Lecture Notes in Mathematics 2215

Séminaire de Probabilités

Catherine Donati-Martin
Antoine Lejay · Alain Rouault *Editors*

Séminaire de Probabilités XLIX

Springer

Lecture Notes in Mathematics 2215

More information about this series at http://www.springer.com/series/304

Catherine Donati-Martin • Antoine Lejay •
Alain Rouault
Editors

Séminaire de Probabilités XLIX

Springer

Editors
Catherine Donati-Martin
Laboratoire de Mathématiques de Versailles
Université Versailles Saint-Quentin
Versailles, France

Antoine Lejay
Institut Elie Cartan de Lorraine
Vandoeuvre-les-Nancy, France

Alain Rouault
Laboratoire de Mathématiques de Versailles
Université Versailles Saint-Quentin
Versailles, France

ISSN 0075-8434 ISSN 1617-9692 (electronic)
Lecture Notes in Mathematics
ISBN 978-3-319-92419-9 ISBN 978-3-319-92420-5 (eBook)
https://doi.org/10.1007/978-3-319-92420-5

Library of Congress Control Number: 2018950488

Mathematics Subject Classification (2010): 60G, 60J, 60K

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

Preface

In this 49th volume we continue to offer a good sample of the main streams of current research on probability and stochastic processes, in particular those active in France. All the contributions come from spontaneous submissions and their diversity illustrates the good health of this branch of mathematics.

Since the publication of the 48th volume, we have received two sad pieces of news:

Jacques Neveu, former professor at the Université Pierre et Marie Curie and École Polytechnique, passed away on May 15, 2016. His influence on the strong development of probability in France was huge. For details and testimonies, we refer to "Hommages à Jacques Neveu", a supplement to *Matapli 112*, 52 pp., 2017, see the website http://smai.emath.fr/IMG/pdf/Matapli_J_Neveu.pdf.

Ron Getoor, professor at the University of California, San Diego, passed away on October 28, 2017. He was a leader in the growth of probability theory. For details, see the website https://www.math.ucsd.edu/memorials/ronald-getoor/.

Both of them published excellent books and contributions to *The Séminaire*. We would like to remind the reader that the website of the *Séminaire* is http://sites.mathdoc.fr/SemProba/ and that all the articles of the Séminaire from Volume I (1967) to Volume XXXVI (2002) are freely accessible from the website http://www.numdam.org/actas/SPS.

We thank the Cellule Math Doc for hosting these articles within the NUMDAM project.

Versailles, France — Catherine Donati-Martin
Vandoeuvre-lès-Nancy, France — Antoine Lejay
Versailles, France — Alain Rouault

Contents

1 Ornstein-Uhlenbeck Pinball and the Poincaré Inequality in a Punctured Domain 1
Emmnuel Boissard, Patrick Cattiaux, Arnaud Guillin, and Laurent Miclo

2 A Probabilistic Look at Conservative Growth-Fragmentation Equations 57
Florian Bouguet

3 Iterated Proportional Fitting Procedure and Infinite Products of Stochastic Matrices 75
J. Brossard and C. Leuridan

4 Limiting Eigenvectors of Outliers for Spiked Information-Plus-Noise Type Matrices 119
Mireille Capitaine

5 Criteria for Exponential Convergence to Quasi-Stationary Distributions and Applications to Multi-Dimensional Diffusions 165
Nicolas Champagnat, Koléhè Abdoulaye Coulibaly-Pasquier, and Denis Villemonais

6 Bismut-Elworthy-Li Formulae for Bessel Processes 183
Henri Elad Altman

7 Large Deviations for Infectious Diseases Models 221
Peter Kratz and Etienne Pardoux

8 The Girsanov Theorem Without (So Much) Stochastic Analysis 329
Antoine Lejay

9 On Drifting Brownian Motion Made Periodic 363
Paul McGill

10 On the Markovian Similarity 375
Laurent Miclo

11 Sharp Rate for the Dual Quantization Problem......................... 405
Gilles Pagès and Benedikt Wilbertz

12 Cramér's Theorem in Banach Spaces Revisited 455
Pierre Petit

13 On Martingale Chaoses 475
B. Rajeev

14 Explicit Laws for the Records of the Perturbed Random Walk on $\mathbb{Z}$ 495
Laurent Serlet

15 A Potential Theoretic Approach to Tanaka Formula for Asymmetric Lévy Processes 521
Hiroshi Tsukada

Chapter 1
Ornstein-Uhlenbeck Pinball and the Poincaré Inequality in a Punctured Domain

Emmnuel Boissard, Patrick Cattiaux, Arnaud Guillin, and Laurent Miclo

Abstract In this paper we study the Poincaré constant for the Gaussian measure restricted to $D = \mathbb{R}^d - \mathscr{B}$ where $\mathscr{B}$ is the disjoint union of bounded open sets. We will mainly look at the case where the obstacles are Euclidean balls $B(x_i, r_i)$ with radii r_i, or hypercubes with vertices of length $2r_i$, and $d \geq 2$. This will explain the asymptotic behavior of a d-dimensional Ornstein-Uhlenbeck process in the presence of obstacles with elastic normal reflections (the Ornstein-Uhlenbeck pinball).

Keywords Poincaré inequalities · Lyapunov functions · Hitting times · Obstacles

MSC 2010 26D10, 39B62, 47D07, 60G10, 60J60

1.1 Introduction

In order to understand the goal of the present paper let us start with a well known question: how many non overlapping unit discs can be placed in a large square S? This problem of discs packing has a very long history including the following other question: is it possible to perform an algorithm yielding to a perfectly random configuration of N such discs at a sufficiently quick rate (exponential for instance)? This is one of the origin of the Metropolis algorithms as refereed in [16].

The meaning of perfectly random is the following: the configuration space for the model is S^N, describing the location of the N centers of the N discs $B(x_i, 1)$,

---

E. Boissard · P. Cattiaux (✉) · L. Miclo
Institut de Mathématiques de Toulouse, CNRS UMR 5219, Université Paul Sabatier, Toulouse Cedex, France
e-mail: cattiaux@math.univ-toulouse.fr; miclo@math.univ-toulouse.fr

A. Guillin
Laboratoire de Mathématiques, CNRS UMR 6620, Université Blaise Pascal, Aubière, France
e-mail: guillin@math.univ-bpclermont.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_1

but under the constraints $d(x_i, \partial S) \geq 1$ and for all $i \neq j$, $|x_i - x_j| \geq 2$. The remaining domain D is quite complicated, and randomness is described by the uniform measure on D.

The answer to the second question is positive, essentially thanks to compactness, but the exponent in the exponential rate of convergence is strongly connected with the Poincaré constant for the uniform measure on D which is, at the present stage, far to be known (the only known upper bounds are disastrous).

One can of course ask the same questions replacing the square by the whole Euclidean space, and the uniform measure by some natural probability measure, for instance the Gaussian one. But this time even the finiteness of the Poincaré constant is no more clear. A very partial study ($N = 2, 3$) of this problem is done in [15].

In all cases, the probability measure under study, and supported by the complicated state space D is actually an invariant (even reversible) measure for some Markovian dynamics, one can study by itself, and which furnishes a possible algorithm. The boundary of D becomes a reflecting boundary for the dynamics.

