infty}$.

- Let X be an $(\mathbb{F}, \mathbb{Q}^\beta)$-martingale. Then, it is a $(\mathbb{F}, \mathbb{P}^*)$-martingale (from the first point) that writes $dX_t = x_t dS_t$. From the decomposition formula in the change of filtration

$$\widehat{X}_t := X_t - \int_0^t x_u \vartheta_u d\langle S\rangle_u \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*),$$

and from the decomposition formula in the change of probability (Girsanov's theorem)

$$\widetilde{X}_t = \widehat{X}_t - \int_0^t \frac{d\langle X, \eta\rangle_u}{\eta_u} \in \mathscr{M}(\mathbb{G}, \mathbb{Q}^\beta).$$

It remains to note (by definition of η) that

$$\widetilde{X}_t = \widehat{X}_t + \int_0^t x_u \vartheta_u d\langle S\rangle_u = X_t,$$

hence that $X \in \mathscr{M}(\mathbb{G}, \mathbb{Q}^\beta)$. It follows that immersion holds under $\mathbb{Q}^\beta$.

□

It follows the important

Corollary 1 *Assume that* $\mathbb{F}$ *is complete and that* $\mathbb{P}^* \in \Theta_{\mathbb{P}}^{\mathbb{F}}(S)$. *If the* $(\mathbb{F}, \mathbb{P}^*)$-*conditional survival process* G^* *has a non constant martingale part*, $\mathbb{P}^*$ *is not a* $\mathbb{G}$-*e.m.m.*, *i.e.*, $\mathbb{P}^* \notin \Theta_{\mathbb{P}}^{\mathbb{G}}(S)$.

Proof From (3), it follows that under immersion

$$G_t = \int_0^\infty \alpha_{t\wedge u}^u du - \int_0^t \alpha_u^u du = \int_0^\infty \alpha_t^u du - A_t = \mathbb{P}(\tau > 0|\mathscr{F}_t) - A_t = 1 - A_t$$

hence G is decreasing and predictable. □

A broader class of credit event may be reached through a definition where the martingale part of the survival process is not equal to zero (see [23]). Under such a framework, immersion does not hold, which means that the "reference neutral-risk probability" is not a neutral-risk probability. A change of measure has therefore to be performed to re-enter a neutral-risk framework. Next section points out the links of this remark with the credit risk premium.

4.3.3 Credit Risk Premium

The last corollary may be interpreted in the following way. The change from the historical probability $\mathbb{P}$ to a neutral-risk probability $\mathbb{P}^*$ aims at correcting the dynamics from the market risk premium. It is well known that to any financial market can be associated a risk premium. It characterizes the return over the risk-free return (the interest rate) an investor may expect for bearing the risk of taking a long position on a derivative written on this market.

Indeed if N is the martingale modelling the alea of this market, and if the asset's return writes:

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dN_t,$$

the dynamics of any derivatives written on S sold for a price P_t at a date t would write $dP_t/P_{t-} = \kappa dt + \alpha dN_t$. A risk-free portfolio can be set-up in buying a quantity $S\sigma$ of the derivative, for a total value of $S\sigma P$ and selling a quantity $P\alpha$ of the asset (completing by ς_t of money market S_t^0 to remain self financed). The value of this portfolio is $\Pi_t = \varsigma_t S_t^0 + S_t \sigma P_t - P_t \alpha S_t$ and the self-financing condition yields

$$\begin{aligned}
d\Pi_t &= \varsigma_t r S_t^0 dt + S_{t-}\sigma dP_t - P_{t-}\alpha dS_t \\
&= \varsigma_t r S_t^0 dt + \sigma\kappa S_t P_t dt + \alpha\sigma S_{t-}P_{t-}dN_t - \alpha P_t S_t \mu dt - \alpha\sigma P_{t-}S_{t-}dN_t \\
&= ((\Pi_t - (\sigma - \alpha)P_t S_t)r + (\sigma\kappa - \alpha\mu)P_t S_t)dt.
\end{aligned}$$

By absence of arbitrage, the return of this risk-free portfolio must be equal to r, so that:

$$(\sigma\kappa - \alpha\mu)P_t S_t = r(\sigma - \alpha)P_t S_t \Longleftrightarrow \frac{\kappa - r}{\alpha} = \frac{\mu - r}{\sigma} = \lambda_S.$$

On the reference market, the e.m.m. $\mathbb{P}^*$ corrects the historical probability $\mathbb{P}$ from the market risk premium: *If immersion does not hold under $\mathbb{P}^*$, it means the market risk premium does not take into account the jump risk premium, and it is necessary to change to a $\mathbb{G}$-e.m.m. $\mathbb{Q}$ under which S remains a martingale.*

5 Incomplete Markets

The question addressed in this last part is the adaptation of the above results in the case where the reference market is incomplete.

When an incomplete model is chosen—in general for its ability to well reproduce a given class of (calibration) instruments and its dynamics property (regarding to the products to price)—one is often conduced to focus on one particular martingale probability, even if the set of $\mathbb{F}$-e.m.m. is not unique. The selection of the e.m.m. is performed by the calibration procedure. The law of the price process is then uniquely determined, and a change of probability within the set of $\mathbb{F}$-e.m.m. will change the price of the selected options or break the imposed constraints.

We assume therefore in this section that an $\mathbb{F}$-e.m.m has been chosen (for pricing the default-free derivatives), hence we restrict our attention on a given $\mathbb{F}$-e.m.m. $\mathbb{P}^*$ defined on $\mathscr{A}$ and equivalent to the historical probability $\mathbb{P}$, such that $S \in \mathscr{M}(\mathbb{F}, \mathbb{P}^*)$.

The purpose of this section is:

- To prove that there exists a unique probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}^*$ such that $\widetilde{S}$ (defined in the last part as the $(n+2)$-uplet composed of the reference asset S and S^{n+1}) is a $(\mathbb{G}, \mathbb{Q})$-martingale, and that preserves the "reference pricing", i.e.,
$$\mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^*(X_T), \quad \text{for any } X_T \in L^2(\mathscr{F}_T).$$
- To prove that immersion property holds under this measure.

Recall that, from hypothesis $\mathbf{H}_1$ and $\mathbf{H}_2$, the random time τ is an initial time that avoids the $\mathbb{F}$-stopping times.

Recall that if X is a real valued local-martingale and Y a $\mathbb{R}^d$-valued local-martingale, a Galtchouck Kunita Watanabe decomposition of X is a decomposition of the form $X = X_0 + H * Y + L$ with $H \in L^2_{loc}(Y)$ and L a local-martingale such that $L_0 = 0$ and strongly orthogonal[23] to Y. It is classical that if Y is continuous or locally square integrable, X admits a Galtchouck-Kunita-Watanabe decomposition (Kunita and Watanabe [26], Galtchouck [15], Jacod [17]).

5.1 The Risk-Neutral Probabilities of the Full Market

Reference assets and $\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S)$. As in the complete case, we assume that S is continuous and that its quadratic variation process is continuous w.r.t. Lebesgue measure, $d\langle S\rangle_t = s_t dt$.

As S is continuous, there exist:

- for any $u \geq 0$, an $\mathbb{F}$-predictable process a^u and a square integrable $\mathbb{F}$-martingale N^u strongly orthogonal to S and
- an $\mathbb{F}$-predictable process z and a square integrable $\mathbb{F}$-martingale $N^Z \in \mathscr{M}(\mathbb{F}, \mathbb{P}^*)$ strongly orthogonal to S

such that the following Galtchouck-Kunita-Watanabe decompositions exist:

$$\begin{cases} \alpha^u = \alpha^u_0 + a^u \star S + N^u \\ Z = Z_0 + z \star S + N^Z \end{cases} \tag{10}$$

where $Z = G + A$. Moreover there exist optional versions of the functions a^u and N^u.

[23] Recall that two local-martingales are strongly orthogonal if their product is a local martingale.

We have seen in (7) that the decomposition of the $(\mathbb{G},\mathbb{P}^*)$-semi-martingale S writes $S=\widehat{S}+C$ with:

$$dC_t=\left(\frac{1-H_t}{G_t}z_t+\frac{H_t}{\alpha^\tau_t}a^\tau_t\right)d\langle S\rangle_t=:\vartheta_t d\langle S\rangle_t,$$

and $\widehat{S}$ continuous. It follows that any $\mathbb{G}$-martingale l admits a Galtchouck-Kunita-Watanabe decomposition of the form

$$l=l_0+\varphi\star\widehat{S}+h$$

with h a local martingale strongly orthogonal to l. Recall that the martingale M is strongly orthogonal to $\widehat{S}$ (it is purely discontinuous) and is a locally square martingale, hence h admits a Galtchouck-Kunita-Watanabe decomposition:

$$h=\psi\star M+N^\perp$$

with $N^\perp$ strongly orthogonal to $(M,\widehat{S})$. It follows that the set of probabilities $\mathbb{G}$-equivalent to $\mathbb{P}^*$ writes:

$$\left\{\mathbb{Q},\left.\frac{d\mathbb{Q}}{d\mathbb{P}^*}\right|_{\mathscr{G}_t}=\eta_t=\mathscr{E}(\varphi\star\widehat{S})_t\mathscr{E}(\psi\star M)_t\mathscr{E}(N^\perp)_t\right\},$$

with (φ,ψ) a pair of $\mathbb{F}$-predictable processes and $N^\perp$ a $\mathbb{G}$-martingale that is strongly orthogonal to $\widehat{S}$ and M.

Finally, the set of $\mathbb{G}$-e.m.m. $\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S)$ writes:

$$\Theta^{\mathbb{G}}_{\mathbb{P}^*}(S)=\left\{\mathbb{Q},\left.\frac{d\mathbb{Q}}{d\mathbb{P}^*}\right|_{\mathscr{G}_t}=\eta_t=\mathscr{E}(-\vartheta\star\widehat{S})_t\mathscr{E}(\psi\star M)_t\mathscr{E}(N^\perp)_t\right\}$$

with $N^\perp\in\mathscr{M}(\mathbb{G},\mathbb{P}^*)$ strongly orthogonal to the pair $(\widehat{S},M)$ and $-1<\psi\in\mathscr{P}(\mathbb{G})$. Indeed the "Girsanov's drift" under each such probability writes $\langle-\vartheta\star\widehat{S},\widehat{S}\rangle$ because of the *strong* orthogonality of the other terms w.r.t. $\widehat{S}$. This set is parameterized by the pair $(\psi,N^\perp)$.

Default sensitive asset. If μ_t denotes the drift of the default sensitive asset S^{n+1}, the general dynamics of the price process can be written as:

$$dS^{n+1}_t=\mu_t dt+\varepsilon_t d\widehat{S}_t+\zeta_t dM_t+dN^{n+1}_t,$$

with $N^{n+1}\in\mathscr{M}^2(\mathbb{G},\mathbb{P}^*)$ strongly orthogonal to the pair $(\widehat{S},M)$ (same argument as before). As before, we assume $d\langle S^{n+1}\rangle_t\ll dt$.

Finally, it follows that the set $\Theta^{\mathbb{G}}_{\mathbb{P}^*}(\widetilde{S})$ writes:

$$\Theta^{\mathbb{G}}_{\mathbb{P}^*}(\widetilde{S})=\left\{\mathbb{Q},\left.\frac{d\mathbb{Q}}{d\mathbb{P}^*}\right|_{\mathscr{G}_t}=\eta_t=\mathscr{E}(-\vartheta\star\widehat{S})_t\mathscr{E}(\psi\star M)_t\mathscr{E}(N^\perp)_t\right\}$$

with $N^{\perp} \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*)$ strongly orthogonal to the pair $(\widehat{S}, M)$ and $\psi \in \mathscr{P}(\mathbb{G})$, defined[24] by

$$\psi_t \zeta_t \alpha_t^t dt = \mu_t dt + \varepsilon_t \vartheta_t s_t dt - d\langle N^{n+1}, N^{\perp}\rangle_t. \tag{11}$$

5.2 Default-Free Pricing Invariance

In both cases, the martingale probability is going to be uniquely defined thanks to the constraint of $\mathbb{F}$-pricing invariance.

Let $X_T \in L^2(\mathscr{F}_T)$, such that $\mathbb{E}^*(X_T) = 0$. The $(\mathbb{F}, \mathbb{P}^*)$-martingale X can be decomposed as $X = x \star S + N$, with $(S, N) \in \mathscr{M}(\mathbb{F}, \mathbb{P}^*)$ strongly orthogonal. The decomposition of this $(\mathbb{G}, \mathbb{P}^*)$-semi-martingale writes

$$X = x \star \widehat{S} + x \star C + \widehat{N} + K,$$

with $(\widehat{S}, \widehat{N}) \in \mathscr{M}(\mathbb{G}, \mathbb{P}^*)$, $dC_t = \vartheta_t d\langle S\rangle_t = \vartheta_t s_t dt$ and:

$$\begin{aligned} dK_t &= \frac{1 - H_t}{G_t} d\langle N, G\rangle_t + \frac{H_t}{\alpha_t^u} d\langle N, \alpha^u\rangle_t \Big|_{u=\tau} \\ &= \frac{1 - H_t}{G_t} d\langle N, N^Z\rangle_t + \frac{H_t}{\alpha_t^u} d\langle N, N^u\rangle_t \Big|_{u=\tau}, \end{aligned}$$

where N^u and N^Z are defined in (10). It follows that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^*(X_T \eta_T) &= \mathbb{E}^*\bigg(\int_0^T \eta_t dX_t + [X, \eta]_T\bigg) \\ &= \mathbb{E}^*\bigg(\int_0^T \eta_t (x_t dC_t + dK_t) + [X, \eta]_T\bigg) \\ &= \mathbb{E}^*\bigg(\int_0^T \eta_t (x_t \vartheta_t d\langle \widehat{S}\rangle_t + dK_t) + [X, \eta]_T\bigg). \end{aligned}$$

Moreover we can write:

$$\mathbb{E}^*([X, \eta]_T) = \mathbb{E}^*([x \star \widehat{S} + \widehat{N}, \mathscr{E}(-\vartheta \star \widehat{S})\mathscr{E}(\psi \star M)\mathscr{E}(N^{\perp})]_T)$$

since $x \star C + K$ is $\mathbb{G}$-predictable with finite variation,

$$\begin{aligned} \mathbb{E}^*([X, \eta]_T) &= \mathbb{E}^*([x \star \widehat{S}, \mathscr{E}(-\vartheta \star \widehat{S})\mathscr{E}(\psi \star M)\mathscr{E}(N^{\perp})]_T) \\ &\quad + \mathbb{E}^*([\widehat{N}, \mathscr{E}(-\vartheta \star \widehat{S})\mathscr{E}(\psi \star M)\mathscr{E}(N^{\perp})]_T) \end{aligned}$$

[24] Notice that $d\langle N^{n+1}, N^{\perp}\rangle_t \ll dt$, from Kunita Watanabe $(d\langle S^{n+1}\rangle_t \ll dt$ implies $d\langle N^{n+1}\rangle_t \ll dt)$.

$$= \mathbb{E}^*\left(\int_0^T -\eta_t \vartheta_t x_t d[\widehat{S}]_t\right) + \mathbb{E}^*\left(\int_0^T \eta_t \psi_t d[\widehat{N}, M]_t\right)$$
$$+ \mathbb{E}^*\left(\int_0^T \eta_t d[\widehat{N}, N^{\perp}]_t\right)$$

and since K is $\mathbb{G}$-predictable with finite variation, $[K, M] = 0$, and $[\widehat{N}, M] = [N, M] = 0$ (recall τ avoids the $\mathbb{F}$-stopping times—M is purely discontinuous and jumps at τ). It follows

$$\mathbb{E}^*([X, \eta]_T) = \mathbb{E}^*\left(\int_0^T -\eta_t \vartheta_t x_t d\langle \widehat{S}\rangle_t + \int_0^T \eta_t d[\widehat{N}, N^{\perp}]_t\right)$$

hence

$$\mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^* \int_0^T \eta_t d(K_t + [\widehat{N}, N^{\perp}]_t). \tag{12}$$

We shall prove that there exists a unique $N^{\perp}$ such that $\mathbb{E}^{\mathbb{Q}}(X_T) = 0$ for any $X_T \in L^2(\mathscr{F}_T)$, such that $\mathbb{E}^*(X_T) = 0$.

Let us define the $\mathbb{F} \otimes \mathscr{B}(\mathbb{R}^+)$-optional process

$$R_t^x = -\int_0^t 1_{\{x \le u\}} \frac{dN_u^x}{\alpha_{u-}^x},$$

which is for any $x \ge 0$ an $\mathbb{F}$-martingale. We start with the following

Lemma 1 *If X^x is an $\mathbb{F} \otimes \mathscr{B}(\mathbb{R}^+)$-optional family of $\mathbb{F}$-semi-martingales, $X = (X_t^{\tau \wedge t}, t \ge 0)$ is a $\mathbb{G}$-semi-martingale.*

Proof 1. The process X is $\mathbb{G}$-adapted since $\tau \wedge t \in \mathscr{G}_t$, and the function X is $\mathbb{F} \otimes \mathscr{B}(\mathbb{R}^+)$-optional.

2. We shall prove that for any sequence of càg piecewise constant $\mathbb{F}$-adapted processes K^n that tends to zero uniformly in (t, ω), $K^n \star S$ tends to zero in probability.[25]

For any subdivision $(t_1, \dots, t_n)$ of $[0, T]$, and any $(K_1^n, \dots, K_n^n)$ such that for any $i \le n$ $K_i^n \in \mathscr{G}_{t_i}$, and $K_i^n = F_i^n h_i^n(t_i \wedge \tau)$ with h_i^n Borel measurable and $F_i^n \in \mathscr{F}_{t_i}$, we introduce

$$K^n \star X_T = \sum_i K_i^n (X_{i+1} - X_i) = \sum_i K_i^n (X_{t_{i+1}}^{\tau \wedge t_{i+1}} - X_{t_i}^{\tau \wedge t_i})$$

[25] Recall that an $(\mathbb{F}, \mathbb{P})$-semi-martingale S can whether be define as the sum of an $\mathbb{F}$-adapted process of finite variation process and an $(\mathbb{F}, \mathbb{P})$-local-martingale, or as an $\mathbb{F}$-adapted processes such that for any sequence of càg piecewise constant $\mathbb{F}$-adapted processes K_n that tends to zero uniformly in (t, ω), $K_n \star S$ tends to zero in probability (Bichteler-Dellacherie theorem). We use this second definition in the proof.

$$= \sum_i F_i^n h_i^n (t_i \wedge \tau)(X_{t_{i+1}}^{\tau \wedge t_{i+1}} - X_{t_i}^{\tau \wedge t_i}).$$

It is classical that a sequence of random variables I_n tends to zero in probability iff the real sequence $\mathbb{E}(1 \wedge |I_n|)$ tends to zero. Following this trail, we have

$$\begin{aligned} \mathbb{E}(1 \wedge |K^n \star X_T|) &= \mathbb{E}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge \tau)(X_{t_{i+1}}^{\tau \wedge t_{i+1}} - X_{t_i}^{\tau \wedge t_i})\right|\right) \\ &= \int_0^\infty \mathbb{E}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge u)(X_{t_{i+1}}^{u \wedge t_{i+1}} - X_{t_i}^{u \wedge t_i})\right| \alpha_T^u\right) du \end{aligned}$$

and from Lebesgue's theorem, since

$$\mathbb{E}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge u)(X_{t_{i+1}}^{u \wedge t_{i+1}} - X_{t_i}^{u \wedge t_i})\right| \alpha_T^u\right) \leq \mathbb{E}(1 * \alpha_T^u) \leq 1,$$

we can write

$$\lim_n \mathbb{E}(1 \wedge |K^n \star X_T|) = \int_0^\infty \lim_n \mathbb{E}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge u)(X_{t_{i+1}}^{u \wedge t_{i+1}} - X_{t_i}^{u \wedge t_i})\right| \alpha_T^u\right) du.$$

Moreover, if the probability $\mathbb{Q}$ is defined on $\mathscr{F}_\infty$ by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathscr{F}_t} = \frac{\alpha_t^u}{\alpha_0^u},$$

we have

$$\begin{aligned} &\lim_n \mathbb{E}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge u)(X_{t_{i+1}}^{u \wedge t_{i+1}} - X_{t_i}^{u \wedge t_i})\right| \alpha_T^u\right) \\ &= \lim_n \mathbb{E}^{\mathbb{Q}}\left(1 \wedge \left|\sum_i F_i^n h_i^n (t_i \wedge u)(X_{t_{i+1}}^{u \wedge t_{i+1}} - X_{t_i}^{u \wedge t_i})\right|\right) = 0 \end{aligned}$$

since X^u is an $\mathbb{F}$-martingale, hence an $\mathbb{F}$-semi-martingale. It follows by a class monotone argument that X is a $\mathbb{G}$-semi-martingale. □

We deduce from this lemma that the process $X = HR^\tau$ is a $\mathbb{G}$-semi-martingale, since $X = X^{\tau \wedge t}$ with $X_t^x = 1_{x \leq t} R_t^x$ (on $\{\tau \leq t\}$, $\tau = \tau \wedge t$). So that to be able to develop the equality (12), we need to prove the

Lemma 2 *If Y^x is an $\mathbb{F} \otimes \mathscr{B}(\mathbb{R}^+)$-optional family of $\mathbb{G}$-semi-martingales, and $Y = (Y_t^{\tau \wedge t}, t \geq 0)$ we have*

$$\left([Y, HR^\tau]_t - \int_0^t \frac{H_u}{\alpha_{u-}^x} d\langle Y^x, N^x \rangle_u \Big|_{x=\tau}, t \geq 0\right) \in \mathscr{M}(\mathbb{G}).$$

Proof 1. Let us first mention that the angle bracket $\langle Y^x, N^x\rangle$ is well defined as

$$\langle Y^x, N^x\rangle = \langle Y^x, a^u \star S\rangle - \langle Y^x, \alpha^x\rangle,$$

the first bracket being defined since S is continuous and the second from [22]. Moreover, $[Y, HR^\tau]$ exists from the last lemma, since both processes are semi-martingales.

2. Let F_t be an $\mathscr{F}_t$-measurable variable and h be Borel measurable. We have:

$$\begin{aligned}\mathbb{E}(F_t h(t\wedge\tau)[Y, HR^\tau]_t^T) &= \int_0^\infty \mathbb{E}(F_t h(t\wedge x)[Y^x, HR^x]_T \alpha_T^x)dx\\ &\quad - \int_0^\infty \mathbb{E}(F_t h(t\wedge x)[Y^x, HR^x]_t \alpha_t^x)dx\\ &= \int_0^\infty \mathbb{E}\left(F_t h(t\wedge x)\int_t^T \alpha_{u-}^x d[Y^x, HR^x]_u\right)dx\end{aligned}$$

since by Itô's rule (the last integral being an $\mathbb{F}$-martingale):

$$[Y^x, HR^x]_T\alpha_T^x - [Y^x, HR^x]_t\alpha_t^x = \int_t^T \alpha_{u-}^x d[Y^x, HR^x]_u + \int_t^T [Y^x, HR^x]_{u-} d\alpha_u^x,$$

hence finally, by definition of R^x:

$$\begin{aligned}\mathbb{E}(F_t h(t\wedge\tau)[Y, HR^\tau]_t^T) &= \int_0^\infty \mathbb{E}\left(\int_t^T F_t h(t\wedge x)\alpha_{u-}^x \frac{1_{x\le u}}{\alpha_{u-}^x} d[Y^x, N^x]_u\right)dx\\ &= \int_0^\infty \mathbb{E}\left(\int_t^T F_t h(t\wedge x)1_{x\le u} d\langle Y^x, N^x\rangle_u\right)dx.\end{aligned}$$

Moreover

$$\begin{aligned}&\mathbb{E}\left(F_t h(t\wedge\tau)\int_t^T \frac{1_{x\le u}}{\alpha_{u-}^x} d\langle Y^x, N^x\rangle_u\Big|_{x=\tau}\right)\\ &= \int_0^\infty \mathbb{E}\left(\int_t^T F_t h(t\wedge x)\frac{1_{x\le u}}{\alpha_{u-}^x} d\langle Y^x, N^x\rangle_u \alpha_T^x\right)dx\\ &= \int_0^\infty \mathbb{E}\left(\int_t^T F_t h(t\wedge x)1_{x\le u} d\langle Y^x, N^x\rangle_u\right)\eta dx,\end{aligned}$$

since $\mathbb{E}(\alpha_T^x | F_{u-}) = \alpha_{u-}^x$, which concludes the proof, by a class monotone argument. □

We deduce from this lemma the two following points:

- The $\mathbb{G}$-semi-martingale HR^τ is special. Indeed, the lemma applied to $Y^x = HR^x$ leads to the existence of the angle bracket $\langle HR^\tau\rangle$.

- The lemma applied to the $\mathbb{F}$-martingale N leads to

$$\left([N, HR^\tau]_t - \int_0^t \frac{H_u}{\alpha^x_{u-}} d\langle N, N^x\rangle_u \Big|_{x=\tau}, t \geq 0\right) \in \mathscr{M}(\mathbb{G}).$$

We introduce the $\mathbb{G}$-semi-martingale

$$\Gamma_t = \int_0^t \frac{H_u - 1}{G_u} dN^Z_u - H_t R^\tau_t.$$

As the first integral is special (N^Z is an $\mathbb{F}$-martingale, hence a special $\mathbb{G}$-semi-martingale, from 2), Γ is a special $\mathbb{G}$-semi-martingale. We can conclude this part with the

Proposition 7 *There exists a unique $\mathbb{G}$-e.m.m. $\mathbb{Q} \in \Theta^{\mathbb{G}}_{\mathbb{P}^*}(\widetilde{S})$, that preserves $\mathscr{F}_\infty$, i.e., such that*

$$\mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^*(X_T), \quad \textit{for any } X_T \in L^2(\mathscr{F}_T).$$

Proof 1. We have seen that for $X_T \in L^2(\mathscr{F}_T)$ with $\mathbb{E}^*(X_T) = 0$,

$$\mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^{\mathbb{Q}}(X_T) = \mathbb{E}^* \int_0^T \eta_t d(K_t + [\widehat{N}, N^\perp]_t).$$

If we choose for $N^\perp$ the (unique) $\mathbb{G}$-martingale part of Γ (in its special decomposition[26]), we have

$$\begin{aligned}\mathbb{E}^* \int_0^T \eta_t d[\widehat{N}, N^\perp]_t &= \mathbb{E}^* \int_0^T \eta_t \frac{H_u - 1}{G_u} d\langle N, N^Z_u\rangle_t - \mathbb{E}^* \int_0^T \eta_t \frac{H_t}{\alpha^u_t} d\langle N, N^u\rangle_t \Big|_{u=\tau} \\ &= -\mathbb{E}^* \int_0^T \eta_t dK_t,\end{aligned}$$

where the first equality comes from the last lemma. It follows that $\mathbb{E}^{\mathbb{Q}}(X_T) = 0$.

2. For any probability $\mathbb{Q}$ defined with another martingale $\widetilde{N}^\perp = N^\perp + \mu$ where $\mu \in \mathscr{M}^2(\mathbb{G}, \mathbb{P}^*)$ is strongly orthogonal to $N^\perp$ non constant, when computing $\mathbb{E}^{\mathbb{Q}}(X_T)$ for a variable such that $\widehat{N} = \mu$,

$$\mathbb{E}^* \int_0^T \eta_t d(n_t + \langle \widehat{N}, \widetilde{N}^\perp\rangle_t) = \mathbb{E}^* \int_0^T \eta_t d\langle \widehat{N}, \mu\rangle_t = \mathbb{E}^* \int_0^T \eta_t d\langle \mu\rangle_t \neq 0.$$

It follows that this probability measure is unique. □

[26]The decomposition martingale plus predictable process with finite variation paths.

5.3 Immersion Property

We have seen that given a reference risk-neutral probability there exists a unique risk-neutral probability on the full market that preserve the "reference pricing". We conclude this survey by the important

Proposition 8 *Under the* $\mathbb{G}$*-e.m.m.* $\mathbb{Q} \in \Theta^{\mathbb{G}}_{\mathbb{P}^*}(\widetilde{S})$ *that preserves* $\mathscr{F}_\infty$, *immersion holds.*

Proof Let $X \in \mathscr{M}(\mathbb{F}, \mathbb{Q})$, $X_t = \mathbb{E}^{\mathbb{Q}}(X_T|\mathscr{F}_t)$. As $\mathbb{Q}|_{\mathscr{F}_\infty} = \mathbb{P}^*|_{\mathscr{F}_\infty}$, $X_t = \mathbb{E}^*(X_T|\mathscr{F}_t)$. Indeed, for $F_t \in \mathscr{F}_t$,

$$\mathbb{E}^*(X_T F_t) = \mathbb{E}^{\mathbb{Q}}(X_T F_t) = \mathbb{E}^{\mathbb{Q}}(X_t F_t) = \mathbb{E}^*(X_t F_t).$$

Moreover, if $X = X_0 + x \star S + N$ with $N \in \mathscr{M}(\mathbb{F}, \mathbb{P}^*)$ strongly orthogonal to S, the $(\mathbb{G}, \mathbb{P}^*)$-decomposition of this process writes:

$$X = X_0 + x \star \widehat{S} + x \star C + \widehat{N} + K.$$

Under $\mathbb{Q}$, from Girsanov's theorem:

$$X = X_0 + x \star \widetilde{S} + x \star \langle \widehat{S}, \log \eta \rangle + x \star C + \widetilde{N} + \langle \widehat{N}, \log \eta \rangle + K,$$

and by definition of $\mathbb{Q}$, $\langle \widehat{S}, \log \eta \rangle = -C$ and $\langle \widehat{N}, \log \eta \rangle = -K$ (see above), hence

$$X = x \star \widetilde{S} + \widetilde{N} \in \mathscr{M}(\mathbb{G}, \mathbb{Q}),$$

which concludes the proof. □

6 Conclusion

In this paper, we have given some arguments that show that it is natural to assume that immersion hypothesis holds for a study of a single default, and proved its deep link with completeness and martingale decomposition.

However, it is well known that it is usually impossible to assume this hypothesis in case of (non-ordered) multi-defaults,and that the martingale parts of the survival probabilities reflects the correlation between the different default times.

Acknowledgements The authors wish to thank Stéphane Crépey for pointing out some mistakes in a preliminary version and the anonymous referee for all his/her accurate remarks.

References

1. Albanese, C., Chen, O.: Pricing equity default swap. Risk **6** (2005)
2. Atlan, M., Leblanc, B.: Hybrid equity-credit modelling. Risk 61–66 (August, 2005)

3. Bélanger, A., Shreve, S.E., Wong, D.: A general framework for pricing credit risk. Math. Financ. **14**, 317–350 (2004)
4. Bielecki, T., Rutkowski, M.: Credit Risk: Modeling, Valuation and Hedging. Springer, Berlin (2002)
5. Black, F., Cox, J.C.: Valuing corporate securities: Some effects of bond indenture provisions. J. Finance **31**, 351–367 (1976)
6. Blanchet-Scalliet, C., Jeanblanc, M.: Hazard rate for credit risk and hedging defaultable contingent claims. Finance Stoch. **8**, 145–159 (2004)
7. Carr, P., Wu, L.: Stock options and credit default swaps: a joint framework for valuation and estimation. Working paper (2005)
8. Carr, P., Wu, L.: Theory and evidence on the dynamic interactions between sovereign credit default swaps and currency options. J. Bank. Financ. **31**, 2383–2403 (2007)
9. Cremers, M., Driessen, M., Maenhout, J., Weinbaum, P.J.: Individual stock options and credit spreads. Working paper, Yale University
10. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
11. El Karoui, N., Jeanblanc, M., Jiao, Y.: Density model for single default. Working paper (2008)
12. Ehlers, P., Schonbucher, P.: Pricing interest rate sensitive portfolio derivatives. Working paper (2007)
13. Ehlers, P., Schonbucher, P.: The influence of FX risk on credit spreads. Working paper (2007)
14. Elliott, R.J., Jeanblanc, M., Yor, M.: On models of default risk. Math. Financ. **10**, 179–195 (2000)
15. Galtchouk, L.I.: The structure of a class of martingales. In: Proceedings Seminar on Random Processes. Druskininkai, Academy of Sciences of Lithuanian SSP I, pp. 7–32 (1975)
16. Heston, S.I.: A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. **6**, 327–343 (1993)
17. Jacod, J.: Calcul Stochastique et Problèmes de Martingales. Lecture Notes in Mathematics, vol. 714. Springer, Berlin (1979)
18. Jacod, J.: A general theorem of representation for martingales. In: Proc. AMS Pro. Symp. Urbana, pp. 37–53 (1976)
19. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, 2nd edn. Springer, Berlin (2003)
20. Jacod, J., Yor, M.: Etude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales. Z. Wahrscheinlichkeitstheor. Verw. Geb. **38**, 83–125 (1977)
21. Jiao, Y.: Le risque de crédit: la modèlisation et la simulation numérique. Thèse de doctorat, Ecole Polytechnique (2006)
22. Jeanblanc, M., Le Cam, Y.: Progressive enlargement of filtrations with Initial times. Stoch. Process. Appl. **119**(8), 2523–2543 (2009)
23. Jeanblanc, M., Le Cam, Y.: Reduced form modeling. Preprint, Evry University (2007, submitted)
24. Jeanblanc, M., Rutkowski, M.: Default risk and hazard processes. In: Geman, H., Madan, D., Pliska, S.R., Vorst, T. (eds.) Mathematical Finance—Bachelier Congress 2000, pp. 281–312. Springer, Berlin (2000)
25. Kabanov, Yu.: Arbitrage theory. In: Jouini, E., Cvitanić, J., Musiela, M. (eds.) Option Pricing, Interest Rates and Risk Management, pp. 3–42. Cambridge University Press, Cambridge (2001)
26. Kunita, H., Watanabe, S.: On square integrable martingales. Nagoya Math. J. **30**, 29–245 (1967)
27. Kusuoka, S.: A remark on default risk models. Adv. Math. Econ. **1**, 69–82 (1999)
28. Linetsky, V.: Pricing equity derivatives subject to bankrupcy. Math. Financ. **16**, 255–282 (2005)
29. Merton, R.: On the pricing of corporate debt: The risk structure of interest rates. J. Finance **29**, 449–470 (1974)

30. Protter, P.E.: A partial introduction to financial asset pricing theory. Stoch. Process. Appl. **91**, 169 (2001)
31. Protter, P.E.: Stochastic Integration and Differential Equations, 2nd edn. Springer, Berlin (2003)
32. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)

Optimal Consumption and Investment with Bounded Downside Risk for Power Utility Functions

Claudia Klüppelberg
and Serguei Pergamenchtchikov

Abstract We investigate optimal consumption and investment problems for a Black-Scholes market under uniform restrictions on Value-at-Risk and Expected Shortfall. We formulate various utility maximisation problems, which can be solved explicitly. We compare the optimal solutions in form of optimal value, optimal control and optimal wealth to analogous problems under additional uniform risk bounds. Our proofs are partly based on solutions to Hamilton-Jacobi-Bellman equations, and we prove a corresponding verification theorem.

Keywords Portfolio optimisation · Stochastic optimal control · Risk constraints · Value-at-Risk · Expected shortfall

Mathematics Subject Classification (2000) 91B28 · 93E20

1 Introduction

We consider an investment problem aiming at optimal consumption during a fixed investment interval $[0, T]$ in addition to an optimal terminal wealth at maturity T. Such problems are of prime interest for the institutional investor, selling asset funds to their customers, who are entitled to certain payment during the duration of an investment contract and expect a high return at maturity. The classical approach to this problem goes back to Merton [10] and involves utility functions, more precisely, the expected utility serves as the functional which has to be optimised.

We adapt this classical utility maximisation approach to today's industry practice: investment firms customarily impose limits on the risk of trading portfolios.

This work was supported by the European Science Foundation through the AMaMeF programme.

C. Klüppelberg (✉)
Center for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, 85747 Garching, Germany
e-mail: cklu@ma.tum.de

S. Pergamenchtchikov
Laboratoire de Mathématiques, Raphaël Salem Université de Rouen, BP. 12, 76801 Saint Etienne du Rouvray, France
e-mail: Serge.Pergamenchtchikov@univ-rouen.fr

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_7,

These limits are specified in terms of downside risk measures as the popular Value-at-Risk (VaR) or Expected Shortfall (ES). We briefly comment on these two risk measures.

As Jorion [5], p. 379 points out, VaR creates a common denominator for the comparison of different risk activities. Traditionally, position limits of traders are set in terms of notional exposure, which may not be directly comparable across treasuries with different maturities. In contrast, VaR provides a common denominator to compare various asset classes and business units. The popularity of VaR as a risk measure has been endorsed by regulators, in particular, the Basel Committee on Banking Supervision, which resulted in mandatory regulations worldwide. One of the well-known drawbacks of VaR is due to its definition as a quantile. This means that only the probability to exceed a VaR bound is considered, the values of the losses are not taken into account. Artzner et al. [1] proposes as an alternative risk measure the Expected Shortfall, defined as the conditional expectation of losses above VaR.

Our approach combines the classical utility maximisation with risk limits in terms of VaR and ES. This leads to control problems under restrictions on uniform versions of VaR or ES, where the risk bound is supposed to be in vigour throughout the duration of the investment. To our knowledge such problems have only been considered in dynamic settings which reduce intrinsically to static problems. Emmer, Klüppelberg and Korn [4] consider a dynamic market, but maximise only the expected wealth at maturity under a downside risk bound at maturity. Basak and Shapiro [2] solve the utility optimisation problem for complete markets with bounded VaR at maturity. Gabih, Gretsch and Wunderlich [3] solve the utility optimisation problem for constant coefficients markets with bounded ES at maturity.

In the present paper we aim now at a truly dynamic portfolio choice of a trader subject to a risk limit specified in terms of VaR or ES. We shall start with Merton's consumption and investment problem for a pricing model driven by Brownian motion with càdlàg drift and volatility coefficients. Such dynamic optimisation problems for standard financial markets have been solved in Karatzas and Shreve [7] by martingale methods. In order to obtain the optimal strategy in "feedback form" basic assumption in [7] on the coefficients is Hölder continuity of a certain order (see e.g. Assumption 8.1, p. 119). In the present paper we use classical optimisation methods from stochastic control. This makes it possible to formulate optimal solutions to Merton's consumption and investment problem in "explicit feedback form" for different power consumption and wealth utility functions. We also weaken the Hölder continuity assumption to càdlàg coefficients satisfying weak integrability conditions.

In a second step we introduce uniform risk limits in terms of VaR and ES into this optimal consumption and investment problem. Our risk measures are specified to represent the required Capital-at-Risk of the institutional investor. The amount of required capital increases with the corresponding loss quantile representing the security of the investment. This quantile is for any specific trader an exogenous variable, which he/she cannot influence. Additionally, each trader can set a specific portfolio's risk limit, which may affect the already exogenously given risk limit of the portfolio. A trader, who has been given a fixed Capital-at-Risk, can now use risk limits for different portfolios categorising the riskiness of his/her portfolios in this way.

It has been observed by Basak and Shapiro [2] that VaR limits only applied at maturity can actually increase the risk. In contrast to this observation, when working with a power utility function and a uniform risk limit throughout the investment horizon, this effect disappears; indeed the optimal strategy for the constrained problem of Theorem 5 given in (3.21) is riskless for sufficiently small risk bound: For a HARA utility function, in order to keep within a sufficiently small risk bound, it is not allowed to invest anything into risky assets at all, but consume everything. This is in contrast to the optimal strategy, when we optimise the linear utility, which recommends to invest everything into risky assets and consume nothing; see (3.12) of Theorem 3.

Within the class of admissible control processes we identify subclasses of controls, which allow for an explicit expression of the optimal strategy. We derive results based on certain utility maximisation strategies, choosing a power utility function for both, the consumption process and the terminal wealth. The literature to utility maximisation is vast; we only mention the books by Karatzas and Shreve [6, 7], Korn [8] and Merton [10]. Usually, utility maximisation is based on concave utility functions. The assumption of concavity models the idea that the infinitesimal utility decreases with increasing wealth. Within the class of power utility functions this corresponds to parameters $\gamma < 1$. The case $\gamma = 1$ corresponds to linear utility functions, meaning that expected utility reduces to expected wealth.

Our paper is organised as follows. In Sect. 2 we formulate the problem. In Sect. 2.1 the Black-Scholes model for the price processes and the parameter restrictions are presented. We also define the necessary quantities like consumption and portfolio processes, also recall the notion of a self-financing portfolio and a trading strategy. Section 2.2 is devoted to the control processes; here also the different classes of controls to be considered later are introduced. The cost functions are defined in Sect. 2.3 and the risk measures in Sect. 2.4. In Sect. 3 all optimisation problems and their solutions are given. Here also the consequences for the trader are discussed. All proofs are summarised in Sect. 4 with a verification theorem postponed to the Appendix.

2 Formulating the Problem

2.1 The Model

We consider a Black-Scholes type financial market consisting of one *riskless bond* and several *risky stocks*. Their respective prices $(S_0(t))_{0\le t\le T}$ and $(S_i(t))_{0\le t\le T}$ for $i = 1, \dots, d$ evolve according to the equations:

$$\begin{cases} \mathrm{d}S_0(t) = r_t S_0(t)\mathrm{d}t, & S_0(0) = 1, \\ \mathrm{d}S_i(t) = S_i(t)\mu_i(t)\mathrm{d}t + S_i(t)\sum_{j=1}^{d}\sigma_{ij}(t)\mathrm{d}W_j(t), & S_i(0) = s_i > 0. \end{cases} \tag{2.1}$$

Here $W_t = (W_1(t), \dots, W_d(t))'$ is a standard d-dimensional Brownian motion; $r_t \in \mathbb{R}$ is the *riskless interest rate*, $\mu_t = (\mu_1(t), \dots, \mu_d(t))' \in \mathbb{R}^d$ is the vector of *stock-appreciation rates* and $\sigma_t = (\sigma_{ij}(t))_{1 \le i,j \le d}$ is the matrix of *stock-volatilities*. We assume that the coefficients r_t, μ_t and σ_t are deterministic functions, which are right continuous with left limits (càdlàg). We also assume that the matrix σ_t is non-singular for Lebesgue-almost all $t \ge 0$.

We denote by $\mathscr{F}_t = \sigma\{W_s, s \le t\}$, $t \ge 0$, the filtration generated by the Brownian motion (augmented by the null sets). Furthermore, $|\cdot|$ denotes the Euclidean norm for vectors and the corresponding matrix norm for matrices. For $(y_t)_{0 \le t \le T}$ square integrable over the fixed interval $[0, T]$ we define $\|y\|_T = (\int_0^T |y_t|^2 \mathrm{d}t)^{1/2}$.

For $t \ge 0$ let $\phi_t \in \mathbb{R}$ denote the amount of investment into bond and

$$\varphi_t = (\varphi_1(t), \dots, \varphi_d(t))' \in \mathbb{R}^d$$

the amount of investment into risky assets. We recall that a *trading strategy* is an $\mathbb{R}^{d+1}$-valued $(\mathscr{F}_t)_{0 \le t \le T}$-progressively measurable process $(\phi_t, \varphi_t)_{0 \le t \le T}$ and that

$$X_t = \phi_t S_0(t) + \sum_{j=1}^{d} \varphi_j(t) S_j(t), \quad t \ge 0,$$

is called the *wealth process*. Moreover, an $(\mathscr{F}_t)_{0 \le t \le T}$-progressively measurable nonnegative process $(c_t)_{0 \le t \le T}$ satisfying for the investment horizon $T > 0$

$$\int_0^T c_t \mathrm{d}t < \infty \quad \text{a.s.}$$

is called *consumption process*.

The trading strategy $((\phi_t, \varphi_t))_{0 \le t \le T}$ and the consumption process $(c_t)_{0 \le t \le T}$ are called *self-financing*, if the wealth process satisfies the following equation

$$X_t = x + \int_0^t \phi_u \mathrm{d}S_0(u) + \sum_{j=1}^{d} \int_0^t \varphi_j(u) \mathrm{d}S_j(u) - \int_0^t c_u \mathrm{d}u, \quad t \ge 0, \tag{2.2}$$

where $x > 0$ is the initial endowment.

In this paper we work with relative quantities, i.e. with the fractions of the wealth process, which are invested into bond and stocks; i.e., we define for $j = 1, \dots, d$

$$\pi_j(t) := \frac{\varphi_j(t) S_j(t)}{\phi_t S_0(t) + \sum_{j=1}^d \varphi_i(t) S_i(t)}, \quad t \ge 0.$$

Then $\pi_t = (\pi_1(t), \dots, \pi_d(t))'$, $t \ge 0$, is called the *portfolio process* and we assume throughout that it is $(\mathscr{F}_t)_{0 \le t \le T}$-progressively measurable. We assume that for the fixed investment horizon $T > 0$

$$\|\pi\|_T^2 := \int_0^T |\pi_t|^2 \mathrm{d}t < \infty \quad \text{a.s.}$$

We also define with $\mathbf{1} = (1, \ldots, 1)' \in \mathbb{R}^d$ the quantities

$$y_t = \sigma_t' \pi_t \quad \text{and} \quad \theta_t = \sigma_t^{-1}(\mu_t - r_t \mathbf{1}), \quad t \geq 0, \tag{2.3}$$

where it suffices that these quantities are defined for Lebesgue-almost all $t \geq 0$. Taking these definitions into account we rewrite (2.2) for X_t as

$$\mathrm{d}X_t = X_t(r_t + y_t'\theta_t)\mathrm{d}t - c_t \mathrm{d}t + X_t y_t' \mathrm{d}W_t, \quad t > 0, \quad X_0 = x > 0. \tag{2.4}$$

This implies in particular that any optimal investment strategy is equal to $\pi_t^* = \sigma_t'^{-1} y_t^*$, where y_t^* is the optimal control process for (2.4). We also require for the investment horizon $T > 0$

$$\|\theta\|_T^2 = \int_0^T |\theta_t|^2 \mathrm{d}t < \infty. \tag{2.5}$$

Besides the already defined Euclidean norm we shall also use for arbitrary $q \geq 1$ the notation $\|f\|_{q,T}$ for the q-norm of (f_t), i.e.

$$\|f\|_{q,T} = \left(\int_0^T |f_t|^q \mathrm{d}ts \right)^{1/q}. \tag{2.6}$$

2.2 *The Control Processes*

Now we introduce the set of control processes $(y_t, c_t)_{0 \leq t \leq T}$. First we choose the consumption process $(c_t)_{0 \leq t \leq T}$ as a proportion of the wealth process; i.e.

$$c_t = v_t X_t,$$

where $(v_t)_{0 \leq t \leq T}$ is a deterministic non-negative function satisfying

$$\int_0^T v_t \mathrm{d}t < \infty.$$

For this consumption we define the *control process* $\varsigma = (\varsigma_t)_{0 \leq t \leq T}$ as $\varsigma_t = (y_t, v_t X_t)$, where $(y_t)_{0 \leq t \leq T}$ is a deterministic function taking values in $\mathbb{R}^d$ such that

$$\|y\|_T^2 = \int_0^T |y_t|^2 \mathrm{d}t < \infty. \tag{2.7}$$

The process $(X_t)_{0 \leq t \leq T}$ is defined by (2.4), which in this case has the following form (to emphasise that the wealth process corresponds to some control process ς we write X^ς)

$$\mathrm{d}X_t^\varsigma = X_t^\varsigma(r_t - v_t + y_t'\theta_t)\mathrm{d}t + X_t^\varsigma y_t' \mathrm{d}W_t, \quad t > 0, \quad X_0^\varsigma = x. \tag{2.8}$$

We denote by $\mathscr{U}$ the set of all such control processes ς.

Note that for every $\varsigma \in \mathscr{U}$, by Itô's formula, (2.8) has solution

$$X_t^{\varsigma} = x e^{R_t - V_t + (y,\theta)_t} \mathscr{E}_t(y), \tag{2.9}$$

where

$$R_t = \int_0^t r_u \mathrm{d}u, \qquad V_t = \int_0^t v_u \mathrm{d}u \quad \text{and} \quad (y,\theta)_t = \int_0^t y_u' \theta_u \mathrm{d}u. \tag{2.10}$$

Moreover, $\mathscr{E}(y)$ denotes the stochastic exponential defined as

$$\mathscr{E}_t(y) = \exp\left(\int_0^t y_u' \mathrm{d}W_u - \frac{1}{2}\int_0^t |y_u|^2 \mathrm{d}u\right) \quad t \geq 0.$$

Therefore, for every $\varsigma \in \mathscr{U}$ the process $(X_t^{\varsigma})_{0\leq t\leq T}$ is positive and continuous.

We consider $\mathscr{U}$ as a first class of control processes for (2.4), for which we can solve the control problem explicitly and interpret its solution. This is due to the fact, as we shall see in Sect. 2.4, that because of the Gaussianity of the log-process we have explicit representations of the risk measures.

It is clear that the behaviour of investors in the model (2.1) depends on the coefficients $(r_t)_{0\leq t\leq T}$, $(\mu_t)_{0\leq t\leq T}$ and $(\sigma_t)_{0\leq t\leq T}$ which in our case are nonrandom known functions and as we will see below (Corollary 3) for the "equilibrate utility functions" case optimal strategies are deterministic, i.e. belong to this class.

A natural generalisation of $\mathscr{U}$ is the following set of controls.

Definition 1 Let $T > 0$ be a fixed investment horizon. A stochastic control process $\varsigma = (\varsigma_t)_{0\leq t\leq T} = ((y_t, c_t))_{0\leq t\leq T}$ is called *admissible* if it is $(\mathscr{F}_t)_{0\leq t\leq T}$-progressively measurable with values in $\mathbb{R}^d \times [0,\infty)$, and (2.4) has a unique strong a.s. positive continuous solution $(X_t^{\varsigma})_{0\leq t\leq T}$ on $[0,T]$. We denote by $\mathscr{V}$ the class of all *admissible control processes*.

2.3 The Cost Functions

We investigate different cost functions, each leading to a different optimal control problem. We assume that the investor wants to optimise expected utility of consumption over the time interval $[0,T]$ and wealth X_T^{ς} at the end of the investment horizon. For initial endowment $x > 0$ and a control process $(\varsigma_t)_{0\leq t\leq T}$ in $\mathscr{V}$, we introduce the *cost function*

$$J(x,\varsigma) := \mathrm{E}_x\left(\int_0^T U(c_t)\mathrm{d}t + h(X_T^{\varsigma})\right),$$

where U and h are *utility functions*. This is a classical approach to the problem; see Karatzas and Shreve [7], Chapter 6.

Here E_x is the expectation operator conditional on $X_0^\varsigma = x$. For both utility functions we choose $U(z) = z^{\gamma_1}$ and $h(z) = z^{\gamma_2}$ for $z \geq 0$ with $0 < \gamma_1, \gamma_2 \leq 1$, corresponding to the cost function

$$J(x,\varsigma) := \mathrm{E}_x\left(\int_0^T c_t^{\gamma_1}\,\mathrm{d}t + (X_T^\varsigma)^{\gamma_2}\right). \tag{2.11}$$

For $\gamma < 1$ the utility function $U(z) = z^\gamma$ is concave and is called a power (or HARA) utility function. We include the case of $\gamma = 1$, which corresponds to simply optimising expected consumption and terminal wealth. In combination with a downside risk bound this allows us in principle to dispense with the utility function, where in practise one has to choose the parameter γ. In the context of this paper it also allows us to separate the effect of the utility function and the risk limit.