In this paper we intend to study the asymptotic behavior of a d-dimensional Ornstein-Uhlenbeck process in the presence of bounded obstacles with elastic normal reflections (looking like a random pinball). The choice of an Ornstein-Uhlenbeck (hence of an invariant measure of Gaussian type) is made for simplicity as it captures already all the new difficulties of this setting, but a general gradient drift diffusion process (satisfying an ordinary Poincaré inequality) could be considered.

Of course for the packing problem in the whole space the obstacles are not bounded, but it seems interesting to look first at the present setting. Our model is also motivated by others considerations we shall give later.

All over the paper we assume that $d \geq 2$. We shall mainly consider the case where the obstacles are non overlapping Euclidean balls or smoothed l^∞ balls (hence smoothed hypercubes) of radius r_i and centers $(x_i)_{1\leq i\leq N\leq +\infty}$, as overlapping obstacles could produce disconnected domains and thus non uniqueness of invariant measures (as well as no Poincaré inequality). We shall also look at different forms of obstacles when it can enlighten the discussion.

To be more precise, consider for $1 \leq N \leq +\infty$, $\mathscr{X} = (x_i)_{1\leq i\leq N\leq +\infty}$ a locally finite collection of points, and $(r_i)_{1\leq i\leq N\leq +\infty}$ a collection of non negative real numbers, satisfying

$$|x_i - x_j| > r_i + r_j \text{ for } i \neq j\,. \tag{1.1}$$

The Ornstein-Uhlenbeck pinball will be given by the following stochastic differential system with reflection

$$\begin{cases} dX_t = dW_t \; - \; \lambda\, X_t\, dt \; + \; \sum_i \,(X_t - x_i)\, dL^i_t\,, \\ L^i_t = \int_0^t \, \mathbb{1}_{|X_s - x_i| = r_i}\, dL^i_s. \end{cases} \tag{1.2}$$

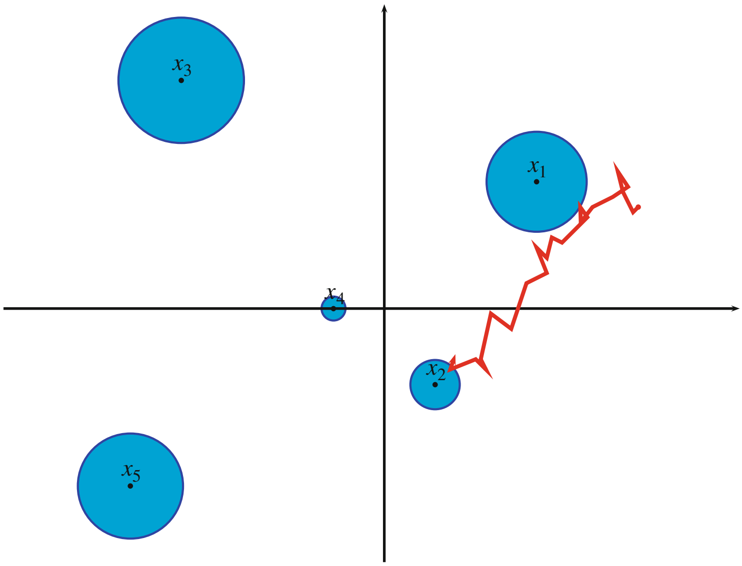


Fig. 1.1 An Ornstein-Uhlenbeck particle in a billiard

Here W is a standard Wiener process and we assume that $\mathbb{P}(|X_0 - x_i| \geq r_i \text{ for all } i) = 1$. L^i is the local time description of the elastic and normal reflection of the process when it hits $B(x_i, r_i)$ (Fig. 1.1).

Existence and non explosion of the process, which is especially relevant for $N = +\infty$, will be discussed in Appendix 1. The process lives in

$$\bar{D} = \mathbb{R}^d - \{x \, ; \, |x - x_i| < r_i \text{ for some } i\}\,, \tag{1.3}$$

that is, we have removed a collection of non overlapping balls (or more generally non overlapping obstacles).

It is easily seen that the process admits an unique invariant (actually reversible) probability measure $\mu_{\lambda,\mathscr{X}}$, which is simply the Gaussian measure restricted to D, i.e.

$$\mu_{\lambda,\mathscr{X}}(dx) = Z_{\lambda,\mathscr{X}}^{-1} \, \mathbb{1}_D(x) \, e^{-\lambda |x|^2} \, dx\,, \tag{1.4}$$

where $Z_{\lambda,\mathscr{X}}$ is of course a normalizing constant. *Hence the process is positive recurrent.*

The question is to describe the rate of convergence for the distribution of the process at time t to its equilibrium measure.

To this end we shall look at the Poincaré constant of $\mu_{\lambda,\mathscr{X}}$ since it is well known that this Poincaré constant captures the exponential rate of convergence to equilibrium for symmetric processes (see e.g. [14] lemma 2.14 and [6] theorem 2.1). Other functional inequalities (logarithmic Sobolev inequality, transportation inequality, …) could be equally considered and the techniques developed here could also prove to be useful in these cases (for examples Lyapunov techniques have been introduced in the study of Super Poincaré inequalities in [13], including logarithmic Sobolev inequalities).

When the number of obstacles N is finite, one can see, using Down, Meyn and Tweedie results [18] and some regularity results for the process following [9, 10], that the process is exponentially ergodic. It follows from [6] theorem 2.1, that $\mu_{\lambda,\mathscr{X}}$ satisfies some Poincaré inequality, i.e. for all smooth f (defined on the whole $\mathbb{R}^d$)

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) \leq C_P(\lambda,\mathscr{X}) \int |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} \,. \tag{1.5}$$

But the above method furnishes non explicit bounds for the Poincaré constant $C_P(\lambda,\mathscr{X})$.

Our first goal is thus to obtain reasonable and *explicit* upper and lower bounds for the Poincaré constant. Surprisingly enough (or not) the case of one hard obstacle already contains non trivial features.

Our second goal is to look at the case of *infinitely many* obstacles, for which the finiteness of the Poincaré constant is not even clear.

Part of the title of the paper is taken from a paper by Lieb et al. [26] which is one of the very few papers dealing with Poincaré inequality in a sub-domain. Of course, one cannot get any general result due to the fact that one can always remove an, as small as we want, subset disconnecting the whole space; so that the remaining sub-domain cannot satisfy some Poincaré inequality. Hence doing this breaks the ergodicity of the process.

The method used in [26] relies on the extension of functions defined in D to the whole space. But the inequality they obtain, involves the energy of this extension (including the part inside D^c), so that it is not useful to get a quantitative rate of convergence for our process.

Our model can be used (or modified) as a model for crowds displacements (involving several particles in the obstacles environment). In particular the design of small obstacles that should kill the Poincaré constant is interesting.

Let us now describe the main results and main methods contained in the paper.

First, it is easily seen, thanks to homogeneity, that

$$C_P(\lambda,\mathscr{X}) = \frac{1}{\lambda}\, C_P(1, \sqrt{\lambda}\,\mathscr{X}) \,. \tag{1.6}$$

where $\sqrt{\lambda}\,\mathscr{X}$ is the homotetic of $\mathscr{X}$, i.e. the collection of $B(\sqrt{\lambda}\,x_i, \sqrt{\lambda}\,r_i)$. Hence we have one degree of freedom in the use of all parameters. This homogeneity property will be used in the paper to improve some bounds.

The first section is peculiar. We look at a single spherical obstacle centered at the origin. We show that the Poincaré constant is given by

$$C_P(\lambda, B(0,r)) \;\approx\; \frac{1}{\lambda} + \frac{r^2}{d}$$

i.e. is up to some universal constant the sum of the Poincaré constant of the Gaussian distribution $1/2\lambda$ and the one of the uniform measure on the sphere of radius r i.e. r^2/d. For the process this reflects the fact that it hits a neighborhood of the origin with an exponential rate given by λ but turns around the sphere with an exponential rate given by r^2/d. This is also in accordance with what is expected when $\lambda \to +\infty$ ($\mu_{\lambda,\mathscr{X}}$ is close to the uniform measure on the sphere) or $r \to 0$ where the obstacle disappears.

We also look at the usual perturbation method for Poincaré inequalities when the center is no more located at the origin (see Propositions 1.2 and 1.3) with results that are not entirely satisfactory. The result for the obstacle $B(0,r)$ can be used, through the decomposition of variance method, to obtain results for a general single ball $B(y,r)$. This is explained in Appendix 3.

The next two Sects. 1.3 and 1.4 are devoted to our main goals in the case of spherical obstacles: obtain explicit controls for the Poincaré constant in the presence of a single obstacle, extend it to a finite number of obstacles, prove that it is still finite in the case of an infinite number of obstacles.

In Sect. 1.3 we develop a "local" Lyapunov method (in the spirit of [5]) around the obstacle. Under a restriction to small sizes, it is possible to give some explicit Lyapunov function. As in recent works [4, 7] the difficulty is then to piece together the Lyapunov functions we may build near the obstacle and far from the obstacle and the origin. Let us describe the main results and methods.

First we are able to find explicit Lyapunov functions in the neighborhood of the obstacles provided

$$\forall i\,,\; r_i < r\sqrt{\lambda} = \sqrt{(d-1)/2} \,-\, 2^{-\frac{3}{4}}\,.$$

This implies some limitation for the dimension namely

$$d \geq 7\,.$$

If in addition the balls $B(x_i, r_i + b(\lambda)))$ are non-overlapping (here $b(\lambda)$ is some explicit constant), then one obtains an explicit upper bound for the Poincaré constant. This is explained in Sect. 1.3.2 in particular in Proposition 1.4.

The remaining of Sect. 1.3 is then dedicated to get rid of the dimension restriction still for small obstacles i.e. provided

$$\forall i\,,\; r_i \;\leq\; \frac{1}{2}\sqrt{(d-1)/2}\,.$$

In Sects. 1.3.3 and 1.3.4 we show how to control the variance of functions compactly supported in the exterior of a large ball containing the origin. As a consequence we get in Sects. 1.3.5 and 1.3.6 a general result for the Poincaré constant when there is only one obstacle, gathering all what was done in these subsections and the previous section.

Finally we prove the finiteness of the Poincaré constant for an infinite number of small obstacles uniformly disconnected, that is such the distance between two distinct obstacles is uniformly larger than some $\varepsilon > 0$ in Corollary 1.1. If we are not able to give a precise description of the Poincaré constant in general, we can give some provided all obstacles are far enough from the origin i.e. if

$$\forall i\ ,\ |x_i|\sqrt{\lambda} > c\sqrt{d}$$

for some constant c (see Proposition 1.5 and the explanations at the beginning of Sect. 1.3.7).

We close Sect. 1.3 by a subsection explaining what happens if we replace Euclidean balls by hypercubes.

In Sect. 1.4 we use the results in [14] in order to build new Lyapunov functions near the obstacles, this time without restriction on the radius. To this end, we study in details how the process avoids a spherical obstacle, using stochastic calculus. This allows us to build a new Lyapunov function near the obstacle, which is given by some exponential moment of the time needed to go around the obstacle. Useful results on the Laplace transform of exit times for some linear processes are recalled in Appendix 2. This new Lyapunov function is then used in Sect. 1.4.2 to obtain an upper estimate for the Poincaré constant in a shell around a spherical obstacle. Together with the method in Sect. 1.3, we can then show (see Proposition 1.9) that provided

$$\forall i\ ,\ |x_i| > r_i + m \quad \text{and} \quad r_i > \frac{1}{2}$$

for some large enough m the Poincaré constant is finite and obtain an upper bound for it. Finally we can extend the result in the case of infinitely many large obstacles. Hence Sects. 1.4 and 1.3 are complementary.

Gathering all this, we have the following key result: for a spherical obstacle located far from the origin, the Poincaré constant does not depend on the radius (contrary to what we conjectured in a previous version of this work). This result allows us to show the following general result in the case of an infinite number of spherical obstacles.

Theorem 1.1 *Let $\mathscr{X} = (x_i)_{1\leq i<+\infty}$ a locally finite collection of distinct points, ordered such that $|x_i| \leq |x_{i+1}|$ for all i, and $\mathscr{R} = (r_i)_{1\leq i<+\infty}$ a collection of non-negative numbers. Assume that there exists $\varepsilon > 0$ with $|x_i - x_j| > r_i + r_j + \varepsilon$ for all $i \neq j$.*

Then for any $\lambda > 0$*, the measure* $\mu_{\lambda,\mathscr{X}}$ *defined in* (1.4) *has a finite Poincaré constant and the reflected Ornstein-Uhlenbeck process in D (defined in* (1.3)*) is exponentially ergodic.*

Section 1.5 is devoted to obtain lower bounds. We show in particular that if we replace Euclidean balls by hypercubes, the situation is drastically changed since each obstacle (in a particular configuration) gives some contribution e^{cr^2} where r denotes the length of an edge of the hypercube. In particular large obstacles far from the origin can make the Poincaré constant go to ∞. We give two approaches of this result: one using exit times for the stochastic process, the second one using isoperimetric ideas. The same isoperimetric ideas are used to give a lower bound for the Poincaré constant in the case of spherical obstacles. To conclude the section we show that replacing balls by some non convex small and far obstacles can kill the exponential ergodicity. This situation is analogous to the one obtained with "touching" spherical obstacles.

The conclusion is that, presumably for uniformly convex obstacles (with an uniform curvature bounded from below uniformly in the location of the obstacles too) a similar result as for spherical obstacles holds true and our method can be used. The only difficulty is to find the good Lyapunov functions. A lack of uniform convexity has some disastrous consequences on the Poincaré constant, even for small and far obstacles.

Dedication During the revision of the paper, we learned about the death of Marc Yor. Everybody knows what a tragedy it is for Probability theory. It turns out that some beautiful results of Marc Yor on exit times for general squared radial Ornstein-Uhlenbeck processes recalled in an Appendix, are crucial in the present paper.