2.4 The Downside Risk Measures

As risk measures we use modifications of the Value-at-Risk and the Expected Shortfall as introduced in Emmer, Klüppelberg and Korn [4]. They can be summarised under the notion of Capital-at-Risk and limit the possibility of excess losses over the riskless investment. In this sense they reflect a capital reserve. If the resulting risk measure is negative (which can happen in certain situations) we interpret this as an additional possibility for investment. For further interpretations we refer to [4].

To avoid non-relevant cases we consider only $0 < \alpha < 1/2$.

Definition 2 (Value-at-Risk (VaR)) Define for initial endowment $x > 0$, a control process $\varsigma \in \mathscr{U}$ and $0 < \alpha \leq 1/2$ the *Value-at-Risk* (*VaR*) by

$$\mathrm{VaR}_t(x,\varsigma,\alpha) := x e^{R_t} - \lambda_t, \quad t \geq 0,$$

where $\lambda_t = \lambda_t(x,\varsigma,\alpha)$ is the α-quantile of X_t^ς, i.e.

$$\lambda_t = \inf\{\lambda \geq 0 : \mathrm{P}(X_t^\varsigma \leq \lambda) \geq \alpha\}.$$

Corollary 1 *In the situation of Definition 2, for every $\varsigma \in \mathscr{U}$ the α-quantile λ_t is given by*

$$\lambda_t = x \exp\left(R_t - V_t + (y,\theta)_t - \frac{1}{2}\|y\|_t^2 - |z_\alpha|\,\|y\|_t\right), \quad t \geq 0,$$

where z_α is the α-quantile of the standard normal distribution, and the other quantities are defined in (2.3) *and* (2.10).

We define the *level risk function* for some coefficient $0 < \zeta < 1$ as

$$\zeta_t(x) = \zeta x e^{R_t}, \quad t \in [0,T]. \tag{2.12}$$

We consider only controls $\varsigma \in \mathcal{U}$ for which the Value-at-Risk is bounded by the level function (2.12) over the interval $[0, T]$; i.e. we require

$$\sup_{0\le t\le T} \frac{\mathrm{VaR}_t(x,\varsigma,\alpha)}{\zeta_t(x)} \le 1. \tag{2.13}$$

We have formulated the time-dependent risk bound in the same spirit as we have defined the risk measures, which are based on a comparison of the minimal possible wealth in terms of a low quantile to the pure bond investment. The risk bound now limits the admissible risky strategies to those, whose risk compared to the pure bond portfolio, represented by ζ, remains uniformly bounded over the investment interval.

Our next risk measure is an analogous modification of the *Expected Shortfall* (ES).

Definition 3 (Expected Shortfall (ES)) Define for initial endowment $x > 0$, a control process $\varsigma \in \mathcal{U}$ and $0 < \alpha \le 1/2$

$$m_t(x,\varsigma,\alpha) = \mathrm{E}_x(X_t^{\varsigma} | X_t^{\varsigma} \le \lambda_t), \quad t \ge 0,$$

where $\lambda_t(x,\varsigma,\alpha)$ is the α-quantile of X_t^{ς}. The *Expected Shortfall* (*ES*) is then defined as

$$\mathrm{ES}_t(x,\varsigma,\alpha) = x e^{R_t} - m_t(x,\varsigma,\alpha), \quad t \ge 0.$$

The following result is an analogon of Corollary 1.

Corollary 2 *In the situation of Definition 3, for any $\varsigma \in \mathcal{U}$ the quantity $m_t = m_t(x,\varsigma,\alpha)$ is given by*

$$m_t(x,\varsigma,\alpha) = x F_\alpha(|z_\alpha| + \|y\|_t) e^{R_t + (y,\theta)_t - V_t}, \quad t \ge 0,$$

where z_α is the α-quantile of the standard normal distribution and

$$F_\alpha(z) = \frac{\int_z^\infty e^{-t^2/2}\mathrm{d}t}{\int_{|z_\alpha|}^\infty e^{-t^2/2}\mathrm{d}t}, \quad z \ge 0.$$

We shall consider all controls $\varsigma \in \mathcal{U}$, for which the Expected Shortfall is bounded by the level function (2.12) over the interval $[0, T]$, i.e. we require

$$\sup_{0\le t\le T} \frac{\mathrm{ES}_t(x,\varsigma,\alpha)}{\zeta_t(x)} \le 1. \tag{2.14}$$

Remark 1 (i) The coefficient ζ introduces some risk aversion behaviour into the model. In that sense it acts similarly as a utility function does. The difference, however, is that ζ has a clear interpretation, and every investor can choose and understand the influence of ζ with respect to the corresponding risk measures.

(ii) If $\|y\|_t = 0$ for all $t \in [0, T]$, then $\mathrm{VaR}_t(x, \varsigma, \alpha) = \mathrm{ES}_t(x, \varsigma, \alpha) = xe^{R_t}(1 - e^{-V_t})$, $0 \le t \le T$. On the other hand, if $\|y\|_t > 0$ for $t \in [0, T]$, then

$$\lim_{\alpha \to 0} \mathrm{VaR}_t(x, \varsigma, \alpha) = \lim_{\alpha \to 0} \mathrm{ES}_t(x, \varsigma, \alpha) = xe^{R_t}.$$

This means that the choice of α influences the risk bounds (2.13) and (2.14). Note, however, that α is chosen by the regulatory authorities, not by the investor. The investor only chooses the value ζ. If ζ is near 0 the risk level is rather low, whereas for ζ close to 1 the risk level is rather high, indeed in such case the risk bounds may not be restrictive at all.

3 Problems and Solutions

In the situation of Sect. 2 we are interested in the solutions to different optimisation problems. Throughout we assume a fixed investment horizon $T > 0$.

In the following we first present the solution to the unconstrained problem and then study the constrained problems. The constraints are in terms of risk bounds with respect to downfall risks like VaR and ES defined by means of a quantile.

3.1 The Unconstrained Problem

We consider two regimes with cost functions (2.11) for $0 < \gamma_1, \gamma_2 < 1$ and for $\gamma_1 = \gamma_2 = 1$. We include the case of $\gamma_1 = \gamma_2 = 1$ for further referencing, although it makes economically not much sense without a risk constraint. The mathematical treatment of the two cases is completely different by nature.

Problem 1

$$\max_{\varsigma \in \mathcal{V}} J(x, \varsigma).$$

Theorem 1 *Consider Problem 1 with $\gamma_1 = \gamma_2 = 1$. Assume a riskless interest rate $r_t \ge 0$ for all $t \in [0, T]$. If $\|\theta\|_T > 0$, then*

$$\max_{\varsigma \in \mathcal{U}} J(x, \varsigma) = \infty.$$

If $\|\theta\|_T = 0$, then a solution exists and the optimal value of $J(x, \varsigma)$ is given by

$$\max_{\varsigma \in \mathcal{U}} J(x, \varsigma) = J(x, \varsigma^*) = xe^{R_T},$$

corresponding to the optimal control $\varsigma_t^* = (y_t^*, 0)$ for all $0 \le t \le T$ with arbitrary deterministic square integrable function $(y_t^*)_{0 \le t \le T}$. In this case the optimal wealth process $(X_t^*)_{0 \le t \le T}$ satisfies the following equation

$$\mathrm{d}X_t^* = X_t^* r_t \mathrm{d}t + X_t^* (y_t^*)' \mathrm{d}W_t, \qquad X_0^* = x. \tag{3.1}$$

Consider now Problem 1 for $0 < \gamma_1, \gamma_2 < 1$. To formulate the solution we define functions

$$A_1(t) = \gamma_1^{q_1} \int_t^T e^{\int_t^s \beta_1(u)\mathrm{d}u} \mathrm{d}s \quad \text{and} \quad A_2(t) = \gamma_2^{q_2} e^{\int_t^T \beta_2(u)\mathrm{d}u}, \quad 0 \le t \le T, \tag{3.2}$$

where $q_i = (1-\gamma_i)^{-1}$ and $\beta_i(t) = (q_i - 1)(r_t + \frac{q_i}{2}|\theta_t|^2)$. Moreover, for all $0 \le t \le T$ and $x > 0$ we define the function $g(t,x) > 0$ as solution to

$$A_1(t) g^{-q_1}(t,x) + A_2(t) g^{-q_2}(t,x) = x \tag{3.3}$$

and

$$p(t,x) = q_1 A_1(t) g^{-q_1}(t,x) + q_2 A_2(t) g^{-q_2}(t,x).$$

Theorem 2 *Consider Problem 1 for $0 < \gamma_1, \gamma_2 < 1$. The optimal value of $J(x,\varsigma)$ is given by*

$$\max_{\varsigma \in \mathcal{V}} J(x,\varsigma) = J(x,\varsigma^*) = \frac{A_1(0)}{\gamma_1} g^{1-q_1}(0,x) + \frac{A_2(0)}{\gamma_2} g^{1-q_2}(0,x),$$

where the optimal control $\varsigma^* = (y^*, c^*)$ is for all $0 \le t \le T$ of the form

$$\begin{cases} y_t^* = \frac{p(t,X_t^*)}{X_t^*}\theta_t \quad (\pi_t^* = \frac{p(t,X_t^*)}{X_t^*}(\sigma_t\sigma_t')^{-1}(\mu_t - r_t\mathbf{1})); \\ c_t^* = (\frac{\gamma_1}{g(t,X_t^*)})^{q_1}. \end{cases} \tag{3.4}$$

The optimal wealth process $(X_t^*)_{0\le t\le T}$ is the solution to

$$\mathrm{d}X_t^* = a^*(t, X_t^*)\mathrm{d}t + (b^*(t, X_t^*))'\mathrm{d}W_t, \qquad X_0^* = x, \tag{3.5}$$

where

$$a^*(t,x) = r_t x + p(t,x)|\theta_t|^2 - \left(\frac{\gamma_1}{g(t,x)}\right)^{q_1} \quad \textit{and} \quad b^*(t,x) = p(t,x)\theta_t.$$

The following result can be found Example 6.7 on p. 106 in Karatzas and Shreve [7]; its proof here is based on the martingale method.

Corollary 3 *Consider Problem 1 for $\gamma_1 = \gamma_2 = \gamma \in (0,1)$ and define*

$$\widetilde{g}_\gamma(t) = \exp\left(\gamma R_t + \frac{q-1}{2}\|\theta\|_t^2\right) \quad \textit{and} \quad q = \frac{1}{1-\gamma}. \tag{3.6}$$

Then the optimal value of $J(x,\varsigma)$ is given by

$$J^*(x) = \max_{\varsigma\in\mathcal{V}} J(x,\varsigma) = J(x,\varsigma^*) = x^\gamma (\|\widetilde{g}_\gamma\|_{q,T}^q + \widetilde{g}_\gamma^q(T))^{1/q},$$

where the optimal control $\varsigma^* = (y^*, c^*)$ *is for all* $0 \le t \le T$ *of the form*

$$\begin{cases} y_t^* = \frac{\theta_t}{1-\gamma} \quad (\pi_t^* = \frac{(\sigma_t\sigma_t')^{-1}(\mu_t - r_t\mathbf{1})}{1-\gamma}); \\ c_t^* = v_t^* X_t^* \quad \textit{and} \quad v_t^* = \frac{\widetilde{g}_\gamma^q(t)}{\widetilde{g}_\gamma^q(T) + \int_t^T \widetilde{g}_\gamma^q(s)\mathrm{d}s}. \end{cases} \tag{3.7}$$

The optimal wealth process $(X_t^*)_{0\le t\le T}$ *is given by*

$$\mathrm{d}X_t^* = X_t^*\left(r_t - v_t^* + \frac{|\theta_t|^2}{1-\gamma}\right)\mathrm{d}t + X_t^*\frac{\theta_t'}{1-\gamma}\mathrm{d}W_t, \qquad X_0^* = x. \tag{3.8}$$

Remark 2 Note that Problem 1 for different $0 < \gamma_1 < 1$ and $0 < \gamma_2 < 1$ was also investigated by Karatzas and Shreve [7]. For Hölder continuous market coefficients they find by the martingale method an implicit "feedback form" of the optimal solution in their Theorem 8.8. In contrast, Theorem 2 above gives the optimal solution in "explicit feedback form" for quite general market coefficients. Our proof is based on a special version of a verification theorem for stochastic optimal control problems, which allows for càdlàg coefficients.

3.2 Value-at-Risk as Risk Measure

For the Value-at-Risk we consider again the cost function (2.11) and, as before, we consider different regimes for $0 < \gamma_1, \gamma_2 < 1$ and $\gamma_1, \gamma_2 = 1$.

Problem 2

$$\max_{\varsigma\in\mathcal{U}} J(x,\varsigma) \quad \text{subject to} \quad \sup_{0\le t\le T} \frac{\mathrm{VaR}_t(x,\varsigma,\alpha)}{\zeta_t(x)} \le 1.$$

To formulate the solution let z_α be the normal α-quantile for $0 < \alpha \le 1/2$ and the constant $\zeta \in (0, 1)$ as in (2.12). Obviously, for $\alpha \to 0$ we have $|z_\alpha| \to \infty$ and, hence, the quotient in (2.13) tends to $1/\zeta > 1$. This means that the bound can be restrictive. We define for θ as in (2.3) the following quantity

$$\rho_{\mathrm{VaR}}^* = \sqrt{(|z_\alpha| - \|\theta\|_T)^2 - 2\ln(1-\zeta)} - (|z_\alpha| - \|\theta\|_T). \tag{3.9}$$

Theorem 3 *Consider Problem 2 for* $\gamma_1 = \gamma_2 = 1$. *Assume a riskless interest rate* $r_t \ge 0$ *for all* $t \in [0, T]$. *Then for*

$$\max(0, 1 - e^{z_\alpha^2/2 - |z_\alpha|\|\theta\|_T}) < \zeta < 1 \tag{3.10}$$

the optimal value of $J(x,\varsigma)$ *is given by*

$$\max_{\varsigma\in\mathcal{U}} J(x,\varsigma) = J(x,\varsigma^*) = x e^{\rho_{\mathrm{VaR}}^*\|\theta\|_T + R_T}. \tag{3.11}$$

If $\|\theta\|_T > 0$, *then the optimal control* $\varsigma^* = (y^*, v^* X^*)$ *is for all* $0 \le t \le T$ *of the form*

$$y_t^* = \rho_{\mathrm{VaR}}^* \frac{\theta_t}{\|\theta\|_T} \quad \left(\pi_t^* = \rho_{\mathrm{VaR}}^* \frac{(\sigma_t \sigma_t')^{-1}}{\|\theta\|_T}(\mu_t - r_t \mathbf{1})\right) \quad \textit{and} \quad v_t^* = 0. \tag{3.12}$$

The optimal wealth process $(X_t^*)_{0 \le t \le T}$ *is given by*

$$\mathrm{d}X_t^* = X_t^* \left(r_t + \rho_{\mathrm{VaR}}^* \frac{|\theta_t|^2}{\|\theta\|_T} \right) \mathrm{d}t + X_t^* \rho_{\mathrm{VaR}}^* \frac{\theta_t'}{\|\theta\|_T} \mathrm{d}W_t, \qquad X_0^* = x.$$

If $\|\theta\|_T = 0$, *then the optimal value of* $J(x, \varsigma)$ *is given by*

$$\max_{\varsigma \in \mathscr{U}} J(x, \varsigma) = J(x, \varsigma^*) = x e^{R_T}, \tag{3.13}$$

corresponding to the optimal control $\varsigma_t^* = (y_t^*, 0)$ *for* $0 \le t \le T$ *with arbitrary deterministic function* $(y_t^*)_{0 \le t \le T}$ *such that*

$$\|y^*\|_T \le \rho_{\mathrm{VaR}}^* = \sqrt{z_\alpha^2 - 2\ln(1 - \zeta)} - |z_\alpha|.$$

In this case the optimal wealth process $(X_t^*)_{0 \le t \le T}$ *satisfies* (3.1).

Remark 3 (i) For $|z_\alpha| \ge 2\|\theta\|_T$ condition (3.10) gives a lower bound 0; i.e. $0 < \zeta < 1$. If $|z_\alpha| < 2\|\theta\|_T$, then condition (3.10) translates to

$$1 - e^{z_\alpha^2/2 - |z_\alpha| \|\theta\|_T} < \zeta < 1;$$

i.e. we obtain a positive lower bound.

(ii) The optimal strategy implies that there will be no consumption throughout the investment horizon. This is due to the fact that the wealth we expect by investment is so attractive that we continue to invest everything. Note that the solution is the same as the solution to the problem without possible consumption.

Now we present a sufficient condition for which the optimal unconstrained strategy (3.7)–(3.8) is solution for Problem 2 in the case $\gamma_1 = \gamma_2 = \gamma \in (0, 1)$. For this we introduce the following functions:

$$\widetilde{\kappa}(\gamma) = \frac{\|\widetilde{g}_\gamma\|_{q,T}^q}{\|\widetilde{g}_\gamma\|_{q,T}^q + \widetilde{g}_\gamma^q(T)} = 1 - e^{-V_T^*} = 1 - e^{-\int_0^T v_t^* \mathrm{d}t},$$

where $(v_t^*)_{0 \le t \le T}$ is the optimal consumption rate introduced in (3.7). By setting $\widetilde{l}(\gamma) = \ln(1 - \widetilde{\kappa}(\gamma))$ we define

$$l_*(\gamma) = \begin{cases} -q\|\theta\|_T |z_\alpha| + \widetilde{l}(\gamma) & \text{for } 0 < \gamma \le 1/2; \\ -q\|\theta\|_T |z_\alpha| + \widetilde{l}(\gamma) - \frac{q(q-2)}{2}\|\theta\|_T^2 & \text{for } 1/2 < \gamma < 1. \end{cases}$$

Theorem 4 *Consider Problem 2 with $\gamma_1 = \gamma_2 = \gamma \in (0,1)$. Assume a riskless interest rate $r_t \geq 0$ for all $t \in [0,T]$ and*

$$1 - e^{l_*(\gamma)} \leq \zeta < 1. \tag{3.14}$$

Then the optimal solution is given by (3.7)–(3.8); i.e. it is equal to the solution of the unconstrained problem.

Remark 4 Theorem 4 does not hold for $\gamma_1 \neq \gamma_2$, since the solution (3.4) does not belong to $\mathscr{U}$.

To formulate the result for different γ_i $(i = 1,2)$ we introduce the following function for $0 \leq \kappa \leq 1$

$$G(x,\kappa) := x^{\gamma_1}\kappa^{\gamma_1}\|\widehat{g}_1\|_{q,T} + x^{\gamma_2}(1-\kappa)^{\gamma_2}\widehat{g}_2(T), \quad x > 0, \tag{3.15}$$

where $q = (1-\gamma_1)^{-1}$, $\widehat{g}_i = \widehat{g}_{\gamma_i}$ and

$$\widehat{g}_\gamma = e^{\gamma R_t} = e^{\gamma \int_0^t r_u \mathrm{d}u}.$$

Moreover, for $x > 0$ we set

$$\kappa_*(x) = \arg\max_{0\leq\kappa\leq 1} G(x,\kappa). \tag{3.16}$$

Note that for $0 < \gamma_1 < 1$ and $0 < \gamma_2 \leq 1$ this function is strictly positive for all $x > 0$; i.e. $0 < \kappa_*(x) \leq 1$. It is easy to see that in the case $\gamma_1 = \gamma_2 =: \gamma$ the function $\kappa_*(x)$ is independent of x and equals to

$$\widehat{\kappa}(\gamma) = \frac{\|\widehat{g}_\gamma\|^q_{q,T}}{\|\widehat{g}_\gamma\|^q_{q,T} + \widehat{g}^q_\gamma(T)}. \tag{3.17}$$

Theorem 5 *Consider Problem 2 with $0 < \gamma_1 < 1$ and $0 < \gamma_2 \leq 1$. Assume a riskless interest rate $r_t \geq 0$ for all $t \in [0,T]$ and*

$$0 < \zeta < \min\{\kappa_*(x), \widehat{\kappa}(\gamma_1)\}. \tag{3.18}$$

Moreover, assume that

$$|z_\alpha| \geq \left(1 + \frac{\max\{\gamma_1,\gamma_2\}}{1-\zeta}\frac{1}{\frac{\partial}{\partial\zeta}\ln G(x,\zeta)}\right)\|\theta\|_T. \tag{3.19}$$

Then the optimal value of $J(x,\varsigma)$ is given by

$$\max_{\varsigma\in\mathscr{U}} J(x,\varsigma) = J(x,\varsigma^*) = x^{\gamma_1}\zeta^{\gamma_1}\|\widehat{g}_1\|_{q,T} + x^{\gamma_2}(1-\zeta)^{\gamma_2}\widehat{g}_2(T), \tag{3.20}$$

where the optimal control $\varsigma^* = (y^*, v^* X^*)$ *is for all* $0 \le t \le T$ *of the form*

$$y_t^* = 0 \quad (\pi_t^* = 0) \quad \textit{and} \quad v_t^* = \frac{\zeta \widehat{g}_1^q(t)}{\|\widehat{g}_1\|_{q,T}^q - \zeta \|\widehat{g}_1\|_{q,t}^q}. \tag{3.21}$$

The optimal wealth process $(X_t^*)_{0 \le t \le T}$ *is given by the deterministic function*

$$X_t^* = x e^{R_t} \frac{\|\widehat{g}_1\|_{q,T}^q - \zeta \|\widehat{g}_1\|_{q,t}^q}{\|\widehat{g}_1\|_{q,T}^q} = x \frac{\zeta}{v_t^*} e^{R_t}, \quad 0 \le t \le T. \tag{3.22}$$

Remark 5 We compare now conditions (3.18)–(3.19) for $\gamma_1 = \gamma_2 = \gamma \in (0, 1)$ with condition (3.14). Making use of the notation in (3.6) we obtain

$$\widetilde{g}_\gamma(t) = \widehat{g}_\gamma(t) e^{\frac{q-1}{2}\|\theta\|_t^2} \ge \widehat{g}_\gamma(t).$$

Taking this inequality into account we find that in the case $0 < \gamma \le 1/2$ (i.e. $1 < q \le 2$), the function $e^{l_*(\gamma)}$ is bounded above by

$$e^{l_*(\gamma)} = \frac{\widetilde{g}_\gamma^q(T) e^{-q\|\theta\|_T |z_\alpha|}}{\|\widetilde{g}_\gamma\|_{q,T}^q + \widetilde{g}_\gamma^q(T)} \le \frac{\widehat{g}_\gamma^q(T) e^{-q(\|\theta\|_T |z_\alpha| - \frac{q-1}{2}\|\theta\|_T^2)}}{\|\widehat{g}\|_{q,T}^q + \widehat{g}_\gamma^q(T)}.$$

Moreover, condition (3.19) implies $|z_\alpha| \ge \|\theta\|_T$. Therefore, taking into account that $1 < q \le 2$ we obtain

$$e^{-q(\|\theta\|_T |z_\alpha| - \frac{q-1}{2}\|\theta\|_T^2)} \le 1.$$

Hence,

$$e^{l_*(\gamma)} \le \frac{\widehat{g}_\gamma^q(T)}{\|\widehat{g}\|_{q,T}^q + \widehat{g}_\gamma^q(T)} = 1 - \widehat{\kappa}(\gamma).$$

Similarly, for $1/2 < \gamma < 1$ (i.e. $q > 2$),

$$e^{l_*(\gamma)} \le \frac{\widehat{g}_\gamma^q(T) e^{-\frac{q}{2}\|\theta\|_T^2}}{\|\widehat{g}\|_{q,T}^q + \widehat{g}_\gamma^q(T)} \le 1 - \widehat{\kappa}(\gamma).$$

So we have shown that $1 - e^{l_*(\gamma)} \ge \widehat{\kappa}(\gamma)$, i.e. condition (3.14) is complementary to conditions (3.18)–(3.19).

We present an example for further illustration.

Example 1 To clarify conditions (3.18)–(3.19) consider again $\gamma_1 = \gamma_2 = \gamma \in (0, 1)$ and $r_t \equiv r > 0$. We shall investigate what happens for $T \to \infty$. First we calculate

$$\kappa_*(x) = \widehat{\kappa}(\gamma) = \frac{\int_0^T e^{q\gamma r t} \mathrm{d}t}{\int_0^T e^{q\gamma r t} \mathrm{d}t + e^{q\gamma r T}} = \frac{1 - e^{-q\gamma r T}}{1 + q\gamma r - e^{-q\gamma r T}} \sim \frac{1}{1 + q\gamma r}$$

as $T \to \infty$, where $q = (1-\gamma)^{-1}$. Thus, condition (3.18) yields for $T \to \infty$ approximately

$$0 < \zeta < \frac{1}{1 + q\gamma r}.$$

The function (3.15) has the following form

$$G(x,\kappa) = x^\gamma e^{\gamma r T}(\kappa^\gamma A(T) + (1-\kappa)^\gamma) \quad \text{with } A(T) = \left(\int_0^T e^{-q\gamma r t} \mathrm{d}t\right)^{1/q}.$$

For the partial derivative with respect to ζ we calculate

$$\frac{\partial}{\partial \zeta} \ln G(x,\zeta) = \gamma \frac{\zeta^{\gamma-1} A(T) - (1-\zeta)^{\gamma-1}}{\zeta^\gamma A(T) + (1-\zeta)^\gamma}.$$

Since

$$\begin{aligned} &\frac{\max\{\gamma_1,\gamma_2\}}{1-\zeta} \frac{1}{\frac{\partial}{\partial \zeta}\ln G(x,\zeta)} \\ &\quad = \frac{\zeta^{\gamma+1} A(T) + \zeta(1-\zeta)^\gamma}{\zeta^\gamma (1-\zeta) A(T) - \zeta(1-\zeta)^\gamma} = O(\zeta) \quad \text{as } \zeta \to 0, \end{aligned}$$

condition (3.19) implies $|z_\alpha| > \|\theta\|_T$ approximately for $\zeta \to 0$. Moreover, the optimal consumption (3.21) is given by

$$v_t^* = \zeta \frac{\gamma q r}{e^{\gamma q r (T-t)} - \zeta - (1-\zeta) e^{-\gamma q r t}}$$

and the optimal wealth process (3.22) is

$$X_t^* = x \frac{e^{rt}}{\gamma q r}(e^{\gamma q r (T-t)} - \zeta - (1-\zeta) e^{-\gamma q r t}), \quad 0 \le t \le T.$$

Conclusion 6 The preceding results allow us to compare the optimal strategies of the unconstrained problems and the constrained problems with VaR bound. We consider a riskless interest rate $r_t \ge 0$ for all $t \in [0, T]$.

When simply optimising expectation, i.e. $\gamma_1 = \gamma_2 = 1$, the VaR constrain puts a limit to the investment strategy and also influences the optimum wealth. On the other hand, there is no change in the consumption, which is zero throughout the investment horizon in both cases.

For $0 < \gamma_1, \gamma_2 \le 1$ the optimal strategy for the utility maximisation problem involves investment and consumption during the investment horizon; cf. Theorem 3. The influence of a VaR bound is dramatic, when it is valid, as it recommends the optimal strategy of no investment, but consumption only; cf. Theorem 5.

3.3 Expected Shortfall as Risk Measure

The next problems concern bounds on the Expected Shortfall.

Problem 3

$$\max_{\varsigma \in \mathcal{U}} J(x,\varsigma) \quad \text{subject to} \quad \sup_{0\le t\le T} \frac{ES_t(x,\varsigma,\alpha)}{\zeta_t(x)} \le 1.$$

To formulate the solution for Problem 3 we define for $\rho \ge 0$ and $0 \le u \le 1$

$$\psi(\rho,u) = \|\theta\|_T \rho u^2 + \ln F_\alpha(|z_\alpha| + \rho u). \tag{3.23}$$

Moreover, we set

$$\rho^*_{\mathrm{ES}} = \sup\{\rho > 0 : \psi(\rho,1) \ge \ln(1-\zeta)\}, \tag{3.24}$$

where we define $\sup\{\emptyset\} = \infty$. We formulate some properties of ψ which will help us to calculate ρ^*_{ES}.

Lemma 1 *Let $0 < \alpha < 1/2$ such that $|z_\alpha| \ge 2\|\theta\|_T$. Then ψ satisfies the following properties.*

(1) *For every $\rho > 0$ the function $\psi(\rho,u)$ is strictly decreasing for $0 \le u \le 1$.*
(2) *The function $\psi(\cdot,1)$ is strictly decreasing.*
(3) *For every $a \le 0$ the equation $\psi(\rho,1) = a$ has a unique positive solution. The equation $\psi(\rho,1) = \ln(1-\zeta)$ has solution ρ^*_{ES} as defined in (3.24). For $|z_\alpha| > 1$ we have*

$$\rho^*_{\mathrm{ES}} \le \frac{-\ln(1-z_\alpha^{-2}) - \ln(1-\zeta)}{|z_\alpha| - \|\theta\|_T}. \tag{3.25}$$

Now we present the solution of Problem 3, where we start again with the situation of a small α, where the risk bound is restrictive.

Theorem 7 *Consider Problem 3 for $\gamma_1 = \gamma_2 = 1$. Assume also that the riskless interest rate $r_t \ge 0$ for all $t \in [0,T]$. Then for every $0 < \zeta < 1$ and for $0 < \alpha < 1/2$ such that $|z_\alpha| \ge 2\|\theta\|_T$ the solution ρ^*_{ES} of $\psi(\rho,1) = \ln(1-\zeta)$ is finite, and the optimal solution is given by (3.12) after replacing ρ^*_{VaR} by ρ^*_{ES}.*

Now we consider Problem 3 with $\gamma_1 = \gamma_2 = \gamma \in (0,1)$. Our next theorem concerns the case of a loose risk bound, where the solution is the same as in the unconstrained case.

Theorem 8 *Consider Problem 3 for $\gamma_1 = \gamma_2 = \gamma \in (0,1)$. Assume that the riskless interest rate $r_t \ge 0$ for all $t \in [0,T]$. Assume also that $|z_\alpha| \ge 2\|\theta\|_T$ and*

$$1 - (1-\widetilde{\kappa}(\gamma)) e^{q\|\theta\|_T^2} F_\alpha(|z_\alpha| + q\|\theta\|_T) \le \zeta < 1. \tag{3.26}$$

Then the optimal solution ς^* *is given by* (3.7)–(3.8); *i.e. it is equal to the solution of the unconstrained problem.*

Now we turn to the general case of $0 < \gamma_1, \gamma_2 \leq 1$, the analogon of Theorem 5.

Theorem 9 *Consider Problem* 3 *for* $0 < \gamma_1 < 1$ *and* $0 < \gamma_2 \leq 1$. *Assume a riskless interest rate* $r_t \geq 0$ *for all* $t \in [0, T]$. *Take* $\kappa_*(x)$ *as in* (3.16). *Assume* (3.18) *and*

$$|z_\alpha| \geq \left(2 + \frac{\max\{\gamma_1, \gamma_2\}}{1-\zeta} \frac{1}{\frac{\partial}{\partial \zeta} \ln G(x, \zeta)}\right) \|\theta\|_T. \tag{3.27}$$

Then the optimal solution ς^* *is given by* (3.21)–(3.22).

Remark 6 For $|z_\alpha| \geq 2\|\theta\|_T$ we calculate

$$F_\alpha(|z_\alpha| + q\|\theta\|_T) = \frac{\int_{|z_\alpha|}^\infty \exp(-\frac{(t+q\|\theta\|_T)^2}{2})\mathrm{d}t}{\int_{|z_\alpha|}^\infty e^{-\frac{t^2}{2}}\mathrm{d}t} \leq \exp\left(-2q\|\theta\|_T^2 - \frac{q^2\|\theta\|_T^2}{2}\right).$$

Recalling from Remark 5 that $\widetilde{g}_\gamma(t) = \widehat{g}_\gamma(t) e^{\frac{q-1}{2}\|\theta\|_t^2}$ we obtain

$$\begin{aligned}(1-\widetilde{\kappa}(\gamma))e^{q\|\theta\|_T^2} F_\alpha(|z_\alpha| + q\|\theta\|_T) &\leq \frac{\widetilde{g}_\gamma^q(T) e^{-\frac{q(q+4)}{2}\|\theta\|_T^2}}{\|\widetilde{g}_\gamma\|_{q,T}^q + \widetilde{g}_\gamma^q(T)} \\ &\leq \frac{\widehat{g}_\gamma^q(T)}{\|\widehat{g}\|_{q,T}^q + \widehat{g}_\gamma^q(T)} e^{-\frac{5q}{2}\|\theta\|_T^2} \leq 1 - \widehat{\kappa}(\gamma),\end{aligned}$$

i.e. condition (3.26) is complementary to condition (3.18).

Remark 7 (i) It should be noted that the optimal solution (3.21)–(3.22) for Problems 2 and 3 does not depend on the coefficients $(\mu_t)_{0\leq t\leq T}$ and $(\sigma_t)_{0\leq t\leq T}$ of the stock price. These parameters only enter into (3.18), (3.19) and (3.27). Consequently, in practice it is not necessary to know these parameters precisely, an upper bound for $\|\theta\|_T$ suffices.

(ii) If $\theta \equiv 0$, then conditions (3.19) and (3.27) are trivial, i.e. the optimal solutions for Problems 2 and 3 for $0 < \gamma_1 < 1$ and $0 < \gamma_2 \leq 1$ are given by (3.21)–(3.22) for every $0 < \alpha < 1/2$ and ζ satisfying (3.18).

Conclusion 10 The preceding results again allow us to compare the optimal strategies of the utility maximisation problems and the constrained problems with ES bound. The structures of the solutions are the same as for a VaR constrain, only certain values have changed.

4 Proofs

4.1 Proof of Theorem 1

First we consider $\|\theta\|_T > 0$. Define for $n \in \mathbb{N}$ the sequence of strategies $\varsigma^{(n)} = (y^{(n)}, v^{(n)} X^{(n)})$ for which $v^{(n)} = 0$ and $y^{(n)} = n\theta$. For this strategy (2.9) implies

$$J(x, \varsigma^{(n)}) = x e^{R_T + n\|\theta\|_T} \to \infty \quad \text{as } n \to \infty.$$

Let now $\|\theta\|_T = 0$. Then the cost function can be estimated above by

$$\begin{aligned} J(x, \varsigma) &= x\left(\int_0^T e^{R_t - V_t} v_t \mathrm{d}t + e^{R_T - V_T}\right) \\ &\le x e^{R_T}\left(\int_0^T e^{-V_t} v_t \mathrm{d}t + e^{-V_T}\right) \\ &= x e^{R_T}. \end{aligned}$$

Thus, every control ς with $v = 0$ matches this upper bound.

4.2 Proof of Theorem 2

We apply the Verification Theorem 11 to Problem 1 for the stochastic control differential equation (2.4). For fixed $\vartheta = (y, c)$, where $y \in \mathbb{R}^d$ and $c \in [0, \infty)$, the coefficients in model (5.26) are defined as

$$\begin{aligned} &a(t, x, \vartheta) = x(r_t + y'\theta_t) - c, \\ &b(t, x, \vartheta) = x|y|, \qquad f(t, x, \vartheta) = c^{\gamma_1}, \qquad h(x) = x^{\gamma_2}, \quad 0 < \gamma_1, \gamma_2 < 1. \end{aligned}$$

This implies immediately $\mathbf{H}_1$. Moreover, by Definition 1 the coefficients are continuous, hence (5.27) holds for every $\varsigma \in \mathcal{V}$.

To check $\mathbf{H}_1$–$\mathbf{H}_3$ we calculate the Hamilton function (5.29) for Problem 1. We have

$$H(t, x, z_1, z_2) = \sup_{\vartheta \in \mathbb{R}^d \times [0,\infty)} H_0(t, x, z_1, z_2, \vartheta),$$

where

$$H_0(t, x, z_1, z_2, \vartheta) = (r_t + y'\theta_t) x z_1 + \frac{1}{2} x^2 |y|^2 z_2 + c^{\gamma_1} - c z_1.$$

For $z_2 \le 0$ we find (recall that $q_i = (1 - \gamma_i)^{-1}$)

$$H(t, x, z_1, z_2) = H_0(t, x, z_1, z_2, \vartheta_0) = r_t x z_1 + \frac{1}{2|z_2|} z_1^2 |\theta_t|^2 + \frac{1}{q_1}\left(\frac{\gamma_1}{z_1}\right)^{q_1 - 1},$$

where $\vartheta_0 = \vartheta_0(t,x,z_1,z_2) = (y_0(t,x,z_1,z_2), c_0(t,x,z_1,z_2))$ with

$$y_0(t,x,z_1,z_2) = \frac{z_1}{x|z_2|}\theta_t \quad \text{and} \quad c_0(t,x,z_1,z_2) = \left(\frac{\gamma_1}{z_1}\right)^{q_1}. \tag{4.1}$$

Now we solve the HJB equation (5.30), which has for our problem the following form:

$$\begin{cases} z_t(t,x) + r_t x z_x(t,x) + \frac{z_x^2(t,x)|\theta_t|^2}{2|z_{xx}(t,x)|} + \frac{1}{q_1}(\frac{\gamma_1}{z_x(t,x)})^{q_1-1} = 0, \\ z(T,x) = x^{\gamma_2}. \end{cases} \tag{4.2}$$

We make the following ansatz:

$$z(t,x) = \frac{A_1(t)}{\gamma_1} g^{1-q_1}(t,x) + \frac{A_2(t)}{\gamma_2} g^{1-q_2}(t,x), \tag{4.3}$$

where the function g is defined in (3.3). One can now prove directly that this function satisfies (4.2) using the following properties of g

$$(-A_1(t)q_1 g^{-q_1} - A_2(t)q_2 g^{-q_2})\frac{\partial}{\partial x}g(t,x) = g(t,x),$$

$$\dot{A}_1(t)g^{-q_1}(t,x) + \dot{A}_2(t)g^{-q_2}(t,x)$$
$$- A_1(t)q_1 g^{-q_1-1}\frac{\partial}{\partial t}g(t,x) - A_2(t)q_2 g^{-q_2-1}\frac{\partial}{\partial t}g(t,x) = 0$$
$$\dot{A}_1(t)g^{-q_1}(t,x) + \dot{A}_2(t)g^{-q_2}(t,x) + \frac{1}{\frac{\partial}{\partial x}g(t,x)}\frac{\partial}{\partial t}g(t,x) = 0.$$

This implies that

$$z_t(t,x) = -\frac{\dot{A}_1(t)}{1-q_1}g^{1-q_1}(t,x) - \frac{\dot{A}_2(t)}{1-q_2}g^{1-q_2}(t,x). \tag{4.4}$$

Moreover, $z_x(t,x) = g(t,x)$ and $z_{xx}(t,x) = -g(t,x)/p(t,x)$. Equation (4.2) implies the following differential equations for the coefficients A_i:

$$\begin{cases} \dot{A}_1(t) = -\beta_1(t)A_1(t) - \gamma_1^{q_1}, \qquad A_1(T) = 0, \\ \dot{A}_2(t) = -\beta_2(t)A_2(t), \qquad A_2(T) = \gamma_2^{q_2}. \end{cases} \tag{4.5}$$

The solution of this system is given by the functions (3.2) in all points of continuity of $(\beta_i(t))_{0\le t\le T}$. We denote this set Γ. By our conditions (all coefficients in the model (2.1) are càdlàg functions) the Lebesgue measure of Γ is equal to T. Note that conditions (2.5) and (4.5) imply that

$$\int_0^T |\dot{A}_i(t)|\mathrm{d}t < \infty$$

for $i=1,2$. Moreover, the definition of $g(t,x)$ in (3.3) implies that $g(\cdot,\cdot)$ is continuous on $[0,T]\times(0,\infty)$. Invoking (4.4) we obtain property (5.32). Hence condition $\mathbf{H}_2$ holds.

Now by (4.1) we find that

$$H(t,x,z_x(t,x),z_{xx}(t,x)) = H_0(t,x,z_x(t,x),z_{xx}(t,x),\vartheta^*(t,x)),$$

where $\vartheta^*(t,x)=(y^*(t,x),c^*(t,x))$ with

$$y^*(t,x)=\frac{p(t,x)}{x}\theta_t \quad \text{and} \quad c^*(t,x)=\left(\frac{\gamma_1}{g(t,x)}\right)^{q_1}.$$

Hence $\mathbf{H}_2$ holds.

Now we check condition $\mathbf{H}_3$. First note that (5.33) is identical to (3.5). By Itô's formula one can show that this equation has a unique strong positive solution given by

$$X_t^* = A_1(t)g^{-q_1}(0,x)e^{-q_1\xi_t} + A_2(t)g^{-q_2}(0,x)e^{-q_2\xi_t} \tag{4.6}$$

with

$$\xi_t = -\int_0^t \left(r_u + \frac{1}{2}|\theta_u|^2\right)\mathrm{d}u - \int_0^t \theta_u'\,\mathrm{d}W_u.$$

This implies $\mathbf{H}_3$.

To check the final condition $\mathbf{H}_4$ note that by definitions (3.3) and (4.6)

$$g(t,X_t^*) = g(0,x)e^{\xi_t}.$$

Therefore, taking into account that

$$X_s^* = A_1(s)g^{-q_1}(s,X_s^*) + A_2(s)g^{-q_2}(s,X_s^*)$$

we obtain for $s\ge t$

$$X_s^* = A_1(s)g^{-q_1}(t,X_t^*)e^{-q_1(\xi_s-\xi_t)} + A_2(s)g^{-q_2}(t,X_t^*)e^{-q_2(\xi_s-\xi_t)}.$$

Hence, for $s\ge t$ we can find an upper bound of the process $z(s,X_s^*)$ given by

$$z(s,X_s^*) \le \frac{g(t,X_t^*)}{\min(\gamma_1,\gamma_2)} e^{\xi_s-\xi_t} X_s^* \le M_*(X_t^*)(e^{(1-q_1)(\xi_s-\xi_t)} + e^{(1-q_2)(\xi_s-\xi_t)}),$$

where

$$M_*(x) = \frac{\sup_{0\le t\le T}(A_1(t)+A_2(t))(g^{1-q_1}(t,x)+g^{1-q_2}(t,x))}{\min(\gamma_1,\gamma_2)}.$$

Moreover, note that the random variables $\xi_s-\xi_t$ and X_t^* are independent. Therefore, for every $m>1$ we calculate ($\mathrm{E}_{t,x}$ is the expectation operator conditional on $X_t^\varsigma = x$)

$$\mathrm{E}_{t,x} \sup_{t\le s\le T} z^m(s, X_s^*) \le 2^{m-1} M_*^m(x) \Big(\mathrm{E} \sup_{t\le s\le T} e^{m_1(\xi_s-\xi_t)} + \mathrm{E} \sup_{t\le s\le T} e^{m_2(\xi_s-\xi_t)}\Big),$$

where $m_1 = m(1-q_1)$ and $m_2 = m(1-q_2)$. Therefore, to check condition $\mathbf{H}_4$ it suffices to show that for every $\lambda \in \mathbb{R}$

$$\mathrm{E} \sup_{t\le s\le T} e^{\lambda(\xi_s-\xi_t)} < \infty. \tag{4.7}$$

Indeed, for every $t \le s \le T$ we set $\mathscr{E}_{t,s} = e^{-\lambda \int_t^s \theta_u' \mathrm{d}W_u - \frac{\lambda^2}{2}\int_t^s |\theta_u|^2 \mathrm{d}u}$, then

$$e^{\lambda(\xi_s-\xi_t)} \le e^{|\lambda| R_T + \frac{|\lambda|+\lambda^2}{2}\|\theta\|_T^2} \mathscr{E}_{t,s}.$$

We recall from (2.3) that $(\theta_s)_{0\le s\le T}$ is a deterministic function. This implies that the process $(\mathscr{E}_{t,s})_{t\le s\le T}$ is a martingale. Hence applying the maximal inequality for positives submartingales (see e.g. Theorem 3.2 in [9]) we obtain that

$$\mathrm{E} \sup_{t\le s\le T} \mathscr{E}_{t,s}^2 \le 4\mathrm{E}\mathscr{E}_{t,T}^2 = 4e^{\lambda^2 \int_t^T |\theta_u|^2 \mathrm{d}u} \le 4e^{\lambda^2 \|\theta\|_T^2}.$$

From this inequality (4.7) follows, which implies $\mathbf{H}_4$. Therefore, by Theorem 11 we get Theorem 2.

4.3 Proof of Theorem 3

First note that restriction (2.13) is equivalent to

$$\inf_{0\le t\le T} L_t(\varsigma) \ge \ln(1-\zeta), \tag{4.8}$$

where

$$L_t(\varsigma) = (y,\theta)_t - V_t - \frac{1}{2}\|y\|_t^2 - |z_\alpha|\|y\|_t \tag{4.9}$$

with notations as in (2.3) and (2.10). Inequality (4.8) and the Cauchy-Schwartz inequality imply that

$$\|y\|_T\|\theta\|_T - \frac{1}{2}\|y\|_T^2 - |z_\alpha|\|y\|_T \ge \ln(1-\zeta)$$

and, consequently,

$$\|y\|_T \le \rho_{\mathrm{VaR}}^*, \tag{4.10}$$

where ρ^*_{VaR} has been defined in (3.4) and satisfies the equation

$$\|\theta\|_T \rho^*_{\text{VaR}} - \frac{1}{2}(\rho^*_{\text{VaR}})^2 - |z_\alpha|\rho^*_{\text{VaR}} = \ln(1-\zeta). \tag{4.11}$$

Moreover, for every $\varsigma \in \mathscr{U}$ (2.9) yields

$$\mathrm{E}_x X^\varsigma_t = x e^{R_t - V_t + (y,\theta)_t}.$$

For every $y \in \mathbb{R}^d$ the upper bound (4.10) and the Cauchy-Schwartz inequality yield

$$\sup_{0\le t\le T} e^{(y,\theta)_t} \le e^{\rho^*_{\text{VaR}}\|\theta\|_T}.$$

Therefore, the cost function (2.11) has an upper bound given by

$$\begin{aligned} J(x,\varsigma) &= x\left(\int_0^T e^{R_t - V_t + (y,\theta)_t} v_t \mathrm{d}t + e^{R_T - V_T + (y,\theta)_T}\right) \\ &\le x e^{\rho^*_{\text{VaR}}\|\theta\|_T + R_T}\left(\int_0^T e^{-V_t} v_t \mathrm{d}t + e^{-V_T}\right) \\ &= x e^{\rho^*_{\text{VaR}}\|\theta\|_T + R_T}. \end{aligned}$$

It is easy to see that the control ς^* defined in (3.12) matches this upper bound, i.e. $J(x,\varsigma^*) = x e^{\rho^*_{\text{VaR}}\|\theta\|_T + R_T}$. To finish the proof we have to check condition (4.8) for this control. If $\|\theta\|_T = 0$ then by (4.9)

$$\begin{aligned} L_t(\varsigma^*) &= -\frac{1}{2}\|y^*\|_t^2 - |z_\alpha|\|y^*\|_t \ge -\frac{1}{2}\|y^*\|_T^2 - |z_\alpha|\|y^*\|_T \\ &\ge -\frac{1}{2}(\rho^*_{\text{VaR}})^2 - |z_\alpha|\rho^*_{\text{VaR}} = \ln(1-\zeta). \end{aligned}$$

Let now $\|\theta\|_T > 0$. Note that condition (3.10) implies $|z_\alpha| \ge 2\|\theta\|_T - \rho^*_{\text{VaR}}$. Moreover, we can represent $L_t(\varsigma^*)$ as

$$L_t(\varsigma^*) = \rho^*_{\text{VaR}} f(\|\theta\|_t/\|\theta\|_T)$$

with

$$f(\eta) = (2\|\theta\|_T - \rho^*_{\text{VaR}})\frac{\eta^2}{2} - |z_\alpha|\eta, \quad 0 \le \eta \le 1.$$

Then

$$\inf_{0\le t\le T} L_t(\varsigma^*) = \rho^*_{\text{VaR}} \inf_{0\le\eta\le 1} f(\eta).$$

Taking into account that for $|z_\alpha| \ge 2\|\theta\|_T - \rho^*_{\text{VaR}}$ this infimum equals $f(1)$ we obtain together with (4.11)

$$\inf_{0\le t\le T} L_t(\varsigma^*) = \rho^*_{\text{VaR}} f(1) = \ln(1-\zeta).$$

This proves Theorem 3.

4.4 Proof of Theorem 4

We have to prove condition (4.8) for the strategy (3.7)–(3.8):

$$L_t(\varsigma^*) = \left(q - \frac{q^2}{2}\right)\|\theta\|_t^2 - V_t^* - q|z_\alpha|\|\theta\|_t$$
$$\geq \left(q - \frac{q^2}{2}\right)\|\theta\|_T^2 \mathbf{1}_{\{q>2\}} - V_T^* - q|z_\alpha|\|\theta\|_T = l_*(\gamma).$$

Now condition (4.8) follows immediately from the restrictions on ζ and the definition of $l_*(\gamma)$.

4.5 Proof of Theorem 5

We prove this theorem as Theorem 3. Firstly, we find an upper bound for the cost function $J(x,\varsigma)$ and, secondly, we show that the optimal control (3.20) matches this bound and satisfies condition (4.8). To this end note that from (2.9) we find that for $\varsigma \in \mathscr{U}$

$$\mathrm{E}_x(X_t^\varsigma)^\gamma = x^\gamma \widehat{g}_\gamma(t) e^{-\gamma V_t + \gamma(y,\theta)_t - \frac{\gamma(1-\gamma)}{2}\|y\|_t^2}. \tag{4.12}$$

This implies for $\varsigma \in \mathscr{U}$ that the cost function (2.11) has the form

$$J(x,\varsigma) = x^{\gamma_1}\int_0^T (e^{-V_t}v_t)^{\gamma_1}\widehat{g}_1(t)\widehat{h}_1(t,y)\mathrm{d}t + x^{\gamma_2}\widehat{g}_2(T)e^{-\gamma_2 V_T}\widehat{h}_2(T,y),$$

where

$$\widehat{h}_i(t,y) = e^{\gamma_i(y,\theta)_t - \frac{\gamma_i(1-\gamma_i)}{2}\|y\|_t^2}.$$

Hölder's inequality with $p = 1/\gamma_1$ and $q = (1-\gamma_1)^{-1}$ yields

$$J(x,\varsigma) \leq \sup_{0\leq t\leq T}\widehat{h}(t,y)\left(x^{\gamma_1}\int_0^T (e^{-V_t}v_t)^{\gamma_1}\widehat{g}_1(t)\mathrm{d}t + x^{\gamma_2}\widehat{g}_2(T)e^{-\gamma_2 V_T}\right)$$
$$\leq \sup_{0\leq t\leq T}\widehat{h}(t,y)(x^{\gamma_1}(1-e^{-V_T})^{\gamma_1}\|\widehat{g}_1\|_{q,T} + x^{\gamma_2}\widehat{g}_2(T)e^{-\gamma_2 V_T}),$$

where $\widehat{h}(t,y) = \max\{\widehat{h}_1(t,y), \widehat{h}_2(t,y)\}$. We abbreviate as before $\|\widehat{g}_1\|_{q,T} := (\int_0^T e^{q\gamma_1 R_t}\mathrm{d}t)^{1/q}$. By setting $\kappa = 1 - e^{-V_T}$ we obtain that

$$J(x,\varsigma) \leq \max_{0\leq t\leq T}\widehat{h}(t,y)G(x,\kappa), \tag{4.13}$$

where $G(\cdot,\cdot)$ is given in (3.15). Moreover, condition (4.8) implies

$$\|y\|_T \le \sqrt{(|z_\alpha| - \|\theta\|_T)^2 + 2\ln\frac{1-\kappa}{1-\zeta}} - (|z_\alpha| - \|\theta\|_T) := \rho(\kappa) \tag{4.14}$$

and $0 \le \kappa \le \zeta < 1$. It is easy to see that $\rho(\kappa) \le \rho(0) = \rho^*_{\mathrm{VaR}}$ for every $0 \le \kappa \le \zeta$. From this inequality follows that for $i = 1, 2$ the functions $\widehat{h}_i(t, y)$ with $0 < \gamma_i \le 1$ can be bounded above by

$$\sup_{0\le t\le T} \widehat{h}_i(t,y) \le \exp\left\{\gamma_i \max_{0\le x\le\rho(\kappa)}\left(x\|\theta\|_T - \frac{(1-\gamma_i)x^2}{2}\right)\right\}$$
$$= \exp\left\{\gamma_i\rho_i(\kappa)\|\theta\|_T - \frac{\gamma_i(1-\gamma_i)}{2}\rho_i^2(\kappa)\right\} := M_i(\rho_i(\kappa)), \tag{4.15}$$

where $\rho_i(\kappa) = \min(\rho(\kappa), x_i)$ with $x_i = q_i\|\theta\|_T$ for $0 < \gamma_i < 1$ and $\rho_i(\kappa) = \rho(\kappa)$ for $\gamma_i = 1$. Therefore, from (4.13) we obtain the following upper bound for the cost function

$$J(x,\varsigma) \le \max_{1\le i\le 2} M_i(\rho_i(\kappa))G(x,\kappa) \le \max_{1\le i\le 2}\sup_{0\le\kappa\le\zeta} M_i(\rho_i(\kappa))G(x,\kappa).$$

If $\rho(0) \le x_i$ then

$$\sup_{0\le\kappa\le\zeta} M_i(\rho_i(\kappa))G(x,\kappa) = \sup_{0\le\kappa\le\zeta} M_i(\rho(\kappa))G(x,\kappa).$$

We calculate this supremum by means of Lemma 2 with $a = 0$ and $b = \zeta$. Note that condition (3.18) guarantees that $\zeta < \kappa_*(x)$, which is defined in (3.16). Therefore, the function $G(x,\cdot)$ has positive first derivative and negative second on $[0,\zeta]$. Moreover, from (4.14) we find the derivative of $\rho(\cdot)$ as

$$\dot{\rho}(\kappa) = -\frac{1}{(1-\kappa)\sqrt{(|z_\alpha| - \|\theta\|_T)^2 + 2\ln(1-\kappa) - 2\ln(1-\zeta)}}$$

and, therefore,

$$\sup_{0\le\kappa\le\zeta} |\dot{\rho}(\kappa)| \le \frac{1}{(1-\zeta)(|z_\alpha| - \|\theta\|_T)}.$$

By (3.19) we obtain that

$$\sup_{0\le\kappa\le\zeta} |\dot{\rho}(\kappa)| \le \frac{1}{\max\{\gamma_1,\gamma_2\}\|\theta\|_T}\frac{\partial \ln G(x,\zeta)}{\partial\zeta}.$$

Now Lemma 2 yields

$$\max_{0\le\kappa\le\zeta} M_i(\rho(\kappa))G(x,\kappa) = M_i(\rho(\zeta))G(x,\zeta) = G(x,\zeta). \tag{4.16}$$

Consider now $x_i < \rho(0)$. We recall that $\rho(\cdot)$ is decreasing on $[0, \zeta]$ with $\rho(\zeta) = 0$. Therefore, there exists $0 \le \kappa_i < \zeta$ such that $\rho(\kappa_i) = x_i$. As $G(x, \cdot)$ is increasing on $[0, \zeta]$ we obtain

$$\max_{0\le\kappa\le\kappa_i} M_i(\rho_i(\kappa))G(x,\kappa) = M_i(\rho(\kappa_i))G(x,\kappa_i).$$

This in combination with (4.16) yields

$$\sup_{0\le\kappa\le\zeta} M_i(\rho_i(\kappa))G(x,\kappa) = \sup_{\kappa_i\le\kappa\le\zeta} M_i(\rho(\kappa))G(x,\kappa) = G(x,\zeta).$$

This implies the following upper bound for the cost function

$$J(x,\varsigma) \le G(x,\zeta). \tag{4.17}$$

Now we find a control to obtain the equality in (4.17). It is clear that we have to take a consumption such that

$$\int_0^T \widehat{g}_1(t)(e^{-V_t}v_t)^{\gamma_1}\mathrm{d}t = (1-e^{-V_T})^{\gamma_1}\|\widehat{g}_1\|_{q_1,T}$$

and $V_T = -\ln(1-\zeta)$. To find this consumption we solve the differential equation on $[0, T]$

$$\dot{V}_t e^{-V_t} = \frac{\zeta}{\|\widehat{g}_1\|^{q_1}_{q_1,T}}\,\widehat{g}^{\,q_1}_1(t), \qquad V_0 = 0.$$

The solution of this equation is given by

$$V^*_t = -\ln\left(1-\zeta\,\frac{\|\widehat{g}_1\|^{q_1}_{q_1,t}}{\|\widehat{g}_1\|^{q_1}_{q_1,T}}\right)$$

and the optimal consumption rate is

$$v^*_t = \dot{V}^*_t = \frac{\zeta\,\widehat{g}^{\,q_1}_1(t)}{\|\widehat{g}_1\|^{q_1}_{q_1,T} - \zeta\,\|\widehat{g}_1\|^{q_1}_{q_1,t}}.$$

We recall that $r_t \ge 0$, therefore, for every $0 \le t \le T$

$$v^*_t \le v^*_T = \frac{\zeta\,\widehat{g}^{\,q_1}_1(T)}{(1-\zeta)\|\widehat{g}_1\|^{q_1}_{q_1,T}}.$$

The condition $0 < \zeta \le \widehat{\kappa}(\gamma_1)$ implies directly that the last upper bound less than 1, i.e. the strategy ς^* defined in (3.21) belongs to $\mathscr{U}$. Moreover, from (4.14) we see that for the value $V^*_T = -\ln(1-\zeta)$ (i.e. $\kappa = \zeta$) the only control process, which satisfies this condition is identical zero; i.e. $y^*_t = 0$ for all $0 \le t \le T$. In this case $\widehat{h}(t, y^*) = 1$ for every $t \in [0, T]$ and, therefore, $J(x, \varsigma^*) = G(x, \zeta)$.

4.6 Proof of Lemma 1

(1) Recall the following well known inequality for the Gaussian integral

$$(1-x^{-2})e^{-x^2/2} < x\int_x^\infty e^{-t^2/2}\mathrm{d}t < e^{-x^2/2}, \quad x \ge 0. \tag{4.18}$$

We use this to check directly that $\psi(\rho,\cdot)$ is for every fixed $\rho > 0$ decreasing for $|z_\alpha| \ge 2\|\theta\|_T$. This implies for $0 \le u \le 1$

$$\frac{\partial\psi(\rho,u)}{\partial u} = 2\|\theta\|_T\rho u - \rho\frac{e^{-(|z_\alpha|+\rho u)^2/2}}{\int_{|z_\alpha|+\rho u}^\infty e^{-t^2/2}\mathrm{d}t} \le \rho(2\|\theta\|_T - |z_\alpha|) < 0.$$

(2) Similarly, we can show that $\psi(\cdot,1)$ is strictly decreasing for $|z_\alpha| \ge \|\theta\|_T$.
(3) From (4.18) we obtain

$$\psi(\rho,1) \le \|\theta\|_T\rho - \ln\int_{|z_\alpha|}^\infty e^{-t^2/2}\mathrm{d}t - \frac{1}{2}(|z_\alpha|+\rho)^2 - \ln(|z_\alpha|+\rho). \tag{4.19}$$

This implies that $\lim_{\rho\to\infty}\psi(\rho,1) = -\infty$. As $\psi(0,1) = 0$ we conclude that the equation $\psi(\rho,1) = a$ has a unique root for every $a \le 0$. Thus ρ^*_{ES} is equal to the root of this equation for $a = \ln(1-\zeta)$. Now for $|z_\alpha| > 1$ inequalities (4.18)–(4.19) imply directly the upper bound for ρ^*_{ES} as given in (3.25).

4.7 Proof of Theorem 7

Note that Lemma 1 implies immediately that $\rho^*_{\mathrm{ES}} < \infty$ and $\psi(\rho^*_{\mathrm{ES}},1) = \ln(1-\zeta)$. Furthermore, inequality (2.14) is equivalent to

$$\inf_{0\le t\le T} L^*_t(\varsigma) \ge \ln(1-\zeta), \tag{4.20}$$

where

$$L^*_t(\varsigma) = (y,\theta)_t - V_t + \ln(F_\alpha(|z_\alpha| + \|y\|_t)).$$

First note that

$$\begin{aligned} L^*_T(\varsigma) &= (y,\theta)_T - V_T + \ln(F_\alpha(|z_\alpha| + \|y\|_T)) \\ &\le \|y\|_T\|\theta\|_T + \ln(F_\alpha(|z_\alpha| + \|y\|_T)) = \psi(\|y\|_T,1). \end{aligned}$$

Therefore, for every strategy $\varsigma \in \mathscr{U}$ satisfying inequality (4.20) for $t = T$ we obtain

$$\ln(1-\zeta) = \psi(\rho^*_{\mathrm{ES}},1) \le L^*_T(\varsigma) \le \psi(\|y\|_T,1).$$

By Lemma 1(2) $\psi(\cdot,1)$ is decreasing, hence $\|y\|_T \le \rho^*_{\mathrm{ES}}$. Therefore, to conclude the proof we have to show (4.20) for the strategy ς^* as defined in (3.12) with

$\rho^*_{\mathrm{VaR}} = \rho^*_{\mathrm{ES}}$. If $\|\theta\|_T = 0$, then $\varsigma^* = (y^*, 0)$ with every function y^* for which $\|y^*\|_T \le \rho^*_{\mathrm{ES}}$. Therefore, if $\|\theta\|_T = 0$, then

$$L^*_t(\varsigma^*) = \psi(\|y^*\|_t, 1) \ge \psi(\|y^*\|_T, 1) \ge \ln(1-\zeta).$$

If $\|\theta\|_T > 0$, then

$$\inf_{0\le t\le T} L^*_t(\varsigma^*) = \inf_{0\le t\le T} \psi\left(\rho^*_{\mathrm{ES}}, \frac{\|\theta\|_t}{\|\theta\|_T}\right) = \psi(\rho^*_{\mathrm{ES}}, 1) = \ln(1-\zeta).$$

This proves Theorem 7.

4.8 *Proof of Theorem 8*

It suffices to prove condition (4.20) for the strategy (3.7)–(3.8). We have

$$\begin{aligned} L^*_t(\varsigma^*) &= \int_0^t (y^*_u)'\theta_u \mathrm{d}u - V^*_t + \ln(F_\alpha(|z_\alpha| + \|y^*\|_t)) \\ &= q\|\theta\|^2_t - V^*_t + \ln(F_\alpha(|z_\alpha| + q\|\theta\|_t)) \\ &\ge \psi_0(\|\theta\|_t) - V^*_T, \end{aligned} \tag{4.21}$$

where

$$\psi_0(u) = qu^2 + \ln F_\alpha(|z_\alpha| + qu) \quad \text{with } q = \frac{1}{1-\gamma}.$$

It is clear that ψ_0 is continuously differentiable. Moreover, by inequality (4.18) we obtain for $0 \le u \le \|\theta\|_T$

$$\begin{aligned} \frac{\mathrm{d}\psi_0(u)}{\mathrm{d}u} &= 2qu - q\frac{e^{-(|z_\alpha|+qu)^2/2}}{\int_{|z_\alpha|+qu}^\infty e^{-t^2/2}\mathrm{d}t} \\ &\le 2qu - q|z_\alpha| - q^2u \le q(2\|\theta\|_T - |z_\alpha|). \end{aligned}$$

Since $|z_\alpha| \ge 2\|\theta\|_T$, $\psi_0(u)$ decreases in $[0, \|\theta\|_T]$. Hence, inequality (4.21) implies

$$L^*_t(\varsigma^*) \ge \psi_0(\|\theta\|_T) - V^*_T = q\|\theta\|^2_T + \ln e^{-V^*_T} F_\alpha(|z_\alpha| + q\|\theta\|_T).$$

Applying condition (3.26) yields (4.20). This proves Theorem 8.

4.9 *Proof of Theorem 9*

We recall that $\psi(\rho, 1) \le 0$ for $\rho \ge 0$. Therefore condition (4.20) implies

$$\ln(1-\zeta) \le -V_T + \psi(\|y\|_T, 1) \le -V_T. \tag{4.22}$$

As in the proof of Theorem 5 we set $\kappa = 1 - e^{-V_T}$ and conclude from this inequality that $0 \le \kappa \le \zeta$. Moreover, from (4.22) we obtain also that

$$\ln(1-\zeta) - \ln(1-\kappa) \le \psi(\|y\|_T, 1).$$

Since, by Lemma 1(2) $\psi(\cdot, 1)$ is decreasing, we get $\|y\|_T \le \rho(\kappa)$, where $\rho(\kappa)$ is the solution of the equation

$$\psi(\rho, 1) = \ln(1-\zeta) - \ln(1-\kappa). \tag{4.23}$$

By Lemma 1(3) the root of (4.23) exists for every $0 \le \kappa \le \zeta$ and is decreasing in κ giving $\rho(\kappa) \le \rho(0) = \rho^*_{\mathrm{ES}}$. Consequently, we estimate the cost function as in Sect. 4.5 and obtain

$$J(x, \varsigma) \le \max_{1\le i\le 2} \max_{\kappa\in[0,\zeta]} M_i(\rho_i(\kappa)) G(x, \kappa), \tag{4.24}$$

where $G(x,\kappa)$ is as in (3.15), $M_i(\cdot)$ is defined in (4.15) and $\rho_i(\kappa) = \min(x_i, \rho(\kappa))$ for $x_i = \|\theta\|_T/(1-\gamma_i)$ for $0 < \gamma_i < 1$ with $\rho_i(\kappa) = \rho(\kappa)$ for $\gamma_i = 1$.

To finish the proof we have to show condition (5.25) of Lemma 2. From (4.23) we find that

$$\dot\rho(\kappa) = \frac{1}{1-\kappa}\left(\frac{\mathrm{d}\psi(\rho,1)}{\mathrm{d}\rho}\right)^{-1}.$$

Now from the definition of ψ in (3.23) and inequality (4.18) follows

$$\frac{\mathrm{d}\psi(\rho,1)}{\mathrm{d}\rho} = \|\theta\|_T - \frac{e^{-(|z_\alpha|+\rho)^2/2}}{\int_{|z_\alpha|+\rho}^{\infty} e^{-t^2/2}\mathrm{d}t} \le \|\theta\|_T - |z_\alpha|.$$

Therefore (3.27) yields (we set $G_1(x,\zeta) = \frac{\partial G(x,\zeta)}{\partial \zeta}$)

$$\sup_{0\le\kappa\le\zeta} |\dot\rho(\kappa)| \le \frac{1}{(1-\zeta)(|z_\alpha| - \|\theta\|_T)} \le \frac{G_1(x,\zeta)}{\max\{\gamma_1,\gamma_2\}\|\theta\|_T G(x,\zeta)}.$$

We apply Lemma 2, and the same reasoning as in the proof of Theorem 5 implies that

$$\max_{0\le\kappa\le\zeta} M_i(\rho_i(\kappa)) G(x,\kappa) \le G(x,\zeta)$$

for $i = 1, 2$. Therefore from the upper bound (4.24) follows

$$J(x,\varsigma) \le G(x,\zeta).$$

The remainder of the proof is the same as for Theorem 5.

Appendix

5.1 A Technical Lemma

Lemma 2 *Let G be some positive two times continuously differentiable function on $[a, b]$ such that $\dot{G}(x) \geq 0$ and $\ddot{G}(x) \leq 0$ for all $a \leq x \leq b$. Moreover, let $\rho : [a, b] \to \mathbb{R}_+$ be continuously differentiable with negative derivative $\dot{\rho}$ satisfying*

$$\sup_{a \leq \kappa \leq b} |\dot{\rho}(\kappa)| \leq \frac{(\ln G(b))'}{\max\{\gamma_1, \gamma_2\} \|\theta\|_T}. \tag{5.25}$$

Recall the definitions of $M_i(\cdot)$ in (4.15). Then the functions $M_1(\rho(\cdot))G(\cdot)$ and $M_2(\rho(\cdot))G(\cdot)$ are increasing in $[a, b]$.

Proof For $\|\theta\|_T = 0$ the result is obvious. Consider now $\|\theta\|_T > 0$. We prove that for $i = 1, 2$ the functions $l_i(x) = \ln M_i(\rho(x)) + \ln G(x)$ are increasing in $[a, b]$. As derivative we obtain

$$\dot{l}_i(\kappa) = \gamma_i \dot{\rho}(\kappa)(\|\theta\|_T - (1 - \gamma_i)\rho(\kappa)) + \frac{\dot{G}(x)}{G(x)}.$$

Since the derivative of the function $\dot{G}(\cdot)/G(\cdot)$ is negative on $[a, b]$, $\dot{G}(\cdot)/G(\cdot)$ is decreasing on $[a, b]$, hence

$$\frac{\dot{G}(x)}{G(x)} \geq \frac{\dot{G}(b)}{G(b)} > 0$$

for $x \in [a, b]$. Therefore, as $\rho > 0$ and $\dot{\rho} < 0$ we find

$$\dot{l}_i(x) \geq (\ln G(b))' - \gamma_i \|\theta\|_T |\dot{\rho}(\kappa)| \geq 0, \quad a \leq \kappa \leq b. \qquad \square$$

5.2 The Verification Theorem

We prove a special form of the verification theorem (see e.g. Touzi [11], p. 16). Consider on the interval $[0, T]$ the stochastic control process given by the Itô process

$$\mathrm{d}X_t^\varsigma = a(t, X_t^\varsigma, \varsigma_t)\mathrm{d}t + b(t, X_t^\varsigma, \varsigma_t)\mathrm{d}W_t, \quad t \geq 0, \quad X_0^\varsigma = x > 0. \tag{5.26}$$

We assume that the control process ς takes values in some set $\mathscr{K} \subseteq \mathbb{R}^d \times [0, \infty)$. Moreover, assume that the coefficients a and b satisfy the following conditions

(1) for all $t \in [0, T]$ the functions $a(t, \cdot, \cdot)$ and $b(t, \cdot, \cdot)$ are continuous on $(0, \infty) \times \mathscr{K}$;
(2) for every deterministic vector $\upsilon \in \mathscr{K}$ the stochastic differential equation

$$\mathrm{d}X_t^\upsilon = a(t, X_t^\upsilon, \upsilon)\mathrm{d}t + b(t, X_t^\upsilon, \upsilon)\mathrm{d}W_t, \quad X_0^\upsilon = x > 0,$$

has an unique strong solution.

Now we introduce admissibles control processes for (5.26). We set $\mathscr{F}_t = \sigma\{W_u, 0 \le u \le t\}$ for any $0 < t \le T$.

Definition 4 A stochastic control process $\varsigma = (\varsigma_t)_{0\le t\le T} = ((y_t, c_t))_{0\le t\le T}$ is called *admissible* on $[0, T]$ with respect to (5.26) if it is $(\mathscr{F}_t)_{0\le t\le T}$-progressively measurable with values in $\mathbb{R}^d \times [0,\infty)$, and (5.26) has a unique strong a.s. positive continuous solution $(X_t^\varsigma)_{0\le t\le T}$ on $[0, T]$ such that

$$\int_0^T (|a(t, X_t^\varsigma, \varsigma_t)| + b^2(t, X_t^\varsigma, \varsigma_t))\mathrm{d}t < \infty \quad \text{a.s.} \tag{5.27}$$

In this context $\mathscr{V}$ is the set of all admissible control processes with respect to (5.26); cf. Definition 1.

Moreover, assume that $f : [0, T] \times (0,\infty) \times \mathscr{K} \to [0,\infty)$ and $h : (0,\infty) \to [0,\infty)$ are continuous utility functions. We define the cost function by

$$J(t, x, \varsigma) := \mathrm{E}_{t,x}\left[\int_t^T f(s, X_s^\varsigma, \varsigma_s)\mathrm{d}s + h(X_T^\varsigma)\right], \quad 0 \le t \le T,$$

where $\mathrm{E}_{t,x}$ is the expectation operator conditional on $X_t^\varsigma = x$. Our goal is to solve the optimisation problem

$$J^*(t, x) := \sup_{\varsigma\in\mathscr{V}} J(t, x, \varsigma). \tag{5.28}$$

To this end we introduce the Hamilton function

$$H(t, x, z_1, z_2) := \sup_{\vartheta\in\mathscr{K}} H_0(t, x, z_1, z_2, \vartheta), \tag{5.29}$$

where

$$H_0(t, x, z_1, z_2, \vartheta) := a(t, x, \vartheta)z_1 + \frac{1}{2}b^2(t, x, \vartheta)z_2 + f(t, x, \vartheta).$$

In order to find the solution to (5.28) we investigate the Hamilton-Jacobi-Bellman equation

$$\begin{cases} z_t(t, x) + H(t, x, z_x(t, x), z_{xx}(t, x)) = 0, & t \in [0, T], \\ z(T, x) = h(x), & x > 0. \end{cases} \tag{5.30}$$

Here z_t denotes the partial derivative of z with respect to t, analogous notation applies to all partial derivatives.

We assume that the following conditions hold:

($\mathbf{H}_1$) *There exists some function* $z : [0, T] \times (0,\infty) \to [0,\infty)$*, which satisfies the following conditions.*

- *For all* $0 \le t_1, t_2 \le T$ *there exists a* $\mathscr{B}[0,T] \otimes \mathscr{B}(0,\infty)$ *measurable function* $z_t(\cdot,\cdot)$ *such that*

$$z(t_2,x) - z(t_1,x) = \int_{t_2}^{t_2} z_t(u,x)\mathrm{d}u, \quad x > 0. \tag{5.31}$$

- *Moreover, we assume that for every* $u \in [0,T]$ *the function* $z_t(u,\cdot)$ *is continuous on* $(0,\infty)$ *such that for every* $N > 1$

$$\lim_{\varepsilon\to 0}\int_0^T \sup_{x,y\in K_N, |x-y|<\varepsilon} |z_t(u,x) - z_t(u,y)|\mathrm{d}u = 0, \tag{5.32}$$

where $K_N = [N^{-1}, N]$.
- *The function* z *has second partial derivative* z_{xx}, *which is continuous on* $[0,T] \times (0,\infty)$.
- *There exists a set* $\Gamma \subset [0,T]$ *of Lebesgue measure* $\lambda(\Gamma) = T$ *such that* $z(t,x)$ *satisfies* (5.30) *for all* $t \in \Gamma \subset [0,T]$ *and for all* $x > 0$.

($\mathbf{H}_2$) *There exists a measurable function* $\vartheta^* : [0,T] \times (0,\infty) \to \mathscr{K}$ *such that*

$$H(t,x,z_x(t,x),z_{xx}(t,x)) = H_0(t,x,z_x(t,x),z_{xx}(t,x),\vartheta^*(t,x))$$

for all $t \in \Gamma$ *and for all* $x \in (0,\infty)$.

($\mathbf{H}_3$) *There exists a unique a.s. strictly positive strong solution to the Itô equation*

$$\mathrm{d}X_t^* = a^*(t,X_t^*)\mathrm{d}t + b^*(t,X_t^*)\mathrm{d}W_t, \quad t \ge 0, \quad X_0^* = x, \tag{5.33}$$

where $a^*(t,x) = a(t,x,\vartheta^*(t,x))$ *and* $b^*(t,x) = b(t,x,\vartheta^*(t,x))$. *Moreover, the optimal control process* $\varsigma_t^* = \vartheta^*(t,X_t^*)$ *for* $0 \le t \le T$ *belongs to* $\mathscr{V}$.

($\mathbf{H}_4$) *There exists some* $\delta > 1$ *such that for all* $0 \le t \le T$ *and* $x > 0$

$$\mathrm{E}_{t,x} \sup_{t\le s\le T} (z(s,X_s^*))^\delta < \infty.$$

Theorem 11 *Assume that* $\mathscr{V} \neq \emptyset$ *and* $\mathbf{H}_1 - \mathbf{H}_4$ *hold. Then for all* $t \in [0,T]$ *and for all* $x > 0$ *the solution to the Hamilton-Jacobi-Bellman equation* (5.30) *coincides with the optimal value of the cost function, i.e.* $z(t,x) = J^*(t,x) = J^*(t,x,\varsigma^*)$, *where the optimal strategy* ς^* *is defined in* $\mathbf{H}_2$ *and* $\mathbf{H}_3$.

Proof For $\varsigma \in \mathscr{V}$ let X^ς be the associated wealth process with initial value $X_0^\varsigma = x$. Define stopping times

$$\tau_n = \inf\left\{ s \ge t : \int_t^s b^2(u,X_u^\varsigma,\varsigma_u) z_x^2(u,X_u^\varsigma)\mathrm{d}u \ge n \right\} \wedge T.$$

Note that condition (5.27) implies that $\tau_n \to T$ as $n \to \infty$ a.s. By continuity of $z(\cdot,\cdot)$ and of $(X_t^{\varsigma})_{0\le t\le T}$ we obtain

$$\lim_{n\to\infty} z(\tau_n, X^{\varsigma}_{\tau_n}) = z(T, X^{\varsigma}_T) = h(X^{\varsigma}_T) \quad \text{a.s.} \tag{5.34}$$

Theorem 12 guarantees that we can invoke Itô's formula, and we conclude from (5.26)

$$\begin{aligned} z(t,x) = \int_t^{\tau_n} f(s, X_s^{\varsigma}, \varsigma_s)\mathrm{d}s + z(\tau_n, X^{\varsigma}_{\tau_n}) - \int_t^{\tau_n} (z_t(s, X_s^{\varsigma}) \\ + H_1(s, X_s^{\varsigma}, \varsigma_s))\mathrm{d}s - \int_t^{\tau_n} b(u, X_u^{\varsigma}, \varsigma_u) z_x(u, X_u^{\varsigma})\mathrm{d}W_u, \end{aligned} \tag{5.35}$$

where

$$H_1(s,x,\vartheta) = H_0(t, x, z_x(t,x), z_{xx}(t,x), \vartheta).$$

Condition $\mathbf{H}_1$ implies

$$z(t,x) \ge \mathrm{E}_{t,x} \int_t^{\tau_n} f(s, X_s^{\varsigma}, \varsigma_s)\mathrm{d}s + \mathrm{E}_{t,x} z(\tau_n, X^{\varsigma}_{\tau_n}).$$

Moreover, by monotone convergence for the first term and Fatou's lemma for the second, and by observing (5.34) we obtain

$$\begin{aligned} &\lim_{n\to\infty} \mathrm{E}_{t,x} \int_t^{\tau_n} f(s, X_s^{\varsigma}, \varsigma_s)\mathrm{d}s + \lim_{n\to\infty} \mathrm{E}_{t,x} z(\tau_n, X^{\varsigma}_{\tau_n}) \\ &\ge \mathrm{E}_{t,x} \int_t^{T} f(s, X_s^{\varsigma}, \varsigma_s)\mathrm{d}s + \mathrm{E}_{t,x} h(X^{\varsigma}_T) := J(t,x,\varsigma), \quad 0 \le t \le T. \end{aligned} \tag{5.36}$$

Therefore, $z(t,x) \ge J^*(t,x)$ for all $0 \le t \le T$.

Similarly, replacing ς in (5.35) by ς^* as defined by $\mathbf{H}_2$–$\mathbf{H}_3$ we obtain

$$z(t,x) = \mathrm{E}_{t,x} \int_t^{\tau_n} f(s, X_s^{*}, \varsigma_s^*)\mathrm{d}s + \mathrm{E}_{t,x} z(\tau_n, X^{*}_{\tau_n}).$$

Condition $\mathbf{H}_4$ implies that the sequence $(z(\tau_n, X^*_{\tau_n}))_{n\in\mathbb{N}}$ is uniformly integrable. Therefore, by (5.34),

$$\lim_{n\to\infty} \mathrm{E}_{t,x} z(\tau_n, X^*_{\tau_n}) = \mathrm{E}_{t,x} \lim_{n\to\infty} z(\tau_n, X^*_{\tau_n}) = \mathrm{E}_{t,x} h(X^*_T),$$

and we obtain

$$\begin{aligned} z(t,x) &= \lim_{n\to\infty} \mathrm{E}_{t,x} \int_t^{\tau_n} f(s, X_s^{*}, \varsigma_s^*)\mathrm{d}s + \lim_{n\to\infty} \mathrm{E}_{t,x} z(\tau_n, X^*_{\tau_n}) \\ &= \mathrm{E}_{t,x}\left(\int_t^{T} f(s, X_s^{*}, \varsigma_s^*)\mathrm{d}s + h(X^*_T) \right) = J(t,x,\varsigma^*). \end{aligned}$$

Together with (5.36) we arrive at $z(t,x) = J^*(t,x)$. This proves Theorem 11. □

Remark 8 Note that in contrast to the usual verification theorem (see e.g. Touzi [11], Theorem 1.4) we do not assume that (5.30) has a solution for all $t \in [0, T]$, but only for almost all $t \in [0, T]$. This provides the possibility to consider market models as in (2.1) with discontinuous functional coefficients. Moreover, in the usual verification theorem the function $f(t, x, \vartheta)$ is bounded with respect to $\vartheta \in \mathscr{K}$ or integrable with all moments finite. This is an essential difference of our situation as for the optimal consumption problem f is not bounded over $\vartheta \in \mathscr{K}$ and we do not assume that f is integrable.

5.3 A Special Version of Itô's Formula

We prove Itô's formula for functions satisfying $\mathbf{H}_1$, an extension, which to the best of our knowledge can not be found in the literature. Consider the Itô equation

$$\mathrm{d}\xi_t = a_t \mathrm{d}t + b_t \mathrm{d}W_t,$$

where the stochastic processes $a = (a_t)_{0 \le t \le T}$ and $b = (b_t)_{0 \le t \le T}$ are measurable, adapted and satisfy for the investment horizon $T > 0$

$$\int_0^T (|a_t| + b_t^2)\mathrm{d}t < \infty \quad \text{a.s.} \tag{5.37}$$

Theorem 12 *Let $f : [0, T] \times (0, \infty) \to [0, \infty)$ satisfy $\mathbf{H}_1$. Assume that the process ξ is a.s. positive on $0 \le t \le T$. Then $(f(t, \xi_t))_{0 \le t \le T}$ is the solution to*

$$\begin{aligned} \mathrm{d}f(t, \xi_t) = & \left(f_t(t, \xi_t) + f_x(t, \xi_t) a_t + \frac{1}{2} f_{xx}(t, \xi_t) \right) b_t^2 \mathrm{d}t \\ & + f_x(t, \xi_t) b_t \mathrm{d}W_t. \end{aligned} \tag{5.38}$$

Remark 9 Note that in contrast to the usual Itô formula we do not assume that f has a continuous derivative with respect to t and continuous derivatives with respect to x on the whole of $\mathbb{R}$. For example, the function (4.3) for $\gamma_1 = \gamma_2 = \gamma \in (0, 1)$ factorises into $z(t, x) = Z(t)x^\gamma$, i.e. is not continuously differentiable with respect to x on $\mathbb{R}$.

Proof First we prove (5.38) for bounded processes a and b, i.e. we assume that for some constant $L > 0$

$$\sup_{0 \le t \le T} (|a_t| + |b_t|) \le L \quad \text{a.s.} \tag{5.39}$$

Let $(t_k)_{1 \le k \le n}$ be a partition of $[0, T]$, more precisely, take $t_k = kT/n$, and consider the telescopic sums

$$\begin{aligned} & f(T,\xi_T)-f(0,\xi_0) \\ & = \sum_{k=1}^{n}(f(t_k,\xi_{t_k})-f(t_{k-1},\xi_{t_k}))+\sum_{k=1}^{n}(f(t_{k-1},\xi_{t_k})-f(t_{k-1},\xi_{t_{k-1}})) \\ & := \sum_{1,n}+\sum_{2,n}. \end{aligned}$$

Taking condition (5.31) into account we can represent the first sum as

$$\Sigma_{1,n}=\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} f_t(u,\xi_{t_k})\mathrm{d}u=\int_0^T f_t(u,\xi_u)\mathrm{d}u+r_{1,n},$$

where

$$r_{1,n}=\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k}(f_t(u,\xi_{t_k})-f_t(u,\xi_u))\mathrm{d}u.$$

Now we prove that $r_{1,n}\xrightarrow{P}0$ as $n\to\infty$. To this end we introduce the stopping time,

$$\tau_N=\inf\{t\geq 0:\xi_t+\xi_t^{-1}\geq N\}\wedge T,\quad N>0. \tag{5.40}$$

As the process ξ is continuous and a.s. positive,

$$\lim_{N\to\infty}\mathrm{P}(\tau_N<T)=0, \tag{5.41}$$

and, hence, $\tau_N\xrightarrow{P}T$ as $N\to\infty$. Moreover, the modulus of continuity of the process ξ satisfies

$$\Delta_\varepsilon(\xi,[0,T]):=\sup_{|t-s|\leq\varepsilon,s,t\in[0,T]}|\xi_t-\xi_s|\xrightarrow{\text{a.s.}}0,\quad \varepsilon\to 0. \tag{5.42}$$

Note now that condition (5.32) implies that for every $N>1$

$$F^*(\eta,N):=\int_0^T\sup_{x,y,\in K_N,|x-y|<\eta}|f_t(u,x)-f_t(u,y)|\mathrm{d}u\to 0\quad\text{as }\eta\to 0,$$

where $K_N=[N^{-1},N]$. This implies that for every $\delta>0$ there exists $\eta_\delta>0$ such that $F^*(\eta_\delta,N)<\delta$. Moreover, taking into account that for $\varepsilon=T/n$ the random variable $r_{1,n}$ is bounded on the ω-set

$$\{\Delta_\varepsilon(\xi,[0,T])\leq\eta_\delta\}\cap\{\tau_N=T\}$$

by $|r_{1,n}|\leq F^*(\eta_\delta,N)<\delta$, we obtain that

$$\mathrm{P}(|r_{1,n}|>\delta)\leq\mathrm{P}(\Delta_\varepsilon(\xi,[0,T])>\eta_\delta)+\mathrm{P}(\tau_N<T).$$

Relations (5.41) and (5.42) imply $r_{1,n} \xrightarrow{P} 0$ as $n \to \infty$. Now define

$$r_{2,n} := \Sigma_{2,n} - \int_0^T f_x(t,\xi_t)\mathrm{d}\xi_t - \frac{1}{2}\int_0^T f_{xx}(t,\xi_t) b_t^2 \mathrm{d}t.$$

We show that $r_{2,n} \xrightarrow{P} 0$ as $n \to \infty$. A Taylor expansion gives

$$\Sigma_{2,n} = \sum_{k=1}^n f_x(t_{k-1},\xi_{t_{k-1}})\Delta\xi_{t_k} + \frac{1}{2}\sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\int_{t_{k-1}}^{t_k} b_u^2 \mathrm{d}u$$
$$+ \frac{1}{2}\sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\alpha_k + \frac{1}{2}\sum_{k=1}^n \widehat{f}_k (\Delta\xi_{t_k})^2, \tag{5.43}$$

where $\alpha_k = (\Delta\xi_{t_k})^2 - \int_{t_{k-1}}^{t_k} b_u^2 \mathrm{d}u$, $\widehat{f}_k = f_{xx}(t_{k-1},\widehat{\xi}_{t_k}) - f_{xx}(t_{k-1},\xi_{t_{k-1}})$ and $\widehat{\xi}_{t_k} = \xi_{t_{k-1}} + \theta_k \Delta\xi_{t_k}$ with $\theta_k \in [0,1]$. Now taking into account that as $n \to \infty$

$$\sum_{k=1}^n f_x(t_{k-1},\xi_{t_{k-1}})\Delta\xi_{t_k} \xrightarrow{P} \int_0^T f_x(t,\xi_t)\mathrm{d}\xi_t,$$

$$\sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\int_{t_{k-1}}^{t_k} b_u^2 \mathrm{d}u \xrightarrow{\text{a.s.}} \int_0^T f_{xx}(t,\xi_t) b_t^2 \mathrm{d}t$$

it suffices to show that the last two terms in (5.43) tend to zero in probability. To this end we represent the first sum as

$$\sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\alpha_k = M_n + R_n,$$

where

$$M_n = \sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\eta_k \quad \text{with } \eta_k = \left(\int_{t_{k-1}}^{t_k} b_u \mathrm{d}W_u\right)^2 - \int_{t_{k-1}}^{t_k} b_u^2 \mathrm{d}u,$$

$$R_n = \sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}})\alpha_k^* \quad \text{with } \alpha_k^* = (\Delta\xi_{t_k})^2 - \left(\int_{t_{k-1}}^{t_k} b_u \mathrm{d}W_u\right)^2.$$

First we estimate the martingale part in this representation. Note that on the set $\{\tau_N = T\}$ the martingale part coincides with the bounded martingale

$$M_n = \sum_{k=1}^n f_{xx}(t_{k-1},\xi_{t_{k-1}\wedge\tau_N})\eta_k.$$

Taking into account that

$$|f_{xx}(t_{k-1},\xi_{t_{k-1}\wedge\tau_N})| \le \sup_{t\in[0,T],\, y\in[N^{-1},N]} |f_{xx}(t,y)| := M_*$$

we obtain

$$\mathrm{E}M_n^2 = \mathrm{E}\sum_{k=1}^{n} f_{xx}^2(t_{k-1}, \xi_{t_{k-1}\wedge\tau_N})\eta_k^2 \leq M_*^2 \sum_{k=1}^{n} \mathrm{E}\left(\int_{t_{k-1}}^{t_k} b_u \mathrm{d}W_u\right)^4$$
$$\leq 3L^4 M_*^2 \sum_{k=1}^{n} (\Delta t_k)^2 = 3L^4 M_*^2 T^2 \frac{1}{n} \to 0, \quad n \to \infty.$$

In the last inequality we used the bound (5.39) for b. We conclude

$$M_n \xrightarrow{P} 0, \quad n \to \infty. \tag{5.44}$$

Using the convergence (5.42) also for $I(t) = \int_0^t b_u \mathrm{d}W_u$ and the upper bound (5.39) for a we obtain

$$|\alpha_k^*| \leq \left(\int_{t_{k-1}}^{t_k} a_u \mathrm{d}u\right)^2 + 2\int_{t_{k-1}}^{t_k} |a_u| \mathrm{d}u \left|\int_{t_{k-1}}^{t_k} b_u \mathrm{d}W_u\right|$$
$$\leq L^2(\Delta t_k)^2 + 2L\Delta_\varepsilon(I, [0, T])\Delta t_k,$$

where $\varepsilon = \Delta t_k = T/n$. This yields $\lim_{n\to\infty} \sum_{k=1}^n |\alpha_k^*| = 0$ a.s. We use analogous arguments as for (5.44) to show that $R_n \xrightarrow{P} 0$. Taking also into account that $\sum_{k=1}^n (\Delta\xi_{t_k})^2$ is bounded in probability, i.e.

$$\lim_{m\to\infty} \mathrm{P}\left(\sum_{k=1}^{n} (\Delta\xi_{t_k})^2 \geq m\right) = 0,$$

it is easy to see that the last sum in (5.43) tends to zero in probability. This proves Ito's formula (5.38) for bounded coefficients (a_t) and (b_t).

To prove Ito's formula under condition (5.37) we introduce for $L \in \mathbb{N}$ the sequence of processes $(\xi_t^L)_{0\leq t\leq T}$ by

$$\mathrm{d}\xi_t^L = a_t^L \mathrm{d}t + b_t^L \mathrm{d}W_t, \qquad \xi_0^L = \xi_0,$$

where $a_t^L := a_t \chi_{\{|a_t|\leq L\}}$ and $b_t^L := b_t \chi_{\{|b_t|\leq L\}}$. For each of these processes we already proved (5.38). Therefore we can write

$$f(T, \xi_T^L) = f(0, \xi_0) + \int_0^T A_t^L \mathrm{d}t + \int_0^T B_t^L \mathrm{d}W_t, \tag{5.45}$$

where $A_t^L = f_t(t, \xi_t^L) + f_x(t, \xi_t^L)a_t^L + f_{xx}(t, \xi_t^L)(b_t^L)^2/2$ and $B_t^L = f_x(t, \xi_t^L)b_t^L$. Note that (5.37) implies immediately

$$\lim_{L\to\infty} \int_0^T (|a_t^L - a_t| + (b_t^L - b_t)^2)\mathrm{d}t = 0 \quad \text{a.s.}$$

Taking this into account we show that

$$\sup_{0\le t\le T} |\xi_t^L - \xi_t| \xrightarrow{P} 0, \quad L \to \infty. \tag{5.46}$$

Indeed, from the definitions of ξ and ξ^L we obtain that

$$\sup_{0\le t\le T} |\xi_t^L - \xi_t| \le \int_0^T |a_t^L - a_t| \mathrm{d}t + \sup_{0\le t\le T} \left| \int_0^t (b_t^L - b_t) \mathrm{d}W_t \right|.$$

Thus for (5.46) it suffices to show that the last term in this inequality tends to zero as $L \to \infty$. By Lemma 4.6, p. 102 in Liptser and Shiryaev [9]) we obtain for every $\varepsilon > 0$

$$\mathrm{P}\left(\sup_{0\le t\le T} \left| \int_0^t (b_t^L - b_t) \mathrm{d}W_t \right| \ge \delta \right) \le \frac{\varepsilon}{\delta^2} + \mathrm{P}\left(\int_0^T (b_t^L - b_t)^2 \mathrm{d}t \ge \varepsilon \right).$$

This implies (5.46). Taking now the limit in (5.45) for L to infinity we obtain (5.38). □

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Financ. **9**, 203–228 (1999)
2. Basak, S., Shapiro, A.: Value at Risk based risk management: optimal policies and asset prices. Rev. Financ. Stud. **14**, 371–405 (1999)
3. Gabih, A., Grecksch, W., Wunderlich, R.: Dynamic portfolio optimization with bounded shortfall risks. Stoch. Anal. Appl. **23**, 579–594 (2005)
4. Emmer, S., Klüppelberg, C., Korn, R.: Optimal portfolios with bounded Capital-at-Risk. Math. Financ. **11**, 365–384 (2001)
5. Jorion, P.: Value at Risk. McGraw-Hill, New York (2001)
6. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, Berlin (1988)
7. Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance. Springer, Berlin (2001)
8. Korn, R.: Optimal Portfolios. Singapore, World Scientific (1997)
9. Liptser, R.S., Shirayev, A.N.: Statistics of Random Processes I. General Theory. Springer, New York (1977)
10. Merton, R.C.: Continuous-Time Finance. Blackwell, Cambridge (1990)
11. Touzi, N.: Stochastic Control Problems, Viscosity Solutions and Applications to Finance. Publications of the Scuola Normale Superiore of Pisa. Scuola Normale Superiore, Pisa (2004)

On Comparison Theorem and its Applications to Finance

Vladislav Y. Krasin and Alexander V. Melnikov

Abstract This paper studies a comparison theorem for solutions of stochastic differential equations and its generalization to the multi-dimensional case. We show, that even though the proof of the generalized theorem follows that of the one-dimensional comparison theorem, the multi-dimensional case requires a different condition on the drift coefficient, known in the theory of differential equations as Kamke-Wazewski condition. We also present several examples of possible applications to option price estimation in finance.

Keywords Comparison theorem · Option pricing · Stochastic differential equations

Mathematics Subject Classification (2000) 60G07 · 65C50 · 91B28

1 Introduction

Stochastic domination theorems play an important role in the theory of stochastic processes and their applications. We focus our attention on results, establishing path-wise almost surely dominance (that is when one process is greater than or equal to another with probability one), which in the theory of stochastic differential equations are referred to as the comparison theorems. These type of theorems are used in a wide range of mathematical problems: from existence and uniqueness of solutions of SDE's to asymptotic behavior (see, for example, [13]). The first comparison theorem for diffusions was proven in [17], and later generalized in [20]. That result was extended to continuous semimartingale case in [11], and in [5] to semimartingales with jumps. The discontinuous case was also considered in [1] and [16].

All the above mentioned works study one-dimensional processes. Extension of the comparison theorem to the multi-dimensional case demands an additional condition on drift coefficient. It was done in [12] for SDE's with respect to a continuous semimartingale. Similar results have been proven in [7] for diffusion equations and in [3] for inequalities.

V.Y. Krasin · A.V. Melnikov (✉)
Department of Mathematical and Statistical Sciences, University of Alberta,
632 Central Academic Building, Edmonton, AB T6G 2G1, Canada
e-mail: melnikov@ualberta.ca

V.Y. Krasin
e-mail: vkrasin@math.ualberta.ca

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_8,

Our first goal is to give the proof of the multi-dimensional comparison theorem, where we add jumps to the above mentioned cases. Presence of jumps can plays an important role in domination/comparison theorems: for example, as shown in [19], a well-known Hajek's mean comparison theorem (see [8]) does not hold in the Poisson case.

One of the main disadvantages of these theorems is that they only allow comparison of processes with identical diffusion coefficients. Moreover, as shown in [15], if initial conditions are not specified, then identical diffusion coefficient is a necessity for comparison of one-dimensional SDE's. The papers [14] and later [6] develop a comparison theorem for SDE's with different diffusions, which requires more specific conditions on the initial values. For a given stochastic process X_t with diffusion coefficient $g(X_t)$ they consider $F(X_t)$, where $F(x) = \int_{X_0}^{x} \frac{1}{g(u)} du$. Using Ito's formula it is easy to show that the new processes will have diffusion coefficient equal to 1. We use that technique in Example 1 in Sect. 3 of this paper.

Our second goal is to show how the comparison theorem can be used in mathematical finance, mainly to estimate option prices. There are few publications, where this question was studied before, for example the above mentioned Hajek's mean comparison theorem was used for this purpose in [9]. We present several examples illustrating how the comparison theorem can be used in practice.

2 Comparison Theorem

As mentioned above, the comparison theorem was proven for discontinuous semimartingales in one-dimensional case, as well as continuous semimartingales in the multi-dimensional case. In this section we combine the two results to prove the multi-dimensional comparison theorem for strong solutions of stochastic differential equations with respect to semimartingales which contain a jump component.

Let $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t)_{t\geq 0}, \mathbf{P})$ be a standard stochastic basis (see [10] for details). Denote $\mathscr{A}$ as a set of increasing processes of finite, integrable variation and $\mathscr{A}_{loc}$ a set of processes, whose localizations belong to $\mathscr{A}$.

Let all processes below be cadlag and adapted to the given filtration. Defined are a d-dimensional non-decreasing continuous process $A = (A^1, A^2, \dots, A^d)$ with $A_i \in \mathscr{A}_{loc}$, d-dimensional continuous local martingale $M = (M^1, M^2, \dots, M^d)$ and a d-dimensional jump measure $\mu = (\mu^1, \mu^2, \dots, \mu^d)$ with compensators $(\nu^1, \nu^2, \dots, \nu^d)$. To simplify our considerations we assume that ν^i are continuous.

Consider stochastic differential equations:

$$dX_t^i = \sum_{j=1}^{d} f_{ij}(X_{t-}) dA_t^j + \sum_{j=1}^{d} g_{ij}(X_{t-}^i) dM_t^j$$

$$+ \mathbf{I}_{\{|u|\leq 1\}} h_i(u, X_{t-}^i) d(\mu_t^i - \nu_t^i) + \mathbf{I}_{\{|u|>1\}} k_i(u, X_{t-}^i) d\mu_t^i \quad (1)$$

$$d\tilde{X}_t^i = \sum_{j=1}^{d} \tilde{f}_{ij}(\tilde{X}_{t-})dA_t^j + \sum_{j=1}^{d} g_{ij}(\tilde{X}_{t-}^i)dM_t^j$$
$$+ \mathbf{I}_{\{|u|)\le 1\}} h_i(u, \tilde{X}_{t-}^i)d(\mu_t^i - \nu_t^i) + \mathbf{I}_{\{|u|>1\}} \tilde{k}_i(u, \tilde{X}_{t-}^i)d\mu_t^i \tag{2}$$

where f_{ij}, g_{ij}, h_i and k_i depend on x, t and ω and are continuous in (t, x) (we write $f(x)$ instead of $f(x, t, \omega)$ to shorten notations). From now on we assume (without going into details) that all considered SDE's admit unique strong solutions.

For further considerations we need a non-negative nondecreasing function $\rho(x)$ such that $\int_0^\varepsilon \rho^2(x)dx = \infty$ for any $\varepsilon > 0$. Define $\{a_n\}$ as a decreasing sequence of positive numbers with $a_0 = 1$ such that $\int_{a_{n+1}}^{a_n} \rho^{-2}(x)dx = n + 1$, $n = 0, 1, 2, \ldots$. We assume that there exists a sequence of positive numbers $\{\varepsilon_n\}$ such that $\varepsilon_n \le a_{n-1} - a_n$ and

$$n^{-1}\rho^2(a_{n-1})\rho^{-2}(a_{n-1} - \varepsilon_n) \to 0 \quad \text{as } n \to \infty. \tag{3}$$

We note that any Hölder function of order $\alpha > 1/2$ will satisfy the above assumption.

Theorem 1 *Assume that for all i and j let functions f_{ij}, g_{ij}, h_i and k_i satisfy*

$$\tilde{f}_{ij}(\tilde{X}_0) > f_{ij}(X_0) \tag{4}$$

$$\tilde{f}_{ij}(\tilde{x_1}, \ldots, \tilde{x}_{i-1}, x_i, \tilde{x}_{i+1}, \ldots) > f_{ij}(x_1, \ldots, x_i, \ldots)$$
$$\text{for } \tilde{x}_k \ge x_k,\ k = 1, \ldots, d. \tag{5}$$

There exist non-negative predictable processes G_t and $H_t(u)$ such that

$$|g_{ij}(y, t) - g_{ij}(x, t)| \le G_t \rho(|y - x|) \tag{6}$$
$$|h_i(u, x, t) - h_i(u, y, t)| \le \rho(|x - y|)H_t(u). \tag{7}$$

For all $y \ge x$

$$h_i(u, y) \ge h_i(u, x) \tag{8}$$
$$\tilde{k}_i(u, y) \ge k_i(u, x) \tag{9}$$

along with a technical conditions: the processes $\int_0^t |g_{kl}(X_{s-}^k)|d\langle M_s^i, M_s^j\rangle$, $\int_0^t |f_{kl}(X_{s-})|dA_s^i$, $\int_0^t |\tilde{f}_{kl}(\tilde{X_{s-}})|dA_s^i$, as well as $\int_0^t |G_s|^2 d\langle M_s^i, M_s^j\rangle$ and $\int_0^t \int_E |H_s(u)|^2 \nu^i(ds, du)$ belong to $\mathscr{A}_{loc}$ (here and everywhere below $E = R^d \setminus \{0\}$).

Then $\tilde{X}_t \ge X_t$ for all t (a.s.).

Proof Without loss of generality assume that the above mentioned processes belong to $\mathscr{A}$, otherwise we can provide their localization.