1.2 Some Results When $N = 1$

1.2.1 The Case of One Centered Ball, i.e. $y = 0$

Assume $N = 1$ and the obstacle is the Euclidean ball $B(y, r)$ with $y = 0$. In this case $\mu_{\lambda,\mathscr{X}} = \nu^0_{\lambda,r}$ is the standard Gaussian measure with variance $\frac{1}{2\lambda}$ restricted to $D = \mathbb{R}^d - B(0, r)$. More generally we will denote by $\nu^y_{\lambda,r}$ the Gaussian measure with mean y and variance $\frac{1}{2\lambda}$ restricted to $\mathbb{R}^d - B(0, r)$.

$\mu_{\lambda,\mathscr{X}}$ is spherically symmetric. Though it is not log-concave, its radial part, proportional to

$$\mathbb{1}_{\rho>r}\, \rho^{d-1}\, e^{-\lambda\,\rho^2}$$

is log concave in ρ so that we may use the results in [8], yielding

Proposition 1.1 *When* $\mathscr{X} = B(0,r)$*, the measure* $\mu_{\lambda,\mathscr{X}}$ *satisfies a Poincaré inequality* (1.5) *with*

$$\frac{1}{2}\left(\frac{1}{2\lambda}+\frac{r^2}{d}\right) \leq \max\left(\frac{1}{2\lambda}\,,\,\frac{r^2}{d}\right) \leq C_P(\lambda,\,B(0,r)) \leq \frac{1}{\lambda}+\frac{r^2}{d}\,.$$

Proof For the upper bound, the only thing to do in view of [8] is to estimate $\mathbb{E}(\xi^2)$ where ξ is a random variable on $\mathbb{R}^+$ with density

$$\rho \mapsto A_\lambda^{-1}\,\mathbb{1}_{\rho>r}\,\rho^{d-1}\,e^{-\lambda\,\rho^2}\,. \tag{1.7}$$

But

$$A_\lambda = \int_r^{+\infty} \rho^{d-1}\,e^{-\lambda\,\rho^2}\,d\rho \geq r^{d-2}\int_r^{+\infty} \rho\,e^{-\lambda\,\rho^2}\,d\rho = \frac{r^{d-2}\,e^{-\lambda\,r^2}}{2\lambda}\,.$$

A simple integration by parts yields

$$\mathbb{E}(\xi^2) = \frac{d}{2\lambda} + \frac{r^d\,e^{-\lambda\,r^2}}{2\,\lambda\,A_\lambda} \leq \frac{d}{2\lambda} + r^2\,.$$

The main result in [8] says that

$$C_P(\lambda,\,B(0,r)) \leq \frac{13}{d}\,\mathbb{E}(\xi^2)\,,$$

hence the result with a constant 13.

Instead of directly using Bobkov's result, one can look more carefully at its proof. The first part of this proof consists in establishing a bound for the Poincaré constant of the law given by (1.7). Here, again, we may apply Bakry-Emery criterion (which holds true on an interval), which furnishes $1/(2\lambda)$. The second step uses the Poincaré constant of the uniform measure on the unit sphere, i.e. $1/d$, times the previous bound for $\mathbb{E}(\xi^2)$. Finally these two bounds have to be summed up, yielding the result.

For the lower bound it is enough to consider the function $f(z) = \sum_{j=1}^d z_j$. Indeed, the energy of f is equal to d. Furthermore on one hand

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) = \frac{\int_r^{+\infty} \rho^{d+1}\,e^{-\lambda\,\rho^2}\,d\rho}{\int_r^{+\infty} \rho^{d-1}\,e^{-\lambda\,\rho^2}\,d\rho} \geq r^2\,,$$

while on the other hand, an integration by parts shows that

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) = \frac{d}{2\lambda} + \frac{r^d\,e^{-\lambda r^2}}{2\lambda\,\int_r^{+\infty} \rho^{d-1}\,e^{-\lambda\,\rho^2}\,d\rho} \geq \frac{d}{2\lambda}$$

yielding the lower bound since the maximum is larger than the half sum.

This result is satisfactory since we obtain the good order. Notice that when r goes to 0 we recover (up to some universal constant) the Gaussian Poincaré constant, and when λ goes to $+\infty$ we recover (up to some universal constant) the Poincaré constant of the uniform measure on the sphere $r S^{d-1}$ which is the limiting measure of $\mu_{\lambda,\mathscr{X}}$. Also notice that the obstacle is really an obstacle since the Poincaré constant is larger than the Gaussian one.

Remark 1.1 It is immediate that the same upper bound is true (with the same proof) for $\nu^0_{\lambda,r,R}(dx) = Z^{-1}_{\lambda,r,R}\, \mathbb{1}_{R>|x|>r}\, e^{-\lambda\,|x|^2}$ i.e. the Gaussian measure restricted to a spherical shell $\{R > |x| > r\}$. For the lower bound some extra work is necessary.

1.2.2 A First Estimate for a General y Using Perturbation

An intuitive idea to get estimates on the Poincaré constant relies on the Lyapunov function method developed in [5] which requires a local Poincaré inequality usually derived from Holley-Stroock perturbation's argument. To be more precise, let us introduce $\nu^y_{\lambda,r}$ which is the Gaussian measure with mean $-y \in \mathbb{R}^d$ restricted to $\mathbb{R}^d - B(0,r)$, and its natural generator

$$L_y = \frac{1}{2}\,\Delta - \lambda\,\langle x + y, \nabla\rangle\,.$$

If we consider the function $x \mapsto h(x) = |y + x|^2$ we see that

$$L_y h(x) = d - 2\lambda |x + y|^2 \leq -\lambda\, h(x) \quad \text{if} \quad |x| \geq |y| + (d/\lambda)^{1/2}\,.$$

So we can use the method in [5]. Consider, for $\varepsilon > 0$, the ball

$$U = B\left(0, \left(|y| + (d/\lambda)^{1/2}\right) \vee (r + \varepsilon)\right)\,.$$

h is a Lyapunov function satisfying

$$L_y h \leq -\lambda h + d\,\mathbb{1}_U\,.$$

Since U^c does not intersect the obstacle $B(0,r)$, we may follow [14] and obtain that

$$C_P(\nu^y_{\lambda,r}) \leq \frac{2}{\lambda} + \left(\frac{2}{\lambda} + 2\right) C_P(\nu^y_{\lambda,r}, U + 1)\,,$$

where $C_P(\nu^y_{\lambda,r}, U+1)$ is the Poincaré constant of the measure $\nu^y_{\lambda,r}$ restricted to the shell

$$S = \left\{r < |x| < 1 + \left(\left(|y| + (d/\lambda)^{1/2}\right) \vee (r + \varepsilon)\right)\right\}\,.$$

Actually since h may vanish, we first have to work with $h+\eta$ for some small η (and small changes in the constants) and then let η go to 0 for the dust to settle.