Following arguments presented in Lemma 1 of [5] and using condition (9) we can see that it is sufficient to consider processes without big jumps (or make $k_i = \tilde{k}_i = 0$).

Consider stopping times

$$T_i = \inf\big(t > 0 : f_{ij}(X_t) > \tilde{f}_{ij}(\tilde{X}_t) \text{ for at least one } j\big).$$

Denote $T = \min(T_i)$ and $\tau = T \wedge t$. Since f_{ij} are continuous, X_t and $\tilde{X}_t$ are right-continuous and $f_{ij}(X_0) < \tilde{f}_{ij}(\tilde{X}_0)$, then $T_i > 0$ and, therefore $\tau > 0$ (a.s.).

Now we consider a sequence of twice continuously differentiable functions $\varphi_n(x)$ such that $\varphi_n(x) \uparrow |x|$ as $n \to \infty$. Such sequence was originally proposed in [20]. Here we follow the procedure presented in [5]:

let $\{\psi_n(x)\}_{n\in N}$ be a sequence of non-negative functions with $\operatorname{supp}\psi_n \subseteq (a_n, a_{n-1})$ and

$$\int_{a_n}^{a_{n-1}} \psi_n(x)dx = 1,$$

$$\psi_n(x) \le (2/n) \times \rho^{-2}(x)$$

$$\arg\max \psi_n = a_{n-1} - \varepsilon_n.$$

Define $\varphi_n(x) = \int_0^{|x|} \int_0^y \psi_n(s)dsdy$. It is easy to see that $|\varphi_n'(x)| \le 1$

Applying Itô's formula to $\varphi_n(\tilde{X}^i_\tau - X^i_\tau)$ we get

$$\varphi_n(\tilde{X}^i_\tau - X^i_\tau) = \text{ local martingale } + I_1 + I_2 + I_3, \tag{10}$$

where

$$I_1 = \int_0^\tau \varphi_n'(\tilde{X}^i_{s-} - X^i_{s-}) \sum_{j=1}^d (\tilde{f}_{ij}(\tilde{X}_{s-}) - f_{ij}(X_{s-}))dA^j_s$$

$$I_2 = 1/2 \int_0^\tau \varphi_n''(\tilde{X}^i_{s-} - X^i_{s-})$$

$$\times \sum_{j,k=1}^d (g_{ij}(\tilde{X}^i_{s-}) - g_{ij}(X^i_{s-}))(g_{ik}(\tilde{X}^i_{s-}) - g_{ik}(X^i_{s-}))d\langle M^j_s, M^k_s\rangle$$

$$I_3 = \int_0^\tau \int_E \mathbf{I}_{\{|u|>1\}}\big[\varphi_n(\tilde{X}^i_{s-} - X^i_{s-} + h_i(u, \tilde{X}^i_{s-}) - h_i(u, X^i_{s-}))$$

$$- \varphi_n(\tilde{X}^i_{s-} - X^i_{s-}) - \varphi_n'(\tilde{X}^i_{s-} - X^i_{s-})(h_i(u, \tilde{X}^i_{s-}) - h_i(u, X^i_{s-}))\big]\nu^i(ds, du).$$

We do not specify the structure of the local martingale in (10) because it is irrelevant for the proof.

Without resorting to additional localization we assume that all terms in (10) admit expectations. Taking expectations of both sides of (10) yields

$$\mathbf{E}\varphi_n(\tilde{X}^i_\tau - X^i_\tau) = \mathbf{E}I_1 + \mathbf{E}I_2 + \mathbf{E}I_3. \tag{11}$$

Moreover,

$$\mathbf{E}I_1 \le \mathbf{E}\int_0^\tau \sum_{j=1}^d (\tilde{f}_{ij}(\tilde{X}_{s-}) - f_{ij}(X_{s-}))dA_s^j = \mathbf{E}(\tilde{X}_\tau^i - X_\tau^i)$$

$$\mathbf{E}I_2 \le 1/2 \sum_{j,k=1}^d \max_{a_n \le x \le a_{n-1}} [\psi_n(|x|)\rho^2(|x|)]\mathbf{E}\int_0^\tau |G_s^2| d\langle M_s^j, M_s^k\rangle$$

$$\le 1/n \sum_{j,k=1}^d \mathbf{E}\int_0^\infty |G_s^2| d\langle M_s^j, M_s^k\rangle \to 0 \quad \text{as } n \to \infty.$$

Denote $h(u, \tilde{X}_{s-}^i) - h(u, X_{s-}^i)$ by Δh_i. The Taylor approximation formula implies that there exists $0 \le \alpha \le 1$ such that

$$\mathbf{E}I_3 = 1/2\mathbf{E}\int_0^\tau \int_E \mathbf{I}_{\{|u|>1\}}\varphi_n''(\tilde{X}_{s-}^i - X_{s-}^i + \alpha\Delta h_i)(\Delta h_i)^2 \nu^i(ds, du)$$

$$\le 1/2\mathbf{E}\int_0^\tau \int_E \mathbf{I}_{\{|u|>1\}}|H_s(u)|^2 \rho^2(|\tilde{X}_{s-}^i - X_{s-}^i|)$$

$$\times \varphi_n''(\tilde{X}_{s-}^i - X_{s-}^i + \alpha\Delta h_i)\nu^i(ds, du).$$

It is possible to show (see Theorem 1 of [5] for details) that

$$\rho^2(|\tilde{X}_{s-}^i - X_{s-}^i|)\varphi_n''(\tilde{X}_{s-}^i - X_{s-}^i + \alpha\Delta h_i) \le n^{-1}\rho^2(a_{n-1})\rho^{-2}(a_{n-1} - \varepsilon_n),$$

and condition (3) implies that $\mathbf{E}I_3$ converges to 0 as $n \to \infty$.

Using the above considerations and taking limits as $n \to \infty$ in (11) we get

$$\mathbf{E}|\tilde{X}_\tau^i - X_\tau^i| \le \mathbf{E}(\tilde{X}_\tau^i - X_\tau^i),$$

and therefore

$$\tilde{X}_\tau^i \ge X_\tau^i \quad \text{(a.s.)} \tag{12}$$

Now consider stopping times $\theta_i = \inf(t > \tau : \tilde{X}_t^i < X_t^i)$, and denote $\theta = \min(\theta_i)$.

If $\theta(\omega) = \infty$ then the proof is complete. Otherwise $\tilde{X}_{\theta-}^i \ge X_{\theta-}^i$ for all i by definition of θ. If there is no jump at time θ, then $\tilde{X}_{\theta-}^i = \tilde{X}_\theta^i \ge X_\theta^i$. Otherwise,

$$\tilde{X}_\theta^i = \tilde{X}_{\theta-}^i + I_{\{|u|\le 1\}}h_i(u, \tilde{X})$$

$$X_\theta^i = X_{\theta-}^i + I_{\{|u|\le 1\}}h_i(u, X),$$

and it follows from (8) that $\tilde{X}_\theta^i \ge X_\theta^i$ for all $i = 1, \ldots, d$.

Now fix i and limit the considerations below to a set $B_i = \{\omega | \theta(\omega) = \theta_i(\omega) < \infty\}$. It follows from definition of θ_i and right-continuity of $\tilde{X}_t$ and X_t that $\tilde{X}_{\theta_i}^i \le X_{\theta_i}^i$ for all $\omega \in B_i$. Therefore, $\tilde{X}_{\theta_i}^i = X_{\theta_i}^i$ (a.s.).

For $j \neq i$ we have that $\theta_j \geq \theta_i$ and, therefore, $\tilde{X}^j_{\theta_i} \geq X^j_{\theta_i}$. Condition (5) implies that $\tilde{f}_{ij}(\tilde{X}_{\theta_i}) > f_{ij}(X_{\theta_i})$ (a.s.) for all j. Define $\eta_i = \inf(t > \theta_i : \tilde{f}_{ij}(\tilde{X}_t) < f_{ij}(X_t)$ for at least one j) and it follows from above that $\eta_i > \theta_i$ (a.s.).

Reproducing the above argument for $\varphi_n(\tilde{X}^i - X^i)$ on time interval $[\theta_i, \eta_i]$ we get that there exist a stopping time $\tau_i > \theta_i$, such that $\tilde{X}^i_t \geq X^i_t$ on $\theta_i < t \leq \tau_i$ which contradicts the definition of θ_i. Therefore, $B_i = \emptyset$ (a.s.) and thus $\theta = \infty$ (a.s.). □

The multi-dimensional comparison theorem has to be used when dealing with processes, whose dynamics are influenced by each other. As an example of such processes one can consider voting and non-voting stocks of the same corporation. It seems very reasonable to assume that dynamics of voting stock prices is influenced by non-voting stock prices and vice versa.

The question of price dependencies for stocks with different voting rights has been studied before. For example in [4] the following discrete model is tested:

$$\begin{aligned} s^v_t &= \ln S^v_t \\ s^n_t &= \ln S^n_t \\ s^n_t &= \alpha + \beta s^v_t + \gamma t + u_t, \quad t = 1, 2, \ldots, \end{aligned} \tag{13}$$

where S^v_t and S^n_t are voting and non-voting stock prices respectively, α, β and γ are constants and u_t is a long-memory process.

For continuous case let the stock prices satisfy the following system of SDE's

$$dS^v_t = S^v_t(\mu^v(t, S^v_t, S^n_t)dt + \sigma^v(t, S^v_t)dW^v_t) \tag{14}$$

$$dS^n_t = S^n_t(\mu^n(t, S^v_t, S^n_t)dt + \sigma^n(t, S^n_t)dW^n_t). \tag{15}$$

As an analogy of (13) in this case one can assume that:

$$ds^n_t = \beta ds^v_t + \gamma dt + \sigma dW_t \tag{16}$$

where W_t is a new Wiener process. Applying Itô's formula to (14) and (15) we get:

$$\begin{aligned} ds^v_t &= \big(\mu^v(t, S^v_t, S^n_t) - 1/2(\sigma^v(t, S^v_t))^2\big)dt + \sigma^v(t, S^v_t)dW^v_t \\ ds^n_t &= \big(\mu^n(t, S^v_t, S^n_t) - 1/2(\sigma^n(t, S^n_t))^2\big)dt + \sigma^n(t, S^n_t)dW^n_t. \end{aligned}$$

Equation (16) then becomes

$$ds^n_t = \big(\beta\mu^v(t, S^v_t, S^n_t) - \beta/2(\sigma^v(t, S^v_t))^2 + \gamma\big)dt + \beta\sigma^v(t, S^v_t)dW^v_t + \sigma dW_t.$$

Equating the dt terms in two representations of ds^n_t we get:

$$\beta\mu^v(t, S^v_t, S^n_t) - \beta/2(\sigma^v(t, S^v_t))^2 + \gamma = \mu^n(t, S^v_t, S^n_t) - 1/2(\sigma^n(t, S^n_t))^2,$$

which shows, that the stock price dynamics (14) and (15) are indeed related.

Above considerations show that two or more dimensional processes can be used in mathematical finance. The multi-dimensional comparison theorem can then be used to estimate option prices in this model using some other pair of processes.

3 Applications to Mathematical Finance

In this section we give specific illustrations showing how the stochastic domination/comparison theorems can be used in mathematical finance. For the purpose of demonstration we shall restrict our attention to the simplest cases only.

Example 1 (Constant elasticity of variance (CEV) model) The model was proposed by Cox and Ross [2] and is often used in mathematical finance. The stock price is said to satisfy the following SDE

$$dS_t = S_t(\mu dt + \sigma S_t^{\alpha-1} dW_t), \qquad S_0 = s \tag{17}$$

where $0 < \alpha < 1$ and W_t is a Wiener process.

Let $F(x) = \sigma^{-1}\int_s^x u^{-\alpha} du = (x^{1-\alpha} - s^{1-\alpha})/(\sigma(1-\alpha))$. Denote $X_t = F(S_t)$ and applying Itô's formula we get

$$\begin{aligned} dX_t &= F'(S_t)dS_t + 1/2F''(S_t)(dS_t)^2 \\ &= (S_t^{1-\alpha}\mu/\sigma - \alpha\sigma/2S_t^{\alpha-1})dt + dW_t \\ &= \mu\big((1-\alpha)X_t + \sigma^{-1}s^{1-\alpha}\big)dt - \alpha\big(2(1-\alpha)X_t + 2\sigma^{-1}s^{1-\alpha}\big)^{-1}dt + dW_t. \end{aligned}$$

The comparison theorem is used to estimate the process X_t from above by a new process Y_t, satisfying the equation

$$dY_t = \mu(\sigma(1-\alpha)Y_t + \sigma^{-1}s^{1-\alpha})dt + dW_t, \qquad Y_0 = 0$$

which is an Ornstein-Uhlenbeck process.

The process of this type is also used in [18] as an interest rate model and has normal distribution.

Applying the comparison theorem to X_t and Y_t yields

$$Y_t \geq X_t = F(S_t) \quad \text{(a.s.)}$$

and so

$$S_t \leq F^{-1}(Y_t) \quad \text{(a.s.)} \tag{18}$$

Consider an increasing function f and an option with payoff $f(S_T)$. Assuming zero interest rates, the price of such option is given by $\tilde{\mathbf{E}}f(S_T)$ for an appropriate martingale measure $\tilde{P}$. But, inequality (18) implies that $\tilde{\mathbf{E}}f(S_T) \leq \tilde{\mathbf{E}}f(F^{-1}(Y_T))$ and thus we obtain an estimate for the option price. Due to normality of Y_T this estimate is easily computable.

Example 2 (A complete model with stochastic volatility) Consider the processes

$$\begin{aligned} dS_t &= S_t\sigma_t dW_t \\ d\sigma_t &= a(\sigma_t)dt + \delta\sigma_t dW_t \end{aligned}$$

where $a(x)$ is a continuous function, such that:

$$a(x) + \delta/2x \geq 0 \quad \text{for } x \geq 0. \tag{19}$$

We also require that $a(0) > 0$ to ensure that $\sigma_t \geq 0$ (a.s.).

Denote $Y_t = \delta^{-1}(\sigma_t - \sigma_0) - \ln(S_t/S_0)$ and applying Itô's formula we have

$$dY_t = \delta^{-1} d\sigma_t - \sigma_t dW_t + 1/2\sigma_t dt = (\delta^{-1} a(\sigma_t)\delta + 1/2\sigma_t)dt.$$

Consider the process $Z_t = \sigma_t \exp[(\sigma_t - \sigma_0) - \ln(S_t/S_0)] = \sigma_t \exp[\delta Y_t]$ and using again Itô's formula we find that

$$\begin{aligned} dZ_t &= \exp[\delta Y_t] d\sigma_t + \delta Z_t dY_t \\ &= \{\exp[\delta Y_t] a(\sigma_t) + Z_t(a(\sigma_t) + 1/2\delta\sigma_t)\}dt + \delta Z_t dW_t \\ &= Z_t f(\sigma_t)dt + \delta Z_t dW_t \end{aligned}$$

where $f(y) = a(y)/y + a(y) + 1/2\delta y$.

If $Z_t f(\sigma_t)|_{Z_t=\sigma_t} > a(\sigma_t)$, then by the comparison theorem $Z_t \geq \sigma_t$ (a.s.).

Calculating $Z_t f(\sigma_t)|_{Z_t=\sigma_t} = a(\sigma_t) + \sigma_t(a(\sigma_t) + 1/2\delta\sigma_t)$ we observe that inequality (19) implies $Z_t \geq \sigma_t$, which is equivalent to $\delta^{-1}(\sigma_t - \sigma_0) - \ln(S_t/S_0) \geq 0$, or $\sigma_t \geq (\sigma_0 + \delta \ln(S_t/S_0))$.

Now consider a new process X_t such that

$$dX_t = (\sigma_0 + \delta \ln(S_t/S_0))dW_t, \qquad X_0 = S_0.$$

Recall (see [8]) that for two processes satisfying

$$\begin{aligned} dR_t &= \gamma(R_t)dW_t \\ dY_t &= \beta_t dW_t \end{aligned}$$

the Hajek's comparison theorem implies that $\mathbf{E}h(R_t) \leq \mathbf{E}h(Y_t)$ for any convex function $h(x)$ if $\beta_t \geq \gamma(Y_t)$. Applying the theorem to X_t and S_t we obtain that $\mathbf{E}h(X_T) \leq \mathbf{E}h(S_T)$ for any convex payoff function $h(x)$.

Example 3 (Pricing stock option using forwards) Consider a market with stochastic interest rate r_t and the stock price satisfying the following SDE's

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t^s) \tag{20}$$

$$dr_t = \alpha_t dt + \beta_t dW_t^r, \qquad dW_t^s dW_t^r = \rho dt \tag{21}$$

where $\rho \in (-1, 1)$ and the bank account is given by $D_t = \exp(\int_0^t r_s ds)$. To price an option with payoff $f_\tau = f(S_\tau)$ we will use bond $B_t(T)$ and a forward contact on the stock, whose price is $F_t(T) = S_t/B_t(T)$ with the same maturity $T > \tau$. The following question arises: How does the choice of T affect the option price?

To answer this question assume that dynamics of the bond price is given by

$$dB_t(T) = B_t(T)(r_t dt + \delta_t(T) dW_t^r). \tag{22}$$

According to Ito's formula we have

$$\begin{aligned} dF_t(T) = F_t(T)[(\mu_t - r_t - \sigma_t\delta_t(T)\rho + 1/2\delta_t^2(T))dt \\ + \sigma_t dW_t^s - \delta_t(T) dW_t^r]. \end{aligned} \tag{23}$$

The option price then has to be computed under an equivalent measure, for which both discounted tradable processes ($B_t(T)/D_t$ and $F_t(T)/D_t$) are martingales. Since $B_t T/D_t$ is already a martingale, the unique equivalent martingale measure $\tilde{P}(T)$ is defined by

$$\sigma_t dW_t^s + (\mu_t - r_t - \sigma_t\delta_t(T)\rho + 1/2\delta_t^2(T))dt = r_t dt + \sigma_t d\tilde{W}_t^s,$$

where $\tilde{W}_t^S$ is a $\tilde{P}(T)$-Wiener process.

Under this new measure the stock price dynamics can be expressed as

$$dS_t = S_t\big((2r_t + \sigma_t\delta_t(T)\rho - 1/2\delta_t^2(T))dt + \sigma_t d\tilde{W}_t^s\big).$$

We note that the option price will be equal to $\tilde{\mathbf{E}}_T(f(S_\tau)/B_\tau)$ where $\tilde{\mathbf{E}}_T$ denotes expectation under $\tilde{P}(T)$.

We see that as far as option pricing goes, changing bond and forward maturity T is equivalent to changing drift coefficient of the stock price.

Suppose $T_1 > \tau$, $T_2 > \tau$ and

$$\sigma_t\delta_t(T_1)\rho - 1/2\delta_t^2(T_1) > \sigma_t\delta_t(T_2)\rho - 1/2\delta_t^2(T_2). \tag{24}$$

Consider two processes

$$dS_t^1 = S_t^1\big((2r_t + \sigma_t\delta_t(T_1)\rho - 1/2\delta_t^2(T_1))dt + \sigma_t dW_t\big)$$
$$dS_t^2 = S_t^2\big((2r_t + \sigma_t\delta_t(T_2)\rho - 1/2\delta_t^2(T_2))dt + \sigma_t dW_t\big)$$

with initial conditions $S_0^1 = S_0^2 = S_0$. Applying the comparison theorem we see that $S_T^1 \geq S_T^2$ (a.s.) and thus $\mathbf{E}(f(S_T^1)/B_T) > \mathbf{E}(f(S_T^1)/B_T)$ for an increasing function $f(x)$.

It is clear that the distributions of S_t^i are the same as the $\tilde{P}(T_i)$-distributions of S_t. Therefore,

$$\tilde{\mathbf{E}}_{T_1} f(S_T) = \mathbf{E} f(S_T^1) > \mathbf{E} f(S_T^2) = \tilde{\mathbf{E}}_{T_2} f(S_T)$$

and we can compare two option prices for different bond and forward maturities T_1 and T_2.

As a particular example we consider the Vasicek interest rate model:

$$dr_t = a(b - r_t)dt + \sigma^r dW_t^r.$$

The bond price will be equal to

$$B_t(T) = \exp[\gamma(t,T) - r_t a^{-1}(1 - e^{-a(t-T))})]$$

for some deterministic function $\gamma(t,T)$ (see [18] for details)

In this case the expression for $\delta_t(T)$ is

$$\delta_t(T) = -a^{-1}\sigma^r(1 - e^{-a(t-T)}).$$

Taking $\rho = 0$ for simplicity, it is easy to see that $-1/2\delta_t^2(T)$ is an increasing function of T. Thus, inequality (24) holds for $T_1 > T_2$. Therefore, increasing the bond and forward maturity T is equivalent to increasing the option price.

Acknowledgements We would like to thank two anonymous referees for their remarks, which were a great help to us.

References

1. Bassan, B., Çinlar, E., Scarsini, M.: Stochastic comparison of Itô processes. Stoch. Process. Appl. **45**, 1–11 (1993)
2. Cox, J.C., Ross, S.A.: The valuation of options for alternative stochastic processes. J. Financ. Econ. **3**, 145–166 (1976)
3. Ding, X., Wu, R.: A new proof for comparison theorems for stochastic differential inequalities with respect to semimartingales. Stoch. Process. Appl. **78**, 155–171 (1998)
4. Dittmann, I.: Fractional cointegration of voting and non-voting shares. Appl. Financ. Econ. **11**(3), 321–332 (2001)
5. Gal'chuk, L.I.: A comparison theorem for stochastic equations with integral with respect to martingales and random measures. Theory Probab. Appl. **27**(3), 450–460 (1982)
6. Gal'chuk, L.I., Davis, M.H.A.: A note on a comparison theorem for equations with different diffusions. Stochastics **6**(2), 147–149 (1982)
7. Geiß, C., Manthey, R.: Comparison theorems for stochastic differential equations in finite and infinite dimensions. Stoch. Process. Appl. **53**, 23–35 (1994)
8. Hajek, B.: Mean stochastic comparison of diffusions. Z. Wahrscheinlichkeitstheor. Verw. Geb. **68**, 315–329 (1985)
9. Henderson, V.: Price comparison results and superreplication: an application to passport options. Appl. Stoch. Models Bus. Ind. **16**, 297–310 (2000)
10. Jacod, J., Shiryayev, A.N.: Limit Theorems for Stochastic Processes. Springer, Berlin (1987)
11. Melnikov, A.V.: On the theory of stochastic equations in components of semimartingales. Mat. Sb. **110**(3), 414–427 (1979). In English: Sb. Math. **38**(3), 381–394 (1981)
12. Melnikov, A.V.: On solutions of stochastic equations with driving semimartingales. In: Proceedings of the Third European Young Statisticians Meeting, Catholic University, Leuven, pp. 120–124 (1983)
13. Melnikov, A.V.: Stochastic differential equations: singularity of coefficients, regression models and stochastic approximation. Russ. Math. Surv. **51**(5), 43–136 (1996)
14. O'Brien Ci, L.: A new comparison theorem for solution of stochastic differential equations. Stochastics **3**, 245–249 (1980)
15. Peng, S., Zhu, X.: Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations. Stoch. Process. Appl. **116**, 370–380 (2006)
16. Situ, R.: Reflecting Stochastic Differential Equations with Jumps and Applications. Chapman Hall/CRC, Boca Raton (2000)

17. Skorokhod, A.V.: Studies in the Theory of Random Process (1961). Addison-Wesley, Reading (1965) (in English)
18. Vasicek, O.A.: An equilibrium characterization of the term structure. J. Financ. Econ. **5**(2), 177–188 (1977)
19. Večeř, J., Xu, M.: The mean comparison theorem cannot be extended to the Poisson case. J. Appl. Probab. **41**(4), 1199–1202 (2004)
20. Yamada, T.: On comparison theorem for solutions of stochastic differential equations and its applications. J. Math. Kyoto Univ. **13**, 497–512 (1973)

Examples of FCLT in Random Environment

R. Liptser

Abstract We consider a diffusion type process being a weak solution of Itô's equation

$$dX_t^\varepsilon = b(\omega, X_t^\varepsilon/\varepsilon)dt + \sigma(\omega, X_t^\varepsilon/\varepsilon)dB_t$$

relative to fixed initial condition $X_0^\varepsilon = x_0$, Brownian motion $(B_t)_{t\geq 0}$, and ergodic stationary random process $(b(\omega,u),\sigma(\omega,u))_{u\in\mathbb{R}}$, treated as a "random environment", where ε is a small positive parameter. Random environment and Brownian motions are independent random objects. Functions $b(\omega,u),\sigma(\omega,u)$ are uniformly bounded and function $\sigma^2(\omega,u)$ is uniformly positive. Random environment obeys some "weak dependence" property (see (9)). We show that the family $\{(X_t^\varepsilon)_{t\leq T}\}_{\varepsilon\to 0}$ converges in law to a continuous Gaussian process $X=(X_t)_{t\leq T}$ with the expectation $\mathsf{E}X_t = x_0 + \mathbf{b}t$ and covariance $\operatorname{cov}(X_t, X_s) = \mathbf{a}(t\wedge s)$, where

$$\mathbf{a} = 1\Big/\mathsf{E}\frac{1}{\sigma^2(\omega,0)}, \qquad \mathbf{b} = \mathsf{E}\frac{b(\omega,0)}{\sigma^2(\omega,0)}\Big/\mathsf{E}\frac{1}{\sigma^2(\omega,0)}.$$

Keywords Diffusion approximation · Oscillating environment · Random environment

Mathematics Subject Classification (2000) 60K37

1 Introduction

1. We study a convergence in law, in the uniform metric on $[0,T]$, of the family $\{(X_t^\varepsilon)_{0\leq t\leq T}\}_{\varepsilon\to 0}$ being a weak solutions of Itô's differential equation in "random environment $b(\omega,u),\sigma(\omega,u)$":

$$X_t^\varepsilon = x_0 + \int_0^t b(\omega, X_s^\varepsilon/\varepsilon)ds + \int_0^t \sigma(\omega, X_s^\varepsilon/\varepsilon)dB_s, \tag{1}$$

where ε is a small positive parameter and $(b(\omega,u),\sigma(\omega,u))_{u\in\mathbb{R}}$ is a stationary random processes independent of a Brownian motion $(B_t)_{0\leq t\leq T}$.

The author gratefully acknowledges the anonymous referee who pointed out author's attention on Papanicolau and Varadhan paper [8].

R. Liptser (✉)
Department of Electrical Engineering Systems, Tel Aviv University, 69978 Tel Aviv, Israel
e-mail: liptser@eng.tau.ac.il

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_9, © Springer-Verlag Berlin Heidelberg 2009

It is well known different type of assumptions providing limit theorems in random environment, see Papanicolau and Varadhan [8], Yurinsky [11], and Sznitman and Zeitouni [9], etc. (see also Fannjiang and Komorowski [4], Fannjiang and Papanicolaou [5], Di Masi et al. [3]).

In this paper, we propose an approach for diffusion approximation in random environment serving the case of $b(\omega,u),\sigma(\omega,u)$ with discontinuous paths (see Sect. 5) which is not completely compatible with existing techniques. Closely related results can be found in Butov and Krichagina [2].

2. The prelimit process X^ε_t is a continuous semimartingale, defined on some stochastic basis, with fixed initial point x_0, drift $B^\varepsilon_t=\int_0^t b(X^\varepsilon_s/\varepsilon)ds$ and martingale $M^\varepsilon_t=\int_0^t\sigma(X^\varepsilon_s/\varepsilon)dB_s$ having the quadratic variation process $\langle M^\varepsilon\rangle_t=\int_0^t\sigma^2(X^\varepsilon_s/\varepsilon)ds$. Taking into account well known homogenization ideas, we may assume for a moment that a limit process X_t is also semimartingale with the same initial condition x_0, drift $\mathbf{b}t$ and continuous martingale with the quadratic variation process $\mathbf{a}t$. Assume more that $\mathbf{a}$, $\mathbf{b}$ are known. Then, due to [7] (Theorem 1, Chap. 8, §3, adapted to the case considered) $(X^\varepsilon_t)_{t\le T}\xrightarrow[\varepsilon\to0]{\text{law}}(X_t)_{t\le T}$ in Skorokhod's topology provided that for any $\eta>0$,

$$\lim_{\varepsilon\to0}\mathsf{P}\Big(\sup_{t\le T}|B^\varepsilon_t-\mathbf{b}t|>\eta\Big)=0$$

and

$$\lim_{\varepsilon\to0}\mathsf{P}\Big(\sup_{t\le T}|\langle M^\varepsilon\rangle_t-\mathbf{a}t|>\eta\Big)=0$$

or, equivalently,

$$\begin{aligned}&\lim_{\varepsilon\to0}\mathsf{P}\bigg(\sup_{t\le T}\bigg|\int_0^t[b(X^\varepsilon_s/\varepsilon)-\mathbf{b}]ds\bigg|>\eta\bigg)=0\\&\lim_{\varepsilon\to0}\mathsf{P}\bigg(\sup_{t\le T}\bigg|\int_0^t[\sigma^2(X^\varepsilon_s/\varepsilon)-\mathbf{a}]ds\bigg|>\eta\bigg)=0.\end{aligned}\tag{2}$$

Since in reality $\mathbf{a}$, $\mathbf{b}$ are unknown, in order to guess them we shall verify (2) for "oscillating environment", when $b(\omega,u)\equiv b(u),\sigma(\omega,u)\equiv\sigma(u)$ with smooth periodic functions $b(u),\sigma(u)$ of the period one and uniformly positive $\sigma^2(u)$. We find new presentations for $\int_0^t[b(X^\varepsilon_s/\varepsilon)-\mathbf{b}]ds$ and $\int_0^t[\sigma^2(X^\varepsilon_s/\varepsilon)-\mathbf{a}]ds$ compatible with the proof of (2) and $\mathbf{a}$, $\mathbf{b}$ determination. To this end, we introduce the function

$$H(x)=\int_0^x\int_0^v\theta(s)dsdv,$$

where $\theta(s)$ is a transform of $b(s),\sigma(s)$, and analyze the random process $\varepsilon^2H(X^\varepsilon_t/\varepsilon)$. By the Itô formula applied to $\varepsilon^2H(X^\varepsilon_t/\varepsilon)$, we find that

$$
\begin{aligned}
&\varepsilon^2 \int_0^{X_t^\varepsilon/\varepsilon} \int_0^v \theta(s)dsdv \\
&\quad = \varepsilon^2 \int_0^{x_0/\varepsilon} \int_0^v \theta(s)dsdv + \varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)dsb(X_v^\varepsilon/\varepsilon)dv \\
&\qquad + \varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds\sigma(X_v^\varepsilon/\varepsilon)dB_v + \int_0^t \theta(X_s^\varepsilon/\varepsilon)\sigma^2(X_s^\varepsilon/\varepsilon)ds.
\end{aligned} \tag{3}
$$

We choose $\theta(s)$ such that $\theta(X_s^\varepsilon/\varepsilon)\sigma^2(X_s^\varepsilon/\varepsilon) = \begin{cases} b(X_s^\varepsilon/\varepsilon)-\mathbf{b}, \\ \sigma^2(X_s^\varepsilon/\varepsilon)-\mathbf{a}, \end{cases}$ that is,

$$
\theta(s) = \begin{cases} \frac{\mathbf{b}-b(s)}{\sigma^2(s)} \\ 1 - \frac{\mathbf{a}}{\sigma^2(s)}, \end{cases}
$$

and derive from (3) that

$$
\begin{aligned}
&\begin{cases} \int_0^t [\sigma^2(X_s^\varepsilon/\varepsilon) - \mathbf{a}]ds \\ \int_0^t [b(X_s^\varepsilon/\varepsilon) - \mathbf{b}]ds \end{cases} \\
&\quad = \varepsilon^2 \int_0^{X_t^\varepsilon/\varepsilon} \int_0^v \theta(s)dsdv - \varepsilon^2 \int_0^{x_0/\varepsilon} \int_0^v \theta(s)dsdv \\
&\qquad - \varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)dsb(X_v^\varepsilon/\varepsilon)dv - \varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds\sigma(X_v^\varepsilon/\varepsilon)dB_v.
\end{aligned} \tag{4}
$$

Denote by Ψ_t^ε any term in the right-hand side of (4). Obviously, (2) holds true if for any $\eta > 0$,

$$
\lim_{\varepsilon \to 0} \mathsf{P}\Big(\sup_{t \le T} |\Psi_t^\varepsilon| > \eta\Big) = 0. \tag{5}
$$

Ψ_t^ε contains an integral $\int_0^{\bullet/\varepsilon} \theta(s)ds$ playing a crucial role in the proof of (5) and in $\mathbf{a}$, $\mathbf{b}$ determination.

It is readily to verify that (5) holds if

$$
\sup_{t>0} \left| \int_0^t \theta(s)ds \right| < \infty. \tag{6}
$$

Since $\theta(s)$ is periodic function of the period one, for $\sup_{t>0} |\int_0^t \theta(s)ds| < \infty$ to be valid it suffices to have

$$
\int_0^1 \theta(s)ds = 0. \tag{7}
$$

So, (7) guarantees (5) and simultaneously provides

$$
\mathbf{a} = 1 \Big/ \int_0^1 \frac{1}{\sigma^2(s)}ds, \qquad \mathbf{b} = \int_0^1 \frac{b(s)}{\sigma^2(s)}ds \Big/ \int_0^1 \frac{1}{\sigma^2(s)}ds.
$$

3. Turning back to the random environment, we continue to keep the uniform boundedness of $b(\omega,u),\sigma(\omega,u)$ and the uniform positiveness of $\sigma^2(\omega,u)$. Both formulas for $\theta(s)$ with $b(u),\sigma(u)$ are replaced by $b(\omega,u),\sigma(\omega,u)$ respectively:

$$\theta(s)=\begin{cases}\frac{\mathbf{b}-b(\omega,s)}{\sigma^2(\omega,s)}\\ 1-\frac{\mathbf{a}}{\sigma^2(\omega,s)}.\end{cases}$$

We compute constants **a** and **b** replacing (7) by

$$\mathsf{E}\theta(0)=0. \tag{8}$$

Unfortunately, under the random environment, (8) does not guarantee (5). Therefore, we introduce "weak dependence property" of $\theta(s)$ relative to the filtration $(\mathscr{F}_u)_{u\in\mathbb{R}}$ generated by $(\sigma(\omega,u),b(\omega,u))_{u\in\mathbb{R}}$:

$$\int_0^\infty\sqrt{\mathsf{E}|\mathsf{E}(\theta(u)|\mathscr{F}_0)|^2}du<\infty. \tag{9}$$

4. The paper is organized as follows. The main result is formulated in Sect. 2 and is proved in Sect. 3. The existence of weak solution for (1) is discussed in Sect. 4. Two examples are given in Sects. 5 and 6.

2 Assumptions, Notations and Main Result

The following notations and assumptions are fixed.

2.1 Notations

- * is the transposition symbol for matrices and vectors.
- $\perp\!\!\!\perp$ symbolizes the independence of random objects.
- $(\Omega,\mathscr{F},(\mathscr{F}_u)_{u\in\mathbb{R}},\mathsf{P})$ and $(\Omega,\mathscr{F},(\mathscr{G}_t)_{t\geq0},\mathsf{P})$ are stochastic bases' with general conditions such that $\bigvee_{u\in\mathbb{R}}\mathscr{F}_u\in\mathscr{G}_0$.

2.2 Assumptions

(i) $(\sigma(\omega,u),b(\omega,u))_{u\in\mathbb{R}}$ is stationary ergodic random processes with right continuous paths possessing left limits defined on $(\Omega,\mathscr{F},(\mathscr{F}_u)_{u\in\mathbb{R}},\mathsf{P})$.

(ii) Functions $b(\omega,u),\sigma(\omega,u)$ are uniformly bounded and function $\sigma^2(\omega,u)$ is uniformly positive.

(iii) With

$$\mathbf{a}=1\Big/\mathsf{E}\frac{1}{\sigma^2(\omega,0)},\qquad \mathbf{b}=\mathsf{E}\frac{b(\omega,0)}{\sigma^2(\omega,0)}\Big/\mathsf{E}\frac{1}{\sigma^2(\omega,0)}$$

zero mean random process

$$\theta(s)=\begin{cases}1-\frac{\mathbf{a}}{\sigma^2(\omega,s)}\\ \frac{b(\omega,s)-\mathbf{b}}{\sigma^2(\omega,s)}\end{cases} \tag{10}$$

obeys weak dependence property (9).

(iv) Brownian motion $B=(B_t)_{t\geq 0}$ is defined on $(\Omega,\mathscr{F},(\mathscr{G}_t)_{t\geq 0},\mathsf{P})$; $(B_t)_{t\geq 0}\perp\!\!\!\perp \mathscr{G}_0$, P-a.s.

Let $X=(dX_t)_{t\leq T}$ be a Gaussian random process with the expectation and covariance

$$\mathsf{E}X_t=x_0+\mathbf{b}t \quad\text{and}\quad \operatorname{cov}(X_t,X_s)=\mathbf{a}(t\wedge s).$$

Theorem 1 *Assume* (i)–(iv) *and let* $(X^\varepsilon_t)_{t\leq T}$ *be any weak solution of* (1).
Then, for any $T>0$,

$$(X^\varepsilon_t)_{t\leq T}\xrightarrow[\varepsilon\to 0]{\text{law}}(X_t)_{t\leq T}$$

in Skorokhod's and uniform metrics on $[0,T]$.

3 The Proof of Theorem 1

Notice that (2) provides the theorem statement in the Skorokhod metric. The limit Gaussian process X_t has continuous drift and covariance and, so that, X_t is continuous process. Moreover, $X_t-x_0-\mathbf{b}t$ is Gaussian process with $\operatorname{cov}(X_t,X_s)=\mathbf{a}(t\wedge s)$. Hence, $X_t-x_0-\mathbf{b}t$ is a Wiener process with diffusion parameter $\mathbf{a}$, that is, X_t is a semimartingale on some stochastic basis, i.e., $X_t=x_0+\mathbf{b}t+\sqrt{\mathbf{a}}W_t$, where W_t is a standard Wiener process. Hence, the convergence in law in Skorokhod' metric with this limit process is also valid in uniform metric.

3.1 Auxiliary Lemma

Lemma 1 (Compare with (6))

$$\mathsf{E}\sup_{0<v\leq C}\left|\varepsilon\int_0^{v/\varepsilon}\theta(s)ds\right|^2\leq\varepsilon\,\text{const}.$$

Proof Due to (iii), Lemma 1, §2, Chap. 9 and the corollary to Theorem 3, §11, Chap. 4 in [7] provide Poisson's type decomposition,

$$\int_0^v\theta(s)ds=U_v-U_0-L_v,$$

where $U_v = -\int_v^\infty \mathsf{E}(\theta(s)|\mathscr{F}_v)ds$ is the stationary process and $(L_v)_{v\ge 0}$ is a local martingale on $(\Omega, \mathscr{F}, (\mathscr{F}_u)_{u\in\mathbb{R}}, \mathsf{P})$ with stationary increments and paths from the Skorokhod space $\mathbb{D}_{[0,\infty)}$. Moreover

(1) $\mathsf{E}|U_0|^2 < \infty$,
(2) $\mathsf{E}L_v^2 = 2v\int_0^\infty \mathsf{E}(\theta(s)\theta(0))ds$.

The use of

$$\sup_{0<v\le C}\left|\int_0^{v/\varepsilon}\theta(s)ds\right|^2 \le 3\Big[U_0^2 + \sup_{0<v\le C}U_{v/\varepsilon}^2 + \sup_{0<v\le C}L_{v/\varepsilon}^2\Big]$$

enables us to reduce the proof to

$$\begin{aligned} &\mathsf{E}\varepsilon^2\sup_{0<v\le C}L_{v/\varepsilon}^2 \le \varepsilon\,\text{const} \quad \text{and} \\ &\mathsf{E}\varepsilon^2\sup_{0<v\le C}U_{v/\varepsilon}^2 \le \varepsilon\,\text{const}. \end{aligned} \tag{11}$$

The first part in (11) follows from the maximal Doob inequality

$$\mathsf{E}\varepsilon^2\sup_{v\le C}|L_{v/\varepsilon}|^2 \le 4\varepsilon^2\mathsf{E}L_{C/\varepsilon}^2 = 8C\varepsilon\int_0^\infty \mathsf{E}(\theta(s)\theta(0))ds = \varepsilon\ \text{const}.$$

In order to verify the second part in (11), set $\alpha_i = \sup_{i-1<v\le i}U_v^2$ and notice that $(\alpha_i)_{i\ge 1}$ forms a stationary sequence of random variables. Assume for a moment that $\mathsf{E}\alpha_1 < \infty$. Then, the desired property holds true since

$$\mathsf{E}\sup_{v\le C}|\varepsilon U_{v/\varepsilon}|^2 = \varepsilon^2\mathsf{E}\max_{i\le C/\varepsilon}\alpha_i \le \varepsilon^2\sum_{i\le C/\varepsilon}\mathsf{E}\alpha_i = \varepsilon C\mathsf{E}\alpha_1.$$

Thus, it is left to prove $\mathsf{E}\alpha_1 < \infty$. Write

$$\begin{aligned} \mathsf{E}\alpha_1 &\le 3\Bigg[\mathsf{E}|U_0|^2 + \mathsf{E}\underbrace{\left(\int_0^1|\theta(s)|ds\right)^2}_{\le\int_0^1\theta^2(s)ds} + \underbrace{\mathsf{E}\sup_{v\le 1}|L_v|^2}_{\le 4\mathsf{E}L_1^2}\Bigg] \\ &\le 3\big[\mathsf{E}|U_0|^2 + \mathsf{E}\theta^2(0) + 4\mathsf{E}|L_1|^2\big] \\ &= 3\left[\mathsf{E}|U_0|^2 + \mathsf{E}\theta^2(0) + 8\int_0^\infty \mathsf{E}\theta(s)\theta(0)ds\right] < \infty. \end{aligned}$$

□

3.2 The Proof of (2)

Set $H(x) = \int_0^x\int_0^v\theta(s)dsdv$. If $\theta(s)$ is continuous function, for any fixed ω, the random function $H(x) = H(\omega, x)$ is twice continuously differentiable in x. Then, by the Itô-Wentzell formula, we find that

$$\varepsilon^2 H(X_t^\varepsilon/\varepsilon) = \varepsilon^2 H(x_0/\varepsilon) + \varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds dX_v^\varepsilon + \int_0^t \theta(X_s^\varepsilon/\varepsilon)\sigma^2(X_s^\varepsilon/\varepsilon)ds. \tag{12}$$

Otherwise, when $\theta(s)$ is measurable (bounded) function only, $H(x)$ has Sobolev's second derivative only. Then, taking into account the independence of random environment and Brownian motion, we obtain (12) by Krylov's version, [6], of Itô's (Itô-Wentzell's) formula.

Due to (10) with chosen **a**, **b** we have

$$\left.\begin{array}{l} \theta(s) = 1 - \frac{\mathbf{a}}{\sigma^2(\omega,s)} : \int_0^t [\sigma^2(X_s^\varepsilon/\varepsilon) - \mathbf{a}]ds \\ \theta(s) = \frac{b(\omega,s)-\mathbf{b}}{\sigma^2(\omega,s)} : \int_0^t [b(X_s^\varepsilon/\varepsilon) - \mathbf{b}]ds \end{array}\right\} = \int_0^t \theta(X_s^\varepsilon/\varepsilon)\sigma^2(X_s^\varepsilon/\varepsilon)ds.$$

Hence, according to (12), it is left to prove that for any $\eta > 0$,

$$\lim_{\varepsilon\to 0} \mathsf{P}\Big(\varepsilon^2 \sup_{t\le T} |H(X_t^\varepsilon/\varepsilon)| \ge \eta\Big) = 0 \tag{13}$$

$$\lim_{\varepsilon\to 0} \mathsf{P}\Big(\varepsilon^2 \sup_{t\le T} |H(x_0/\varepsilon)| \ge \eta\Big) = 0 \tag{14}$$

$$\lim_{\varepsilon\to 0} \mathsf{P}\bigg(\varepsilon \sup_{t\le T} \bigg|\int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds dX_v^\varepsilon\bigg| \ge \eta\bigg) = 0. \tag{15}$$

3.2.1 The Tightness $\{\sup_{t\le T} |X_t^\varepsilon|\}_{\varepsilon\to 0}$

Lemma 2 $\lim_{C\to\infty} \overline{\lim}_{\varepsilon\to 0} \mathsf{P}(\sup_{t\le T} |X_t^\varepsilon| > C) = 0$.

Proof Since x_0 is a fixed number and $b(\omega, u)$ is bounded function, the proof is reduced to

$$\lim_{C\to\infty} \overline{\lim_{\varepsilon\to 0}} \mathsf{P}\bigg(\sup_{t\le T}\bigg|\int_0^t \sigma(X_s^\varepsilon/\varepsilon)dB_s\bigg| > C\bigg) = 0$$

and verified with the help of Chebyshev's and maximal Doob's inequalities:

$$\mathsf{P}\bigg(\sup_{t\le T}\bigg|\int_0^t \sigma(X_s^\varepsilon/\varepsilon)dB_s\bigg| > C\bigg) \le \frac{4}{C^2}\mathsf{E}\sup_{t\le T}\bigg|\int_0^t \sigma(X_s^\varepsilon/\varepsilon)dB_s\bigg|^2$$

$$\le \frac{T}{C^2}\mathsf{E}\int_0^T \sigma^2(X_s^\varepsilon/\varepsilon)ds \le \frac{4lT}{C^2} \xrightarrow[C\to\infty]{} 0. \qquad \square$$

3.2.2 The Proof of (13)–(15)

Denote by $\mathfrak{B}^\varepsilon$ any of sets:

- $\{\varepsilon^2 \sup_{t\le T} |H(X_t^\varepsilon/\varepsilon)| \ge \eta\}$
- $\{\varepsilon^2 \sup_{t\le T} |H(x_0/\varepsilon)| \ge \eta\}$
- $\{\varepsilon \sup_{t\le T} | \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds dX_v^\varepsilon| \ge \eta\}$

and by

- $\mathfrak{A}_C = \{\sup_{t\le T} |X_t^\varepsilon| \le C\}$
- $\tau_C = \inf\{t \le T : |X_t^\varepsilon| > C\}$.

Write, $\mathsf{P}(\mathfrak{B}_\varepsilon) \le \mathsf{P}(\mathfrak{B}_\varepsilon \cap \mathfrak{A}_C) + \mathsf{P}(\mathfrak{A}_C)$. Using the estimates below (recall $\tau_C = T$ on the set $\mathfrak{A}_C$), being valid on the set $\mathfrak{A}_C$,

$$\sup_{t\le T} |\varepsilon^2 H(X_t^\varepsilon/\varepsilon) \le \varepsilon^2 T \sup_{|v|\le C}\left|\int_0^{v/\varepsilon} \theta(s)ds\right|$$

$$\sup_{t\le T} |\varepsilon^2 H(x_0/\varepsilon) \le \varepsilon^2 T \sup_{|v|\le |x_0|}\left|\int_0^{v/\varepsilon} \theta(s)ds\right|$$

$$\sup_{t\le T}\left|\varepsilon \int_0^t \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds dX_v^\varepsilon\right| \le \text{const}\, T \sup_{|v|\le C}\left|\varepsilon \int_0^{v/\varepsilon} \theta(s)ds\right|$$
$$+ \sup_{t\le T}\left|\varepsilon \int_0^{t\wedge\tau_C} \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds \sigma(X_v^\varepsilon/\varepsilon)dB_v\right|,$$

we find that $\mathfrak{B}_\varepsilon \cap \mathfrak{A}_C \subseteq \bigcup_{i=1}^2 \mathfrak{B}_\varepsilon^{(i)}$, where (here η' is any positive constant)

$$\mathfrak{B}_\varepsilon^{(1)} = \left\{\varepsilon \sup_{|v|\le C\vee|x_0|}\left|\int_0^{v/\varepsilon} \theta(s)ds\right| \ge \eta'\right\}$$

$$\mathfrak{B}_\varepsilon^{(2)} = \left\{\sup_{t\le T}\left|\varepsilon \int_0^{t\wedge\tau_C} \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds \sigma(X_v^\varepsilon/\varepsilon)dB_v\right| \ge \eta'\right\}.$$

Due to Lemma 2, (13)–(15) are valid provided that for any $C > 0$, $\eta > 0$,

$$\lim_{\varepsilon\to 0} \mathsf{P}\left(\sup_{|v|\le C}\left|\varepsilon \int_0^{v/\varepsilon} \theta(s)ds\right| \ge \eta\right) = 0 \tag{16}$$

$$\lim_{\varepsilon\to 0} \mathsf{P}\left(\sup_{t\le T}\left|\varepsilon \int_0^{t\wedge\tau_C} \int_0^{X_v^\varepsilon/\varepsilon} \theta(s)ds \sigma(X_v^\varepsilon/\varepsilon)dB_v\right| \ge \eta\right) = 0. \tag{17}$$

Since $\theta(s)$ is the stationary process,

$$\sup_{0\le v\le C}\left|\varepsilon \int_0^{v/\varepsilon} \theta(s)ds\right| \overset{\text{law}}{=} \sup_{-C\le v\le 0}\left|\varepsilon \int_0^{v/\varepsilon} \theta(s)ds\right|, \tag{18}$$

so that, (16) is implied by Lemma 1. The proof of (17) also uses (18) and Lemma 1 combining with the maximal Doob's inequality,

$$\mathsf{P}\left(\sup_{t\le T}\left|\varepsilon\int_0^{t\wedge\tau_C}\int_0^{X_v^\varepsilon/\varepsilon}\theta(s)ds\sigma(X_v^\varepsilon/\varepsilon)dB_v\right|\ge\eta\right)$$
$$\le 4\varepsilon^2\mathsf{E}\int_0^{T\wedge\tau_C}\left(\int_0^{X_v^\varepsilon/\varepsilon}\theta(s)ds\right)^2\sigma^2(X_v^\varepsilon/\varepsilon)dv$$
$$\le \text{const}\,\mathsf{E}\sup_{|v|\le C}\left|\varepsilon\int_0^{v/\varepsilon}\theta(s)ds\right|^2\underbrace{\xrightarrow[\varepsilon\to 0]{0}}_{\text{by Lemma 1}}.$$

The proof is done.

4 Diffusion in Random Environment

For function $b(\omega,u),\sigma(\omega,u)$, Lipschitz continuous in u uniformly in ω, the existence and uniqueness of (1) is proved in a standard way (henceforth, the symbol "ω" is omitted).

We consider the case of measurable function b,σ and restrict ourselves by considering a weak solution explicitly constructed by time scaling and change of probability measure (for other approaches see [1, 10]). This approach imposes the boundedness of $|b|$, $|\sigma|$ and the uniform positiveness of σ^2.

4.1 $b(\omega,u)\equiv 0$

Let β_t be a Brownian motion independent of $\sigma(\omega,u)$. Assume that the pair (σ,β) is defined on a probability space $(\Omega,\mathscr{F},\mathsf{Q})$ supplied by two filtration's (satisfying the usual conditions): $(\mathscr{F}_u)_{u\in\mathbb{R}}$ and $(\mathscr{G}_t)_{t\ge 0}$ such that $\bigvee_{u\in\mathbb{R}}\mathscr{F}_u\subset\mathscr{G}_0$ and σ is $\mathscr{G}_0$-measurable while $\beta\perp\!\!\!\perp\mathscr{G}_0$. Since σ^2 is uniformly positive, a stopping time

$$\tau_t=\inf\left\{r:\int_0^r\frac{1}{\sigma^2(\omega,(\beta_s+x_0)/\varepsilon)}ds\ge t\right\}$$

strictly increases in t, that is, $\int_0^{\tau_t}\frac{1}{\sigma^2(\omega,(\beta_s+x_0)/\varepsilon)}ds\equiv t$ and

$$\tau_t=\int_0^t\sigma^2(\omega,(\beta_{\tau_s}+x_0)/\varepsilon)ds.$$

Then, obviously, $(\beta_{\tau_t},\mathscr{G}_{\tau_t})$ is a continuous martingale with the quadratic variation process τ_t and, by the Levy-Doob theorem, the random process

$$B_t=\int_0^t\frac{1}{\sigma(\omega,(\beta_{\tau_s}+x_0)/\varepsilon)}d\beta_{\tau_s}$$

is Brownian motion. Since $B=(B_t)_{t\ge 0}$ is the process with independent increments for non-overlapping intervals and $B_0=0$, we have "$B \perp\!\!\!\perp \mathscr{G}_0$" ⇒ "$B \perp\!\!\!\perp \bigvee_{u\in\mathbb{R}} \mathscr{F}_u$". Set $Y_t := x_0 + \beta_{\tau_t}$. The definition of B_t implies

$$Y_t = x_0 + \int_0^t \sigma(\omega, Y_s/\varepsilon) dB_s, \tag{19}$$

that is, a weak solution of (1) with zero drift exists for $\mathsf{P}=\mathsf{Q}$.

4.2 $b(\omega,u)\not\equiv 0$

A weak solution with nonzero drift is constructed by applying the Girsanov theorem. With Y_t, defined in (19), set

$$\Upsilon_T = \exp\bigg(\int_0^T \frac{b(\omega, Y_s/\varepsilon)}{\sigma(\omega, Y_s/\varepsilon)} dB_s - \frac{1}{2}\int_0^T \frac{b^2(\omega, Y_s/\varepsilon)}{\sigma^2(\omega, Y_s/\varepsilon)} ds\bigg).$$

Since $\frac{b(\omega,Y_s/\varepsilon)}{\sigma(\omega,Y_s/\varepsilon)}$ is bounded and $T<\infty$, we have $\int_\Omega \Upsilon_T d\mathsf{Q}=1$ and define a probability measure P by letting $d\mathsf{P} := \Upsilon_T d\mathsf{Q}$. Then, by the Girsanov theorem,

$$\widehat{B}_t = B_t - \int_0^t \frac{b(\omega, Y_s/\varepsilon)}{\sigma(\omega, Y_s/\varepsilon)} ds$$

is the Brownian motion on the stochastic basis $(\Omega, \mathscr{F}, (\mathscr{G}_t)_{t\ge 0}, \mathsf{P})$. In other words, the process Y_t, defined on $(\Omega, \mathscr{F}, (\mathscr{G}_t)_{t\ge 0}, \mathsf{P})$, admits the following representation

$$Y_t = x_0 + \int_0^t b(\omega, Y_s/\varepsilon) ds + \int_0^t \sigma(\omega, Y_s/\varepsilon) d\widehat{B}_s.$$

So, it is left to prove that $(\widehat{B}_t)_{t\le T} \perp\!\!\!\perp (b(u),\sigma(u))_{\mathbb{R}}$ relative to P. We mention that $(\widehat{B}_t)_{t\le T} \perp\!\!\!\perp \mathscr{G}_0$ with respect to P while $(b(u),\sigma(u))_{\mathbb{R}} \in \mathscr{G}_0$ relative to Q and, due to $\mathsf{P} \ll \mathsf{Q}$, with respect to P as well.

5 Markov Chain as Random Environment

We simulate the random environment as follow. Let $(Z_u)_{u\in\mathbb{R}}$ be a Markov stationary random processes with values in the finite alphabet $\{a_1, a_2, \dots, a_n\}$ and the transition intensity matrix Λ having simple zero eigenvalue. It is well known that, then, z is the ergodic process.

Let $b(u) = g(Z_u)$ and $\sigma(u) = h(Z_u)$ for some bounded measurable function $g(\cdot), h(\cdot)$, where $\min_{1\le n} h^2(a_i) > 0$. Denote by $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ the distribution of z_0 and compute

$$\mathbf{a} = 1\Big/\sum_{i=1}^n \frac{1}{h^2(a_i)}\pi_i, \qquad \mathbf{b} = \sum_{i=1}^n \frac{g(a_i)}{h^2(a_i)}\pi_i \Big/ \sum_{i=1}^n \frac{1}{h^2(a_i)}\pi_i.$$

For these **a** and **b**, $\theta(s) = 1 - \frac{\mathbf{a}}{\sigma^2(s)}$ and $\theta(s) = \frac{b(s)-\mathbf{b}}{\sigma^2(s)}$ are zero mean random processes with finite second moments.

It is left to verify weak dependence condition (9) which in the Markov case becomes

$$\int_0^\infty \sqrt{\mathsf{E}(\mathsf{E}(\theta(s)|Z_0))^2}ds < \infty. \tag{20}$$

Further,

$$\mathsf{E}(\theta(s)|Z_0) = \mathsf{E}(f(Z_s)|Z_0), \quad \text{where } f(x) = \begin{cases} 1 - \frac{1}{h^2(x)}, \\ \frac{\mathbf{b}-g(x)}{h^2(x)}. \end{cases}$$

On the other hand, $\mathsf{E}(f(Z_s)|Z_0) = \sum_{i=1}^n f(a_i)\mathsf{P}(Z_s = a_i|Z_0)$, a.s., while by the ergodic property of (Z_u) there exist $c > 0$, $\lambda > 0$ such that

$$\sum_{i=1}^n |f(a_i)\mathsf{P}(Z_s = a_i|Z_0)| \le ce^{-\lambda s}, \quad \exists c > 0,\ \lambda > 0.$$

Hence, (20) follows.

6 Langevin Random Environment

Let $Z_u = \binom{q_u}{p_u}_{u\in\mathbb{R}}$ solves stochastic Langevin equation

$$\begin{aligned} \dot{q}_u &= p_u \\ \dot{p}_u &= -(p_u + q_u) + \dot{W}_u, \end{aligned} \tag{21}$$

relative to Gaussian white noise $\dot{W}_u$ is a Gaussian white noise independent of Brownian motion $(B_t)_{t\ge 0}$. It can be verified directly that the matrix $A = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ has eigenvalues with negative real parts. Hence Z_u is stationary and ergodic zero mean Gaussian process on $\mathbb{R}$ with continuous paths. Moreover, Z_u is Markov process too.

Set $b(u) = g(p_u)$ a $\sigma(u) = h(q_u)$, where g, h are bounded and h^2 is uniformly positive. In particular, then, (1) possesses a weak solution.

Let

$$f(x) = \begin{cases} 1 - \frac{\mathbf{a}}{h^2(x_1)}, \\ \frac{\mathbf{b}-g(x_2)}{h^2(x_1)}, \end{cases} \qquad x = (x_1, x_2).$$

Set $\theta(s) = f(q_s, p_s)$. To simplify further analysis, let us assume that g and h are Lipschitz continuous functions and, therefore, f is Lipschitz continuous function too.

The Markov property of Z_u enables us to verify (20) only. To this end, we show that

$$\sqrt{\mathsf{E}(\mathsf{E}(\theta(s)|Z_0))^2} \le ce^{-s\lambda}, \quad \exists c, \lambda > 0. \tag{22}$$

Parallel to process Z_u let us introduce its version $(Z_t)_{t\ge 0}$:

$$\begin{aligned} dq_t &= p_t dt \\ dp_t &= -(p_t + q_t)dt + dW_t \end{aligned} \tag{23}$$

subject to the initial condition q_0, p_0. Denote by $Z_t^x = \binom{q_t^x}{p_t^x}$ the solution of (23) subject to $q_0^x = x_1$, $p_0^x = x_2$. Set $\theta^x(s) = \theta(s)_{|Z_0=x}$ and notice that

$$\mathsf{E}(\mathsf{E}(\theta(s)|Z_0))^2 = \int_{\mathbb{R}^2} (\mathsf{E}\theta^x(s))^2 dF(X),$$

where $F(x)$ is the distribution function of Z_0. Now, the use of boundedness for $\theta(s), \theta^x(s)$, and $\mathsf{E}\theta(s) \equiv 0$, and the Lipschitz property of $\theta(s)$, imply

$$\begin{aligned} (\mathsf{E}\theta^x(s))^2 &= (\mathsf{E}[\theta^x(s) - \theta(s)])^2 \\ &\le \text{const}\,\mathsf{E}|\theta^x(s) - \theta(s)| \\ &\le \text{const}\,\mathsf{E}[|q_s^{x_1} - q_s| + |p_s^{x_2} - p_s|]. \end{aligned}$$

On the other hand, $Z_t^x - Z_t$ solves the linear differential equation

$$\frac{d}{dt}[Z_t^x - Z_t] = A[Z_t^x - Z_t]$$

subject to $[Z_0 - x]$. Hence $[Z_t^x - Z_t] = e^{tA}[Z_0^x - x]$ and, in a view of eigenvalues of matrix A have strictly negative real parts, there exist positive constants $c(x)$, depending on x, and λ such that $\mathsf{E}[|q_s^{x_1} - q_s| + |p_s^{x_2} - p_s|] \le c(x)e^{-s\lambda}$ and $c(x) = \text{const}\,(1 + |x_1| + |x_2|)$. Thus, (22) is valid, with $c = \int_{\mathbb{R}^2} c(x)dF(x)$, and Theorem 1 is applicable with

$$\mathbf{a} = 1 \Big/ \int_{\mathbb{R}^2} \frac{1}{h^2(x_1)} dF(x) \quad \text{and} \quad \mathbf{b} = \int_{\mathbb{R}^2} \frac{g(x_2)}{h^2(x_1)} dF(x) \Big/ \int_{\mathbb{R}^2} \frac{1}{h^2(x_1)} dF(x).$$

References

1. Brox, Th.: A one-dimensional diffusion process in a Wiener medium. Ann. Probab. **14**(4), 1206–1218 (1986)
2. Butov, A.A., Krichagina, E.V.: A functional limit theorem for a symmetric walk in a random environment. Russ. Math. Surv. **43**(2), 163–164 (1988)
3. Di Masi, A., Ferrari, P.A., Goldstein, S., Wick, W.D.: An invariance principle for reversible Markov processes. Applications to random motions in random environments. J. Stat. Phys. **55**(3–4), 787–855 (1989)

4. Fannjiang, A., Komorowski, T.: An invariance principle for diffusion in turbulence. Ann. Probab. **27**(2), 751–781 (1999)
5. Fannjiang, A., Papanicolaou, G.: Diffusion in turbulence. Probab. Theory Relat. Fields **105**, 279–334 (1996)
6. Krylov, N.V.: Controlled Diffusion Processes. Applications of Mathematics, vol. 14. Springer, Berlin (1980)
7. Liptser, R.Sh., Shiryayev, A.N.: Theory of Martingales. Mathematics and its Applications (Soviet Series), vol. 49. Kluwer Academic, Dordrecht (1989)
8. Papanicolau, G.C., Varadhan, S.R.S.: Boundary value problems with rapidly oscillating coefficients. In: Proceedings of the Colloquium on Random Fields, Esztergom (1979). Coll. Math. Soc. J. Bolyai, vol. 27, pp. 836–874. North-Holland, Amsterdam (1981)
9. Sznitman, A.S., Zeitouni, O.: An invariance principle for isotropic diffusions in random environment. Invent. Math. **164**(3), 455–567 (2006)
10. Schumacher, S.: Diffusions with random coefficients. In: Particle Systems, Random Media and Large Deviations, Brunswick, Maine, 1984. Contemp. Math., vol. 41, pp. 351–356. Am. Math. Soc., Providence (1985)
11. Yurinsky, V.V.: Averaging of an elliptic boundary value problem with random coefficients. Sib. Mat. Zh. **21**(3), 209–223 (1980)

The Optimal Time to Exchange one Asset for Another on Finite Interval

Yuliya Mishura and Georgiy Shevchenko

Abstract Let S_t^1, S_t^2 be correlated geometric Brownian motions. We consider the following problem: find the stopping time $\tau^* \le T$ such that

$$\sup_{\tau \in [0,T]} \mathsf{E}[S_\tau^1 - S_\tau^2] = \mathsf{E}[S_{\tau^*}^1 - S_{\tau^*}^2]$$

where the supremum is taken over all stopping times from $[0, T]$. A similar problem, but on infinite interval, was studied by MacDonald and Siegel (Int. Econ. Rev. 26:331–349, 1985), and by Hu and Oksendal (Finance Stoch. 2(3):295–310, 1998), who also considered multiple assets. For a finite time horizon, the problem gets considerably more complicated and cannot be solved explicitly. In this paper we study generic properties of the optimal stopping set and its boundary curve, and derive an integral equation for the latter.

Keywords Optimal stopping · Geometric Brownian motion · Finite horizon · Free boundary problem

Mathematics Subject Classification (2000) Primary 60G40 · 60J65 · Secondary 35R35

1 Introduction

Let $(\Omega, \mathscr{F}, \mathsf{P})$ be a complete probability space. Let also S_t^1, S_t^2 be stochastic processes modeling prices of two stocks $\mathscr{S}^1$ and $\mathscr{S}^2$, possibly correlated. We adopt the most common model for a stock price process, namely, the geometric Brownian motion. Precisely, we assume that the processes S_t^1 and S_t^2 solve the following stochastic differential equations:

$$dS_t^i = \delta_i S_t^i dt + \sigma_i S_t^i dW_t^i, \quad i = 1, 2, \tag{1}$$

where W_t^1 and W_t^2 are standard Brownian motions (w.r.t. the filtration $\mathscr{F}_t$ generated by them) with a correlation coefficient ρ, i.e. $\mathsf{E}[W_t^1 W_t^2] = \rho t$. Here δ_i, σ_i, ρ are

Y. Mishura (✉) · G. Shevchenko
Kyiv National Taras Shevchenko University, 64 Volodymyrska, 01033 Kyiv, Ukraine
e-mail: myus@univ.kiev.ua

G. Shevchenko
e-mail: zhora@univ.kiev.ua

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_10,

some constants. Suppose that investor holding stock $\mathscr{S}^2$ should exchange it for stock $\mathscr{S}^1$ within given time interval $[0, T]$, and the choice of exchange time is up to the investor; one may regard this as a futures exchange contract. Clearly, the cost of the contract if exchanging at time t is $S_t^1 - S_t^2$. It is natural to assume that the investor intends to maximize the expected discounted cost of the contract, so we have the following optimization problem:

$$\mathsf{E}[e^{-r\tau}(S_\tau^1 - S_\tau^2)] \to \max, \tag{2}$$

where τ is an $\mathscr{F}_t$-stopping time bounded by T, and r is a risk-free rate of return. Observe that $e^{-rt}S_t^i$, $i = 1, 2$ are geometric Brownian motions as well with δ_i replaced by $\delta_i - r$, so we can assume without loss of generality that $r = 0$. Thus, we want to find an $\mathscr{F}_t$-stopping time τ^* such that

$$\sup_{\tau \in [0,T]} \mathsf{E}[S_\tau^1 - S_\tau^2] = \mathsf{E}[S_{\tau^*}^1 - S_{\tau^*}^2].$$

In the case of infinite time interval (i.e. $T = \infty$) the problem was solved by MacDonald and Siegel [8]: the optimal stopping time τ^* is given by

$$\tau^* = \inf\{t \geq 0 : S_t^1 \geq \mu S_t^2\},$$

where the constant μ can be calculated explicitly in terms of δ_i and σ_i, $i = 1, 2$.

The first paper, where *options* to exchange one asset for another were considered, was by Margrabe [7]; since that such options are frequently called Margrabe options. Ideas of this paper, which help to reduce the dimensionality of a pricing problem, were further developed in Shiryaev and Shepp [9]. We refer to [3], where a beautiful exposition of this so-called duality approach is given. In [4] this approach is applied to investigate the problems of valuing contingent claims with homogeneous payoffs in Lévy market model.

The purpose of this paper is to approach the solution of this problem for a finite time horizon T. In such setting, problem (2) gets much more complicated. The reason is that for infinite time horizon, it is time-homogeneous: starting from any time moment is the same as starting from zero; however, for a finite time horizon it is no longer time-homogeneous.

The paper is organized as follows. In Sect. 2 we give essential properties of the stopping domain and the premium function. In Sect. 3, we show how integral equations for the premium function and the optimal stopping domain boundary are derived. In Sect. 4, we give some useful properties of the optimal stopping domain boundary.

2 Basic Properties of Premium Function and Stopping Domain

One can regard optimization problem (2) as a problem of optimal exercise of an American type option. Thus, we can use the theory of American options pricing, omitting the details which can be found e.g. in [1, 2].

For $t \in [0, T]$, $s_1, s_2 \in \mathbb{R}_+ := (0, \infty)$ define the *premium function*

$$f(t; s_1, s_2) = \sup_{\tau \in [t,T]} \mathsf{E}[S_\tau^1 - S_\tau^2 \mid S_t^1 = s_1, S_t^2 = s_2]$$

which identifies the maximal expected cost of continuation (i.e. of not exchanging stock $\mathscr{S}^2$) at time t. This is also can be treated as the value, or the price, of the exchange contract at time t if the stock prices are s_1 and s_2. The cost of exchanging at time t, as we have already mentioned, is

$$g(S_t^1, S_t^2) = S_t^1 - S_t^2.$$

Exchange at time t is optimal if $f(t; S_t^1, S_t^2) \le g(S_t^1, S_t^2)$. Thus we can define the *stopping region*

$$\mathscr{G} = \{(t; s_1, s_2) \in [0, T] \times \mathbb{R}_+^2 : f(t; s_1, s_2) \le g(s_1, s_2)\}$$

and the *continuation region*

$$\mathscr{C} = \{(t, s_1, s_2) \in [0, T] \times \mathbb{R}_+^2 : f(t; s_1, s_2) > g(s_1, s_2)\}.$$

We will also use $\mathscr{G}_t$ and $\mathscr{C}_t$ for t-sections of these sets.

Then the optimal stopping time τ^* is

$$\tau^* = \inf\{t \in [0, T] : (t; S_t^1, S_t^2) \in \mathscr{G}\},$$

and $\tau^* = T$ if the latter set is empty.

Theorem 1 *The stopping set $\mathscr{G}$ has the following properties*:

1. $\forall (t, s_1, s_2) \in \mathscr{G}\ \forall \lambda > 0 : (t, \lambda s_1, \lambda s_2) \in \mathscr{G}$.
2. $\forall (t, s_1, s_2) \in \mathscr{G}\ \forall t' \ge t : (t', s_1, s_2) \in \mathscr{G}$.

3a. *If* $\delta_1 \le 0, \delta_2 \ge 0$, *then* $\mathscr{G} = [0, T] \times \mathbf{R}_+^2$.
3b. *If* $\delta_1 \ge 0, \delta_2 \le 0$ *and* $\delta_1 - \delta_2 \ne 0$, *then* $\mathscr{G} = \{T\} \times \mathbb{R}_+^2$.
3c. *If* $\delta_1 > 0, \delta_2 > 0$, *then* $\forall t \mathscr{G}_t \ne \varnothing$ *and* $\forall (t, s_1, s_2) \in \mathscr{G} \forall s_1' \le s_1 \colon (t, s_1', s_2) \in \mathscr{G}$.
3d. *If* $\delta_1 < 0, \delta_2 < 0$, *then* $\forall t \mathscr{G}_t \ne \varnothing$ *and* $\forall (t, s_1, s_2) \in \mathscr{G} \forall s_1' \ge s_1 \colon (t, s_1', s_2) \in \mathscr{G}$.

4. $\forall t \in [0, T] \{(s_1, s_2) \in \mathbb{R}_+^2 : \delta_1 s_1 > \delta_2 s_2\} \subset \mathscr{C}_t$.

Proof 1. This is because $g(\lambda s_1, \lambda s_2) = \lambda g(s_1, s_2)$ and $f(t; \lambda s_1, \lambda s_2) = \lambda f(t; s_1, s_2)$.

2. This is general property for American options on stocks with prices modeled by homogeneous processes, see e.g. [10].

3. Let $t \in [0, T]$ be fixed. Assume $(t, s_1, s_2) \in \mathscr{G}$. This amounts to

$$\mathsf{E}[S_\tau^1 - S_\tau^2 \mid S_t^1 = s_1, S_t^2 = s_2] = \mathsf{E}[s_1 S_\tau^1 - s_2 S_\tau^2 \mid S_t^1 = S_t^2 = 1] \le s_1 - s_2$$

for all stopping times $\tau \in [t, T]$, or, equivalently,

$$s_1(\mathsf{E}[S_\tau^1 \mid S_t^1 = 1] - 1) \le s_2(\mathsf{E}[S_\tau^2 \mid S_t^2 = 1] - 1). \tag{3}$$

One can write by Doob's optional sampling theorem:

$$\mathsf{E}[S_\tau^i \mid S_t^i = 1] - 1 = \mathsf{E}\left[\int_t^\tau S_u^i(\delta_i du + \sigma_i dW_u^i) \mid S_t^i = 1\right]$$
$$= \mathsf{E}\left[\int_t^\tau \delta_i S_u^i du \mid S_t^i = 1\right].$$

Hence, $\mathsf{E}[S_\tau^i \mid S_t^i = 1] - 1 \ge 0$ for $\delta_i \ge 0$ and $\mathsf{E}[S_\tau^i \mid S_t^i = 1] - 1 \le 0$ for $\delta_i \le 0$; moreover the inequalities are strict if $\delta_i \ne 0$ and $\mathsf{P}(\tau > t) > 0$. Thus in the case 3a inequality (3) holds for all stopping times τ, in the case 3b it holds only for $\tau \equiv t$ (hence, if $(t, s_1, s_2) \in \mathscr{G}$, any stopping time $\tau \in [t, T]$ should be identical to t, implying $t = T$), and the statements 3a, 3b follow. Further, the inequality (3) will remain true if s_1 is replaced by $s_1' \le s_1$ in the case 3c and if it is replaced by $s_1' \ge s_1$ in the case 3d. Moreover, the expressions $\mathsf{E}[S_\tau^i \mid S_t^i = 1]$ are bounded (one can use $\sup_{u\in[t,T]} |W_t^i|$ to get a bound), thus for some s_1 and s_2 (3) holds. Hence the statements of 3c and 3d follow.

4. Now write for arbitrary stopping time $\tau \ge t$

$$Z(\tau) := \mathsf{E}[S_\tau^1 - S_\tau^2 \mid S_0^1 = s_1, S_0^2 = s_2]$$
$$= s^1 - s^2 + \mathsf{E}\left[\int_t^\tau S_u^1(\delta_1 du + \sigma_1 dW_u^1)\right.$$
$$\left. - \int_t^\tau S_u^2(\delta_2 du + \sigma_2 dW_u^2) \,\middle|\, S_0^1 = s_1, S_0^2 = s_2\right]$$
$$= g(s_1, s_2) + \mathsf{E}\left[\int_0^\tau (S_u^1\delta_1 - S_u^2\delta_2) du \,\middle|\, S_0^1 = s_1, S_0^2 = s_2\right].$$

Define $D = \{(s_1, s_2) \in R_+^2 : \delta_1 s_1 > \delta_2 s_2\}$. Suppose $(S_t^1, S_t^2) = (s_1, s_2) \in D$. We can define the exit time of (S^1, S^2) from D:

$$\tau_D = \inf\{u \ge t : (S_u^1, S_u^2) \notin D\}$$

($\tau = T$ if the set is empty). Since the process (S_1, S_2) is continuous and D is open, we have $\tau_D > t$, and $Z(\tau_D) > g(s_1, s_2)$. On the other hand, $f(t, s_1, s_2) \ge Z(\tau_D)$, therefore we have $(t, s_1, s_2) \in \mathscr{C}$. The statement 4 is proved. □

Remark 1 Taking the stopping time $\tau \equiv T$, we can moreover write that

$$\forall t \in [0, T]\{(s_1, s_2) \in \mathbb{R}_+^2 : (e^{\delta_1(T-t)} - 1)s_1 > (e^{\delta_2(T-t)} - 1)s_2\} \subset \mathscr{C}_t,$$

which clearly gives a better result in the case $\delta_1 > \delta_2$.

As we see, the stopping set is trivial whenever δ_1 and δ_2 have different signs. Thus we can assume that $\delta_1\delta_2 > 0$. We will moreover assume that

$$\delta_1 > 0, \qquad \delta_2 > 0.$$

The financial background of this assumption is that on average returns on risky assets are greater than on non-risky ones (recall that we assumed zero interest rate, that is, asset prices are already discounted, and δ_i are discounted returns).

Corollary 1 *The stopping set $\mathscr{G}$ has the structure*:

$$\mathscr{G} = \{(t, s_1, s_2) \in [0, T) \times \mathbb{R}_+^2 \mid s_1 \le \beta(t)s_2\} \cup (\{T\} \times \mathbb{R}_+^2),$$

where $\beta(t) : [0, T] \to (0, \delta_2/\delta_1]$ is certain increasing function. Moreover, $\beta(t) \le (e^{\delta_2(T-t)} - 1)/(e^{\delta_1(T-t)} - 1)$.

Remark 2 As one can see, $\beta(T)$ can be chosen freely. We put $\beta(T) = \delta_2/\delta_1$ to make β continuous: it will be shown in Sect. 4 that $\beta(T-) = \delta_2/\delta_1$.

3 Integral Equations for the Premium Function and the Threshold Curve

Recall that the optimal stopping time τ^* is the time when the process (t, S_t^1, S_t^2) enters the stopping region $\mathscr{G}$:

$$\tau^* = \inf\{t : (t, S_t^1, S_t^2) \in \mathscr{G}\},$$

and $\tau^* = T$ if the latter set is empty. Similarly, let

$$\tau^*(t) = \inf\{u \in [t, T] : (u, S_u^1, S_u^2) \in \mathscr{G}\},$$

and $\tau^*(t) = T$ if the latter set is empty.

Then the premium function is, of course,

$$f(t; s_1, s_2) = \mathsf{E}[S^1_{\tau^*(t)} - S^2_{\tau^*(t)} \mid S_t^1 = s_1, S_t^2 = s_2]. \tag{4}$$

We denote Φ and ϕ the standard normal distribution function and probability density function respectively.

Theorem 2 *The premium function f solves the equation*

$$\begin{aligned} f(t; s_1, s_2) = s_1 - s_2 + \delta_1 s_1 \int_t^T & e^{\delta_1(u-t)}\Phi(D_1(s_1, s_2, \beta(u), u-t))du \\ - \delta_2 s_2 \int_t^T & e^{\delta_2(u-t)}\Phi(D_2(s_1, s_2, \beta(u), u-t))du, \end{aligned} \tag{5}$$

$$D_{1,2}(s_1, s_2, \beta(u), u-t) = \frac{1}{\sigma\sqrt{u-t}}\left[\log\left(\frac{s_1}{\beta(u)s_2}\right) + (\delta_1 - \delta_2 \pm \sigma^2/2)(u-t)\right],$$

where $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$, *and the optimal boundary* β *is a solution of the equation*

$$\beta(t) = \frac{\delta_2 \int_t^T e^{\delta_2(u-t)} \Phi(d_2(\beta(t), \beta(u), u-t))du}{\delta_1 \int_t^T e^{\delta_1(u-t)} \Phi(d_1(\beta(t), \beta(u), u-t))du}, \tag{6}$$

$$d_{1,2}(\beta(t), \beta(u), u-t) = \frac{1}{\sigma\sqrt{u-t}}\left[\log\left(\frac{\beta(t)}{\beta(u)}\right) + \left(\delta_1 - \delta_2 \pm \frac{\sigma^2}{2}\right)(u-t)\right].$$

Proof As long as the premium function can be represented as (4), in the continuation region $\mathscr{C}$ we have $\mathscr{L}f = 0$, where

$$\mathscr{L} = \frac{\partial}{\partial t} + \frac{1}{2}\left[\sigma_1^2 s_1^2 \frac{\partial^2}{\partial s_1^2} + \sigma_2^2 s_2^2 \frac{\partial^2}{\partial s_2^2} + \rho\sigma_1\sigma_2 s_1 s_2 \frac{\partial^2}{\partial s_1 \partial s_2}\right] + \delta_1 s_1 \frac{\partial}{\partial s_1} + \delta_2 s_2 \frac{\partial}{\partial s_2}$$

is the generator of diffusion $Y_t := (t, S_t^1, S_t^2)$. Moreover, evidently one has $f(t, s_1, s_2) = g(s_1, s_2) = s_1 - s_2$ for $(t, s_1, s_2) \in \mathscr{G}$. Observe that $\mathscr{L}g = \delta_1 s_1 - \delta_2 s_2 \le 0$ for $(t, s_1, s_2) \in \mathscr{G}$ (item 4 of Theorem 1). Hence, we have

$$\mathscr{L}f = 0 \quad \text{on } \mathscr{C}, \tag{7}$$

$$\mathscr{L}f \le 0 \quad \text{on } \mathscr{G}, \tag{8}$$

or, since $f = g$ on $\mathscr{G}$,

$$\mathscr{L}f \le 0, \qquad (f - g)\mathscr{L}f = 0. \tag{9}$$

Also one has the following "smooth fit" conditions:

$$\frac{\partial f}{\partial s_1} = 1, \qquad \frac{\partial f}{\partial s_2} = -1, \quad \text{on } \partial\mathscr{G}. \tag{10}$$

Equations (7)–(10) make up so-called "free boundary problem". On more information on how they are derived, see e.g. [10].

Now write for $t \in [0, T]$

$$f(t; S_t^1, S_t^2) = f(T; S_T^1, S_T^2) - \int_t^T df(u, S_u^1, S_u^2). \tag{11}$$

One can write by the Itô formula

$$df(t; S_t^1, S_t^2) = a_t dt + b_t^1 dW_t^1 + b_t^2 dW_t^2,$$

where

$$a_t = \mathscr{L}f(t; S_t^1, S_t^2),$$

and b^i_t, $i = 1, 2$ are square integrable processes, whose value is not essential for us. Taking conditional expectation of (11) and using (7), we arrive to

$$f(t; S^1_t, S^2_t) = \mathsf{E}[f(T; S^1_T, S^2_T)|\mathscr{F}_t] + \int_t^T \mathsf{E}[(\delta_2 S^2_u - \delta_1 S^1_u)1_{\mathscr{G}}(u, S^1_u, S^2_u)|\mathscr{F}_t]du \tag{12}$$

We have $(u, s_1, s_2) \in \mathscr{G}$ iff $s_1 \le \beta(u)s_2$. Further,

$$S^i_u = S^i_t \exp\{(\delta_i - \sigma_i^2/2)(u - t) + \sigma_i(W^i_u - W^i_t)\}.$$

Thus,

$$S^1_u \le \beta(u)S^2_u \iff \exp\{\alpha(u - t) + \sigma_1(W^1_u - W^1_t) - \sigma_2(W^2_u - W^2_t)\} \le \beta(u)S^2_t/S^1_t,$$

where $\alpha = \delta_1 - \delta_2 - \sigma_1^2/2 + \sigma_2^2/2$. Note that $\sigma_1(W^1_u - W^1_t) - \sigma_2(W^2_u - W^2_t) = \sigma\sqrt{u - t}Z$, where Z is a standard normal variable independent of $\mathscr{F}_t$, σ is as above. Thus, we can further write

$$S^1_u \le \beta(u)S^2_u \iff Z \le K_t(u),$$

where

$$K_t(u) = \frac{1}{\sigma\sqrt{u - t}}[\log(\beta(u)S^2_t/S^1_t) - \alpha(u - t)]$$

is $\mathscr{F}_t$-measurable.

Now we want to evaluate the integrand in (12). Write

$$W^i_u - W^i_t = \sqrt{u - t}(\rho_i Z + \sqrt{1 - \rho_i^2}Z^i),$$

where the random variables Z^i are standard normal and independent of $\mathscr{F}_t$ and Z; $\rho_1 = (\sigma_1 - \rho\sigma_2)/\sigma$, $\rho_2 = (\rho\sigma_1 - \sigma_2)/\sigma$. Now we can write

$$\begin{aligned}
&\mathsf{E}[\exp\{\sigma_i(W^i_u - W^i_t)1_{\mathscr{G}}(u, S^1_u, S^2_u)\} \mid \mathscr{F}_t] \\
&= \mathsf{E}[\exp\{\sigma_i\rho_i\sqrt{u - t}Z\}\exp\{\sigma_i\sqrt{(1 - \rho_i^2)(u - t)}Z^i\}1_{Z \le K_t(u)} \mid \mathscr{F}_t] \\
&= \mathsf{E}[\exp\{\sigma_i\rho_i\sqrt{u - t}Z\}1_{Z \le K_t(u)} \mid \mathscr{F}_t]\mathsf{E}[\exp\{\sigma_i\sqrt{(1 - \rho_i^2)(u - t)}Z^i\} \mid \mathscr{F}_t] \\
&= \exp\{\sigma_i^2\rho_i^2(u - t)/2\}\Phi(K_t(u) - \sigma_i\rho_i\sqrt{u - t})\exp\{\sigma_i^2(1 - \rho_i^2)(u - t)/2\} \\
&= \exp\{\sigma_i^2(u - t)/2\}\Phi(K_t(u) - \sigma_i\rho_i\sqrt{u - t}).
\end{aligned}$$

Plugging this into (12), after some routine transformations we arrive to (5). We may substitute $s_1 = \beta(t)s_2$ to this equation and, taking into account that $f(t; \beta(t)s_2, s_2) = \beta(t)s_2 - s_2$, get (6).

It is clear that having solved equation (6), one can plug its solution into (5) to get the premium function. But the equation for the threshold curve β is quite complicated and hardly can be solved explicitly. In the next section we approach somehow the solution of this equation. □

Remark 3 The problem can in fact be reduced to one state variable. Indeed, a direct calculation shows that

$$\mathsf{E}[S^1_\tau - S^2_\tau] = \mathsf{E}^{\mathsf{P}^1}[e^{\delta_1\tau}(1 - S_\tau)], \tag{13}$$

where

$$S_t = \frac{S^2_0}{S^1_0}\exp\{(\delta_2 - \delta_1 - \sigma^2/2)t + \sigma W_t\},$$

$d\mathsf{P}^1/d\mathsf{P} = S^1_T/\mathsf{E}[S^1_T]$ and W_t is a standard Brownian motion under P^1. A detailed derivation can be found in [4].

Formula (13) allows to derive (5) and (6) somewhat simpler. It will be used in the next section to derive important properties of the threshold curve.

4 Approaching Solution of Integral Equation for Threshold Curve

First we establish some properties of function β. First, analogously to (3), we can write from (13) that $(t, s_1, s_2) \in \mathscr{G}$ iff

$$\frac{s_2}{s_1}(\mathsf{E}[G_\tau] - 1) \geq \mathsf{E}[e^{\delta_1\tau}] - 1$$

for all stopping times $\tau \in [0, T - t]$, where

$$G_t = \exp\{\alpha t + \sigma W_t\},$$

$\alpha = \delta_2 - \sigma^2/2$, W is Brownian motion. Thus we can write

$$\beta(t) = \inf\frac{\mathsf{E}[G_\tau] - 1}{\mathsf{E}[e^{\delta_1\tau}] - 1},$$

where the infimum is taken over all stopping times $\tau \in [0, T - t]$ (we define the fraction to be equal to δ_2/δ_1 if $\tau \equiv 0$, as $\beta(t) \leq \delta_2/\delta_1$ anyway). Notice that this way β can be defined also for negative values.

We put $\mu(t) = \beta(T - t)$, $t > 0$.

Proposition 1 $\mu(t) = \frac{\delta_2}{\delta_1}(1 + c_1 t^{1/2} + O(t))$, $t \to 0$, *where* c_1 *is a negative constant.*

Proof Take a stopping time $\tau \in [0, t]$. Write

$$e^{\delta_1\tau} - 1 = \delta_1\tau + \frac{1}{2}e^{\delta_1\theta(\tau)}\tau^2 = \delta_1\tau(1 + O(t)), \quad t \to 0,$$

and $O(t)$ is uniform w.r.t. τ, as $\tau \in [0, t]$. Hence,

$$\mathsf{E}[e^{\delta_1\tau}] - 1 = \delta_1\mathsf{E}[\tau](1 + O(t)), \quad t \to 0.$$

Further,

$$G_\tau - 1 = \alpha\tau + \sigma W_\tau + \alpha\sigma\tau W_\tau + \frac{1}{2}\sigma^2(W_\tau)^2 + \frac{1}{6}\sigma^3(W_\tau)^3 + \cdots, \tag{14}$$

We have by the Itô formula, integration by parts and the optional sampling theorem $\mathsf{E}[W_\tau] = 0$, $\mathsf{E}[(W_\tau)^2] = \mathsf{E}[\tau]$, $\mathsf{E}[(W_\tau)^3] = 3\mathsf{E}[\int_0^\tau W_s ds] = 3\mathsf{E}[\tau W_\tau]$. The dots in (14) mean the terms with expectation $\mathsf{E}[\tau]O(t)$ (where O is uniform in τ). This can be checked easily: for instance,

$$\mathsf{E}[(W_\tau)^4] = 6\mathsf{E}\left[\int_0^\tau (W_s)^2 ds\right] = 6\mathsf{E}[\tau(W_\tau)^2] - 3\mathsf{E}[\tau^2] \le 6t\mathsf{E}[(W_\tau)^2] = 6t\mathsf{E}[\tau].$$

Thus, we have

$$\begin{aligned}\mu(t) &= \inf \frac{\delta_2\mathsf{E}[\tau] + \delta_2\sigma\mathsf{E}[\tau W_\tau] + \mathsf{E}[\tau]O(t)}{\delta_1\mathsf{E}[\tau](1 + O(t))} \\ &= \frac{\delta_2}{\delta_1}\left(1 + \sigma \inf \frac{\mathsf{E}[\tau W_\tau]}{\mathsf{E}[\tau]} + O(t)\right)(1 + O(t)), \quad t \to 0,\end{aligned}$$

where the infimum is taken over stopping times $\tau \in [0, t]$. By the scaling property of Brownian motion we have

$$\inf_{\tau\in[0,t]} \frac{\mathsf{E}[\tau W_\tau]}{\mathsf{E}[\tau]} = t^{1/2} \inf_{\tau\in[0,1]} \frac{\mathsf{E}[\tau W_\tau]}{\mathsf{E}[\tau]}.$$

Observe that for $\tau \in [0, 1]$ $|\mathsf{E}[\tau W_\tau]| \le (\mathsf{E}[\tau^2]\mathsf{E}[(W_\tau)^2])^{1/2} \le \mathsf{E}[\tau]$, thus the infimum is finite. On the other hand, one can put $\tau = 1/2$ if $W_{1/2} \ge 0$ and $\tau = 1$ otherwise, with this choice $\mathsf{E}[\tau W_\tau] < 0$ clearly. The statement is proved.

In order to simplify matters, we introduce the following function:

$$\gamma(t) = \frac{\delta_1}{\delta_2}\beta(T - t)\exp\{(\delta_1 - \delta_2)t\}.$$

It is straightforward to check that γ solves

$$\gamma(t) = \frac{\int_0^t e^{-\delta_2 u}\Phi(g_2(\gamma(t), \gamma(u), t, u))du}{\int_0^t e^{-\delta_1 u}\Phi(g_1(\gamma(t), \gamma(u), t, u))du}, \tag{15}$$

$$g_{1,2}(\gamma(t), \gamma(u), t, u) = \frac{1}{\sigma\sqrt{t-u}}\left[\log\left(\frac{\gamma(t)}{\gamma(u)}\right) \pm \frac{\sigma^2}{2}(t - u)\right].$$

Now assume that this equation has a solution which can be found by successive approximations. One can check that if one plugs a constant γ to the right-hand side of (15), then the left-hand side has the form $1 + c_1 t^{1/2} + c_2 t + c_3 t^{3/2} + \cdots$. Moreover, if one plugs γ of this form into the rhs, then the lhs is again of this form, so it is natural to seek γ in this form, i.e. we assume

$$\gamma(t) = 1 + c_1 t^{1/2} + c_2 t + c_3 t^{3/2} + \cdots$$

for t small enough. We moreover assume that

$$\gamma(t) = 1 + c_1 t^{1/2} + \cdots + c_k t^{k/2} + \theta_k(t) t^{(k+1)/2}, \tag{16}$$

where the function θ_k is locally bounded.

Now we are going to calculate first coefficients of this expansion. We will give details only for the coefficient c_1, the others can be computed in similar manner. In what follows $C(t)$ will denote generic function, which is bounded in neighborhood of zero and may change from one line to another; similarly, C will denote generic constant.

First, we write the Taylor expansion for $\log\gamma(t)$ in zero. For this, we use expansion (16) with $k = 3$. It is clear that for t small enough, $\gamma(t) \in [1/2, 1]$, so we can write for such t

$$\log\gamma(t) = (\gamma(t) - 1) + \frac{1}{2}(\gamma(t) - 1)^2 + \frac{1}{3}(\gamma(t) - 1)^3 + C(t)(\gamma(t) - 1)^4,$$

so, by the boundedness of θ_3 in (16)

$$\log\gamma(t) = c_1 t^{1/2} + (c_2 - c_1^2/2)t + (c_3 - c_1 c_2 + c_3^2/3)t^{3/2} + C(t)t^2.$$

Thus, we can write

$$\Phi(g_2(\gamma(t), \gamma(u), t, u)) = \Phi\left(\frac{c_1}{\sigma}\mu\left(\frac{u}{t}\right)\right) + \frac{1}{\sigma}\phi\left(\frac{c_1}{\sigma}\mu\left(\frac{u}{t}\right)\right)m(t, u) + \frac{1}{2\sigma}\phi'(\theta(t, u))m(t, u)^2,$$

where ϕ is the standard normal density function,

$$\mu(v) = \frac{1 - v^{1/2}}{(1 - v)^{1/2}} = \frac{(1 - v)^{1/2}}{1 + v^{1/2}},$$

$$m(t, u) = \left(c_2 - \frac{c_1^2}{2} - \frac{\sigma^2}{2}\right)(t - u)^{1/2} + \left(c_3 - c_1 c_2 + \frac{c_3}{3}\right)\frac{t^{3/2} - u^{3/2}}{(t - u)^{1/2}} + C(t)t^2 + C(u)u^2,$$

and $\theta(t, u)$ is some number, which is not essential for us. Further, write $e^{-\delta_2 t} = 1 + C(t)t$ and evaluate the numerator in (15), changing the variable $u \to vt$ and noticing that functions μ and

$$\frac{1 - v^{3/2}}{(1 - v)^{1/2}} = \mu(v)(1 - v^{1/2} + v)$$

are bounded:

$$
\begin{aligned}
&\int_0^t e^{-\delta_2 u}\Phi(g_2(\gamma(t),\gamma(u),t,u))du \\
&\quad = t\int_0^1 (1+C(t)t)\bigg[\Phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)+\frac{1}{\sigma}\phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)m(t,vt) \\
&\qquad + \frac{1}{2\sigma}\phi'(\theta(t,vt))m(t,vt)^2\bigg]dv \\
&\quad = t\int_0^1 \bigg[\Phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg) \\
&\qquad + \frac{1}{\sigma}\phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)\bigg\{t^{1/2}\bigg(c_2-\frac{c_1^2}{2}-\frac{\sigma^2}{2}\bigg)(1-v)^{1/2}+C(v)t\bigg\} \\
&\qquad + \frac{1}{2\sigma}\phi'(\theta(t,vt))\bigg\{t^{1/2}\bigg(c_2-\frac{c_1^2}{2}-\frac{\sigma^2}{2}\bigg)(1-v)^{1/2}+C(v)t\bigg\}^2 \\
&\qquad + C(t,v)(t+v)\bigg]dv \\
&\quad = t\int_0^1 \Phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)dv \\
&\qquad + t^{3/2}\frac{1}{\sigma}\bigg(c_2-\frac{c_1^2}{2}-\frac{\sigma^2}{2}\bigg)\int_0^1 \phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)(1-v)^{1/2}dv + C(t)t^2.
\end{aligned}
$$

The same can be written for the denominator of (15):

$$
\begin{aligned}
&\int_0^t e^{-\delta_1 u}\Phi(g_1(\gamma(t),\gamma(u),t,u))du \\
&\quad = t\int_0^1 \Phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)dv \\
&\qquad + t^{3/2}\frac{1}{\sigma}\bigg(c_2-\frac{c_1^2}{2}+\frac{\sigma^2}{2}\bigg)\int_0^1 \phi\bigg(\frac{c_1}{\sigma}\mu(v)\bigg)(1-v)^{1/2}dv + C(t)t^2,
\end{aligned}
$$

thus, expanding the fraction, we have

$$
\gamma(t) = 1 - \sigma\xi\bigg(\frac{c_1}{\sigma}\bigg)t^{1/2} + C(t)t,
$$

where

$$
\xi(x) = \frac{\int_0^1 \phi(x\mu(v))(1-v)^{1/2}dv}{\int_0^1 \Phi(x\mu(v))dv}.
$$

Hence we have

$$c_1 = \sigma x_0,$$

where x_0 is a solution of equation $\xi(x) = -x$. Clearly, it should be negative. □

Lemma 1 *The equation $\xi(x) = -x$ has unique negative solution.*

Proof Integrating by parts, we can write $t/\xi(-t)$ as

$$\frac{t}{\xi(-t)} = \frac{\sqrt{2\pi}t + 2t^2 \int_0^1 e^{-t^2\mu(v)^2/2} v\mu'(v)dv}{\frac{4}{3}e^{-t^2/2} - \frac{4}{3}t^2 \int_0^1 e^{-t^2\mu(v)^2/2}\mu(v)\mu'(v)(1-v)^{3/2}dv}.$$

Now consider the difference between the numerator and denominator of the last expression. We can make the change of variable $u = \mu(v)$ (respectively, $v = \nu(u) = (1-u^2)^2/(1+u^2)^2$) and represent it in the form:

$$D(t) = \frac{4}{3}e^{-t^2/2} - \sqrt{2\pi}t + t^2 \int_0^1 e^{-t^2u^2/2}\left[\frac{4}{3}u(1-\nu(u))^{3/2} + 2\nu(u)\right]du.$$

Now note that $0 \le \chi(u) := \frac{4}{3}u(1-\nu(u))^{3/2} + 2\nu(u) \le \frac{4}{3}(1-\nu(u)) + 2\nu(u) \le 2$, thus we can write

$$\begin{aligned} D(t) &\le \frac{4}{3}e^{-t^2/2} - \sqrt{2\pi}t + 2t\int_0^1 e^{-t^2u^2/2} t\,du \\ &= \frac{4}{3}e^{-t^2/2} - \sqrt{2\pi}t + 2t\int_0^t e^{-s^2/2}ds \\ &= \frac{4}{3}e^{-t^2/2} - 2t\int_t^\infty e^{-s^2/2}ds \le \frac{4}{3}e^{-t^2/2} - 2\frac{t^2}{t^2+1}e^{-t^2/2}. \end{aligned}$$

The last inequality is due to a well-know property of the standard normal distribution function. Hence, for $t > \sqrt{2}$ the function $D(t)$ is negative. Similarly, for $t \le \sqrt{2}$

$$\begin{aligned} (2D(t)t^{-1/2})' &\le -\frac{4}{3t^{3/2}}e^{-t^2/2} - \sqrt{2\pi}t^{-1/2} \\ &\quad + 2\int_0^1 \left(\frac{1}{2}t^{-1/2} - t^{3/2}u^2\right)e^{-t^2u^2/2}\chi(u)du \\ &\le t^{-3/2}\left[-\frac{4}{3}e^{-t^2/2} - t\sqrt{2\pi} + 2t\int_0^{1\wedge\frac{1}{t\sqrt{2}}}(1-2t^2u^2)e^{-t^2u^2/2}du\right] \\ &\le t^{-3/2}\left[-\frac{4}{3}e^{-t^2/2} - t\sqrt{2\pi} + 2\int_0^{\frac{1}{\sqrt{2}}} e^{-s^2/2}ds\right]. \end{aligned}$$

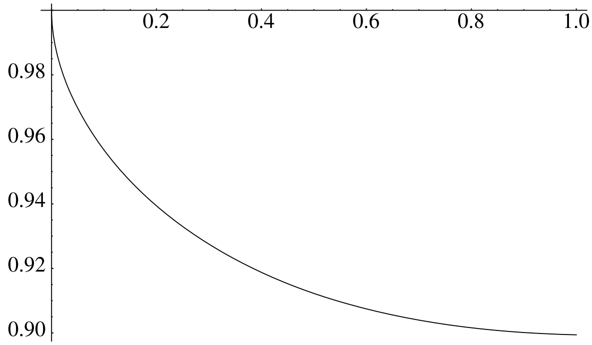


Fig. 1 $\beta(T-t)$ for $\delta_1 = \delta_2 = 0.02, \sigma = 0.2$

It is easy to check that $\frac{4}{3}e^{-t^2/2} + t\sqrt{2\pi}$ increases on $[0, \sqrt{2}]$. Thus, we can write

$$(2D(t)t^{-1/2})' \le -\frac{4}{3} + 2\sqrt{2\pi}\left[\Phi\left(\frac{1}{\sqrt{2}}\right) - \frac{1}{2}\right] < 0.$$

Summing up, $D(t)t^{-1/2}$ decreases on $[0, \sqrt{2}]$, it is equal to $+\infty$ in 0 and it is negative on $(\sqrt{2}, +\infty)$. Hence there is exactly one point t such that $D(t)t^{-1/2} = 0$, and the statement follows. □

Remark 4 Solving $\xi(x) = -x$ numerically by successive approximations, one gets the following value: $x_0 \approx -0.6388332158$. It is worth to mention that recently we found the same constant (but calculated less accurately) in [6] with respect to the problem of valuing American put option on a dividend-paying asset.

Remark 5 Combining the result of the lemma with the proof of Proposition 1, we get the following interesting result:

$$\sup \frac{\mathsf{E}[\tau W_\tau]}{\mathsf{E}[\tau]} = -x_0,$$

where the supremum is taken over all stopping times $\tau \in [0, 1]$.

The rest of coefficients in the expansion (15) can be found in the similar manner. The formulas for them are quite cumbersome, we only give the one for c_2:

$$c_2 = \frac{\delta_2 - \delta_1 + \sigma x_0 I_1/I_2}{1 - x_0^2} + \frac{\sigma^2 x_0^2}{2},$$

where

$$I_1 = \frac{1}{2\sqrt{2\pi}} \int_0^1 e^{-x_0^2\mu^2(u)/2}\mu(u)du, \qquad I_2 = \int_0^1 \Phi(x_0\mu(u))du,$$

and x_0 is the solution of the equation $\xi(x) = -x$.

Figure 1 is the graph of β for some particular values of parameters, to draw it, we cut the expansion of γ at t (i.e., four non-constant terms are used).