Now we apply Holley-Stroock perturbation argument. Indeed

$$\nu^y_{\lambda,r}(dx) = C(y,\lambda)\, e^{-2\lambda\,\langle x,y\rangle}\, \nu^0_{\lambda,r}(dx)$$

for some constant $C(y,\lambda)$. In restriction to the shell S, it is thus a logarithmically bounded perturbation of $\nu^0_{\lambda,r}$ with a logarithmic oscillation less than

$$4\lambda\,|y|\,\left(1+\left(\left(|y|+(d/\lambda)^{1/2}\right)\vee(r+\varepsilon)\right)\right)$$

so that we have obtained

$$C_P(\lambda,\,B(y,r)) \le \frac{2}{\lambda} + \left(2+\frac{2}{\lambda}\right)\left(\frac{1}{\lambda}+\frac{r^2}{d}\right)\, e^{4\lambda\,|y|\,\left(1+\left(\left(|y|+(d/\lambda)^{1/2}\right)\vee(r+\varepsilon)\right)\right)}\,.$$

The previous bound is bad for small λ's but one can use the homogeneity property (1.6), and finally, letting ε go to 0

Proposition 1.2 *For a general y, the measure $\mu_{\lambda,B(y,r)}$ satisfies a Poincaré inequality* (1.5) *with*

$$C_P(\lambda,\,B(y,r)) \le \frac{2}{\lambda}\left(1+2\left(1+\frac{r^2\,\lambda}{d}\right)e^{4\sqrt{\lambda}\,|y|\,\left(1+\left(|y|\sqrt{\lambda}+d^{1/2}\right)\vee r\sqrt{\lambda}\right)}\right)\,.$$

The previous result is not satisfactory for large values of $|y|$, r or λ. In addition it is not possible to extend the method to more than one obstacle. Finally we have some extra dimension dependence when $y=0$ due to the exponential term. Our aim will now be to improve this estimate.

Another possible way, in order to evaluate the Poincaré constant, is to write, for

$$g = f - \frac{\int f(x)\,e^{-\lambda\langle x,y\rangle}\,\nu^0_{\lambda,r}(dx)}{\int e^{-\lambda\langle x,y\rangle}\,\nu^0_{\lambda,r}(dx)}\,,\ \text{ so that } \int g(x)\,e^{-\lambda\langle x,y\rangle}\,\nu^0_{\lambda,r}(dx) = 0$$

$$\mathrm{Var}_{\nu^y_{\lambda,r}}(f) \le \int g^2\,d\nu^y_{\lambda,r} = C(\lambda,y,r)\int\left(g\,e^{-\lambda\langle x,y\rangle}\right)^2 d\nu^0_{\lambda,r} \tag{1.8}$$

$$\le C(\lambda,y,r)\,C_P(\lambda,\,B(0,r))\int\left|\nabla\left(g\,e^{-\lambda\langle x,y\rangle}\right)\right|^2 d\nu^0_{\lambda,r} \tag{1.9}$$

$$\le 2\,C_P(\lambda,\,B(0,r))\left(\int|\nabla g|^2\,d\nu^y_{\lambda,r} + \lambda^2\,|y|^2\int g^2\,d\nu^y_{\lambda,r}\right). \tag{1.10}$$

It follows first that, provided $2\,C_P(\lambda,\,B(0,r))\,\lambda^2\,|y|^2 \le \frac{1}{2}$,

$$\int g^2\,d\nu^y_{\lambda,r} \le 4\,C_P(\lambda,\,B(0,r)) \int |\nabla g|^2\,d\nu^y_{\lambda,r}\,,$$

and finally

Proposition 1.3 *If* $4\,\lambda\,|y|^2\left(1+\frac{r^2\,\lambda}{d}\right) \le 1$, *the measure* $\mu_{\lambda,\mathscr{X}}$ *where* $\mathscr{X} = B(y,r)$ *satisfies a Poincaré inequality* (1.5) *with*

$$C_P(\lambda,\,B(y,r)) \le 4\left(\frac{1}{\lambda}+\frac{r^2}{d}\right).$$

One can note that under the condition $4\,\lambda\,|y|^2\left(1+\frac{r^2\,\lambda}{d}\right) \le 1$, Propositions 1.2 and 1.3 yield, up to some dimension dependent constant, similar bounds. Of course the first proposition is more general.

1.3 Using Lyapunov Functions

In what we did previously we have used Lyapunov functions vanishing in a neighborhood of the obstacle(s). Indeed a Lyapunov function (generally) has to belong to the domain of the generator, in particular its normal derivative (generally) has to vanish on the boundary of the obstacle. Since it seems that a squared distance is a good candidate it is natural to look at the geodesic distance in the punctured domain D (see [2] and also [23] for small time estimates of the density in this situation). Unless differentiability problems (the distance is not everywhere C^2) it seems that this distance does not yield the appropriate estimate (calculations being tedious).

Instead of trying to get a “global” Lyapunov function, we shall build “locally” such functions.

In this section we consider the case $1 \le N \le +\infty$ i.e. we may consider as well an infinite number of obstacles.

To be more precise, consider an open neighborhood (in D) U of the obstacles and some smooth function χ supported in D such that $\mathbb{1}_{U^c} \le \chi \le 1$ (in particular χ vanishes on the boundary of the obstacles). Let f be a smooth (compactly supported) function and m be such that $\int \chi\,(f-m)\,d\mu_{\lambda,\mathscr{X}} = 0$. Then

$$\begin{aligned}\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) &\le \int_D (f-m)^2 d\mu_{\lambda,\mathscr{X}} = \int_U (f-m)^2 d\mu_{\lambda,\mathscr{X}} + \int_{U^c} (f-m)^2 d\mu_{\lambda,\mathscr{X}} \\ &\le \int_U (f-m)^2 d\mu_{\lambda,\mathscr{X}} + \int_{\mathbb{R}^d} \chi^2\,(f-m)^2 d\mu_{\lambda,\mathscr{X}}\end{aligned}$$

$$\leq \int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{2\lambda} \int_{\mathbb{R}^d} |\nabla(\chi\,(f-m))|^2\, d\mu_{\lambda,\mathscr{X}}$$
$$\leq \int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{\lambda} \int_D \left(|\nabla\chi|^2\,(f-m)^2 + \chi^2\,|\nabla f|^2\right) d\mu_{\lambda,\mathscr{X}}\,,$$

where we have used that $\mu_{\lambda,\mathscr{X}}$ is simply the Gaussian measure on the support of χ, introducing the Poincaré constant of the Gaussian $1/2\lambda$. It follows

$$\begin{aligned}\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) &\leq \int_D (f-m)^2 d\mu_{\lambda,\mathscr{X}} \\ &\leq \left(1+\frac{\|\nabla\chi\|_\infty^2}{\lambda}\right) \int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{\lambda}\int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}}\,.\end{aligned} \tag{1.11}$$

We thus see that what we have to do is to get some bound for $\int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}}$ in terms of the energy $\int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}}$ for any smooth f which is exactly what is done by finding a "local" Lyapunov function.

1.3.1 Two Useful Lemmas on Lyapunov Function Method

We may now present two particularly useful lemmas concerning Lyapunov function method and localization. Let us begin by the following remark: in the previous derivation assume that for some $p > 1$ and some constant C,

$$\int_U (f-m)^2 d\mu_{\lambda,\mathscr{X}} \leq \frac{\lambda}{p\,\|\nabla\chi\|_\infty^2} \int_{\mathbb{R}^d} \chi^2\,(f-m)^2 d\mu_{\lambda,\mathscr{X}} + C\int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}}. \tag{1.12}$$

Then, using the Poincaré inequality for the Gaussian measure, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \chi^2\,(f-m)^2 d\mu_{\lambda,\mathscr{X}} &\leq \frac{1}{\lambda}\int_{\mathbb{R}^d} \left(|\nabla\chi|^2\,(f-m)^2 + \chi^2\,|\nabla f|^2\right) d\mu_{\lambda,\mathscr{X}} \\ &\leq \frac{\|\nabla\chi\|_\infty^2}{\lambda}\int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{\lambda}\int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}} \\ &\leq \frac{1}{p}\int_{\mathbb{R}^d} \chi^2\,(f-m)^2 d\mu_{\lambda,\mathscr{X}} \\ &\quad + \frac{1}{\lambda}\,(1 + C\,\|\nabla\chi\|_\infty^2) \int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}}\end{aligned}$$

so that

$$\int_{\mathbb{R}^d} \chi^2 (f-m)^2 d\mu_{\lambda,\mathscr{X}} \leq \frac{p}{(p-1)\lambda} (1 + C \parallel \nabla\chi \parallel_\infty^2) \int_D |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}}$$

and using (1.12)

$$\int_U (f-m)^2 d\mu_{\lambda,\mathscr{X}} \leq \left(C + \frac{1}{(p-1) \parallel \nabla\chi \parallel_\infty^2} (1 + C \parallel \nabla\chi \parallel_\infty^2) \right) \int_D |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}}$$

and finally

Lemma 1.1 *If* (1.12) *holds for some smooth χ supported in D and such that $\mathbb{1}_{U^c} \leq \chi \leq 1$, then*

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) \leq \frac{1}{p-1} \left(Cp + \frac{1}{\parallel \nabla\chi \parallel_\infty^2} + \frac{p(1 + C \parallel \nabla\chi \parallel_\infty^2)}{\lambda} \right) \int_D |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} .$$

From now on we assume that ∂D is smooth enough and we denote by n the normalized inward (pointing into D) normal vector field on ∂D.

Now recall the basic lemma used in [5, 14] we state here in a slightly more general context (actually this lemma is more or less contained in [14] Remark 3.3)

Lemma 1.2 *Let f be a smooth function with compact support in $\bar{D}$ and W a positive smooth function. Denote by $\mu^S_{\lambda,\mathscr{X}}$ the trace (surface measure) on ∂D of $\mu_{\lambda,\mathscr{X}}$. Then the following holds*

$$\int_D \frac{-LW}{W} f^2 \, d\mu_{\lambda,\mathscr{X}} \leq \frac{1}{2} \int_D |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} + \frac{1}{2} \int_{\partial D} \frac{\partial W}{\partial n} \frac{f^2}{W} \, d\mu^S_{\lambda,\mathscr{X}} .$$

Proof We recall the proof for the sake of completeness. Using the first Green formula we have (recall that n is pointing inward)

$$\begin{aligned}
\int_D \frac{-2\,LW}{W} f^2 \, d\mu_{\lambda,\mathscr{X}} &= \int_D \left\langle \nabla \left(\frac{f^2}{W} \right) , \nabla W \right\rangle d\mu_{\lambda,\mathscr{X}} + \int_{\partial D} \frac{\partial W}{\partial n} \frac{f^2}{W} \, d\mu^S_{\lambda,\mathscr{X}} \\
&= 2 \int_D \frac{f}{W} \langle \nabla f, \nabla W \rangle \, d\mu_{\lambda,\mathscr{X}} - \int_D \frac{f^2}{W^2} |\nabla W|^2 \, d\mu_{\lambda,\mathscr{X}} \\
&\quad + \int_{\partial D} \frac{\partial W}{\partial n} \frac{f^2}{W} \, d\mu^S_{\lambda,\mathscr{X}} \\
&= - \int_D \left| \frac{f}{W} \nabla W - \nabla f \right|^2 d\mu_{\lambda,\mathscr{X}} + \int_D |\nabla f|^2 \, d\mu_{\lambda,r} \\
&\quad + \int_{\partial D} \frac{\partial W}{\partial n} \frac{f^2}{W} \, d\mu^S_{\lambda,\mathscr{X}} .
\end{aligned}$$

1.3.2 Localizing Around the Obstacles

From now on for simplicity we will assume that $D^c = \cup_i B(x_i, r_i)$ where the B's are non overlapping Euclidean balls. We shall indicate at the end how the results extend to others situations, in particular to smoothed hypercubes.

We will construct first Lyapunov functions near the obstacles. Hence we will build open neighborhoods U_i for each ball, and will assume that the U_i's are non overlapping sets too.

Not to introduce immediately too much notations, we shall write things for one ball denoted by $B(y, r)$. Let $h > 0$ and assume that one can find a Lyapunov function W such that $LW \leq -\theta W$ for $|x - y| \leq r + 2h$ and $\partial W/\partial n \leq 0$ on $|x - y| = r$. Choose some smooth function ψ such that $\mathbb{1}_{\{|x-y|\leq r+2h\}} \geq \psi \geq \mathbb{1}_{\{|x-y|\leq r+h\}}$ and, for some $\varepsilon > 0$,

$$\| \nabla\psi \|_\infty \leq (1+\varepsilon)/h\,.$$

Applying Lemma 1.2 to $\psi\, f$ we obtain

$$\begin{aligned}
\int_{r<|x-y|<r+h} f^2\, d\mu_{\lambda,\mathscr{X}} &\leq \int_{r<|x-y|<r+2h} (\psi\, f)^2\, d\mu_{\lambda,\mathscr{X}} \\
&\leq \frac{1}{\theta} \int_{r<|x-y|<r+2h} \frac{-LW}{W}\, (\psi\, f)^2\, d\mu_{\lambda,\mathscr{X}} \\
&\leq \frac{1}{\theta} \int_{r<|x-y|<r+2h} |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}} \\
&\quad + \frac{1}{\theta} \left(\frac{1+\varepsilon}{h}\right)^2 \int_{r+h<|x-y|<r+2h} f^2\, d\mu_{\lambda,\mathscr{X}}\,.
\end{aligned}$$

Me may of course let ε go to 0.

Next choose $U = \cup_i B(x_i, r_i + h_i)$, $\mathbb{1}_D \geq \chi \geq \mathbb{1}_{U^c}$ and assume that the balls $B(x_i, r_i + 2h_i)$ are non overlapping. Assume that one can find Lyapunov functions W_i such that $LW_i \leq -\theta_i W_i$ for $|x-x_i| \leq r_i+2h_i$ and $\partial W_i/\partial n \leq 0$ on $|x-x_i| = r_i$. Let $h = \min h_i$, $\theta = \min \theta_i$. Using a similar argument as before we may assume that actually $\| \nabla\chi \|_\infty = \frac{1}{h}$.

The previous inequality applied to $f - m$ in each ball yields

$$\int_U (f-m)^2\, d\mu_{\lambda,\mathscr{X}} \leq \frac{1}{\theta} \int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{\theta\, h^2} \int_{\mathbb{R}^d} \chi\, (f-m)^2\, d\mu_{\lambda,\mathscr{X}} \tag{1.13}$$

i.e. (1.12) is satisfied with

$$C = \frac{1}{\theta} \quad \text{and} \quad p = \lambda\,\theta\,h^4\,, \tag{1.14}$$

provided the latter is larger than 1.