References

1. Broadie, M., Detemple, J.: The valuation of American options on multiple assets. Math. Financ. **7**(3), 241–286 (1997)
2. Chen, X., Chadam, J.: A mathematical analysis of the optimal exercise boundary for American put options. SIAM J. Math. Anal. **38**(5), 1613–1641 (2007)
3. Eberlein, E., Papapantoleon, A., Shiryaev, A.N.: On the duality principle in option pricing: semimartingale setting. Finance Stoch. **12**(2), 265–292 (2008)
4. Fajardo, J., Mordecki, E.: Pricing derivatives on two-dimensional Lévy processes. Int. J. Theor. Appl. Financ. **9**(2), 185–197 (2006)
5. Hu, Y., Øksendal, B.: Optimal time to invest when the price processes are geometric Brownian motions. Finance Stoch. **2**(3), 295–310 (1998)
6. Lamberton, D., Villeneuve, S.: Critical price near maturity for an American option on a dividend-paying stock. Ann. Appl. Probab. **13**(2), 800–815 (2003)
7. Margrabe, W.: The value of an option to exchange one asset for another. J. Finance **33**(1), 177–186 (1978)
8. McDonald, R.L., Siegel, D.R.: Investment and the valuation of firms when there is an option to shut down. Int. Econ. Rev. **26**, 331–349 (1985)
9. Shepp, L., Shiryaev, A.: A new look at the "Russian option". Teor. Veroyatnost. Primenen. **39**(1), 130–149 (1994)
10. Villeneuve, S.: Exercise regions of American options on several assets. Finance Stoch. **3**(3), 295–322 (1999)

Arbitrage Under Transaction Costs Revisited

Miklós Rásonyi

Abstract We present a novel arbitrage-related notion for markets with transaction costs in discrete time and characterize it in terms of price systems. Pertinence of this concept is demonstrated. A discussion of the case with one risky asset and an outlook on continuous-time models complement the main result.

Keywords Transaction costs · Absence of arbitrage · Consistent price systems

Mathematics Subject Classification (2000) 91B28 · 60G42

1 Introduction

After the pioneering papers [17] and [5], a geometric approach to modelling markets with transaction costs has been initiated in [18] and various versions of the fundamental theorem of asset pricing have been shown for discrete-time market models in [1, 11, 19, 20, 24, 31] and [10].

In the present article we will deviate from the usual notion of arbitrage figuring in previous papers and propose an alternative concept.

After introducing our geometric framework, Theorem 1 in Sect. 2 characterizes the new "no sure gain in liquidation value" **NGV** property (see Definition 3) in terms of the existence of price systems. Section 3 investigates **NGV** in models with one risky asset. Section 4 contains the proofs. Section 5 refers to certain strongly related phenomena in continuous-time models which motivated our investigations. The appendix provides some necessary technical tools.

Scalar product in $\mathbb{R}^d$ will be denoted by $\langle\cdot,\cdot\rangle$; $\mathscr{B}(\mathbb{R}^d)$ stands for the Borel-sets of $\mathbb{R}^d$; we will also need the unit ball $U := \{x \in \mathbb{R}^d : \ |x| \le 1\}$. The closure of a set $H \subset \mathbb{R}^d$ in the Euclidean topology is written as $\overline{H}$. The positive dual cone of a closed cone $K \subset \mathbb{R}^d$ is defined as

$$K^* := \{x \in \mathbb{R}^d : \langle x, c\rangle \ge 0, \ \text{for all } c \in K\},$$

The support received from Hungarian National Science Foundation (OTKA) under grant F 049094, from Austrian Science Foundation (FWF) under grant P 19456 and from Vienna Science and Technology Fund (WWTF) under grant MA 13 is gratefully acknowledged. The author thanks P. Grigoriev, D. Rokhlin and W. Schachermayer for helpful discussions and an anonymous referee for an insightful report.

M. Rásonyi (✉)
Computer and Automation Institute of the Hungarian Academy of Sciences, Kende utca 13-17, 1111 Budapest, Hungary
e-mail: rasonyi@sztaki.hu

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*,
DOI 10.1007/978-3-642-02608-9_11,

this is also a closed cone. We say that K is proper if $\operatorname{int} K^* \neq \emptyset$. For subsets $M \subset \mathbb{R}^d$, $\operatorname{cone}(M)$ denotes the cone generated by M.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t=0}^T, P)$ be a discrete-time stochastic basis with finite time horizon. In this paper we suppose that each σ-algebra we are dealing with contains P-zero sets and that $\mathscr{F}_0$ is trivial. If $\mathscr{G} \subset \mathscr{F}$ is a σ-algebra then an element $A \in \mathscr{G} \otimes \mathscr{B}(\mathbb{R}^d)$ is called a $\mathscr{G}$-measurable random set. If the sections $A(\omega)$ of this random set are cones for almost all ω then we speak about a random cone and similarly about random closed cones, random convex sets, etc.

If A is some (not necessarily $\mathscr{G}$-measurable) random set which is nonempty a.s. then $L^0(A, \mathscr{G})$ denotes the family of $\mathscr{G}$-measurable $\mathbb{R}^d$-valued functions f such that $f(\omega) \in A(\omega)$ for almost all ω. Such an f is called a *$\mathscr{G}$-measurable selector* of A. The set of elements of $L^0(A, \mathscr{G})$ with finite expectation is denoted by $L^1(A, \mathscr{G})$.

We now recall elements of the abstract geometric approach to markets with proportional friction, see [19, 24] or [31] for more detailed descriptions.

We assume that G_t is an $\mathscr{F}_t$-measurable random closed cone in $\mathbb{R}^d$ containing $\mathbb{R}^d_+$ a.s., for $t = 0, \dots, T$. We may think of $G_t(\omega)$ as the set of "nonnegative" positions in d assets at time t and in the state of the world $\omega \in \Omega$.

Define

$$A_s^t := \sum_{u=s}^{t} L^0(-G_u, \mathscr{F}_u).$$

Traders in the market are assumed to have the information structure $\mathscr{F}_t$, $t = 0, \dots, T$ and to act in a self-financing way hence A_s^t corresponds to the set of attainable positions at time t when trading starts at time s with 0 initial endowment.

Starting with initial endowment $\xi \in L^0(\mathbb{R}^d, \mathscr{F}_s)$ at time s an investor may attain by time t the payoffs which are elements of the set

$$A_s^t(\xi) := \{\xi + V : V \in A_s^t\}.$$

Definition 1 We say that there is *efficient friction* **EF** if G_t is proper a.s. for $t = 0, \dots, T$.

In the model of [19] condition **EF** means that there are no freely exchangeable assets (even allowing for indirect transfers), see Proposition 5.3 of [7].

2 Arbitrage and Price Systems

The following notion of arbitrage goes back to [24].

Definition 2 *Weak absence of arbitrage* $\mathbf{NA}^w$ holds if $V \in A_0^T \cap L^0(\mathbb{R}^d_+, \mathscr{F}_T)$ implies $V = 0$ a.s.

This concept is called simply "no arbitrage" **NA** in [31] and [11].

The above notion means that a trader with 0 initial endowment is unable to attain a non-zero nonnegative portfolio. In the present study we shall require that a position with positive liquidation value should be attainable from a given endowment only if the liquidation value of this endowment is itself positive. To put it in another way, arriving at an a.s. solvent position is possible only if the investor was solvent already at the beginning.

Definition 3 There is *no sure gain in liquidation value* **NGV** if for all $0 \le s \le T$, $\xi \in L^0(\mathbb{R}^d, \mathscr{F}_s)$ and $V \in A_s^T(\xi)$ with $V_T \in G_T$ a.s. we necessarily have $\xi \in G_s$ a.s.

Remark 1 Take an arbitrary $V \in A_0^T(\xi)$ for some $\xi \in \mathbb{R}^d$. Write

$$V_t := \xi + \sum_{j=0}^{t} \eta_j, \quad 0 \le t \le T,$$

where $\eta_j \in L^0(-G_j, \mathscr{F}_j)$. If $V = V_T \in G_T$ a.s. then (as $V_T \in A_{T-1}^T(V_{T-1})$) **NGV** implies that $V_{T-1} \in G_{T-1}$ a.s. Repeating the procedure we get $V_t \in G_t$ a.s. for $0 \le t \le T$ and even $\xi \in G_0$.

Thus **NGV** implies that an investor ending up with a solvent position must have been solvent during the whole period of his activity in the market. We do not know if this latter condition is strictly weaker than **NGV**.

Remark 2 The analogous property holds in frictionless arbitrage-free models, too (nonnegative terminal wealth implies that the portfolio value is nonnegative over the whole trading period), and it is a strictly weaker condition than absence of arbitrage (as the example in Remark 4 below witnesses). This shows that **NGV** is a fairly natural and mathematically appealing concept.

Now we introduce the dual variables of this model.

Definition 4 Suppose **EF**. The set of martingales $(Z_t)_{s \le t \le v}$ such that $Z_t \in G_t^* \setminus \{0\}$ (resp. $Z_t \in \operatorname{int} G_t^*$), $s \le t \le v$ is denoted by $\mathscr{M}_s^v(G^* \setminus \{0\})$ (resp. $\mathscr{M}_s^v(\operatorname{int} G^*)$).

Following the terminology of [31], elements of $\mathscr{M}_0^T(G^* \setminus \{0\})$ (resp. $\mathscr{M}_0^T(\operatorname{int} G^*)$) are called *consistent price systems* (resp. *strictly consistent price systems*). Also the expression "shadow price" is sometimes used in the related literature. In the absence of the **EF** property one has to replace interior by the relative interior in the above definition.

At a given time instant s, elements of $L^0(\operatorname{int} G_s^*, \mathscr{F}_s)$ can be interpreted as positive price functionals (as they are positive on the set of positions $G_s \setminus \{0\}$). The following condition requires that given any positive price today, we may prolongate it in a consistent (martingale) way up to the end of the time horizon.

Definition 5 We say that *prices are consistently extendable* **PCE**, if for each $0 \le s \le T$ and $X \in L^1(\operatorname{int} G_s^*, \mathscr{F}_s)$ there exists $Z \in \mathscr{M}_s^T(\operatorname{int} G^*)$ satisfying $Z_s = X$ a.s.

Theorem 1 *Under* **EF**, *the following are equivalent*:

(a) **NGV**
(b) *For each* $0 \le t \le T-1$ *and* $\xi \in L^0(\mathbb{R}^d, \mathscr{F}_t)$ *such that* $\xi \in G_{t+1}$ *a.s. we have* $\xi \in G_t$ *a.s.*
(c) **PCE**

Item (b) above means that it is not possible to make a gain in liquidation value in one step. Thus, as in the classical case of frictionless markets, if it is not possible to produce an "arbitrage" in one step then it is not feasible in several steps either (see e.g. [6]). This property seems fairly convenient and in contrast with the previously used no-arbitrage concepts in the transaction cost case where such a one-step reduction fails.

Remark 3 The notion **PCE** expresses a certain concatenation property for the dual processes. It allows reducing multistep arguments to single step ones (see $*$) in the proof of Theorem 1 in Sect. 4) and may be useful in e.g. dynamic programming type arguments, too. More comments on the role of **PCE** are given in Sect. 5.

For comparison, we recall the main result of [11] about the characterization of $\mathbf{NA}^w$ in the two-dimensional case. Its multidimensional extension is a delicate issue, see [31] and [16].

Theorem 2 *Let* $d := 2$, *then* $\mathbf{NA}^w$ *holds iff*

$$\mathscr{M}_0^T(G^* \setminus \{0\}) \neq \emptyset.$$

Theorem 2 holds for arbitrary d if Ω is finite, see [24], but fails for infinite Ω and $d \ge 3$, see [31] and [11]. In [24] a stronger no-arbitrage concept $\mathbf{NA}^s$ has also been introduced and found equivalent to the existence of a strictly consistent price system, for Ω finite. This result was generalized in [19] for arbitrary Ω under the **EF** hypothesis. Dropping **EF** one needs the $\mathbf{NA}^r$ concept of [31] to get strictly consistent price systems. All these developments are presented in [21].

Remark 4 It is unclear how to generalize Theorem 1 for the case where **EF** fails. The analogous statement for frictionless markets fails (as noticed by D. Rokhlin): take $S_0 = 1$, $P(S_1 = 2) = P(S_1 = 1) = 1/2$. Starting from initial capital $x < 0$ and using strategy $\phi \in \mathbb{R}$ the value of our portfolio will be $x + \phi(S_1 - S_0)$ which is never positive a.s., hence the analogue of **NGV** holds for this model. However, the usual no-arbitrage property fails and there is no equivalent martingale measure for the process S. In contrast, under **EF**, **NGV** implies, in particular, the existence of (strictly) consistent price systems (by Theorem 1) and is thus stronger than any of the previous no-arbitrage concepts $\mathbf{NA}^w, \mathbf{NA}^s, \mathbf{NA}^r$.

3 Markets with One Risky Asset

Let us consider a market with one risky asset (stock) traded at bid and ask prices given by the adapted processes

$$0 < \underline{S}_t \le \overline{S}_t, \quad t = 0, \dots, T. \tag{1}$$

We assume that the riskless asset has constant price 1 (no interest rate).

Positions in the riskless asset and in the stock can be described by elements $(x, y) \in \mathbb{R}^2$. Non-negative ("solvent") positions at time t are those which satisfy either

$$x \ge 0, \qquad y \ge 0, \tag{2}$$

or

$$x \le 0 \quad \text{and} \quad y\underline{S}_t + x \ge 0, \tag{3}$$

or

$$y \le 0 \quad \text{and} \quad y\overline{S}_t + x \ge 0. \tag{4}$$

We may thus define the $\mathscr{F}_t$-measurable random (closed) cones

$$G_t := \{(x, y) \in \mathbb{R}^2 : (2) \text{ or } (3) \text{ or } (4) \text{ holds}\}, \quad t = 0, \dots, T,$$

these sets a.s. contain the positive orthant.

It is easy to check that **EF** holds iff $\underline{S}_t < \overline{S}_t$ a.s. for each t and that

$$G_t^* = \{(w_1, w_2) : \underline{S}_t w_1 \le w_2 \le \overline{S}_t w_1\}.$$

Example 1 Let us consider the deterministic market with $\underline{S}_0 = \underline{S}_1 = 0.9$, $\overline{S}_0 = 1.2$, $\overline{S}_1 = 1.1$. The process $Z_0 = Z_1 = (1, 1)$ is a (strictly) consistent price system, hence this model is arbitrage-free in the sense of previous papers, e.g. [19] or [31].

However, the position $(20/17, -1)$ is such that it has liquidation value $-1 \times 1.2 + 20/17 < 0$ at time 0 and $-1 \times 1.1 + 20/17 > 0$ at time 1 and thus **NGV** fails in this model.

Failure of **NGV** arose because the bid-ask interval had shrunk from time 0 to time 1. We have the following result.

Theorem 3 *In the above setting assume* ***EF***. *Then* ***NGV*** *holds iff for each* $t \le T - 1$,

$$[\underline{S}_t, \overline{S}_t] \subseteq [\min \operatorname{supp} \operatorname{Law}(\underline{S}_{t+1}|\mathscr{F}_t), \sup \operatorname{supp} \operatorname{Law}(\overline{S}_{t+1}|\mathscr{F}_t)] \quad a.s. \tag{5}$$

(Here the sup may be infinity.)

Remark 5 To put it differently, **NGV** fails iff at some time instant either the conditional future bid is a.s. larger than the present bid price or the conditional future ask is a.s. smaller then the present ask price. We explain a possible argument that such events should not occur in a liquid market: let us suppose that e.g. the bid price a.s. will increase. Then nobody would be willing to sell the risky asset which is in contrast with the observed uninterrupted trading. Hence the **NGV** seems to be acceptable from the intuitive point of view, though rather stringent. It can be interpreted as "the situation can not get more favourable a.s.".

In the case of constant proportional transaction costs the picture simplifies even more. Let S_t be an adapted process representing stock price and let $0 < \varepsilon < 1$ be the constant transaction cost coefficient. Define

$$\underline{S}_t := S_t(1-\varepsilon), \qquad \overline{S}_t := S_t(1+\varepsilon). \tag{6}$$

The notation **NGV**-ε refers to the absence of gain in liquidation value property in this market model with parameter ε; **PCE**-ε, $G^*(\varepsilon)$, $\mathbf{NA}^w$ are also self-explanatory.

The rigidity of the transaction cost structure in the present case causes that the notion **NGV** does not depend on the effective value of ε (as long as it is strictly positive).

Corollary 1 *In the above model with proportional transaction costs the following are equivalent*:

(a) *For some* ε, **NGV**-ε.
(b) *For some* ε, **PCE**-ε.
(c) *For* each ε, **NGV**-ε.
(d) *For* each ε, **PCE**-ε.
(e) $\mathbf{NA}^w$-ε *holds for each* ε.
(f) *There exist* (*strictly*) *consistent price systems for each* ε.
(g) $P(S_{t+1} - S_t < 0|\mathscr{F}_t) < 1$ *a.s. and* $P(S_{t+1} - S_t > 0|\mathscr{F}_t) < 1$ *a.s., for all* $0 \le t \le T-1$.

Remark 6 Frictionless absence of arbitrage for the price process S can be characterized as follows (see [6]): for each $0 \le t \le T-1$, for almost all ω, either $P(S_{t+1} = S_t|\mathscr{F}_t)(\omega) = 1$ or both $P(S_{t+1} - S_t \le 0|\mathscr{F}_t)(\omega) < 1$ and $P(S_{t+1} - S_t \ge 0|\mathscr{F}_t)(\omega) < 1$ hold (i.e. price can be neither nonincreasing a.s. nor nondecreasing a.s.).

Note that (g) above is *almost* frictionless absence of arbitrage, but slightly weaker: it prohibits only that the price strictly increases (decreases) with (conditional) probability 1. When (g) holds it may thus happen that $P(S_{t+1} - S_t \ge 0|\mathscr{F}_t) = 1$ a.s. and $P(S_{t+1} - S_t = 0|\mathscr{F}_t) < 1$ a.s. on some event $A \in \mathscr{F}_t$ of positive probability, which cannot hold if there is no arbitrage without transaction costs.

Example 2 If the $\mathscr{F}_t$-conditional support of both $\underline{S}_{t+1}$ and $\overline{S}_{t+1}$ equals $[0,\infty)$ a.s. then (5) and thus **NGV** trivially holds. In the context of continuous-time models the analogous "conditional full support" **CFS** condition was introduced in [13] (inspired by [12]). It was shown that **CFS** implies (f) (in fact, all the statements (a)–(f)) of

Corollary 1 for processes with continuous trajectories. **CFS** is valid for a large class of often used processes, see [13] and [4].

4 Proofs

Instead of showing the closedness of the set of attainable claims in an appropriate topology and relying on consequences of the Hahn-Banach theorem in infinite dimensional spaces, in this paper (as well as in [26, 28, 29] and [30]) finite-dimensional separation theorems are combined with measurable selection. Thus we return to the spirit of the seminal paper [6].

In the sequel we will use the measurable selection theorem as stated in e.g. III. 44–45 of [8]. Standard verifications for the measurability of certain subsets will be omitted.

We need the notion of conditional expectation for random sets (see [25] for further details). The following result is Lemma 4.3 of [26].

Lemma 1 *Let $\mathscr{G} \subset \mathscr{H}$ be σ-algebras. Let $C \subset U$ be an $\mathscr{H}$-measurable random convex compact set. Then there exists a $\mathscr{G}$-measurable random convex compact set $E(C|\mathscr{G}) \subset U$ satisfying*

$$L^0(E(C|\mathscr{G}), \mathscr{G}) = \{E(\vartheta|\mathscr{G}):\ \vartheta \in L^0(C, \mathscr{H})\}.$$

If $0 \in C$ *a.s. then* $0 \in E(C|\mathscr{G})$ *a.s., too.*

We also recall Lemma 4.4 of the same paper.

Lemma 2 *Let $\mathscr{G} \subset \mathscr{H}$ be σ-algebras and let $C \subset U$ be a $\mathscr{H}$-measurable random compact convex set containing* 0 *such that* $\operatorname{int} C \neq \emptyset$ *a.s. Then* $\operatorname{int} E(C|\mathscr{G}) \neq \emptyset$ *a.s. and*

$$\{E(\vartheta|\mathscr{G}):\ \vartheta \in L^0(\operatorname{int} C, \mathscr{H})\} \subset L^0(\operatorname{int} E(C|\mathscr{G}), \mathscr{G}) \subset \{E(\vartheta|\mathscr{G}):\ \vartheta \in L^0(2\operatorname{int} C, \mathscr{H})\}.$$

Proof of Theorem 1 We add one more equivalent statement which is of auxiliary nature.

$(*)$ $H_t := \operatorname{cone}(\operatorname{int} E(G^*_{t+1} \cap U|\mathscr{F}_t)) \supset \operatorname{int} G^*_t$ for each $0 \le t < T$.

(a) $\Rightarrow$ (b): Rather trivial since

$$\xi \in G_{t+1} = G_{t+1} + \mathbb{R}^d_+ = \mathbb{R}^d_+ - (-G_{t+1}),$$

thus (using measurable selection) there is $\eta \in L^0(-G_{t+1}\mathscr{F}_{t+1})$ such that $\xi + \eta \in \mathbb{R}^d_+ \subset G_T$. As $\xi + \eta \in A^T_t(\xi)$, **NGV** implies $\xi \in G_t$ a.s.

(b) $\Rightarrow$ $(*)$: Let us suppose that $(*)$ fails, i.e. for some t, on a set $A \in \mathscr{F}_t$ of positive measure, there exist points $p_\omega \in G^*_t(\omega) \setminus \overline{H}_t(\omega)$, $\omega \in A$. By the measurable selection theorem we may assume that pI_A is an $\mathscr{F}_t$-measurable function. The one point set p_ω can be separated from the closed set $\overline{H}_t(\omega)$ (Corollary 11.4.2 of [27]), i.e. there exist $\xi_\omega \in U$ and $a_\omega \in \mathbb{R}$ such that $\langle \xi_\omega, p_\omega\rangle < a_\omega$ and $\langle \xi_\omega, q\rangle \geq a_\omega$ for all $q \in H_t(\omega)$ and for all $\omega \in A$. As H_t is a cone, we may choose $a_\omega := 0$. Again, we may assume that ξ is an $\mathscr{F}_t$-measurable random variable and $\xi = 0$ on A^c.

We claim that $\xi \in G_{t+1}$ a.s. Indeed, if this failed on a set $B \in \mathscr{F}_{t+1}$ of positive measure then there would be $\zeta \in L^0([\operatorname{int} G^*_{t+1}] \cap U, \mathscr{F}_{t+1})$ such that $\langle \zeta_\omega, \xi_\omega\rangle < 0$ for $\omega \in B$. Choosing $\varepsilon > 0$ small enough,

$$E\langle \zeta I_B + f\varepsilon I_{B^c}, \xi\rangle < 0, \tag{7}$$

where f is an arbitrary $\mathscr{F}_{t+1}$-measurable selector of $[\operatorname{int} G^*_{t+1}] \cap U$. But, by Lemma 1, $E(\zeta I_B + \varepsilon f I_{B^c}|\mathscr{F}_t) \in H_t$, hence $E\langle \xi, E(\zeta I_B + \varepsilon f I_{B^c}|\mathscr{F}_t)\rangle \geq 0$ should hold by the choice of ξ, but this contradicts (7).

To finish the proof of this implication notice that, on A, $\langle \xi, p\rangle < 0$, so necessarily $P(\xi \notin G_t) > 0$, hence (b) cannot hold. We may conclude that (b) $\Rightarrow$ $(*)$.

$(*)$ $\Rightarrow$ (c): Take any $0 \leq s \leq T-1$ and $X \in L^1(\operatorname{int} G^*_s, \mathscr{F}_s)$. We will construct $(Z_u)_{s\leq u\leq T}$ by induction, starting from $Z_s := X$. Let us suppose that the martingale $Z_s, \dots, Z_u$ has been defined, we shall construct Z_{u+1}. By $(*)$, there is $\alpha \in L^0((0,\infty), \mathscr{F}_u)$ such that $\alpha Z_u \in \operatorname{int} E(G^*_{u+1} \cap U|\mathscr{F}_u)$ a.s. Hence Lemma 2 implies that there is $Y \in L^0(2\operatorname{int}[G^*_{u+1} \cap U], \mathscr{F}_{u+1})$ such that $E(Y|\mathscr{G}_u) = \alpha Z_u$. Setting $Z_{u+1} := Y/\alpha$ we see that $E(Z_{u+1}|\mathscr{F}_u) = Z_u$ at least in a generalized sense, but as Z_u was integrable, Z_{u+1} is also integrable, and the inductive step is shown.

(c) $\Rightarrow$ (a): Take $V_k = \xi + \sum_{j=s}^k \eta_j \in A^k_s(\xi)$, $s \leq k \leq T$ where $\eta_j \in L^0(-G_j, \mathscr{F}_j)$ and ξ is $\mathscr{F}_s$-measurable. Then $V_T \in A^T_s(\xi)$. Suppose that also $V_T \in L^0(G_T, \mathscr{F}_T)$. We shall show $\xi \in G_s$ a.s., arguing by contradiction.

If $\xi \in G_s$ failed on some set $B \in \mathscr{F}_s$ of positive measure we could choose $\zeta \in L^1(\operatorname{int} G^*_s, \mathscr{F}_s)$ in such a way that

$$\langle \zeta, \xi\rangle < 0 \quad \text{on } B. \tag{8}$$

Then, by (c), there is $Z \in \mathscr{M}^T_s(\operatorname{int} G^*)$ such that $Z_s = \zeta$. We have $\langle Z_T, V_T\rangle \geq 0$ a.s., hence, as $\eta_T \in -G_T$, also $\langle Z_T, V_{T-1}\rangle \geq 0$ a.s. Taking $\mathscr{F}_{T-1}$-conditional expectations we find that $\langle Z_{T-1}, V_{T-1}\rangle \geq 0$ a.s.

Iterating the same argument, finally $\langle Z_s, \xi\rangle = \langle \zeta, \xi\rangle \geq 0$ a.s., which is nonsense by (8). □

Proof of Theorem 3 Roughly speaking, **NGV** is true iff the $\mathscr{F}_t$-conditional projection of G^*_{t+1} contains G^*_t (see $(*)$) in the previous proof) and this latter is equivalent to (5). The technical difficulty lies in proving that each element of the wedge

$$\{(u, v) \in \mathbb{R}^2_+ : u \operatorname{ess.inf} \underline{S}_1 < v < u \operatorname{ess.sup} \overline{S}_1\}$$

is the expectation of some 2-dimensional random variable (X, Y) satisfying a.s.

$$X\underline{S}_1 < Y < X\overline{S}_1,$$

and then proving the conditional version of this statement (with $\mathscr{F}_t$ replacing the trivial σ-algebra and $\underline{S}_{t+1}, \overline{S}_{t+1}$ replacing $\underline{S}_1, \overline{S}_1$). We need Lemmata 3 and 4 to tackle this.

Let us first suppose (5). Take $\theta \in L^1(\operatorname{int} G_t^*, \mathscr{F}_t)$. Let ν be the (regular) $\mathscr{F}_t$-conditional law of $(\underline{S}_{t+1}, \overline{S}_{t+1})$. As **EF** holds, we have $\nu(W) = 1$ a.s., see the definition of W in Lemma 3 below. Define

$$D(\omega) := \{(u, v) \in \mathbb{R}_+^d : u \min \operatorname{supp} \operatorname{Law}(\underline{S}_{t+1}|\mathscr{F}_t) \le v \le u \sup \operatorname{supp} \operatorname{Law}(\overline{S}_{t+1}|\mathscr{F}_t)\},$$

(if the sup is ∞ we mean $<$ instead of $\le$ in the second inequality) note that the conditional probabilities above are the marginals ν_1, ν_2 of ν.

As (5) implies $G_t^*(\omega) \subset D(\omega)$ a.s. (and hence also $\operatorname{int} G_t^*(\omega) \subset \operatorname{int} D(\omega)$), we may apply Lemma 4. The random variable

$$\tilde{\theta} := f(\underline{S}_{t+1}, \overline{S}_{t+1}, \omega)$$

satisfies $E(\tilde{\theta}|\mathscr{F}_t) = \theta$ in the generalized sense and $\tilde{\theta} \in L^1$ since $\theta \in L^1$. We claim that $\tilde{\theta} \in G_{t+1}^*$ a.s. Indeed, if we had e.g. $\tilde{\theta}_2/\tilde{\theta}_1 > \overline{S}_{t+1}$ on some event of positive probability, we would necessarily have some $A \in \mathscr{F}_t$ with $P(A) > 0$ such that for $\omega \in A$

$$0 < P(\tilde{\theta}_2/\tilde{\theta}_1 > \overline{S}_{t+1}|\mathscr{F}_t) = \nu(\{f_2(x, y)/f_1(x, y) > y\}, \omega),$$

which contradicts (13).

We will now prove that the existence of $\tilde{\theta}$ entails (b) of Theorem 1, which then implies **NGV**. Let $\xi \in L^0(G_{t+1}, \mathscr{F}_t)$. If we had $\xi \notin G_t$ on $A \in \mathscr{F}_t$ then for some $\theta \in L^1(\operatorname{int} G_t^*, \mathscr{F}_t)$ one has $\langle \theta, \xi \rangle < 0$ on A. We clearly have $\langle \tilde{\theta}, \xi \rangle \ge 0$ by $\tilde{\theta} \in G_{t+1}^*$ hence, taking $\mathscr{F}_t$-conditional expectation, also $\langle \theta, \xi \rangle \ge 0$ a.s. This is possible only if $P(A) = 0$.

To see the converse implication, we remark that, by the proof of Theorem 1, **NGV** implies $(*)$. Define $D(\omega)$ as in the previous argument.

If (5) failed on some $A \in \mathscr{F}_t$ with $P(A) > 0$ we could take $\theta \in L^1(\operatorname{int} G_t^*, \mathscr{F}_t)$ such that $\theta \notin D$ on A. $(*)$ implies that, for a suitable $\alpha \in L^0((0, \infty), \mathscr{F}_t)$, we have $\alpha\theta \in \operatorname{int} E(G_{t+1}^* \cap U|\mathscr{F}_t)$ a.s., hence Lemma 1 provides $\tilde{\theta} \in G_{t+1}^* \cap U$ with $E(\tilde{\theta}|\mathscr{F}_t) = \alpha\theta$. Then $E(\tilde{\theta}/\alpha|\mathscr{F}_t) = \theta$ and $\tilde{\theta}/\alpha$ is integrable as θ is.

Clearly, for any $\hat{\theta} \in L^1(G_{t+1}^*, \mathscr{F}_{t+1})$ we have $E(\hat{\theta}|\mathscr{F}_t) \in D$ a.s. But $\hat{\theta} := \tilde{\theta}/\alpha$ contradicts the choice of θ. □

Proof of Corollary 1 The equivalence (e) $\Longleftrightarrow$ (f) follows from Theorem 2, noting that an $\varepsilon/2$-consistent price system is ε-strictly consistent.[1]

[1]This equivalence also follows from the main result of [31], see also [20], and remains true in a multidimensional setting whenever we consider a family of polyhedral cone valued processes $G(\varepsilon), \varepsilon > 0, \operatorname{int} G_t(\varepsilon) \supset \mathbb{R}_+^d \setminus \{0\}$ such that for $0 < \varepsilon' < \varepsilon$ we have $G_t(\varepsilon) \setminus \{0\} \subset \operatorname{int} G_t(\varepsilon')$ a.s. for each t.

Theorem 1 implies (a) $\iff$ (b) and (c) $\iff$ (d). The implication from (f) to (d) has been shown in [14], even for continuous-time models. For completeness we reproduce this proof in Lemma 5 below. (c) $\Rightarrow$ (a) and (d) $\Rightarrow$ (f) are trivial. So it remains to show (a) $\Rightarrow$ (c) and (a) $\iff$ (g). We get from Theorem 3 that (a) is equivalent to

$$[S_t(1-\varepsilon), S_t(1+\varepsilon)] \subseteq [\min \operatorname{supp} \operatorname{Law}(S_{t+1}(1-\varepsilon)|\mathscr{F}_t), \sup \operatorname{supp} \operatorname{Law}(S_{t+1}(1+\varepsilon)|\mathscr{F}_t)], \quad (9)$$

a.s. for each t, but if this holds for one particular ε it holds for *all* $0 < \varepsilon < 1$, showing (a) $\Rightarrow$ (c). Notice also that (9) is equivalent to

$$S_t \in [\min \operatorname{supp} \operatorname{Law}(S_{t+1}|\mathscr{F}_t), \sup \operatorname{supp} \operatorname{Law}(S_{t+1}|\mathscr{F}_t)],$$

a.s. for all t, which is precisely (g). □

5 Conclusion

We now make some comparisons to studies on continuous-time models. The equivalence between (e) and (f) of Corollary 1 has been shown in [14] for processes $(S_t)_{t\in[0,T]}$ with continuous trajectories. Note that (f) $\Rightarrow$ (d) also holds in this model class (see [14]).

It follows from results of [21] (see also [23] and [9]) that for continuous cone-valued processes the so-called condition **B** implies the closedness of the set of payoffs of admissible strategies (bounded from below in a suitable sense), which is crucial when establishing hedging theorems or building up a proper duality theory (e.g. for utility maximization). Condition **B** is very closely related to **PCE**. In fact, it can easily be checked that the above cited results can be derived assuming **PCE**. In discrete time (under **EF**) one can show that condition **B** implies **PCE** (and is a strictly stronger property).

In [2] the closedness of payoffs is proved for price processes with jumps admitting a strictly consistent price system. However, a different class of trading strategies and a different notion of admissibility are used. Under **PCE** one can show (along the lines of Proposition 9 in [2]) that this new notion of admissibility coincides with the usual one in [23].

We expect that (some suitable modification of) Theorem 1 holds for continuous-time models with several assets and with a general bid-ask structure.

Appendix

The following technical material was needed in the above arguments. We denote by $C_0(\overline{\mathbb{R}^2})$ the set of continuous functions on $\mathbb{R}^2$ vanishing at infinity, equipped with the supremum norm. This is a Polish space.

Lemma 3 *Let ν be a probability measure on $\mathbb{R}^2$ such that $\nu(W) = 1$ where*

$$W := \{(u, v) \in \mathbb{R}^2 : \ u, v > 0, u < v\}.$$

Let ν_1, ν_2 denote its marginals. Define the cone D as

$$D := \{(u, v) \in \mathbb{R}^2_+ : u \min \operatorname{supp} \nu_1 \le v \le u \sup \operatorname{supp} \nu_2\}.$$

(*If the sup is ∞ we mean $<$ instead of $\le$ in the second inequality.*) *For any point $p = (p_1, p_2) \in \operatorname{int} D$ there exist ν-a.s.* (*strictly*) *positive functions $f_1, f_2 \in C_0(\overline{\mathbb{R}^2})$ such that*

$$p_i = \int_{\mathbb{R}^2} f_i(u, v)\nu(du, dv), \quad i = 1, 2.$$

and $f_1(u, v)u < f_2(u, v) < f_1(u, v)v$ hold ν-almost surely.

Proof We shall use tricks from the key lemma of [22]. We first remark that $\operatorname{int} D \neq \emptyset$ under the conditions of the present Lemma. Let K be the set of $p \in \operatorname{int} D$ for which the statement holds. The relative interior of a convex set coincides with that of its closure (see III.2.1.8 in [15]). Hence, as K is a convex set, it suffices to show that K is dense in D. It is thus enough to prove, for each $p \in D$, the existence of ν-a.s. strictly positive $g_1^n, g_2^n \in C_0(\overline{\mathbb{R}^2})$ such that

$$\int_{\mathbb{R}^2} g_i^n d\nu \to p_i, \quad n \to \infty, \quad i = 1, 2 \tag{10}$$

and

$$g_1^n(u, v)u < g_2^n(u, v) < g_1^n(u, v)v, \quad \nu\text{-a.s.}$$

We may and will assume $p_1 = 1$. There are $\underline{p} \le p_2 \le \overline{p}$ such that $\underline{p} \in \operatorname{supp} \nu_1$ and $\overline{p} \in \operatorname{supp} \nu_2$. Hence p_2 is a convex combination of $\underline{p}$ and $\overline{p}$ and we may and will assume that e.g. $p_2 = \underline{p}$, i.e. $p_2 \in \operatorname{supp} \nu_1$.

We could try to take e.g.

$$g_1^n(u, v) := \frac{I_{[p_2 - 1/n, p_2 + 1/n]}(u)}{\nu_2([p_2 - 1/n, p_2 + 1/n])},$$
$$g_2^n(u, v) := (u + (1/n)[v - u])g_1^n(u, v),$$

but these fail to be continuous and do not vanish at infinity (in fact, g_1^n does not depend on v). Since for the measurable selection argument in Lemma 4 below we need to insist on $g_i^n \in C_0(\overline{\mathbb{R}^2})$, $i = 1, 2$ a more complicated construction is used.

Define

$$J_n := [p_2 - 1/n, p_2 + 1/n] \times [-N(n), N(n)], \quad c_n = \nu(J_n), \quad n \ge 1,$$

and choose $N(n)$ so large that $c_n > 0$, $n \ge 1$. This is possible by the definition of the support of a measure. Define also

$$K_n := [p_2 - 2/n, p_2 + 2/n] \times [-N(n) - 1, N(n) + 1].$$

Denote by $a_n(u,v), b_n(u,v)$ the distance of the point (u,v) from J_n, K_n^c, respectively. These are continuous functions of (u,v). We further set

$$\ell(u,v) := 1/(u^2+v^2+1), \quad u,v \in \mathbb{R}.$$

Define

$$\begin{aligned} g_1^n(u,v) &= \frac{1}{c_n} I_{J_n}(u,v) + \frac{1}{n}\ell(u,v) I_{K_n^c}(u,v) \\ &\quad + \left[\left\{1 - h_n\left(\frac{a_n(u,v)}{a_n(u,v)+b_n(u,v)}\right)\right\}\frac{1}{n}\ell(u,v)\right. \\ &\quad \left. + h_n\left(\frac{a_n(u,v)}{a_n(u,v)+b_n(u,v)}\right)\frac{1}{c_n}\right] I_{K_n \setminus J_n}, \\ g_2^n(u,v) &:= (u + (1/n)[(v-u)\wedge 1]) g_1^n(u,v), \end{aligned}$$

where the continuous function $h_n : [0,1] \to [0,1]$ is chosen in such a way that $h_n(1) = 0$, $h_n(0) = 1$ and

$$\int_{K_n \setminus J_n} (|u|+1) h_n\left(\frac{a_n(u,v)}{a_n(u,v)+b_n(u,v)}\right)\nu(du,dv) < c_n/n. \tag{11}$$

(This holds for $h_n(x) := (1-x)^{\alpha(n)}$, $x \in [0,1]$ with $\alpha(n) > 0$ suitably large.)

The functions g_1^n, g_2^n are ν-almost surely strictly positive, lie in $C_0(\overline{\mathbb{R}^2})$ and clearly

$$g_1^n(u,v)u < g_2^n(u,v) < g_1^n(u,v)v \quad \nu\text{-a.s.}$$

Let (X,Y) be random variables with joint law ν. The expectation of the first term of $g_1^n(X,Y)$ equals 1, the second term tends to 0 uniformly by boundedness of ℓ and the third term vanishes as $n \to \infty$ due to (11); hence $Eg_1^n(X,Y) \to 1$. Similarly, the first term of $Eg_2^n(X,Y)$ tends to p_2, $(|u|+1)/\ell(u,v)$ is bounded hence the second term goes to 0 uniformly and the third term vanishes again by (11). This shows (10) and thus finishes the proof of this lemma. □

Lemma 4 *Let $\mathscr{G} \subset \mathscr{F}$ a σ-algebra. Let $\nu(A,\omega)$, $A \in \mathscr{B}(\mathbb{R}^2)$, $\omega \in \Omega$ be a $\mathscr{G}$-measurable stochastic kernel*[2] *with marginals ν_1, ν_2 such that a.s. $\nu(W) = 1$ (see the statement of the previous lemma). Define*

$$D(\omega) := \{(u,v) \in \mathbb{R}_+^2 : u \min \operatorname{supp} \nu_1(\omega) \le v \le u \sup \operatorname{supp} \nu_2(\omega)\},$$

with $<$ replacing $\le$ if the sup is infinity.

[2] I.e. for almost all ω, ν is a probability measure on the Boreliens of $\mathbb{R}^2$; for each fixed $A \in \mathscr{B}(\mathbb{R}^2)$ it is a $\mathscr{G}$-measurable function.

Let $\theta \in L^0(\operatorname{int} D, \mathscr{G})$ be arbitrary. Then there exist a $\mathscr{B}(\mathbb{R}^2) \otimes \mathscr{G}$-measurable functions f_1, f_2 such that, for almost all ω,

$$\int_{\mathbb{R}^2} f_i(x, y, \omega)\nu(dx, dy, \omega) = \theta_i(\omega), \quad i = 1, 2 \tag{12}$$

and for almost all ω, $\nu(\cdot, \cdot, \omega)$-a.s.

$$f_1(x, y, \omega)x < f_2(x, y, \omega) < f_1(x, y, \omega)y. \tag{13}$$

Proof By Lemma 3, for (almost all) fixed ω there exist $f_i(\cdot, \cdot, \omega) \in C_0(\overline{\mathbb{R}^2})$, $i = 1, 2$, such that (12) holds true. The measurable selection theorem enables us to choose f_1, f_2 in a $\mathscr{G}$-measurable way.[3] Then $f_i(\cdot, \cdot, \cdot)$, $i = 1, 2$ are continuous in the first two variables and measurable in the third one, hence they are $\mathscr{B}(\mathbb{R}^2) \otimes \mathscr{G}$-measurable (p. 70 of [3]). □

The following Lemma is taken from [14].

Lemma 5 *Let $\mathscr{M}_0^T(G^*(\eta) \setminus \{0\}) \neq \emptyset$ hold for each $0 < \eta < 1$ and take $f \in L^1(\operatorname{int} G_t^*(\varepsilon), \mathscr{F}_t)$. Then there exists $W \in \mathscr{M}_t^T(\operatorname{int} G^*(\varepsilon))$ such that $W_t = f$.*

Proof Let $\tilde{W}_t(\eta)$ be a fixed element of $\mathscr{M}_0^T(G^*(\eta) \setminus \{0\})$, for each η. We have, by definition,

$$1 - \eta \leq g_t(\eta) := \frac{\tilde{W}_t^2(\eta)}{\tilde{W}_t^1(\eta) S_t} \leq 1 + \eta, \tag{14}$$

and for $h_t := f_2/(f_1 S_t)$ we have

$$1 - \varepsilon < h_t < 1 + \varepsilon.$$

We define an $\mathscr{F}_t$-measurable partition of Ω:

$$A_n^+ := \{1 + n\varepsilon/(n+1) > h_t \geq 1 + (n-1)\varepsilon/n\}, \tag{15}$$

$$A_n^- := \{1 - (n-1)\varepsilon/n > h_t \geq 1 - n\varepsilon/(n+1)\}, \quad n \geq 1. \tag{16}$$

Now set

$$W_u^1 := f_1 \sum_{n=1}^{\infty} I_{A_n^+ \cup A_n^-} \frac{\tilde{W}_u^1(\varepsilon/(9n+3))}{\tilde{W}_t^1(\varepsilon/(9n+3))}$$

and

$$W_u^2 := f_1 \sum_{n=1}^{\infty} I_{A_n^+ \cup A_n^-} \frac{h_t}{g_t(\varepsilon/(9n+3))} \frac{\tilde{W}_u^2(\varepsilon/(9n+3))}{\tilde{W}_t^1(\varepsilon/(9n+3))}, \quad t \leq u \leq T.$$

[3]We have again skipped here some standard arguments that the set of suitable pairs $(\omega, f_i(\cdot, \cdot, \omega))$ is measurable with respect to the product σ-algebra.

It is clear that W_u, $t \le u \le T$ is a martingale and

$$W_t = f.$$

Moreover, on A_n^+ we have for $t \le u \le T$,

$$1-\varepsilon < \frac{1-\varepsilon/(9n+3)}{1+\varepsilon/(9n+3)} \le \frac{h_t}{g_t}\left(1-\frac{\varepsilon}{9n+3}\right) \le \frac{W_u^2}{W_u^1 S_u} \le \frac{h_t}{g_t}\left(1+\frac{\varepsilon}{9n+3}\right)$$
$$\le \left(1+\frac{n\varepsilon}{n+1}\right)\frac{1}{1-\varepsilon/(9n+3)}\left(1+\frac{\varepsilon}{9n+3}\right) < 1+\varepsilon,$$

where g_t is an abbreviation for $g_t(\varepsilon/(9n+3))$. Similarly, on A_n^-,

$$1+\varepsilon > \frac{1+\varepsilon/(9n+3)}{1-\varepsilon/(9n+3)} \ge \frac{h_t}{g_t}\left(1+\frac{\varepsilon}{9n+3}\right) \ge \frac{W_u^2}{W_u^1 S_u} \ge \frac{h_t}{g_t}\left(1-\frac{\varepsilon}{9n+3}\right)$$
$$\ge \left(1-\frac{n\varepsilon}{n+1}\right)\frac{1}{1+\varepsilon/(9n+3)}\left(1-\frac{\varepsilon}{9n+3}\right) > 1-\varepsilon,$$

hence $(W_u)_{t \le u \le T} \in \mathscr{M}_t^T(\operatorname{int} G^*)$. □

References

1. Bouchard, B.: No-arbitrage in discrete-time markets with proportional transaction costs and general information structure. Finance Stoch. **10**, 276–297 (2006)
2. Campi, L., Schachermayer, W.: A super-replication theorem in Kabanov's model of transaction costs. Finance Stoch. **10**, 579–596 (2006)
3. Castaing, C., Valadier, M.: Convex Analysis and Measurable Multifunctions. Lecture Notes in Mathematics, vol. 580. Springer, Berlin (1977)
4. Cherny, A.: Brownian moving averages have conditional full support. Ann. Appl. Probab. **18**(5), 1825–1830 (2008)
5. Cvitanić, J., Karatzas, I.: Hedging and portfolio optimization under transaction costs: a martingale approach. Math. Financ. **6**, 133–165 (1996)
6. Dalang, R.C., Morton, A., Willinger, W.: Equivalent martingale measures and no-arbitrage in stochastic securities market models. Stochastics Stoch. Rep. **29**, 185–201 (1990)
7. Delbaen, F., Kabanov, Yu.M., Valkeila, E.: Hedging under transaction costs in currency markets: a discrete-time model. Math. Financ. **12**, 45–61 (2002)
8. Dellacherie, C., Meyer, P.A.: Probabilities and Potential. North-Holland, Amsterdam (1978)
9. Denis, E., De Vallière, D., Kabanov, Yu.M.: Hedging of American options under transaction costs. Preprint (2008)
10. De Vallière, D., Kabanov, Yu.M., Stricker, Ch.: No-arbitrage criteria for financial markets with transaction costs and incomplete information. Finance Stoch. **11**, 237–251 (2007)
11. Grigoriev, P.G.: On low dimensional case in the fundamental asset pricing theorem with transaction costs. Stat. Decis. **23**, 33–48 (2005)
12. Guasoni, P.: No arbitrage with transaction costs, with fractional Brownian motion and beyond. Math. Financ. **16**, 569–582 (2006)
13. Guasoni, P., Rásonyi, M., Schachermayer, W.: Consistent price systems and face-lifting pricing under transaction costs. Ann. Appl. Probab. **18**, 491–520 (2008)

14. Guasoni, P., Rásonyi, M., Schachermayer, W.: The fundamental theorem of asset pricing for continuous processes with small transaction costs. Ann. Finance (2009, forthcoming)
15. Hirriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms, vol. I. Springer, Berlin (1993)
16. Jacka, S., Berkaoui, A., Warren, J.: No arbitrage and closure results for trading cones with transaction costs. Preprint (2007)
17. Jouini, É., Kallal, H.: Martingales and arbitrage in securities markets with transaction costs. J. Econ. Theory **66**, 178–197 (1995)
18. Kabanov, Yu.M.: Hedging and liquidation under transaction costs in currency markets. Finance Stoch. **3**, 237–248 (1999)
19. Kabanov, Yu.M., Rásonyi, M., Stricker, Ch.: No-arbitrage criteria for financial markets with efficient friction. Finance Stoch. **6**, 371–382 (2002)
20. Kabanov, Yu.M., Rásonyi, M., Stricker, Ch.: On the closedness of the sum of convex cones in L^0 and the robust no-arbitrage property. Finance Stoch. **7**, 403–412 (2003)
21. Kabanov, Yu.M., Safarian, M.: Markets with transaction costs. In: Mathematical Theory. Springer, Berlin (2009, forthcoming)
22. Kabanov, Yu.M., Stricker, Ch.: On equivalent martingale measures with bounded densities. In: Séminaire de Probabilités, XXXV. Lecture Notes in Mathematics, vol. 1755, pp. 139–148. Springer, Berlin (2001)
23. Kabanov, Yu.M., Stricker, Ch.: Hedging of contingent claims under transaction costs. In: Sandman, K., Schönbucher, Ph. (eds.) Advances in Finance and Stochastics. Essays in honour of Dieter Sondermann, pp. 125–136. Springer, Berlin (2002)
24. Kabanov, Yu.M., Stricker, Ch.: The Harrison-Pliska arbitrage pricing theorem under transaction costs. J. Math. Econ. **35**, 185–196 (2001)
25. Molchanov, I.: Theory of Random Sets. Springer, Berlin (2005)
26. Rásonyi, M.: New methods in the arbitrage theory of markets with transaction costs. In: Séminaire de Probabilités, XLI. Lecture Notes in Mathematics, vol. 1934, pp. 455–462. Springer, Berlin (2008)
27. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
28. Rokhlin, D.B.: Martingale selection problem and asset pricing in finite discrete time. Electron. Commer. Probab. **12**, 1–8 (2006)
29. Rokhlin, D.B.: A theorem on martingale selection for a random sequence with relatively open convex values. Math. Notes **81**, 543–548 (2007)
30. Rokhlin, D.B.: A constructive no-arbitrage criterion with transaction costs in the case of finite discrete time. Theory Probab. Appl. **52**, 93–107 (2008)
31. Schachermayer, W.: The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. Math. Financ. **14**, 19–48 (2004)

On the Linear and Nonlinear Generalized Bayesian Disorder Problem (Discrete Time Case)

Albert N. Shiryaev and Pavel Y. Zryumov

Abstract This paper considers the generalized Bayesian disorder problem in the discrete time case with two types of the penalty function—the linear and the nonlinear ones. The main results for these cases are given in Theorems 1 and 2, respectively.

Keywords Disorder problem · Optimal stopping problem

Mathematics Subject Classification (2000) Primary 60G40 · Secondary 60J50 · 60J05 · 62C99

1 Linear Penalty Case

1. Let θ be a parameter taking values in the set $\{0, 1, \dots, \infty\}$. Suppose that on the probability space $(\Omega, \mathscr{F}, \mathrm{P})$ we consider the sequence of independent random variables $X = (X_0, X_1, \dots, X_n, \dots)$. For given θ we suppose that random variable X_n with $n < \theta$ has the distribution $\mathrm{F}_\infty(x)$ and for $n \ge \theta$ the distribution function is $\mathrm{F}_0(x)$. Their density (with respect to the distribution $(\mathrm{F}_0 + \mathrm{F}_\infty)/2$) will be denoted by $f_\infty(x)$ and $f_0(x)$, $x \in \mathbb{R}$. For given θ let $\mathrm{P}_\theta = \mathrm{Law}(X \mid \theta, \mathrm{P})$ be the law of X, and let $\mathscr{F}_n = \sigma(X_0, X_1, \dots, X_n)$. For simplicity of considerations we assume that $d\mathrm{F}_0 \ll d\mathrm{F}_\infty$.

Denote by $\mathfrak{M}_T$ the class of finite stopping times (with respect to $(\mathscr{F}_n)_{n\ge 0}$) such that $\mathrm{E}_\infty \tau \ge T$ where $T > 0$.

The *generalized Bayesian problem* (with a *linear* penalty function) consists in finding stopping time τ_T^*, if it exists, such that

$$\sum_{\theta=0}^{\infty} \mathrm{E}_\theta (\tau_T^* - \theta)^+ = \inf_{\tau \in \mathfrak{M}_T} \sum_{\theta=0}^{\infty} \mathrm{E}_\theta (\tau - \theta)^+. \tag{1}$$

A.N. Shiryaev (✉)
Steklov Mathematical Institute, Russian Academy of Sciences, Gubkina str. 8, Moscow 119991, Russia
e-mail: albertsh@mi.ras.ru

P.Y. Zryumov
Department of Mechanics and Mathematics, Moscow State University, Leninskie gory 1, Moscow 119992, Russia
e-mail: pavel.zryumov@gmail.com

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*, DOI 10.1007/978-3-642-02608-9_12,

The similar problem for the case of Brownian motion was formulated and investigated in [1, 3, 4]. It turns out that the methods of these papers (especially of [1, 4]) permit to describe the structure of the optimal stopping time for the generalized Bayesian problem in case of discrete time too.

2. The following theorem plays here the key role.

Theorem 1 *For any finite stopping time τ from $\mathfrak{M}_T$*

$$\sum_{\theta=0}^{\infty} \mathrm{E}_\theta(\tau-\theta)^+ = \mathrm{E}_\infty \sum_{n=0}^{\tau-1} \psi_n, \tag{2}$$

where the Markov sequence $\psi=(\psi_n)_{n\ge 0}$ satisfies the recurrent equations

$$\psi_n = (1+\psi_{n-1})\frac{f_0(X_n)}{f_\infty(X_n)}, \qquad \psi_{-1}=0. \tag{3}$$

Proof It is evident that

$$(\tau-\theta)^+ = \sum_{k=1}^{\infty} \mathrm{I}(\tau-\theta\ge k) = \sum_{k=\theta+1}^{\infty} \mathrm{I}(k\le\tau).$$

So,

$$\mathrm{E}_\theta(\tau-\theta)^+ = \sum_{k=\theta+1}^{\infty} \mathrm{E}_\theta \mathrm{I}(k\le\tau). \tag{4}$$

Since $k-1\ge\theta$ and $\{k\le\tau\}\in\mathscr{F}_{k-1}$ we find

$$\mathrm{E}_\theta \mathrm{I}(k\le\tau) = \mathrm{E}_k \frac{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}{d(\mathrm{P}_k|\mathscr{F}_{k-1})}\mathrm{I}(k\le\tau) = \mathrm{E}_\infty \frac{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}{d(\mathrm{P}_k|\mathscr{F}_{k-1})}\mathrm{I}(k\le\tau), \tag{5}$$

where $\mathrm{P}_\theta|\mathscr{F}_{k-1}$ and $\mathrm{P}_k|\mathscr{F}_{k-1}$ are restrictions of the measures P_θ and P_k onto the σ-algebra $\mathscr{F}_{k-1}$.

Introduce the notation

$$L_n = \frac{d(\mathrm{P}_0|\mathscr{F}_n)}{d(\mathrm{P}_\infty|\mathscr{F}_n)}, \quad n\ge 0, \qquad L_{-1}=1.$$

Then

$$\frac{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}{d(\mathrm{P}_k|\mathscr{F}_{k-1})} = \frac{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}{d(\mathrm{P}_\infty|\mathscr{F}_{k-1})} = \frac{\frac{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}{d(\mathrm{P}_0|\mathscr{F}_{k-1})}}{\frac{d(\mathrm{P}_\infty|\mathscr{F}_{k-1})}{d(\mathrm{P}_0|\mathscr{F}_{k-1})}} = \frac{L_{k-1}}{\frac{d(\mathrm{P}_0|\mathscr{F}_{k-1})}{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})}}$$

$$= \frac{L_{k-1}}{\frac{d(\mathrm{P}_0|\mathscr{F}_{\theta-1})}{d(\mathrm{P}_\infty|\mathscr{F}_{\theta-1})}} = \frac{L_{k-1}}{L_{\theta-1}}, \tag{6}$$

where we used for $k-1 \ge \theta$ the property

$$\frac{d(\mathrm{P}_0|\mathscr{F}_{k-1})}{d(\mathrm{P}_\theta|\mathscr{F}_{k-1})} = \frac{f_0(X_0)\cdots f_0(X_{\theta-1})\cdot f_0(X_\theta)\cdots f_0(X_{k-1})}{f_\infty(X_0)\cdots f_\infty(X_{\theta-1})\cdot f_0(X_\theta)\cdots f_0(X_{k-1})} = \frac{d(\mathrm{P}_0|\mathscr{F}_{\theta-1})}{d(\mathrm{P}_\infty|\mathscr{F}_{\theta-1})}.$$

From (4)–(6) we deduce

$$\sum_{\theta=0}^{\infty} \mathrm{E}_\theta(\tau-\theta)^+ = \mathrm{E}_\infty \sum_{\theta=0}^{\infty}\left[\sum_{n=\theta+1}^{\infty} \mathrm{I}(n\le\tau)\frac{L_{n-1}}{L_{\theta-1}}\right] = \mathrm{E}_\infty \sum_{\theta=0}^{\infty}\sum_{n=\theta+1}^{\tau}\frac{L_{n-1}}{L_{\theta-1}} = \mathrm{E}_\infty \sum_{n=1}^{\tau}\left(\sum_{\theta=0}^{n-1}\frac{L_{n-1}}{L_{\theta-1}}\right) \tag{7}$$

which implies that (2) holds for

$$\psi_n = \sum_{\theta=0}^{n}\frac{L_n}{L_{\theta-1}}, \qquad \psi_{-1}=0. \tag{8}$$

The recurrent equations (3) follow immediately from (8):

$$\psi_n = \frac{L_n}{L_{n-1}}\left(1+\sum_{\theta=0}^{n-1}\frac{L_{n-1}}{L_{\theta-1}}\right) = (1+\psi_{n-1})\frac{f_0(X_n)}{f_\infty(X_n)}.$$

□

3.

Remark 1 Statistical procedures based on the process $\psi = (\psi_n)_{n\ge 0}$ are well known in the statistical literature as "Shiryaev-Roberts procedures".

Remark 2 From Theorem 1 it follows that for solving the conditionally Bayesian problem (1) in the class $\mathfrak{M}_T$ we need to solve the following conditionally optimal stopping problem: to find a stopping time $\tau_T^* \in \mathfrak{M}_T$ such that

$$\mathrm{E}_\infty \sum_{n=0}^{\tau_T^*-1} \psi_n = \inf_{\tau\in\mathfrak{M}_T} \mathrm{E}_\infty \sum_{n=0}^{\tau-1}\psi_n. \tag{9}$$

The standard method of solution of such problems is based on ideas of the Lagrange multipliers: for any $C>0$, to find a stopping time $\tilde{\tau}_C$ such that

$$\mathrm{E}_\infty \sum_{n=0}^{\tilde{\tau}_C-1}(\psi_n - C) = \inf_\tau \mathrm{E}_\infty \sum_{n=0}^{\tau-1}(\psi_n - C) \tag{10}$$

where imfimum is taken over all finite stopping times τ.

If there exist $C = C(T)$ such that $\mathrm{E}_\infty \tilde{\tau}_{C(T)} = T$, then this stopping time is optimal in the class $\mathfrak{M}_T$ and so we may take $\tau_T^* = \tilde{\tau}_{C(T)}$.

Remark 3 The "classical" Bayesian disorder problem consists (see [3]) in finding stopping time $\tau_T^* \in \mathfrak{M}_T$, if it exists, such that

$$\mathrm{P}(\tau_T^* \le \theta) + c\mathrm{E}(\tau_T^* - \theta)^+ = \inf_{\tau \in \mathfrak{M}_T} \big(\mathrm{P}(\tau \le \theta) + c\mathrm{E}(\tau - \theta)^+\big),$$

where θ has a geometric prior distribution with parameter p and $c > 0$. The optimal τ_T^* for this problem was given in [3].

It turns out that when $p \to 0$ this Bayesian problem "converges" to the generalized Bayesian problem. This fact together with the result similar to the statement of Theorem 1 were also obtained in [2].

2 Nonlinear Penalty Case

1. Instead of the "linear case" (1) now we consider the following nonlinear problem with "nonlinear penalty function" $G = G(n)$, $n \ge 0$: to find a stopping time τ_T^* in the class $\mathfrak{M}_T$, $T > 0$, such that

$$\sum_{\theta=0}^{\infty} \mathrm{E}_\theta G((\tau_T^* - \theta)^+) = \inf_{\tau \in \mathfrak{M}_T} \sum_{\theta=0}^{\infty} \mathrm{E}_\theta G((\tau - \theta)^+). \tag{11}$$

In the linear case $G(n) = n$, $n \ge 0$, Theorem 1 claims that in problem (1) there exists one sufficient statistics, namely $\psi = (\psi_n)_{n \ge 0}$, which is a Markov sequence (with respect to P_∞). Now we want to find under what conditions on function $G = G(n)$, $n \ge 0$, there exist a finite number of statistics which form a multidimensional Markov system of sufficient statistics for solving the corresponding stopping time problem.

2. Suppose that $G = G(n)$, $n \ge 0$, is a nondecreasing function with $G(0) = 0$ and

$$G(n) = \sum_{k=1}^{n} g(k), \qquad \text{where } g(k) \ge 0 \text{ for } k > 0.$$

If $\tau \ge \theta$ then

$$\begin{aligned} G(\tau - \theta) &= \sum_{k=1}^{\tau-\theta} g(k) = \sum_{k=1}^{\infty} \mathrm{I}(1 \le k \le \tau - \theta) g(k) \\ &= \sum_{n=\theta+1}^{\infty} \mathrm{I}(n \le \tau) g(n - \theta). \end{aligned}$$

Thus,

$$\begin{aligned}
\mathrm{E}_\theta G((\tau-\theta)^+) &= \mathrm{E}_\theta \mathrm{I}(\tau \ge \theta) G(\tau-\theta) \\
&= \mathrm{E}_\theta \mathrm{I}(\tau \ge \theta) \sum_{n=\theta+1}^{\infty} \mathrm{I}(n \le \tau) g(n-\theta) \\
&= \mathrm{E}_\theta \sum_{n=\theta+1}^{\infty} \mathrm{I}(n \le \tau) g(n-\theta) = \sum_{n=\theta+1}^{\infty} g(n-\theta) \mathrm{E}_\theta \mathrm{I}(n \le \tau). \quad (12)
\end{aligned}$$

Since $\{n \le \tau\} \in \mathscr{F}_{n-1}$, we deduce, using (6), that for $n-1 \ge \theta$

$$\begin{aligned}
\mathrm{E}_\theta \mathrm{I}(n \le \tau) &= \mathrm{E}_n \frac{d\mathrm{P}_\theta}{d\mathrm{P}_n} \mathrm{I}(n \le \tau) = \mathrm{E}_n \frac{d(\mathrm{P}_\theta|\mathscr{F}_{n-1})}{d(\mathrm{P}_n|\mathscr{F}_{n-1})} \mathrm{I}(n \le \tau) \\
&= \mathrm{E}_\infty \frac{d(\mathrm{P}_\theta|\mathscr{F}_{n-1})}{d(\mathrm{P}_\infty|\mathscr{F}_{n-1})} \mathrm{I}(n \le \tau) = \mathrm{E}_\infty \frac{L_{n-1}}{L_{\theta-1}} \mathrm{I}(n \le \tau).
\end{aligned}$$

Substituting this into (12) implies that

$$\mathrm{E}_\theta G((\tau-\theta)^+) = \sum_{n=\theta+1}^{\infty} g(n-\theta) \mathrm{E}_\infty \frac{L_{n-1}}{L_{\theta-1}} \mathrm{I}(n \le \tau).$$

Thus

$$\begin{aligned}
\sum_{\theta=0}^{\infty} \mathrm{E}_\theta G((\tau-\theta)^+) &= \sum_{\theta=0}^{\infty} \left[\sum_{n=\theta+1}^{\infty} g(n-\theta) \mathrm{E}_\infty \frac{L_{n-1}}{L_{\theta-1}} \mathrm{I}(n \le \tau) \right] \\
&= \mathrm{E}_\infty \sum_{n=1}^{\tau} \left[\sum_{\theta=0}^{n-1} g(n-\theta) \frac{L_{n-1}}{L_{\theta-1}} \right] \\
&= \mathrm{E}_\infty \sum_{n=1}^{\tau} \Psi_{n-1}(g) = \mathrm{E}_\infty \sum_{n=0}^{\tau-1} \Psi_n(g) \quad (13)
\end{aligned}$$

where

$$\Psi_n(g) = \sum_{\theta=0}^{n} g(n+1-\theta) \frac{L_n}{L_{\theta-1}}.$$

From (13) we find the following representation:

$$\inf_{\tau \in \mathfrak{M}_T} \sum_{\theta=0}^{\infty} \mathrm{E}_\theta (\tau-\theta)^+ = \inf_{\tau \in \mathfrak{M}_T} \mathrm{E}_\infty \left[\sum_{n=0}^{\tau-1} \Psi_n(g) \right]. \quad (14)$$

3. To get for the problem (11) a finite number of Markovian sufficient statistics let us assume that for $t \geq 0$

$$g(t) = \sum_{m=0}^{M} \sum_{k=0}^{K} c_{mk} e^{\lambda_m t} t^k, \tag{15}$$

where $\lambda_0 = 0$.

Consider first the case $K = 0$:

$$g(t) = \sum_{m=0}^{M} c_{m0} e^{\lambda_m t}. \tag{16}$$

Under this assumption

$$\begin{aligned} \Psi_n(g) &= \sum_{\theta=0}^{n} \sum_{m=0}^{M} c_{m0} e^{\lambda_m (n+1-\theta)} \frac{L_n}{L_{\theta-1}} \\ &= c_{00} \sum_{\theta=0}^{n} \frac{L_n}{L_{\theta-1}} + \sum_{m=1}^{M} \sum_{\theta=0}^{n} c_{m0} e^{\lambda_m (n+1-\theta)} \frac{L_n}{L_{\theta-1}}. \end{aligned} \tag{17}$$

Put

$$\psi_n = \sum_{\theta=0}^{n} \frac{L_n}{L_{\theta-1}}, \qquad \psi_{-1} = 0 \tag{18}$$

and

$$\psi_n^{(m,0)} = \sum_{\theta=0}^{n} e^{\lambda_m (n+1-\theta)} \frac{L_n}{L_{\theta-1}} = \sum_{\theta=0}^{n} \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}}, \tag{19}$$

where $L_n^{(m)} = e^{\lambda_m n} L_n$.

Then

$$\begin{aligned} \psi_n^{(m,0)} &= \sum_{\theta=0}^{n} \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} = \frac{L_n^{(m)}}{L_{n-1}^{(m)}} + \sum_{\theta=0}^{n-1} \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} \\ &= e^{\lambda_m} \frac{L_n}{L_{n-1}} + \frac{L_n^{(m)}}{L_{n-1}^{(m)}} \sum_{\theta=0}^{n-1} \frac{L_{n-1}^{(m)}}{L_{\theta-1}^{(m)}} = e^{\lambda_m} \frac{f_0(X_n)}{f_\infty(X_n)} (1 + \psi_{n-1}^{(m,0)}). \end{aligned}$$

So, we have the following system of equations for ψ_n and $(\psi_n^{(1,0)}, \dots, \psi_n^{(M,0)})$:

$$\begin{cases} \psi_n = \frac{f_0(X_n)}{f_\infty(X_n)} (1 + \psi_{n-1}), \qquad \psi_{-1} = 0, \\ \psi_n^{(m,0)} = e^{\lambda_m} \frac{f_0(X_n)}{f_\infty(X_n)} (1 + \psi_{n-1}^{(m,0)}), \qquad \psi_{-1}^{(m,0)} = 0. \end{cases} \tag{20}$$

It is interesting to note that (with respect to the measure P_∞) all sequences $\psi = (\psi_n)_{n\ge 0}$, $\psi^{(m,0)} = (\psi_n^{(m,0)})_{n\ge 0}$ are Markovian and by (17)

$$\Psi_n(g) = c_{00}\psi_n + \sum_{m=1}^{M} c_{m0}\psi_n^{(m,0)}, \tag{21}$$

i.e. in the case (16) $\Psi_n(g)$ is a sum of the Markovian sequences from the set $(\psi, \psi^{(1,0)}, \dots, \psi^{(M,0)})$.

4. Now assume that $M = 0$. In this case

$$g(t) = \sum_{k=0}^{K} c_{0k}t^k = c_{00} + \sum_{k=1}^{K} c_{0k}t^k. \tag{22}$$

Denote for $1 \le k \le K$

$$\psi_n^{(0,k)} = \sum_{\theta=0}^{n} (n+1-\theta)^k \frac{L_n}{L_{\theta-1}}. \tag{23}$$

Then

$$\begin{aligned}
\psi_n^{(0,k)} &= \sum_{\theta=0}^{n}\sum_{i=0}^{k} C_k^i (n-\theta)^i \frac{L_n}{L_{\theta-1}} \\
&= \sum_{\theta=0}^{n} \frac{L_n}{L_{\theta-1}} + \sum_{i=1}^{k}\sum_{\theta=0}^{n} C_k^i (n-\theta)^i \frac{L_n}{L_{\theta-1}} \\
&= \psi_n + \sum_{i=1}^{k} \frac{L_n}{L_{n-1}} \sum_{\theta=0}^{n-1} C_k^i (n-\theta)^i \frac{L_{n-1}}{L_{\theta-1}} \\
&= \psi_n + \frac{f_0(X_n)}{f_\infty(X_n)} \sum_{i=1}^{k} C_k^i \psi_{n-1}^{(0,i)} \\
&= \frac{f_0(X_n)}{f_\infty(X_n)} \left(\sum_{i=0}^{k} C_k^i \psi_{n-1}^{(0,i)} + 1 \right),
\end{aligned} \tag{24}$$

where $\psi_{n-1}^{(0,0)} = \psi_{n-1}$.

So, for the case (22) the family of statistics $(\psi_n, \psi_n^{(0,1)}, \dots, \psi_n^{(0,K)})_{n\ge 0}$ is Markovian satisfying to the following system:

$$\begin{cases}
\psi_n = \frac{f_0(X_n)}{f_\infty(X_n)}(1+\psi_{n-1}), \qquad \psi_{-1} = 0, \\
\psi_n^{(0,k)} = \frac{f_0(X_n)}{f_\infty(X_n)}(\sum_{i=0}^{k} C_k^i \psi_{n-1}^{(0,i)} + 1), \qquad \psi_{-1}^{(0,k)} = 0, \quad k = 1, \dots, K.
\end{cases} \tag{25}$$

5. Consider finally the general case (15).
Denote for $1 \le m \le M$, $1 \le k \le K$

$$\psi_n^{(m,k)} = \sum_{\theta=0}^{n} e^{\lambda_m (n+1-\theta)} (n+1-\theta)^k \frac{L_n}{L_{\theta-1}} = \sum_{\theta=0}^{n} (n+1-\theta)^k \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} \tag{26}$$

where $L_n^{(m)} = e^{\lambda_m n} L_n$. Then

$$\begin{aligned}
\psi_n^{(m,k)} &= \sum_{\theta=0}^{n} \sum_{i=0}^{k} C_k^i (n-\theta)^i \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} \\
&= \sum_{i=1}^{k} \sum_{\theta=0}^{n} C_k^i (n-\theta)^i \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} + \sum_{\theta=0}^{n} \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} \\
&= \sum_{i=1}^{k} \sum_{\theta=0}^{n-1} C_k^i (n-\theta)^i \frac{L_n^{(m)}}{L_{\theta-1}^{(m)}} + \psi_n^{(m,0)} \\
&= \sum_{i=1}^{k} C_k^i \frac{L_n^{(m)}}{L_{n-1}^{(m)}} \sum_{\theta=0}^{n-1} (n-\theta)^i \frac{L_{n-1}^{(m)}}{L_{\theta-1}^{(m)}} + \psi_n^{(m,0)} \\
&= e^{\lambda_m} \frac{f_0(X_n)}{f_\infty(X_n)} \sum_{i=1}^{k} C_k^i \psi_{n-1}^{(m,i)} + \psi_n^{(m,0)}.
\end{aligned} \tag{27}$$

Together with (20) the formula (27) gives recurrent equations:

$$\psi_n^{(m,k)} = e^{\lambda_m} \frac{f_0(X_n)}{f_\infty(X_n)} \left[\sum_{i=0}^{k} C_k^i \psi_{n-1}^{(m,i)} + 1 \right], \qquad \psi_{-1}^{(m,k)} = 0. \tag{28}$$

Hence, we get the following extension of Theorem 1 for nonlinear penalty functions.

Theorem 2 *For the case of independent observations and the nonlinear penalty function $G(n) = \sum_{k=1}^{n} g(k)$ with g given by (15) the system*

$$(\psi_n, \psi_n^{(m,k)})_{n \ge 0}, \quad 0 \le m \le M, \ 0 \le k \le K \tag{29}$$

is a Markovian family with recurrent equations (20), (25), (28). This family forms a system of sufficient statistics in the sense that they define $\Psi_n(g)$:

$$\Psi_n(g) = \sum_{m=0}^{M} \sum_{k=0}^{K} c_{mk} \psi_n^{(m,k)}. \tag{30}$$

Example 1 If $G(n) = n, n \ge 0$, i.e. $g \equiv 1$, then there exists only one sufficient statistics $\psi = (\psi_n)_{n \ge 0}$.

Example 2 If $G(n) = n^2 + n$, i.e. $g(n) = 2n$, then

$$\Psi_n(g) = \Psi_{n-1}(g)\frac{f_0(X_n)}{f_\infty(X_n)} + 2\psi_n. \tag{31}$$

Remark 4 It is important to note that the recurrent equations of Theorems 1 and 2 can be directly extended on more general cases of nonindependent random variables: in all recurrent equations the ratios $f_0(X_n)/f_\infty(X_n)$ should be changed to L_n/L_{n-1}. (See more details about "θ-models" in [5].)

References

1. Feinberg, E.A., Shiryaev, A.N.: Quickest detection of drift change for Brownian motion in generalized Bayesian and minimax settings. Stat. Decis. **24**, 445–470 (2006)
2. Pollak, M., Tartakovsky, A.G.: On optimality properties of the Shiryaev-Roberts procedure. Private communication (2008)
3. Shiryaev, A.N.: On optimal methods in quickest detection problems. Theory Probab. Appl. **8**, 22–46 (1963)
4. Shiryaev, A.N.: Generalized Bayesian nonlinear quickest detection problems: on Markov family of sufficient statistics. In: Mathematical Control Theory and Finance, pp. 377–386. Springer, Berlin (2008)
5. Shiryaev, A.N.: On stochastic models and optimal methods in the problems of the quickest detection. Teor. Veroyatnost. Primenen. **53**, 417–436 (2008). English translation will appear in Theory Probab. Appl.

Long Time Growth Optimal Portfolio with Transaction Costs

Lukasz Stettner

Abstract Discrete and continuous growth optimal portfolio optimization over long time horizon is studied. Proportional transaction costs consisting of fixed proportional plus proportional to the volume of transaction are considered. An obligatory diversification is imposed, which allows the process of portions of capital invested in assets to be ergodic. Existence of solutions to suitable Bellman equations is proved and the form of optimal strategies is shown. For continuous time model an additional fixed deterministic delay in transactions is assumed.

Keywords Growth of portfolio · Transaction costs · Markov process · Bellman equation · Long term portfolio selection

Mathematics Subject Classification (2000) 91B28 · 93E20

1 Introduction

Assume we are given a market with m risky assets. Denote by $S_i(t)$ the price of the i-th asset at time t. Let

$$\frac{S_i(t+s)}{S_i(t)} = \zeta_i(t, s, z(\cdot), \xi(\cdot)), \tag{1}$$

where $(z(t)) \in D$ forms a right continuous Markov process on a complete separable metric space D with transition operator $P_t(z, dy)$ describing the evolution of economic factors, $(\xi(t))$ stands for a Markov process with independent increments and the processes $(z(t))$ and $(\xi(t))$ are independent. We furthermore assume that the function $\zeta(t, s, z(\cdot), \xi(\cdot))$ is positive and right continuous in s and t, depends on the trajectory of $z(u)$ and on the increments $\xi(u) - \xi(t)$ for $u \in [t, t+s]$, $\zeta(t, 0, z(\cdot), \xi(\cdot)) = 1$ and ζ is a multiplicative functional i.e. $\zeta_i(t, s, z(\cdot), \xi(\cdot))(\omega) = \zeta_i(s, z(\cdot), \xi(\cdot))(\theta_t\omega)$ with θ_t standing for the Markov shift operator (see [4]) of the pair $(z(t), \xi(t))$ and suitably defined $\zeta_i(s, z(\cdot), \xi(\cdot))$ and

$$\zeta_i(t+s, z(\cdot), \xi(\cdot))(\omega) = \zeta_i(t, z(\cdot), \xi(\cdot))(\omega)\zeta_i(s, z(\cdot), \xi(\cdot))(\theta_t\omega).$$

Research supported by MNiSzW grant P03A 01328.

L. Stettner (✉)
Institute of Mathematics, Polish Academy of Sciences, Sniadeckich 8, 00-956 Warsaw, Poland
e-mail: stettner@impan.gov.pl

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*, DOI 10.1007/978-3-642-02608-9_13,

As typical examples of such multiplicative functional ζ one can consider an exponent of sum of Riemann and stochastic integrals with respect to Levy or Brownian motion noise (see e.g. [3, 9] or [11]).

Denote by $X^-(t)$ the wealth process at time t before possible transactions and by $X(t)$ the wealth process after possible transactions. Let $\pi_i^-(t)$ be the portion of the wealth process invested in the i-th asset at time t before possible transactions and $\pi_i(t)$ the portion of the wealth located in the i-th asset after transactions at time t. We shall say that $\pi(t) = (\pi_1(t), \ldots, \pi_m(t))^T$ (where T stands for the transpose) and similarly $\pi^-(t)$ form portfolios at time t after or before possible transactions.

Denote by S^0 the polyhedral set $\{(\nu_1, \ldots, \nu_m)^T : \nu_i \geq 0, \sum_{i=1}^m \nu_i \leq 1\}$ and by S the simplex $\{(\nu_1, \ldots, \nu_m)^T \in S^0 : \sum_{i=1}^m \nu_i = 1\}$.

For given $\pi \in S^0$ let $g(\pi) = (g_1(\pi), \ldots, g_m(\pi))^T$, where $g_i(\pi) = \frac{\pi_i}{\sum_{j=1}^m \pi_j}$. After change of portfolio from π to π' the wealth X is diminished by $c(\hat{\pi} - \pi)X$, where $\hat{\pi}$ is a certain element of S^0 (we shall see in Lemma 1 below that it is defined in a unique way) such that $\pi' = g(\hat{\pi})$ and for $\nu \in S^0 - S^0$ (the algebraic difference of the sets S^0)

$$c(\nu) = \kappa + \sum_{i=1}^m c_{1i}(\nu_i)^+ + \sum_{i=1}^m c_{2i}(\nu_i)^- \tag{2}$$

with $0 < c_{1i}, c_{2i} < 1 - \kappa$ and $1 > \kappa \geq 0$ and where $(\nu_i)^+ = \max\{0, \nu_i\}$ while $(\nu_i)^- = \max\{-\nu_i, 0\}$. Given portfolio π and wealth X we can change portfolio to π' if there exists $\hat{\pi}$ such that

$$X(c(\hat{\pi} - \pi)) = X - X \sum_{i=1}^m \hat{\pi}_i \tag{3}$$

and $g(\hat{\pi}) = \pi'$. Consequently given π we can choose portfolio π' if and only if there is $\hat{\pi} \in S^0$ such that

$$\sum_{i=1}^m \hat{\pi}_i + c(\hat{\pi} - \pi) = 1 \tag{4}$$

where

$$\pi' = g(\hat{\pi}). \tag{5}$$

Given $\pi, \pi' \in S$ define the function

$$F^{\pi,\pi'}(\delta) := \delta + c(\delta\pi' - \pi). \tag{6}$$

We have

Lemma 1 *There is a unique continuous function* $e : S \times S \mapsto [0, 1 - \kappa]$ *such that for* $\pi, \pi' \in S$ *we have*

$$F^{\pi,\pi'}(e(\pi, \pi')) = 1. \tag{7}$$

Furthermore e is bounded away from 0 *and provided that* $\kappa > 0$ *we have*

$$e(\pi, \pi')e(\pi', \pi'') < e(\pi, \pi''). \tag{8}$$

Proof The inequality (8) only has to be proved since the remaining part follows directly from Lemma 1 of [19]. Notice that by the definition of e we have

$$e(\pi, \pi') + c(e(\pi, \pi')\pi' - \pi) = 1$$

and

$$e(\pi', \pi'') + c(e(\pi', \pi'')\pi'' - \pi') = 1.$$

Therefore using homogeneity of c

$$\begin{aligned} &e(\pi, \pi')e(\pi', \pi'') + (e(\pi', \pi'') - 1)\kappa + c(e(\pi, \pi')e(\pi', \pi'')\pi'' - e(\pi, \pi')\pi') \\ &\quad = e(\pi, \pi') = 1 - c(e(\pi, \pi')\pi' - \pi). \end{aligned} \tag{9}$$

Since

$$\begin{aligned} &c(e(\pi, \pi')e(\pi', \pi'')\pi'' - e(\pi, \pi')\pi') + c(e(\pi, \pi')\pi' - \pi) \\ &\quad \geq \kappa + c(e(\pi, \pi')e(\pi', \pi'')\pi'' - \pi) \end{aligned}$$

by (9) we obtain that $F^{\pi,\pi''}(e(\pi, \pi')e(\pi', \pi'')) < 1$ from which (8) immediately follows. □

Consequently, given an initial wealth process $X^-(t)$ and portfolio $\pi^-(t)$ we can choose any post transaction portfolio $\pi(t) \in S$. Then, as a result of transaction costs the wealth process is diminished to $X(t)$, where following (3) and (4)

$$X(t) = e(\pi^-(t), \pi(t))X^-(t). \tag{10}$$

Furthermore provided that there are no transactions in the time interval $[t, t+s)$ we have

$$\begin{aligned} X^-(t+s) &= \sum_{i=1}^{m} \frac{\pi_i(t)X(t)}{S_i(t)} S_i(t+s) \\ &= X(t) \sum_{i=1}^{m} \pi_i(t)\zeta_i(t, s, z(\cdot), \xi(\cdot)) \\ &:= X(t)\pi(t)^T \zeta(t, s, z(\cdot), \xi(\cdot)) \end{aligned} \tag{11}$$

and

$$\pi^-(t+s) = g(\pi(t) \diamond \zeta(t, s, z(\cdot), \xi(\cdot)), \tag{12}$$

with $(\pi(t) \diamond \zeta(t, s, z(\cdot), \xi(\cdot)))_i := \pi_i(t)\zeta_i(t, s, z(\cdot), \xi(\cdot))$.

In the paper we restrict the class of admissible portfolio strategies to those which do not allow multiple transactions at the same moment (due to (8) it is not optimal to make more than one transaction at the moment). We maximize the following growth portfolio rate

$$J((\pi(t)) = \liminf_{N\to\infty} \frac{1}{N} E\{\ln X(N)\}. \tag{13}$$

For this purpose we are interested in the ergodic behaviour of the process $(\pi(t))$. In general this is a typical transient process with absorbing states at the vertices of the simplex S. To "improve" its ergodic properties we obligatorily diversify our portfolio. This means as soon as the process $(\pi(t))$ leaves the set $S_\delta = \{(\nu_1, \nu_2, \ldots, \nu_m) \in S : \forall_{i=1,\ldots,m} \nu_i \geq \delta\}$ with a fixed $\delta > 0$ we make an obligatory transaction changing portfolio to $\pi' \in S_\delta$. In the continuous time case we shall additionally assume that after each transaction the next one can appear only after $h > 0$ units of time so that consecutive impulses are separated by a fixed deterministic value $h > 0$. Consequently we control process $(\pi(t))$ changing it in an impulse way. Each admissible control can be characterized by a sequence $V = (\tau_n, \pi'_n)$ consisting of the stopping times τ_n and new chosen at time τ_n portfolios π'_n.

Since

$$X^-(t) = X^-(0) \prod_{n=0}^{\infty} e(\pi^-(n), \pi(n))\pi(\tau_n)^T \zeta\left(\tau_n, \tau_{n+1} \wedge t - \tau_n \wedge t, z(\cdot), \xi(\cdot)\right), \tag{14}$$

our cost functional is of the form

$$J(V) = \liminf_{N\to\infty} \frac{1}{N} E\left\{ \ln\left(X^-(0) \prod_{n=0}^{\infty} e(\pi^-(n), \pi(n)) \right.\right.$$
$$\left.\left. \times \pi(\tau_n)^T \zeta\left(\tau_n, \tau_{n+1} \wedge N - \tau_n \wedge N, z(\cdot), \xi(\cdot)\right)\right)\right\}. \tag{15}$$

Portfolio selection with transaction costs with discounted cost has been studied in a number of papers starting from [12] followed in [11] and [9] and references therein. Long term portfolio selection has been considered first in [10] was continued for growth optimal portfolio (GOP) in [2] and for lognormal asset prices in [1]. An adaptive approach to such problem with proportional transaction costs was considered in [8, 17] and [6]. Portfolio selection with proportional fixed proportional costs plus proportional costs to the amount of transaction was considered first in [13] and then in [7]. This paper generalizes [13] and [7] in this sense that a very general model for asset prices is admitted. Two cases are solved separately: discrete time model, when we change portfolio at discrete time moments and continuous time model. In the case of a continuous time it is imposed additionally that after each portfolio change next possible portfolio change is allowed after deterministic $h > 0$ units of time. Growth optimal portfolio with fixed plus proportional transaction cost was considered (in this paper we have proportional costs: to the wealth process transaction and to the volume of transaction) in [14] and [15] under strong assumption concerning the asset prices growth. General result in this case seems to be still lacking.

2 Discrete Time Case

In this section we study discrete time case and therefore assume that $t = 0, 1, \dots$

$$\frac{S_i(t+1)}{S_i(t)} = \zeta_i(z(t+1), \xi(t+1)) \tag{16}$$

where $(z(t))$ is a Markov process on locally compact separable metric space D with transition operator $P(x, dy)$, and $(\xi(t))$ is a sequence of i.i.d. random variables, independent of $(z(t))$ with common law η. We assume furthermore that $(z(t))$ is uniformly ergodic in the sense (see [5]) that

$$\sup_{z,z' \in D} \sup_{A \in \mathscr{B}(D)} (P(z, A) - P(z', A)) := d < 1. \tag{17}$$

Our purpose is to maximize the following cost functional (an analog of (15))

$$J(V) = \liminf_{N\to\infty} \frac{1}{N} E\Bigg\{\sum_{t=0}^{N-1} \big(1_{\pi^-(t)\neq\pi(t)} \ln e(\pi^-(t), \pi(t)) + \ln\big(\pi(t)^T \zeta(z(t+1), \xi(t+1))\big)\big)\Bigg\} \tag{18}$$

over all impulse strategies $V = (\tau_n, \pi_n')$, where $\pi(\tau_n) = \pi_n' \in S_\delta$ with fixed $\delta > 0$. We shall approximate the cost functional by discounted cost with $\beta > 0$

$$J^\beta_{\pi,z}(V) = E_{\pi,z}\Bigg\{\sum_{t=0}^{\infty} e^{-\beta t} \big(1_{\pi^-(t)\neq\pi(t)} \ln e(\pi^-(t), \pi(t)) + \ln\big(\pi(t)^T \zeta(z(t+1), \xi(t+1))\big)\big)\Bigg\} \tag{19}$$

where and in the sequel by $E_{\pi,z}$ we denote conditional expectation given the process $(\pi(t), z(t))$ starts with $\pi(0) = \pi$ and $z(0) = z$. Let

$$w^\beta(\pi, z) = \sup_V J^\beta(\pi, z). \tag{20}$$

Assume that the mapping

$$(\pi, z) \mapsto r(\pi, z) := E_{\pi,z}\{\ln(\pi^T \zeta(z(1), \xi(1)))\} \tag{21}$$

is bounded continuous. We have

Proposition 1 *Assume that the transition operator of the pair* $(\pi(t), z(t))$ *is continuous in total variation norm on* $S_\delta \times D$, *i.e. for* $\pi_{(n)} \to \pi \in S_\delta$, $z_n \to z$

$$\sup_A \left| P_{\pi_{(n)},z_n}\{(\pi(1), z(1)) \in A\} - P_{\pi,z}\{(\pi(1), z(1)) \in A\} \right| \to 0 \tag{22}$$

as $n \to \infty$ *with supremum over all Borel subsets* A *of* $S \times D$. *Then the function* w^β *is bounded and continuous on* $S_\delta \times D$ *and on* $(S \setminus S_\delta) \times D$ *separately. Moreover* w^β *is a solution to the following Bellman equation*

$$w^\beta(\pi, z) = \max\left\{Mw^\beta(\pi, z), E_{\pi,z}\left\{\ln(\pi^T \zeta(z(1), \xi(1))) + e^{-\beta} w^\beta(\pi(1), z(1))\right\}\right\} \tag{23}$$

for $\pi \in S_\delta$ *with* $Mw(\pi, z) = \sup_{\pi' \in S_\delta}\{\ln e(\pi, \pi') + w(\pi', z)\}$ *and*

$$w^\beta(\pi, z) = Mw^\beta(\pi, z) \tag{24}$$

whenever $\pi \notin S_\delta$.

Proof Iterating (23) and (24) for all $n \geq 1$ we obtain

$$\begin{aligned} w^\beta(\pi, z) = \sup_\tau E_{\pi,z} \Bigg\{ & \sum_{t=0}^{\tau \wedge \sigma \wedge n-1} e^{-\beta t} \ln(\pi(t)^T \zeta(z(t+1), \xi(t+1))) \\ & + \chi_{\tau \wedge n < \sigma} e^{-\beta \tau \wedge n} Mw^\beta(\pi(\tau \wedge n), z(\tau \wedge n)) \vee w^\beta(\pi(\tau \wedge n), z(\tau \wedge n)) \\ & + \chi_{\sigma \leq \tau \wedge n} Mw^\beta(\pi(\sigma), z(\sigma)) \Bigg\} \end{aligned} \tag{25}$$

where $\sigma = \inf\{n \geq 0 : \pi(n) \notin S_\delta\}$ with supremum over all stopping times τ. Letting $n \to \infty$, taking into account that $\beta > 0$, and then substituting the form of Mw^β we finally obtain that the solution to (23) coincides with the value (20). Consequently it remains to show the existence of a continuous bounded solution to (25). Let for a bounded continuous functions q and w

$$\begin{aligned} F_q^\beta w(\pi, z) := \max\Big\{ & Mq(\pi, z), E_{\pi,z}\Big\{\ln(\pi^T \zeta(z(1), \xi(1))) \\ & + e^{-\beta}\big(\chi_{\pi(1) \in S_\delta} w(\pi(1), z(1)) + \chi_{\pi(1) \notin S_\delta} Mq(\pi(1), z(1))\big)\Big\}\Big\} \end{aligned} \tag{26}$$

for $\pi \in S_\delta$ and $F_q^\beta w(\pi, z) = Mq(\pi, z)$ for $\pi \notin S_\delta$.

Since $|a \vee b - a' \vee b'| \leq |a - a'| \vee |b - b'|$ we have that

$$|F_q^\beta w_1(\pi, z) - F_q^\beta w_2(\pi, z)| \leq e^{-\beta} \|w_1 - w_2\|,$$

where $\|\cdot\|$ stands for the supremum norm, and therefore F_q^β is a contraction in the class of bounded functions on $S \times D$. Moreover under (21) and (22) F_q^β transforms bounded functions into bounded functions continuous separately on $S_\delta \times D$ and on $(S \setminus S_\delta) \times D$. For each bounded q there is a fixed point of F_q^β. Consequently one can define a sequence of solutions $w_0(\pi,z)=F^{\beta}_{q_0}(\pi,z)$, with $q_0\equiv 0$, $w_1(\pi,z)=F^{\beta}_{w_0}(\pi,z)$ and inductively $w_n(\pi,z)=F^{\beta}_{w_{n-1}}(\pi,z)$. Notice furthermore that $w_n(\pi,z)$ corresponds to the optimal value of the at most $n+1$ portfolio changes and stopped after the last transaction. Since the optimal portfolio strategies do not allow more than one transaction at the moment and $\beta>0$ we finally have that w_n converges uniformly to w_β which is a solution to (23) and (24). □

Remark 1 The assumption (22) says that joint density of the pair $(\pi(1),z(1))$, provided it exists, has a version which is continuous pointwise with respect to the $(\pi(0),z(0))$ with $\pi(0)\in S_\delta$, which is equivalent by Scheffe theorem [18] to the L^1 continuity of the transition densities. From Lemma 4 of [20] we have the following form of the density $d_{\pi,z}$ of the first $m-1$ coordinates of $g(\pi\diamond\zeta(z,\xi(1)))$:

$$\begin{aligned} d_{\pi,z}&(x_1,x_2,\ldots,x_{m-1})\\ &=\int_0^\infty x_m^{m-1}\frac{1}{\pi_1\cdots\pi_m}\\ &\quad\times q_z\left(\frac{1}{\pi_1}x_1x_m,\frac{1}{\pi_2}x_2x_m,\ldots,\frac{1}{\pi_{m-1}}x_{m-1}x_m,\frac{1}{\pi_m}\left(1-\sum_{i=1}^{m-1}x_i\right)\right)dx_m \end{aligned}\tag{27}$$

where q_z is the density of $\zeta(z,\xi(1))$. If q_z is continuous function of its coordinates and the density of $z(1)$ is also continuous with respect to the initial state $z(0)$ then we have the continuity of the density of the pair $(\pi(1),z(1))$ from which (22) follows.

Remark 2 Notice that the value function w^β maybe discontinuous at the boundary of S_δ. It happens when at certain $\pi\in S_\delta$ such that a coordinate $\pi_k=\delta$ and $z\in D$ we have

$$Mw^\beta(\pi,z)<E_{\pi,z}\left\{\ln\left(\pi^T\zeta(z(1),\xi(1))\right)+e^{-\beta}w^\beta(\pi(1),z(1))\right\}$$

i.e. when it is optimal at point (π,z) do not make transaction even though we are at the boundary of S_δ.

Proposition 2 *Under* (17) *the family of functions*

$$\left\{w^\beta(\pi,z)-\inf_{\pi'\in S,z'\in D}w^\beta(\pi',z'),\beta>0\right\}$$

is bounded.

Proof Notice first that $\sup_{\pi\in S}w^\beta(\pi,z)>\sup_{\pi\in S}Mw^\beta(\pi,z)$ since e is bounded from above by $1-\kappa$ and as soon as $\pi\notin S_\delta$ we have to change portfolio to have portfolio in S_δ. On the other hand $Mw^\beta(\pi,z)\geq\ln\alpha+\sup_{\pi''\in S_\delta}w^\beta(\pi'',z)$, where α is a lower bound for e. Therefore by (23) we have for $z,z'\in D$

$$\begin{aligned}
&\sup_{\pi\in S} w^\beta(\pi,z) - \inf_{\pi'\in S} w^\beta(\pi',z') \\
&\quad\le \sup_{\pi\in S} E_{\pi,z}\big\{\ln\big(\pi^T\zeta(z(1),\xi(1))\big) + e^{-\beta} w^\beta(\pi(1),z(1))\big\} \\
&\qquad - \inf_{\pi'\in S} \max\Big\{\ln\alpha + \sup_{\pi''\in S_\delta} w^\beta(\pi'',z'), E_{\pi',z'}\big\{\ln\big(\pi'^T\zeta(z(1),\xi(1))\big) \\
&\qquad + e^{-\beta} w^\beta(\pi(1),z(1))\big\}\Big\} \\
&\quad\le \|r\|_{sp} + |\ln\alpha| + \sup_{\pi\in S} |E_{\pi,z}\{w^\beta(\pi(1),z(1))\} - E_{\pi,z'}\{w^\beta(\pi(1),z(1))\}| \\
&\quad\le \|r\|_{sp} + |\ln\alpha| + d\|w^\beta\|_{sp} \qquad (28)
\end{aligned}$$

where in the last line we have used (17) and $\|f\|_{sp} = \sup_y f(y) - \inf_{y'} f(y')$. Consequently

$$\|w^\beta\|_{sp} \le \frac{\|r\|_{sp} + |\ln\alpha|}{1-d} \qquad (29)$$

which completes the proof. □

Let

$$v^\beta(\pi,z) = w^\beta(\pi,z) - \inf_{\pi'\in S, z'\in D} w^\beta(\pi',z'). \qquad (30)$$

Then by (23) we have

$$\begin{aligned}
v^\beta(\pi,z) = \max\big\{&Mv^\beta(\pi,z), E_{\pi,z}\big\{\ln\big(\pi^T\zeta(z(1),\xi(1))\big) \\
&- (1-e^{-\beta}) \inf_{\pi'\in S, z'\in D} w^\beta(\pi',z') + e^{-\beta} v^\beta(\pi(1),z(1))\big\}\big\}. \qquad (31)
\end{aligned}$$

We can now formulate the main result of this section

Theorem 1 *Assume that the transition operator of the pair $(\pi(t),z(t))$ is continuous in variation norm on $S_\delta\times D$. Then under (17) and (21) there is a bounded function v continuous separately on $S_\delta\times D$ and $(S\setminus S_\delta)\times D$ and a constant λ such that*

$$v(\pi,z) = \max\big\{Mv(\pi,z), E_{\pi,z}\big\{\ln\big(\pi^T\zeta(z(1),\xi(1))\big) - \lambda + v(\pi(1),z(1))\big\}\big\} \qquad (32)$$

for $\pi\in S_\delta$ and

$$v(\pi,z) = Mv(\pi,z) \qquad (33)$$

whenever $\pi\notin S_\delta$. Moreover $\lambda = \inf_V J(V)$ and the strategy $\hat V = (\hat\tau_n, \hat\pi_n)$, where

$$\begin{aligned}
\hat\tau_1 &= \inf\{t\ge 0 : v(\pi(t),z(t)) = Mv(\pi(t),z(t))\} \\
\hat\tau_{n+1} &= \hat\tau_n + \hat\tau_1\circ\theta_{\hat\tau_n}
\end{aligned} \qquad (34)$$

for $n = 0, 1, \dots$ *with* θ *standing for the Markov shift operator of the pair* $(\pi(t), z(t))$ *and*

$$\hat{\pi}_n = \hat{\pi}(\pi^-(n), z(n)) \tag{35}$$

where $\hat{\pi}$ *stands for the Borel selector in the operator* M*, is optimal i.e.* $J(\hat{V}) = \lambda$.

Proof We claim that the family $\{v^\beta(\pi, z), \beta > 0\}$ is bounded and equicontinuous separately on $S_\delta \times D$ and $(S \setminus S_\delta) \times D$. The boundedness follows directly from Proposition 2. From (23) we obtain that

$$\begin{aligned}
&v^\beta(\pi, z) - v^\beta(\pi', z') \\
&\quad \le \max\Bigg\{ \sup_{\pi'' \in S_\delta} \Bigg[\Bigg| \ln \frac{e(\pi, \pi'')}{e(\pi', \pi'')} \Bigg| \\
&\qquad + \big| E_{\pi'', z}\big\{\ln\big(\pi''^T \zeta(z(1), \xi(1))\big)\big\} - E_{\pi'', z'}\big\{\ln\big(\pi''^T \zeta(z(1), \xi(1))\big)\big\}\big| \\
&\qquad + \big| E_{\pi'', z}\big\{v^\beta(\pi(1), z(1))\big\} - E_{\pi'', z'}\big\{v^\beta(\pi(1), z(1))\big\}\big| \Bigg], \\
&\qquad \big| E_{\pi, z}\big\{\ln\big(\pi^T \zeta(z(1), \xi(1))\big)\big\} - E_{\pi', z'}\big\{\ln\big(\pi'^T \zeta(z(1), \xi(1))\big)\big\}\big| \\
&\qquad + \big| E_{\pi, z}\big\{v^\beta(\pi(1), z(1))\big\} - E_{\pi', z'}\big\{v^\beta(\pi(1), z(1))\big\}\big| \Bigg\}.
\end{aligned} \tag{36}$$

Equicontinuity of the family $\{v^\beta(\pi, z), \beta > 0\}$ on $S_\delta \times D$ now follows from the continuity of e (Lemma 1), (21) and Proposition 2 together with (22). Using (24) and Lemma 1 we obtain the equicontinuity of the family $\{v^\beta(\pi, z), \beta > 0\}$ on $(S \setminus S_\delta) \times D$. By (21) $(1 - e^{-\beta}) \inf_{\pi' \in S, z' \in D} w^\beta(\pi', z')$ is bounded uniformly in $\beta > 0$. Therefore one can find a subsequence $\beta_n \to 0$ and a constant λ such that $\lim_{n\to\infty}(1 - e^{-\beta_n}) \inf_{\pi' \in S, z' \in D} w^{\beta_n}(\pi', z') \to \lambda$. Since the family $\{v^\beta(\pi, z)\}$ is equicontinuous on $S_\delta \times D$ and on $(S \setminus S_\delta) \times D$ and bounded then by Ascoli Arzela theorem [16] there is a further subsequence n_k of n, and a continuous bounded function v such that $v^{\beta_{n_k}}(\pi, z) \to v(\pi, z)$ uniformly on compact subsets from $S_\delta \times D$ and $(S \setminus S_\delta) \times D$ as $k \to \infty$. Letting $\beta_{n_k} \to 0$ in (31) we obtain (32). Notice that (33) follows directly from (24). The form of optimal strategies (34) and (35) can be obtained in a standard way (see the proof of Theorem 1 in [19] and references therein). □

3 Continuous Time Case

We consider now a continuous time case in the setting from Sect. 1. We control portfolio process $\pi(t)$ so as to maximize the cost functional J defined in (13) with the following restrictions: after each transaction (portfolio change) we are not allowed to make new transactions for fixed $h > 0$ units of time; when

$$\pi(t) \in \tilde{S}_\delta := \{\pi \in S, \exists_k \ \pi_k \le \delta\}$$

and we are allowed to make transaction we immediately change portfolio to enter the set $S_{\delta'}$ with $\delta' > \delta$, when

$$\pi(t) \in S_\delta^0 := \{\pi \in S, \forall_k\ \pi_k > \delta\}$$

we are allowed to change portfolio but we have no such obligation—we make transaction when it is profitable for us taking into account our fixed plus proportional transaction costs c defined in (2). Our portfolio strategy is impulse of the form $V = (\tau_n, \pi_n')$ consisting of transactions at moments τ_n, where $\tau_{n+1} \geq h + \tau_n$, with portfolio $\pi(\tau_n) = \pi_n'$ after transaction at time τ_n.