We may thus apply Lemma 1.1 and obtain

Lemma 1.3 *Let $h > 0$ and $\theta > 0$. Assume that for $h_i \geq h$ the balls $B(x_i, r_i + 2h_i)$ are non overlapping. Assume in addition that one can find Lyapunov functions W_i such that $LW_i \leq -\theta_i W_i$ for $|x - x_i| \leq r_i + 2h_i$, $\partial W_i/\partial n \leq 0$ on $|x - x_i| = r_i$, $\theta_i \geq \theta$.*

Then, provided $\lambda\,\theta\,h^4 > 1$,

$$\text{Var}_{\mu_{\lambda,\mathscr{X}}}(f) \leq \frac{h^2\,(2 + (\theta + \lambda)\,h^2)}{\lambda\,\theta\,h^4\, -\, 1} \int_D |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}}\,.$$

Hence all we have to do is to find a "good" Lyapunov function.

For the moment, U will be an open ball centered at y. Without loss of generality (if necessary) we may assume that $y = (a, 0)$ for some $a \in \mathbb{R}^+$, 0 being the null vector of $\mathbb{R}^{d-1}$. The (non normalized) normal vector field at the boundary of $B(y, r)$, pointing inward D, is thus $x - y = (x^1 - a, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{d-1}$.

We shall exhibit some Lyapunov function W_y near the obstacle. For $|\bar{x}| \leq r + 2h$ define

$$W_y(x^1, \bar{x}) = (r + 2h + \varepsilon)^2 - |\bar{x}|^2\,.$$

Then $\nabla W_y(x^1, \bar{x}) = (0, -2\bar{x})$ and

$$\frac{\partial W_y}{\partial n}(x^1, \bar{x}) = -\,\frac{2|\bar{x}|^2}{|x - y|} \leq 0\,. \tag{1.15}$$

Now $LW_y = -(d - 1) + 2\,\lambda\,|\bar{x}|^2$ so that $LW_y \leq -\,2\,\lambda\,W_y$ provided

$$d - 1 \geq 2\,\lambda\,(r + 2h + \varepsilon)^2\,. \tag{1.16}$$

As before we may let ε go to 0 so that we obtain (1.13) with $\theta = 2\lambda$ and $p = 2\lambda^2\,h^4 > 1$.

Choosing $h = b/\sqrt{\lambda}$, with $p = 2b^4 > 1$, we see that we must have $d \geq 7$ and $r\sqrt{\lambda} \leq \sqrt{(d - 1)/2} - 2b$. Finally we have shown

Proposition 1.4 *Let $b > 0$ and $r > 0$ be such that $2b^4 > 1$ and $r\sqrt{\lambda} \leq \sqrt{(d - 1)/2} - 2b$, so that $d \geq 7$.*

Let $D^c = \cup_i B(x_i, r_i)$ where $r_i \leq r$ for all i. Assume that the balls $B(x_i, r_i + 2b/\sqrt{\lambda})$ are non overlapping.

Then the measure $\mu_{\lambda,\mathscr{X}}$ *satisfies a Poincaré inequality* (1.5) *with*

$$C_P(\lambda, \mathscr{X}) \le \frac{1}{\lambda}\,\frac{b^2(3b^2+2)}{2b^4-1}\,.$$

The dimension dependence clearly indicates that, even for small r's, we presumably did not find the good Lyapunov function. However for large dimensions we see that small enough obstacles do not alternate the finiteness of the Poincaré constant.

Also notice that if we define $\beta = \frac{2b\sqrt{2}}{\sqrt{d-1}}$ the condition on r reads

$$r\sqrt{\lambda} \le (1-\beta)\,\sqrt{(d-1)/2} \quad \text{for some } \beta \text{ such that} \quad 1 > \beta > \frac{2^{5/4}}{\sqrt{d-1}}\,. \tag{1.17}$$

In the next three subsections we shall adapt the previous method in order to cover all dimensions but for far enough obstacles.

1.3.3 Localizing Away from the Obstacles and the Origin

Consider now $W(x) = |x|^2$ so that for $1 > \eta > 0$,

$$LW(x) = d - 2\lambda\, W(x) \le -\,2\lambda\,(1-\eta)\, W(x) \quad \text{for } |x| \ge \sqrt{\frac{d}{2\lambda\eta}}\,.$$

We will obtain some Dirichlet-Poincaré bound, i.e. we look at functions g which are smooth and compactly supported in $|x| \ge \sqrt{\frac{d}{2\lambda\eta}}$ (hence vanish on the boundary of this large ball). But we also have to assume that no obstacle intersects the boundary of this region of the space. Hence we have to replace the sphere $\{|x| = \sqrt{d/2\lambda\,\eta}\}$ by some smooth hypersurface S such that $S \subset D$ and $\sqrt{d/2\lambda\,\eta} \le d(0,S) \le c\sqrt{d/2\lambda\,\eta}$ for some $c > 1$ and for all $x_i \in \mathscr{X}$, $B(x_i, r_i + 3h_i) \cap S = \emptyset$. We also assume that the balls $B(x_i, r_i + 3h_i)$ are non overlapping.

It will be clear in what follows that such an S does exist, but for the moment the existence of S is an assumption. The whole space D in thus divided in two connected components D_0 containing 0 and D_∞ such that S is the boundary of both.

We consider now the $x_i \in \mathscr{X}$ such that $B(x_i, r_i + 3h_i) \subset D_\infty$, in particular $|x_i|$ is large enough. We denote by $\mathscr{X}_\infty$ this set.

Let g be compactly supported in D_∞. For all $1 \le \varepsilon \le 2$ we apply Lemma 1.2 in

$$D_\varepsilon = D_\infty \cap_{x_i \in \mathscr{X}_\infty} \{|x - x_i| \ge r_i + \varepsilon h_i\}\,,$$

since the support of g does not intersect S, i.e.

$$\int_{D_\varepsilon} \frac{-LW}{W}\, g^2 \, d\mu_{\lambda,\mathscr{X}} \;\le\; \frac{1}{2}\int_{D_\varepsilon} |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} \;+\; \frac{1}{2}\int_{\partial D_\varepsilon} \frac{\partial W}{\partial n}\, \frac{g^2}{W}\, d\mu^{\varepsilon}_{\lambda,\mathscr{X}} \,,$$

where $\mu^{\varepsilon}_{\lambda,\mathscr{X}}$ denotes the trace of $\mu_{\lambda,\mathscr{X}}$ on the boundary ∂D_ε.