Consider first the following discounted costs functional

$$J_{\pi,z}^\beta(V) = E_{\pi,z}\Bigg\{\sum_{n=1}^{\infty} e^{-\beta\tau_n}\Big(\ln e(\pi(\tau_n^-), \pi(\tau_n)) + \ln\Big(\pi(\tau_{n-1})^T \zeta(\tau_{n-1}, \tau_n - \tau_{n-1}, z(\cdot), \xi(\cdot))\Big)\Big)\Bigg\}. \tag{37}$$

Let $w^\beta(\pi, z) = \sup_V J_{\pi,z}^\beta(V)$. Assume that for $\pi_{(n)} \to \pi \notin FrS$, where $FrS := \{\pi \in S : \exists_k, \pi_k = 0\}$ and $z_n \to z$

$$\sup_A \Big| P_{\pi_{(n)},z_n}\{(\pi(h), z(h)) \in A\} - P_{\pi,z}\{(\pi(h), z(h)) \in A\}\Big| \to 0 \tag{38}$$

as $n \to \infty$ with supremum over all Borel subsets A of $S \times D$ and

$$D_{FrS} = \inf\{t \geq 0 : \pi(t) \in FrS\} = \infty \tag{39}$$

P a.e. for any portions of wealth process invested in assets $\pi(t)$ starting from $\pi(0) \in S \setminus FrS$, provided that there are no transactions. Assume furthermore that the mapping

$$(\pi, z) \mapsto E_{\pi,z}\{\ln(\pi^T(h)\zeta(0, h, z(\cdot), \xi(\cdot)))\} \tag{40}$$

is bounded continuous and for any bounded function, continuous on $(S \setminus FrS) \times D$ the function w

$$F^\beta w(\pi, z) = \sup_\tau E_{\pi,z}\Big\{e^{-\beta\tau\wedge\sigma}\Big(\ln\Big(\pi^T(\tau\wedge\sigma)\zeta(0, \tau\wedge\sigma, z(\cdot), \xi(\cdot))\Big) + M_h w(\pi(\tau\wedge\sigma), z(\tau\wedge\sigma))\Big)\Big\} \tag{41}$$

where $\sigma = \inf\{t \geq 0 : \pi(t) \in \tilde{S}_\delta\}$ and

$$M_h w^\beta(\pi, z) = \sup_{\pi' \in S_{\delta'}} \Big\{\ln e(\pi, \pi') + e^{-\beta h} E_{\pi',z}\Big\{\ln\Big(\pi^T(h)\zeta(0, h, z(\cdot), \xi(\cdot))\Big) + w^\beta(\pi(h), z(h))\Big\}\Big\} \tag{42}$$

is bounded and continuous on $(S \setminus FrS) \times D$. We have

Proposition 3 *Under* (38)–(41) *the function* w^β *is a solution to the following Bellman equation*

$$w^\beta(\pi, z) = \sup_\tau E_{\pi,z}\big\{e^{-\beta\tau\wedge\sigma}\big(\ln\big(\pi^T(\tau\wedge\sigma)\zeta(0,\tau\wedge\sigma,z(\cdot),\xi(\cdot))\big) + M_h w^\beta(\pi(\tau\wedge\sigma), z(\tau\wedge\sigma))\big)\big\} \tag{43}$$

and w^β *is continuous on* $(S\setminus FrS)\times D$.

Proof Notice first that by (38) and (40) the mapping M_h transforms bounded functions into continuous on $(S\setminus FrS)\times D$. Moreover since $h>0$, F^β defined in (41) is a contraction in the class of bounded functions and the functions $w_0(\pi,z) = F^\beta 0(\pi,z)$, $w_{n+1}(\pi,z) = F^\beta w_n(\pi,z)$ are value functions for the cost functional (37) stopped after first, n plus one transactions. By contractivity of F^β functions w_n converge uniformly to w^β. Consequently there is a solution w^β to (43). If w^β is a solution to (43) that is continuous on $(S\setminus FrS)\times D$ then by iteration (taking into account (39)) we obtain that it coincides with the value function corresponding to the cost functional (37). □

Remark 3 In the case of continuous time model we are thinking about ζ of the form of exponent with Brownian or more general Levy noise $\xi(\cdot)$. Key assumption is (41) which is satisfied for particular models of lognormal asset prices, generally speaking nonsingular diffusion models or regular Levy models of asset prices (see [3] for more details). Notice that we obligatorily diversify portfolio when the process $\pi(t)$ enters the closed set $\tilde{S}_\delta$. After transaction our portfolio is in $S_{\delta'}\subset S_\delta$. The assumption (39) is not strong. Under quite general assumptions (law of large numbers holds) there is q_k such that $\lim_{t\to\infty}\frac{1}{t}\ln\zeta_k(0,t,z(\cdot),\xi(\cdot)) = q_k$ for $k=1,\dots,m$. Then whenever $\pi(0)\in S\setminus FrS$ we have that $\pi(t)\to\pi(\infty)$ where $\pi(\infty)$ is a Dirac measure concentrated on the vertex (or vertices) of S corresponding to maximal q_k. Since the coordinates of $\pi(0)$ and ζ are strictly positive $\pi(t)$ reaches FrS in the limit only.

In the remaining part of this section we follow the arguments used in discrete time case. We assume additionally that

$$\sup_{z,z'\in D}\sup_{A\in\mathscr{B}(D)}(P_h(z,A)-P_h(z',A)) := d < 1 \tag{44}$$

and

$$\sup_\tau\sup_{\pi\in S, z\in D} E_{\pi,z}\big\{|\ln(\pi^T(\tau\wedge\sigma)\zeta(0,\tau\wedge\sigma,z(\cdot),\xi(\cdot)))|\big\} := K < \infty. \tag{45}$$

We have

Proposition 4 *Under* (44) *and* (45) *the family of functions*

$$\Big\{w^\beta(\pi,z) - \inf_{\pi'\in S, z'\in D} w^\beta(\pi',z'), \beta>0\Big\}$$

is bounded.

Proof By (42) and (45) we have that

$$\sup_{\pi,z} w^\beta(\pi,z) - \inf_{\pi',z'} w^\beta(\pi',z') \le 2K + \sup_{\pi,z} M_h w^\beta(\pi,z) - \inf_{\pi',z'} M_h w^\beta(\pi',z'). \quad (46)$$

On the other hand since the difference of supremums is majorized by supremum of the difference, using also (44) we have

$$\begin{aligned}
&\sup_{\pi,z} M_h w^\beta(\pi,z) - \inf_{\pi',z'} M_h w^\beta(\pi',z') \\
&\quad \le 2L + \sup_{\pi\in S_{\delta'}} \Big(\sup_{z\in D} E_{\pi,z}\{w^\beta(\pi(h),z(h))\} - \inf_{z'\in D} E_{\pi,z'}\{w^\beta(\pi(h),z(h))\}\Big) \\
&\quad \le 2L + d\|w^\beta\|_{sp} \qquad (47)
\end{aligned}$$

with L standing for the bound of $|E_{\pi,z}\{\ln(\pi^T(h)\zeta(0,h,z(\cdot),\xi(\cdot)))\}|$. Therefore by (46) and (47) we obtain

$$\|w^\beta\|_{sp} \le \frac{2(K+L)}{1-d} \quad (48)$$

which completes the proof. □

Following (30) let now $v^\beta(\pi,z) = w^\beta(\pi,z) - \inf_{\pi'\in S, z'\in D} w^\beta(\pi',z')$. By (42) we then have

$$\begin{aligned}
v^\beta(\pi,z) = \sup_\tau E_{\pi,z}\Big\{ &-(1-e^{-\beta(\tau\wedge\sigma)}) \inf_{\pi'\in S, z'\in D} w^\beta(\pi',z') \\
&+ e^{-\beta\tau\wedge\sigma}\big(\ln\big(\pi^T(\tau\wedge\sigma)\zeta(0,\tau\wedge\sigma,z(\cdot),\xi(\cdot))\big) \\
&+ M_h v^\beta(\pi(\tau\wedge\sigma), z(\tau\wedge\sigma))\big)\Big\}. \qquad (49)
\end{aligned}$$

We shall need the following two technical assumptions: the family

$$\Big\{F^\beta w(\pi,z);\ \beta>0,\ \sup_{\pi\in S, z\in D} |w(\pi,z)| \le K\Big\} \quad (50)$$

is equicontinuous for $(\pi,z)\in S\setminus FrS$ given that

$$\Big\{M_h w(\pi,z),\ \sup_{\pi\in S, z\in D} |w(\pi,z)| \le K\Big\}$$

is equicontinuous for $(\pi,z)\in S\setminus FrS$, and furthermore

$$\lim_{\beta\to 0} E_{\pi,z}\left\{\frac{1}{\beta}(1-e^{-\beta\sigma})\right\} = E_{\pi,z}\{\sigma\} < \infty. \quad (51)$$

The main result of this section can be formulated as follows

Theorem 2 *Under* (38)–(41) *and* (44), (45), (50) *and* (51) *there is a bounded function v continuous on $(S \setminus FrS) \times D$ and on $FrS \times D$ and a constant λ such*

$$v(\pi, z) = \sup_{\tau} E_{\pi,z}\big\{-\lambda(\tau \wedge \sigma) + \big(\ln\big(\pi^T(\tau \wedge \sigma)\zeta(0, \tau \wedge \sigma, z(\cdot), \xi(\cdot))\big) + M_h v(\pi(\tau \wedge \sigma), z(\tau \wedge \sigma))\big)\big\}. \quad (52)$$

Moreover $\lambda = \inf_V J(V)$ and the strategy $\hat{V} = (\hat{\tau}_n, \hat{\pi}_n)$, where

$$\hat{\tau}_1 = \inf\{t \geq 0 : v(\pi(t), z(t)) = M_h v(\pi(t), z(t))\},$$
$$\hat{\tau}_{n+1} = \hat{\tau}_n + \hat{\tau}_1 \circ \theta_{\hat{\tau}_n} \quad (53)$$

for $n = 0, 1, \dots$ with θ standing for the Markov shift operator of the pair $(\pi(t), z(t))$ and

$$\hat{\pi}_n = \hat{\pi}(\pi^-(n), z(n)) \quad (54)$$

where $\hat{\pi}$ stands for the Borel selector in the operator M_h, is optimal i.e. $J(\hat{V}) = \lambda$.

Proof By (38) and (48) we see that the family $\{M_h v^\beta(\pi, z); \beta > 0\}$ is equicontinuous for $(\pi, z) \in (S \setminus FrS) \times D$. Therefore under (50) the family $\{v^\beta; \beta > 0\}$ is also equicontinuous for $(\pi, z) \in (S \setminus FrS) \times D$. Iterating (42) and using (45) we see that $\beta \inf_{\pi' \in S, z' \in D} w^\beta(\pi', z')$ is bounded. Therefore there is a subsequence $\beta_n \to 0$ and a constant λ such that $\beta_n \inf_{\pi' \in S, z' \in D} w^{\beta_n}(\pi', z') \to \lambda$ as $n \to \infty$. Choosing a further subsequence we can find a function v such that subsequence of $v^{\beta_n}(\pi, z)$ converges uniformly on compact subsets of $(S \setminus FrS) \times D$ to v and by equicontinuity of v^β on $FrS \times D$ (follows directly from the definition of M_h and (44)) the convergence is also uniform on compacts from $FrS \times D$. Consequently using (51) we can let $\beta_n \to 0$ in (49) to obtain finally (52). The remaining part of the proof follows as in discrete time case. □

Remark 4 There are a number of technical assumptions since we wanted to treat a very general scheme. In the case of nondegenerate diffusion asset prices model or when large deviation estimates for $\log \zeta$ hold (see [15]) we have finite moments of σ for any initial state (π, z). Consequently (51) holds and in the case of continuous trajectories of the asset prices (together with continuous dependence of the trajectories on initial conditions) also (41) and (50) are satisfied.

References

1. Akian, M., Sulem, A., Taksar, M.I.: Dynamic optimization of a long-term growth rate for a portfolio with transaction costs and logaritmic utility. Math. Financ. **5**, 153–188 (2001)
2. Algoet, P.H., Cover, T.M.: Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. Ann. Probab. **16**, 876–898 (1988)
3. Applebaum, D.: Lévy Processes and Stochastic Calculus. Cambridge University Press, Cambridge (2004)

4. Blumenthal, R.M., Getoor, R.K.: Markov Processes and Potential Theory. Academic Press, San Diego (1968)
5. Doob, J.L.: Stochastic Processes. Chapman & Hall, London (1953)
6. Györfi, L., Vajda, I.: Growth optimal portfolio selection strategies with proportional transaction costs. Preprint (2007)
7. Irle, A., Sass, J.: Good portfolio strategies under transaction costs: a renewal theoretic approach. In: Grossinho, M.R., Shiryaev, A.N., Esquivel, M., Oliveira, P.E. (eds.) Stochastic Finance, pp. 321–341. Springer, Berlin (2006)
8. Iyengar, G.: Universal investment in markets with transaction costs. Math. Financ. **15**, 359–371
9. Kabanov, Y., Klüppelberg, C.: A geometric approach to portfolio optimization in models with transaction costs. Finance Stoch. **8**, 207–227 (2004)
10. Kelly, J.L.: A new interpretation of information rate. Bell Syst. Tech. J. **35**, 917–926 (1956)
11. Korn, R.: Portfolio optimisation with strictly positive transaction costs and impulse control. Finance Stoch. **2**, 85–114 (1998)
12. Magill, M.J., Constantanides, M.: Portfolio selection with transaction costs. J. Econ. Theory **13**, 245–263 (1976)
13. Morton, A.J., Pliska, S.R.: Optimal portfolio management with fixed transaction costs. Math. Financ. **5**, 337–356 (1995)
14. Palczewski, J., Stettner, L.: Maximization of the portfolio growth rate under fixed and proportional transaction costs. Commun. Inf. Syst. **7**, 31–58 (2007)
15. Palczewski, J., Stettner, L.: Growth-optimal portfolios under transaction costs. Appl. Math. **35**, 1–31 (2008)
16. Royden, H.L.: Real Analysis. Macmillan, New York (1968)
17. Schäfer, D.: Nonparametric Estimation for Financial Investment under Log-Utility. PhD Dissertation, Mathematical Institute University of Stuttgart, Aachen (2002)
18. Scheffe, H.: A useful convergence theorem for probability distributions. Ann. Math. Stat. **18**, 434–438 (1947)
19. Stettner, L.: Discrete time risk sensitive portfolio optimization with consumption and proportional transaction costs. Appl. Math. **32**(4), 395–404 (2005)
20. Stettner, L.: Discrete time infinite horizon risk sensitive portfolio selection with proportional transaction costs. In: Stettner, L. (ed.) Advances in Mathematics of Finance. Banach Center Publications, vol. 86, pp. 231–241. PWN, Warsaw (2008).

On the Approximation of Geometric Fractional Brownian Motion

Esko Valkeila

Abstract We give an approximation to geometric fractional Brownian motion. The approximation is a simple corollary to a 'teletraffic' functional central limit theorem by Gaigalas and Kaj in (Bernoulli 9:671–703, 2003). We analyze the central limit theorem of Gaigalas and Kaj from the point of view of semimartingale limit theorems to have a better understanding of the arbitrage in the limit model. With this approximation we associate the corresponding pricing model sequence, which has the no-arbitrage property and which is complete.

Keywords Arbitrage · Geometric fractional Brownian motion · Approximation

Mathematics Subject Classification (2000) 60F17 · 60H99 · 91B28

1 Introduction

1.1 Geometric Fractional Brownian Motion

In the classical Black-Scholes pricing model the stock price S is modeled by a geometric Brownian motion: $S_t = e^{W_t - \frac{1}{2}t}$; here W is the standard Brownian motion. This model implies that the one dimensional distributions of the stock prices are log-normal, and the log-returns of the stocks are independent normal random variables. But empirical studies show that log-returns often have so-called long-range dependency property (see [18, Chap. IV]). One way to model this observed long-range dependency is to replace the driving standard Brownian motion by fractional Brownian motion. Then one obtains fractional Black-Scholes model, or geometric fractional Brownian motion, given by the price process $S_t = e^{B_t^H}$; here B^H is a fractional Brownian motion.

Fractional Brownian motion (fBm) B^H is a continuous centered Gaussian process. Here $H \in (0,1)$ is the self-similarity index and the covariance of the process B^H is given by

$$E(B_s^H B_t^H) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}).$$

E. Valkeila (✉)
Institute of Mathematics, Helsinki University of Technology, P.O. Box 1100, 02015 Helsinki, Finland
e-mail: esko.valkeila@tkk.fi

F. Delbaen et al. (eds.), *Optimality and Risk—Modern Trends in Mathematical Finance*, DOI 10.1007/978-3-642-02608-9_14,

A Lévy type of characterization theorem for fBm was recently proved in [15]. The parameter H allows to include the standard Brownian motion W to the fBm family: the process $B^{\frac{1}{2}}$ is a standard Brownian motion. The standard Brownian motion is a martingale, but it is well-known that when the parameter $H \neq \frac{1}{2}$, then fBm process B^H is not a semi-martingale.

The parameter H allows us to model dependency in the data, since for $H \in (\frac{1}{2}, 1)$ the increments of the process B^H are positively correlated. However, the fact that the process B^H is not a semimartingale, makes it difficult to use fBm as the source of randomness in Stochastic Finance at least theoretically: the reason is the fact that in the pricing models based on geometric fBm one can also give explicit arbitrage strategies, see [18, p. 659].

But the possibility to model dependency in the data is an attractive feature of the fBm process B^H, also for the market models in Stochastic Finance. In spite of the theoretical difficulties related to the arbitrage, there have been several proposals to use it as a model in stochastic finance. One of them is based on the fact that the arbitrage possibilities depend how one defines continuous trading, i.e. stochastic integrals. One can show that the arbitrage possibilities in the fractional Black-Scholes model disappear, if one uses Skorohod integrals to model trading strategies (Hu and Øksendal [9], Elliot and van der Hoek [5]). But a new problem appears: the continuous trading based on Skorohod integrals is difficult to interpret economically (see Sottinen and Valkeila [20], Björk and Hult [2] for more information on this point). On the other hand, if one goes to more realistic market models, and for example includes transaction costs in the market models, then the ideal continuous time trading strategies turn out to be of bounded variation. In this case one can show that geometric fBm models can be economically meaningful (Guasoni [7], Guasoni et al. [8]). It is also well known, that in the case where one cannot use continuous time trading, the pricing models with geometric fBm are to some extent arbitrage free (see [1, 3, 11]).

1.2 Motivation

The purpose of this note is to study the approximation of geometric fBm $S_t^H = e^{B_t^H}$. The approximation is understood in the sense of weak convergence, more precisely the distributions of the approximating prelimit sequence converge weakly to the distribution of the geometric fBm in the Skorohod space D.

This note has two different motivations. First comes from the fact that there are at least two 'financially' motivated approximations to geometric fBm. The approximation given by Sottinen [19] is based on complicated 'fractional' binomial tree, and as a binary tree this approximation is complete. Surprisingly, this approximation is not arbitrage-free, if the step size in the fractional binomial tree is big enough. Hence there are arbitrage opportunities already in the prelimit model. Klüppelberg and Kühn [13] proposed an alternative approximation, based on Poisson shot noise processes, to geometric fBm. Their approximation is arbitrage free, but not complete. So one can ask, if there is an approximation to geometric fBm, where the prelimit sequence is arbitrage-free and complete? In this note we show how to construct such an approximation. As mentioned, the limit has arbitrage opportunities, and our approximation might give some new insight on the arbitrage in the limit.

The second motivation comes from our recent work with Bender and Sottinen [1]. In this work we consider a class of models, where the randomness of the risky asset comes from mixed Brownian—fractional Brownian motion. Take this process to be $\varepsilon W + B^H$, where W is a standard Brownian motion, B^H is a fBm with index $H \in (\frac{1}{2}, 1)$, and independent of W. If we take the model of the risky asset S^ε to be

$$S_t^\varepsilon = \exp\left\{\varepsilon W_t + B_t^H - \frac{1}{2}\varepsilon^2 t\right\},$$

then there is a unique *hedging price* for the standard European type of options, provided that one uses so-called *allowed* (in the terminology [1]) strategies only. But in this model one can let $\varepsilon \to 0$, and ask weather the limiting prices make sense? It turns out that we have the following limiting price with an European call with strike K: $(S_0 - K)^+$. We get the same limiting price from our approximation, and we give two different explanations for this. One is based on the path-wise approach given in Dzhaparidze [4], and the other one is based on computing the limit price using the martingale measure of the approximating sequence.

1.3 The Structure of the Note

First we introduce the 'teletraffic' approximation from [6], discuss its properties from the point of view semimartingale weak limit theorems (see [10]), and prove that the corresponding geometric processes also converge weakly. We then argue that the prelimit sequence defines a sequence of pricing models, which are complete and have the no-arbitrage property. We conclude with a discussion.

2 Approximation of fBm

2.1 Construction of the Approximation

We will not prove any new approximation to fBm. Instead, we will the use the 'teletraffic' approximation to fBm, interpret this weak limit theorem as a semimartingale limit theorem of a special kind: the approximating sequence is based on semimartingales, but the limit is not a semimartingale.

We start with an approximation given by Gaigalas and Kaj [6]. This goes as follows: let G be a continuous distribution function of interarrival times η_i, $i \geq 2$

for a renewal counting process N. Let $\mu = E\eta_2$. Assume that this distribution has heavy tails:

$$1 - G(t) \sim t^{-(1+\beta)} \tag{1}$$

as $t \to \infty$ with $\beta \in (0,1)$. For the first interarrival time η_1 we assume that it has the distribution

$$G_0(t) = \frac{1}{\mu}\int_0^t (1 - G_s)ds, \tag{2}$$

so that the renewal counting process

$$N_t = \sum_{k=1}^{\infty} 1_{\{\tau_k \le t\}}$$

is stationary, where $\tau_1 = \eta_1$ and $\tau_k := \eta_1 + \cdots + \eta_k$. Take now countably infinite number of independent copies $N^{(i)}$ of N, numbers $a_m > 0$, $a_m \to \infty$ such that

$$\frac{m}{a_m^{\beta}} \to \infty \quad \text{as } m \to \infty; \tag{3}$$

in the terminology of Gaigalas and Kaj [6] this is the case of *fast connection rate*. Define the workload process W_t^m by

$$W_t^m = \sum_{k=1}^{m} N_t^{(k)}.$$

Note that the process W_t^m is again a counting process, since the interarrival distribution is continuous and the components $N^{(k)}$ are independent, and these facts imply that there are no simultaneous jumps of the components $N^{(k)}$. We have that $EW_t^m = \frac{mt}{\mu}$, since W_t^m is a stationary process. For the following proposition see Gaigalas and Kaj [6]:

Proposition 1 *Assume* (1) *and* (3). *Let*

$$Y^m(t) := \mu^{\frac{3}{2}}\sqrt{\frac{\beta(1-\beta)(2-\beta)}{2}}\,\frac{W^m_{a_m t} - m\mu^{-1}a_m t}{m^{\frac{1}{2}}a_m^{1-\frac{\beta}{2}}}. \tag{4}$$

Then Y^m converges weakly in the Skorohod space D to a fBm B^H, where $H = 1 - \frac{\beta}{2}$.

2.2 Further Properties of the Approximation

In order to discuss the application to finance, and to construct an approximation to geometric fBm $S_t^H = e^{B_t^H}$, we will have a look at Proposition 1 from the viewpoint of semimartingale limit theorems. We can write the explicit semimartingale decomposition, but only with respect to a big filtration, where we can keep track of the jumps of individual components $N^{(k)}$.

First we recall how one obtains the compensator of a renewal counting process by keeping track on the jump times and using the interarrival distributions; this is due to Jacod (see [14, Theorem 18.2, p. 270]). Assume that N is a renewal counting process with interarrival times η_j such that $G_j(t) = P(\eta_j \le t)$ has the form

$$G_j(t) = \int_0^t g_j(s)ds, \quad \text{where } g_j(s) > 0 \; \forall s \ge 0. \tag{5}$$

Let τ_j be the jump times and define $b_j(t) = \tau_j \wedge t - \tau_{j-1} \wedge t$, $j \ge 1$. Let H_j be the integrated hazard function of G_j:

$$H_j(t) = \int_0^t \frac{g_j(s)}{1 - G_j(s)} ds.$$

Work with the history $\mathbb{F}^N$ of the process N. Then the $(P, \mathbb{F}^N)$-compensator A of N can be written as

$$A_t = \sum_{j=1}^{\infty} H_j(b_j(t))$$

(see [14, 16]). Note that we have the relation

$$t = \sum_{j=1}^{\infty} b_j(t).$$

Next, consider the workload process W_t^m. Assume that we can keep track of the jumps of the processes $N^{(k)}$, i.e. we work with the filtration $\bar{\mathbb{F}}^m$, where $\bar{F}_t^m = \sigma\{N_s^{(k)} : s \le t, k = 1, \dots, m\}$. Define $b_j^{(k)}(t) = \tau_j^{(k)} \wedge t - \tau_{j-1}^{(k)} \wedge t$, and then by the independence of the processes $N^{(k)}$ the $(\bar{\mathbb{F}}, P)$-compensator of $N^{(k)}$ is

$$A_t^{(k)} = \sum_{j=1}^{\infty} H_j(b_j^{(k)}(t)).$$

Hence we obtain that the $(\bar{\mathbb{F}}^m, P)$-compensator A^m of the workload process W^m is

$$A_t^m = \sum_{k=1}^{m} A_t^{(k)};$$

note also that

$$mt = \sum_{k=1}^{m} \sum_{j=1}^{\infty} b_j^{(k)}(t).$$

The process Y^m given by (4) is a semimartingale, since it has bounded variation on compacts. Let us now write the semimartingale decomposition of the process Y^m with respect to the big filtration $\mathbb{F}^m$, where $F_t^m = \bar{F}_{a_m t}^m$ and probability measure P, associated to the interarrival times given by (1) and (2). To simplify notation put

$$c(\mu,\beta) := \mu^{\frac{3}{2}}\sqrt{\frac{\beta(1-\beta)(2-\beta)}{2}},$$

$c_m := m^{\frac{1}{2}} a_m^{1-\frac{\beta}{2}}$ and $\Lambda_t^m := \frac{m\mu^{-1}a_m t}{c_m}$.

Since the process Y^m is a semimartingale, it has a semimartingale decomposition

$$Y^m = M^m + L^m; \tag{6}$$

here $L_t^m = c(\mu,\beta)\frac{A_{a_m t}^m}{c_m} - \Lambda_t^m$.

The martingale part of the semimartingale Y^m is

$$M_t^m := c(\mu,\beta)\frac{W_{a_m t}^m - A_{a_m t}^m}{c_m}.$$

Note that the compensator of W^m with respect to the big filtration $\bar{\mathbb{F}}$ is continuous, and hence the process L^m is a continuous process with bounded variation.

The square bracket of the martingale part M^m of the semimartingale Y^m is

$$[M^m, M^m]_t = (c(\mu,\beta))^2\frac{W_{a_m t}^m}{c_m^2}.$$

But our assumption (3) implies

$$E[M^m, M^m]_t = (c(\mu,\beta))^2 \mu t a_m^{\beta-1} \to 0,$$

as $m\to\infty$. With the Doob inequality we obtain that $\sup_{s\le t}|M_s^m| \stackrel{P}{\to} 0$.

Denote uniform on compacts convergence in probability by $\stackrel{ucp}{\to}$.

We obtain the following semimartingale interpretation of Proposition 1:

Proposition 2 *Assume* (1) *and* (3). *Let*

$$Y^m = M^m + L^m$$

be the $(\mathbb{F}^m, P)$*-semimartingale decomposition of the process* Y^m *given by* (3). *Then the sequence* L^m *of continuous bounded variation processes converges weakly in the Skorohod space* D *to the fBm* B^H *with* $H = 1-\frac{\beta}{2}$, *and* $M^m \stackrel{ucp}{\to} 0$, *as* $m\to\infty$.

Remark 1 We gave the semimartingale decomposition of the process Y^m with respect to the filtration $\mathbb{F}^m$. Since the process Y^m is adapted to its own filtration $\mathbb{F}^{Y^m} = \mathbb{F}^{W^m}$, it has a different semimartingale decomposition with respect to

$(\mathbb{F}^{Y^m}, P)$, and the new compensator is $\tilde{L}^m$. The process $\tilde{L}^m$ is the $(\mathbb{F}^{Y^m}, P)$ dual predictable projection of the process Y^m. To compute the process $\tilde{L}^m$ explicitly is apparently difficult, because we do not know, which one of the components had a jump.

2.3 Approximation to Geometric fBm

Consider the solution to the equation

$$dS_t^m = S_{t-}^m dY_s^m, \quad \text{with } S_0^m = S_0, \tag{7}$$

where $Y^m = c(\mu,\beta)\frac{W^m}{c_m} - \Lambda^m$, and

$$\Delta Y_t^m = c(\mu,\beta)\frac{\Delta W_t^m}{c_m} = \Delta M_t^m.$$

It is known that (7) has a unique solution of the form

$$S_t^m = S_0 e^{-\Lambda_t^m} \prod_{s \le t}(1 + \Delta Y^m(s)) =: \mathscr{E}(Y^m)_t. \tag{8}$$

Proposition 3 *Assume* (1) *and* (3) *and let Y^m be as in* (4). *Then the solution to* (7), *given by* (8), *converges weakly to the geometric fBm $S_t = S_0 e^{B_t^H}$ in the Skorohod space D.*

Proof Use the inequality $x - \frac{1}{2}x^2 \le \log(1+x) \le x$ valid for $x \ge 0$ to obtain the following:

$$S_0 e^{Y^m(t) - \frac{1}{2}\frac{W^m_{a_m t}}{c_m^2}} \le S_t^m \le S_0 e^{Y^m(t)}.$$

We already know that $\frac{W^m_{a_m t}}{c_m^2} \to 0$ in $L^1(P)$, as $m \to \infty$. Hence the claim follows by Slutskys theorem and the continuous mapping theorem. □

With the notation of Proposition 2 we have

Corollary 1 *The sequence of continuous bounded variation processes e^{L^n} converges weakly to the geometric fBm e^{B^H} in the Skorohod space D.*

3 Some Properties of the Approximation

3.1 Set-Up

Assume that we have (1) and (2), the process Y^m is defined by (4), and S^m is defined by (7). We interpret the prelimit approximation S^m as a stock price. To simplify the discussion we assume that the interest rate for the *bank account* is equal to 0, and that there is no drift on the stock price. So we have a sequence of pricing models

$$(S^m, \mathbb{F}^m, P^m) \overset{w(P^m)}{\longrightarrow} (S_0 e^{B^H}, \mathbb{F}^{B^H}, P), \tag{9}$$

where B^H is a fBm with $H = 1 - \frac{\beta}{2} \in (\frac{1}{2}, 1)$.

We will show that the prelimit market model with S^m and bank account is arbitrage-free model and complete.

3.2 Prelimit Market Models are Arbitrage-Free

The basic randomness of the approximating pricing model sequence S^m comes from the workload process W^m. We shall show that there exists a probability measure Q^m such that W^m is a Poisson process with intensity $\frac{m}{\mu}$. Then the approximation S^m will be a Q^m-martingale, and the prelimit market models are arbitrage free.

We work first with a single component of the workload process.

Assume that N is a renewal counting process with first interarrival time distribution given by (2) and all the rest interarrival times have distribution given by (1). We assume that with respect to the measure our counting process N is a renewal counting process with respect to the measure P. Fix $T > 0$. First we shall show that there exists a probability measure Q such that $Q_T \sim P_T$, where $Q_T = Q|F_T^N$, $P_T = P|F_T^N$, and with respect to the measure Q the counting process N is a Poisson process with intensity μ^{-1}. Put $G_i = G$ when $i \geq 1$ and define

$$\kappa(s, N) := \frac{g_{N_{s-}}(s)}{1 - G_{N_{s-}}(s)}.$$

Define the density between the measures Q and P by

$$\frac{dQ}{dP}\Big|F_t^N = e^{\int_0^t (\kappa(s,N) - \frac{1}{\mu})ds + \int_0^t \log \frac{1}{\mu\kappa(s,N)} dN_s}. \tag{10}$$

Obviously we have

$$\frac{dP}{dQ}\Big|F_t^N = e^{\int_0^t (\frac{1}{\mu} - \kappa(s,N))ds + \int_0^t \log(\mu\kappa(s,N)) dN_s}. \tag{11}$$

The Hellinger process between the measures P and Q is then given by

$$h(P, Q)_t = \frac{1}{2} \int_0^t \left(\sqrt{\kappa(s, N)} - \sqrt{\frac{1}{\mu}} \right)^2 ds.$$

Under our assumptions $h(P, Q)_t \leq A_t + \nu t < \infty$ $(P + Q)$-a.s.; we can now use [10, Theorem IV.2.1] and conclude that the measures P_T and Q_T are equivalent. We have shown the following:

Lemma 1 *Assume that X is a counting process. With respect to measure Q it is a Poisson process with intensity $\frac{1}{\mu}$ and with respect to a measure P is it a renewal counting process with first interarrival time distribution given by* (2) *and all the rest interarrival times have distribution given by* (1). *Moreover, the laws P_T and Q_T are equivalent.*

The next step is to show that the law of the process

$$W_t^m = \sum_{k=1}^{m} N_t^{(k)}$$

is equivalent to the law of Poisson process with intensity $\frac{m}{\mu}$. Note that the process W_t^m is not any more a renewal counting process. We show that the prelimit pricing models driven by the processes Y^m have the no-arbitrage property. For this it is sufficient to show that the original probability measure is equivalent to a probability measure Q such that the process W_t^m is a Poisson process with intensity $\frac{m}{\mu}$. The proof is not very difficult, and will follow from Lemmas 1 and 2.

Lemma 2 *Let X^k, $k = 1, \ldots, m$ be a sequence of counting processes. Assume that with respect to the measure Q they are independent Poisson processes with intensity $\frac{1}{\mu}$, and with respect to the measure P they are independent renewal counting processes, and their interarrival times satisfy* (5). *Then the sum process $W^m = \sum_{k=1}^m X^k$ is a counting process with respect to the measures P and Q, with respect to the measure Q it is a Poisson process with intensity $\frac{m}{\mu}$, and the Q-law of W^m, Q^m is equivalent to the P-law, P^m of W^m on $[0, T]$. Here the filtration is the big filtration $F_t^m := \bigvee_{k=1}^m F_t^{X^k}$.*

Proof Since the processes X^k are stochastically continuous and independent with respect to the measures Q and P, we have that $P(\Delta X_s^k = 1, \Delta X_s^l = 1) = 0$ for $k \neq l$ for all $s \geq 0$, and similarly with respect to the measure Q. Hence the aggregated process W^m is a counting process.

Obviously the sum of independent Poisson processes is again a Poisson process, not only in the big filtration $\mathbb{F}^m$, but in the filtration $\mathbb{F}^{W^m}$, too. If the $(P, \mathbb{F}^{X^k})$ compensator of X^k is A^k, and because the sum of martingales is a martingale again, we have that the $(P, \mathbb{F}^m)$ compensator of W^m is $\sum_{k=1}^m A^k$. We can now repeat the argument given to obtain Lemma 1 and conclude that the measures Q_T^m and P_T^m are equivalent in the filtration $\mathbb{F}^m$. □

Remark 2 If we consider the measures P^m and Q^m restricted to the filtration $\mathbb{F}^{W^m}$ they are also equivalent on $[0, T]$, since $\mathbb{F}^{W^m} \subset \mathbb{F}^m$. But it is difficult to write the $(P^m, \mathbb{F}^{W^m})$-compensator of W^m explicitly (see also Remark 1).

Let us now return to the model driven by (4). We have that the aggregated process W^m is a Poisson process with intensity $\frac{m}{\mu}$. We it has the law Q^m described in Lemma 1 and Lemma 2. With respect to the original measure P, which corresponds to the renewal counting process model with interarrival times given by (1) for interarrivals after the first jump and by (2) for the first interarrival. We have that the measures are equivalent on the interval $[0, a_m T]$. This means that the approximation process Y^m is a martingale with respect to $(Q^m, \mathbb{F}^m)$, or with respect to $(Q^m, \mathbb{F}^{Y^m})$, too.

What happens with the approximation? Recall that $S^m = \mathscr{E}(Y^m)$. But $Y^m = M^m + L^m$, where L^m is a continuous process, and hence $[M^m, L^m] = 0$. So using Yor's formula for stochastic exponents we can write the approximating sequence as

$$S_t^m = S_0 e^{L_t^m} \mathscr{E}(M^m)_t,$$

where $\mathscr{E}(M^m) \overset{ucp}{\to} 1$ with respect to the measure P^m. We know that the approximation $(S^m, \mathbb{F}^m, P^m)$ weakly converges to the geometric fBm. On the other hand, with respect to the martingale measure Q^m the sequence Y^m is a martingale sequence, $Y^m \overset{ucp}{\longrightarrow} 0$ with respect to Q^m, and $S^m \overset{ucp}{\longrightarrow} S_0$ with respect to Q^m. So the price $(S_0 - K)^+$ is a limit

$$(S_0 - K)^+ = \lim_m E_{Q^m}(S_T^m - K)^+$$

for the European call.

3.3 *Prelimit Market Models are Complete*

To show that the prelimit market models are complete, it is enough to show that the market models driven by a Poisson process martingale are complete. We consider the following market model, so-called *Poisson market* according to the terminology of Dzhaparidze [4]. We follow the arguments of Dzhaparidze and show that the prelimit market is complete. Note that the argument given below is pathwise.

Let N be a counting process, $\alpha > 0$ and $\gamma > 0$ are constants, and consider the pathwise solution S to the following linear equation

$$dS_t = S_{t-}(\alpha dN_t - \gamma dt) \quad \text{with } S_0 = s;$$

then the unique solution to this is

$$S_t = s e^{-\gamma t} \prod_{s \le t} (1 + \alpha \Delta N_s) = s e^{-\gamma t} (1 + \alpha)^{N_t}.$$

Denote the jump times N by τ_k, $k = 1, 2, \ldots$. Fix $T > 0$ and assume that there is no jump at time T. Let $M \ge 0$ be such that $\tau_M < T < \tau_{M+1}$. Define $s_k(t)$ by

$$s_k(t) = s(1 + \alpha)^k e^{-\gamma t} 1_{[\tau_k, \tau_{k+1})}(t) = s(1 + \alpha)^k e^{-\alpha \lambda t} 1_{[\tau_k, \tau_{k+1})}(t);$$

with $\lambda = \frac{\gamma}{\alpha}$. The functions s_k describe the states of the price process S_t. Obviously

$$S_t = \sum_{k=0}^{M} s_k(t) 1_{[\tau_k, \tau_{k+1})}(t).$$

We can write the left-hand limit process S_{t-} as follows

$$S_{t-} = s_0(t) 1_{[0,\tau_1]}(t) + \sum_{k=1}^{M} s_k(t) 1_{(\tau_k, \tau_{k+1}]}(t).$$

Let S_{t-} be in the state $s_k(t)$. Then, at time t the stock price either stays in this state or jumps to the state $s_{k+1}(t)$. Define the difference operator D in the state space of S as follows: if S_{t-} is in state $s_k(t)$ then DS_t is in the state

$$D_{k+1}(S_t) = s_{k+1}(t) - s_k(t). \tag{12}$$

We have then the following

Proposition 4 *The states of the stock price process satisfy the following differential equations*

$$\frac{ds_k(t)}{dt} = -\lambda D_{k+1}(S_t), \quad \text{when } t \in (\tau_k, \tau_{k+1}]. \tag{13}$$

Proof See [4, Proposition 4.4.1]. □

The Poisson probabilities $p_j(\lambda)$ are defined by $p_j(\lambda) = \frac{\lambda^j}{j!} e^{-\lambda}$ for $\lambda > 0$. One can include the value $\lambda = 0$ by defining $p_j(0) = \delta_{j0}$, where $\delta_{j0} = 1_0(j)$ is the Kronecker delta.

Consider the following system of differential-difference equations

$$\frac{dx_k(t)}{dt} = -\lambda(x_{k+1}(t) - x_k(t)), \quad k = 0, 1, \ldots \tag{14}$$

subject to boundary conditions

$$x_k(T) = w_k(T), \quad k = 0, 1, \ldots. \tag{15}$$

Proposition 5 *A solution to the system* (14) *with boundary conditions* (15) *is given by*

$$x_k(t) = \sum_{j=0}^{\infty} p_j(\lambda(T-t)) w_{k+j}(T), \tag{16}$$

if the numbers $w_k(T)$ allow differentiation under the summation sign.

Proof See [4, Proposition 4.2.3]. □

Next we indicate how to apply the above results a hedging problem in finance. Let W_T be a functional of the price process path S_t, $0 \le t \le T$. If the price process S has a state $s_k(T)$, then the value functional W_T has a state $w_k(T)$. Recall that a market model is *complete*, if we can find a self-financing strategy π such that

$$W_T = V_T^\pi = v + \int_0^T \pi_s dS_s.$$

It follows from [4, Proposition 4.4.5] that if we define a value process V^π by the formula

$$V_t^\pi = v + \int_0^t \pi_s dS_s, \tag{17}$$

where

$$v = \sum_{j=0}^{\infty} p_j(\lambda T) w_j(T) \tag{18}$$

and when $s \in (\tau_k, \tau_{k+1}]$ we define

$$\pi_s = \sum_{j=0}^{\infty} p_j(\lambda(T-s)) \frac{(1+\alpha)w_{j+k}(T) - w_{j+k+1}(T)}{\alpha}, \tag{19}$$

then we obtain self-financing strategy π, which replicates the claim $W(T)$.

Remark 3 The probabilistic interpretation of (17), (18) and (19) that the process N is a Poisson process with intensity λ. Note that the results were obtained in a path-wise way, without any probability.

We end this subsection by giving an option pricing formula (for European call only, but of course all is valid for a more much bigger class of options).

Recall some properties of Poisson probabilities. Put

$$F(j_0; \lambda) = \sum_{j > j_0} p_j(\lambda). \tag{20}$$

We have the following connection between $F(j_0; \lambda)$ in Gamma integrals

$$\Gamma(c, x) = \int_x^\infty e^{-t} t^{c-1} dt, \quad \text{where } c, x > 0:$$

$$F(j_0; \lambda) = 1 - \frac{\Gamma(j_0 + 1, \lambda)}{j_0!}. \tag{21}$$

Proposition 6 *Consider the pricing of an European option* $(S_T - K)^+$ *in the Poisson market model. Then the fair price* C^E *of this option is given by*

$$C^E = S_0 F(j_0; (1+\alpha)\lambda T) - K F(j_0; \lambda T), \tag{22}$$

where

$$j_0 = \left\lfloor \frac{\log(\frac{K}{S_0}) + \gamma T}{\log(1+\alpha)} \right\rfloor. \tag{23}$$

We can now apply the above to the approximating model S^m. Now from (4) we obtain

$$\alpha = \alpha^m = c(\mu, \beta) \frac{a_m^{\frac{\beta}{2}}}{m^{\frac{1}{2}}},$$

$$\gamma = \gamma^m = c(\mu, \beta)\mu \frac{a_m^{\frac{\beta}{2}}}{m^{\frac{1}{2}}} \quad \text{and}$$

$$\lambda = \lambda^m = \frac{\gamma^m}{\alpha^m} = \mu.$$

From (3) we obtain that $\alpha^m \to 0$, $\beta^m \to 0$, and if $K > S_0$, then $j_0 \to \infty$, and if $K < S_0$, then $j_0 \to -\infty$. Put this in (22) and we obtain that the limiting price is $(S_0 - K)^+$.

4 Discussion and Conclusion

Consider the market model of the following type. The stock price S is driven by a process $X = \varepsilon W + B^H$; here W is a standard Brownian motion, B^H is a fBm with Hurst index $H > \frac{1}{2}$, independent of W; the linear stochastic differential equation defining the stock price is

$$dS_t^\varepsilon = S_t^\varepsilon dX_t, \text{ with } S_0 \tag{24}$$

as the initial value. One can show that the solution to (24) is

$$S_t^\varepsilon = S_0 e^{\varepsilon W_t + B_t^H - \frac{1}{2}\varepsilon^2 t}.$$

It was shown in [17] that the hedging price for standard European type of options is the same as in the model, where we do not have the fBm component B^H at all. Recently we in Bender et al. [1] have extended this argument to a bigger class of options, and also discussed arbitrage possibilities in this kind of models. So the price of an European call $(S_T^\varepsilon - K)^+$ is given by the classical Black & Scholes pricing formula

$$S_0\Phi\left(\frac{\log\frac{S_0}{K}+\frac{1}{2}\varepsilon^2 T}{\varepsilon\sqrt{T}}\right)-K\Phi\left(\frac{\log\frac{S_0}{K}-\frac{1}{2}\varepsilon^2 T}{\varepsilon\sqrt{T}}\right). \tag{25}$$

Take now $\varepsilon_n \to 0$ and define S^{ε_n} by

$$S_t^{\varepsilon_n}=S_0 e^{\varepsilon_n W_t+B_t^H-\frac{1}{2}\varepsilon_n^2 t}.$$

We have that $S_t^{\varepsilon_n} \stackrel{w}{\to} S_0 e^{B_t^H}$, as $n\to\infty$, and we have again an approximation to geometric fBm. It is easy to check that the limit, as $\varepsilon_n \to 0$, of the price in (25) is given by

$$(S_0-K)^+. \tag{26}$$

Recall that we obtained the same limit for European call as a limit of hedging prices in Sect. 3.3 and as a limit risk neutral prices in Sect. 3.2.

Note that if a price process S is continuous and has bounded variation, then we have the following

$$(S_T-K)^+=(S_0-K)^++\int_0^T 1_{\{S_s\geq K\}}dS_s; \tag{27}$$

hence a candidate for the hedging price would be $(S_T-K)^+$. But this makes no sense, since this kind of pricing model has arbitrage opportunities, unless $S_t=S_0$ for all $t\leq T$.

We have shown that in our approximation:

- The prelimit sequence $S^m=\mathscr{E}(Y^m)$ is a P^m semimartingale and Q^m martingale.
- The weak limit along P^m is geometric fBm, which is not a semimartingale, and along Q^m is the constant S_0, which is a martingale.

We can formulate yet another property of our approximation:

- The sequence of measures P^m and Q are entirely separated: there exists events C_m such that $P^m(C_m)\to 1$ and $Q^m(C_m)\to 0$, as $m\to\infty$.

In [12] Kabanov and Kramkov discuss so-called asymptotic arbitrage, which is related to the notions of contiguity and entire separation.

This means the following: let π^m be a sequence of self-financing strategies and S^m a vector valued price process such that

$$(\pi^m\cdot S^m)_t := \sum_{k=1}^m \int_0^t \pi_s^{m,k}dS_s^{m,k}\geq -1. \tag{28}$$

They defined the following three types of asymptotic arbitrage, but we mention only one:

- If in addition to (28), we have $\limsup_m P^m((\pi^m\cdot S^m)\geq C)=1$ as $m\to\infty$ for any $C>0$, then π^m realizes *strong asymptotic arbitrage*.

We refer to Kabanov and Kramkov [12] for more information how asymptotic arbitrage is related to contiguity and entire separation. We mention only that entire separation implies some kind of asymptotic arbitrage.

We end our discussion by reformulating our approximation in the spirit of large financial markets. Define the price process of the ith asset $S^{(i)}$ by

$$dS_t^{(i)} = S_{t-}^{(i)} c(\mu,\beta) \frac{1}{c_m} d(N_t^{(i)} - A_t^{(i)}), \tag{29}$$

where $N^{(i)}$ is the renewal counting process, $A^{(i)}$ is the compensator of $N^{(i)}$ with respect to the filtration $\mathbb{F}^{N^{(i)}}$, and where $c(\mu,\beta) := \mu^{\frac{3}{2}}\sqrt{\frac{\beta(1-\beta)(2-\beta)}{2}}$, and $c_m := m^{\frac{1}{2}} a_m^{1-\frac{\beta}{2}}$, as before. Let $S_0^{(i)} = \frac{S_0}{m}$. Then the model in (4) can be considered as the sum $S^m = \sum_{i=1}^m S^{(i)}$, and as we know already, the martingale measures Q^m and the 'historical' measures P^m are entire separated and the market model $(\tilde{S}^m, \mathbb{F}^m, P^m)$ with $\tilde{S}^m = (S^{(1)}, \dots, S^{(m)})$ admits asymptotic arbitrage.

Acknowledgements The author thanks Ingemar Kaj for an interesting discussion at the Mittag-Leffler institute in November 2007, an anonymous referee for some corrections and the Academy of Finland for the support (grant 212875).

References

1. Bender, C., Sottinen, T., Valkeila, E.: Pricing by hedging beyond semimartingales. Finance Stoch. **12**, 441–468 (2008)
2. Björk, T., Hult, H.: A note on Wick products and the fractional Black-Scholes model. Finance Stoch. **9**, 197–209 (2005)
3. Cheridito, P.: Arbitrage in fractional Brownian motion models. Finance Stoch. **7**, 533–553 (2003)
4. Dzhaparidze, K.O.: Introduction to Option Pricing in a Securities Market. CWI SYLLABUS, Amsterdam (2000)
5. Elliott, R., van der Hoek, J.: A general fractional white noise theory and applications to finance. Math. Financ. **13**, 301–330 (2003)
6. Gaigalas, R., Kaj, I.: Convergence of scaled renewal processes and a packet arrival model. Bernoulli **9**, 671–703 (2003)
7. Guasoni, P.: No arbitrage under transaction costs, with fractional Brownian motion and beyond. Math. Financ. **16**, 569–582 (2006)
8. Guasoni, P., Rsonyi, M., Schachermayer, W.: Consistent price systems and face-lifting pricing under transaction costs. Ann. Appl. Probab. **18**, 491–520 (2008)
9. Hu, Y., Øksendal, B.: Fractional white noise calculus and applications to finance. Infin. Dimens. Anal. Quantum Probab. Relat. Top. **6**, 1–32 (2003)
10. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, 2nd. edn. Springer, New York (2003)
11. Jarrow, R.A., Protter, P., Sayit, H.: No arbitrage without semimartingales. Ann. Appl. Probab. **19**(2), 596–616 (2009)
12. Kabanov, Yu.M., Kramkov, D.O.: Asymptotic arbitrage in large financial markets. Finance Stoch. **2**, 143–172 (1998)
13. Klüppelberg, C., Kühn, C.: Fractional Brownian motion as a weak limit of Poisson shot noise processes—with applications to finance. Stoch. Process. Appl. **113**, 333–251 (2004)

14. Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes. II Applications. Springer, New York (2001)
15. Mishura, Y., Valkeila, E.: An extension of the Lévy characterization to fractional Brownian motion. HUT, Institute of Mathematics, Preprint A514, 24 pages (2006)
16. Nikunen, M., Valkeila, E.: A Prohorov bound for a Poisson process and an arbitrary counting process with some applications. Stoch. Stoch. Rep. **37**, 133–151 (1991)
17. Schoenmakers, J., Kloeden, P.: Robust option replication for a Black-Scholes model extended with nondeterministic trends. J. Appl. Math. Stoch. Anal. **12**, 113–120 (1999)
18. Shiryaev, A.N.: Essentials of Stochastic Finance. Facts, Models, Theory. World Scientific, New Jersey (1999)
19. Sottinen, T.: Fractional Brownian motion, random walks and binary market models. Finance Stoch. **5**, 343–355 (2001)
20. Sottinen, T., Valkeila, E.: On arbitrage and replication in the fractional Black-Scholes pricing model. Stat. Decis. **21**, 93–107 (2003)