It yields for all ε as before

$$\begin{aligned}\int_{D_\varepsilon} g^2 \, d\mu_{\lambda,\mathscr{X}} &\le \frac{1}{2\lambda\,(1-\eta)} \int_{D_\varepsilon} \frac{-LW}{W}\, g^2 \, d\mu_{\lambda,\mathscr{X}} \\ &\le \frac{1}{4\lambda\,(1-\eta)} \int_{D_1} |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} \\ &\quad + \frac{1}{4\lambda\,(1-\eta)} \int_{\partial D_\varepsilon} \frac{\partial W}{\partial n}\, \frac{g^2}{W}\, d\mu^{\varepsilon}_{\lambda,\mathscr{X}} \,.\end{aligned}$$

Remark that $(1/W)\,|\frac{\partial W}{\partial n}|(x) \le 2/|x|$ so that we obtain

$$\begin{aligned}&\int_{D_2} g^2 \, d\mu_{\lambda,\mathscr{X}} \\ &\quad\le \frac{1}{4\lambda\,(1-\eta)} \left(\int_{D_1} |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} + \sum_{x_i \in \mathscr{X}_\infty} \frac{2}{(|x_i| - r_i - 2h_i)} \right. \\ &\quad\quad \left. \int_{|x-x_i| = r_i + \varepsilon h_i} g^2 \, d\mu^{\varepsilon}_{\lambda,\mathscr{X}} \right) .\end{aligned}$$

Integrating the previous inequality with respect to ε for $1 \le \varepsilon \le 2$ we obtain

Lemma 1.4 *With the notations of this subsection, let g be a smooth function compactly supported in D_∞, then*

$$\begin{aligned}&\int_{D_2} g^2 \, d\mu_{\lambda,\mathscr{X}} \\ &\quad\le \frac{1}{4\lambda\,(1-\eta)} \left(\int_{D_1} |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} + \sum_{x_i \in \mathscr{X}_\infty} \frac{2}{h_i\,(|x_i| - r_i - 2h_i)} \right. \\ &\quad\quad \left. \int_{r_i + h_i \le |x-x_i| \le r_i + 2\,h_i} g^2 \, d\mu_{\lambda,\mathscr{X}} \right) .\end{aligned}$$

1.3.4 Localizing Away from the Origin for the Far Enough Obstacles

Now we shall put together the previous two localization procedures.

Remark that, during the proof of Lemma 1.3 (more precisely with an immediate modification), we have shown the following: provided we can find a Lyapunov function in the neighborhood $|x-y| \leq r+3h$ of the obstacle $|x-y| \geq r$,

$$\int_{r<|x-y|<r+2h} f^2\, d\mu_{\lambda,\mathscr{X}} \leq \frac{1}{\theta} \int_{r<|x-y|<r+3h} |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{\theta\, h^2} \int_{r+2h<|x-y|<r+3h} f^2\, d\mu_{\lambda,\mathscr{X}}\,,$$

so that using the Lyapunov function W_y in Sect. 1.3.2 (yielding $\theta = 2\,\lambda$) we have, provided $d-1 \geq 2\lambda\,(r+3h)^2$,

$$\int_{r<|x-y|<r+2h} f^2\, d\mu_{\lambda,\mathscr{X}} \leq \frac{1}{2\lambda} \int_{r<|x-y|<r+3h} |\nabla f|^2\, d\mu_{\lambda,\mathscr{X}} + \frac{1}{2\lambda\, h^2} \int_{r+2h<|x-y|<r+3h} f^2\, d\mu_{\lambda,\mathscr{X}}\,. \tag{1.18}$$

Hence we have to assume that, at least for the far enough obstacles, $d-1 \geq 2\lambda\,(r_i+3h_i)^2$. At the same time, Lemma 1.4 shows that we have to choose h_i as large as possible. So in the sequel we choose

$$\lambda = 1 \quad , \quad b<1 \quad , \quad h_i = h = \frac{b}{3}\sqrt{(d-1)/2} \quad , \eta = \frac{1}{2}\,.$$

In order to fulfill the conditions in the previous subsection, we have to assume that for all far enough x_i, (i.e. all x_i such that $|x_i| > c\sqrt{d} + \sqrt{(d-1)/2}$ for some $c \geq 1$)

$$r_i \leq (1-b)\sqrt{(d-1)/2}\,.$$

We thus make the following assumption

Theorem 1.1 *Ordering the x_i's such that $|x_i| \leq |x_{i+1}|$ for all i, we assume that there exists some $0 \leq n < +\infty$ such that $r_i \leq (1-b)\sqrt{(d-1)/2}$ for some $b<1$ and all $i \geq n$. In addition we assume that for $i \geq n$ the balls $B(x_i, r_i+3h)$ are non overlapping.*

Consider now the smallest $c \geq \frac{1}{h^3\sqrt{d}}$ (this value will be explained below) such that the open ball $B_d = B(0, c\sqrt{d})$ contains all the $B(x_i, r_i+1)$ for $i<n$. B_d can contain or intersect only a finite number of balls $B(x_i, r_i+h)$ for $i \geq n$. If

such a ball is included in B_d there is nothing to do. If such a ball intersects B_d but is not contained in B_d we may smoothly deform the boundary of B_d in order to push $B(x_i, r_i + h)$ in the interior of the modified domain. We can do so for all balls intersecting the boundary and in addition in a such a way that all others $B(x_i, r_i+3h)$ are still in the exterior of the modified domain. The boundary of this deformation of B_d is denoted by S and it is easily seen that with this construction we are in the situation of the previous subsection.

From now on we use the notation D_0, D_∞ and D_ε introduced therein.

Now for a smooth function g with compact support included in D_∞, we denote

$$A = \int_{D_\infty - D_2} g^2 \, d\mu_{\lambda,\mathscr{X}} \, ,$$

$$B = \int_{D_2} g^2 \, d\mu_{\lambda,\mathscr{X}} \, ,$$

and

$$C = \int_{D_\infty} |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} \, .$$

According to (1.18) and to Lemma 1.4, we obtain (recall that $\lambda = 1$)

$$A \leq \frac{1}{2}\left(C + \frac{1}{h^2}\, B\right) \quad \text{and} \quad B \leq \frac{1}{2}\left(C + \frac{2}{h\, c\sqrt{d}}\, A\right) .$$

Hence,

$$A \, \leq \, \frac{1}{2}\left(1 + \frac{1}{2\, h^2}\right) C \, + \, \frac{1}{2\, h^3\, c\, \sqrt{d}}\, A \, ,$$

and thanks to our choice of c we get finally

$$A \, \leq \, \left(1 + \frac{1}{2\, h^2}\right) C \quad , \quad B \leq \left(1 + h^2\right) C \, .$$

This yields

Lemma 1.5 *Let* $0 < b < 1$ *and* $h = \frac{b}{3}\sqrt{(d-1)/2}$. *Assume that* $\lambda = 1$, *and Assumption 1.1 is satisfied. Then, for all smooth function* g, *compactly supported in* D_∞ *(which depends on* b*), it holds*

$$\int g^2 \, d\mu_{\lambda,\mathscr{X}} \, \leq \, K \int |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} \, ,$$

with

$$K = 2 + \frac{1}{2h^2} + h^2 .$$

1.3.5 Localizing Around the Origin for a Far Enough Single Obstacle

Assume that $N = 1$ and that the single obstacle is far enough, i.e. $n = 0$ in Assumption 1.1. Actually, in order to get explicit bounds, we shall take here for D_∞ the exterior of a large ball $B(0, c\sqrt{d})$ but assume that $B(y, r + 3h)$ lies in the complement of $B(0, c\sqrt{d} + 1)$.

To get some bound for the Poincaré constant, it remains now to follow the method in [5, 14]. Let f be a smooth function with compact support. Assume that we are in the situation of Lemma 1.5 (in particular $\lambda = 1$).

Recall that $\mu_{\lambda,\mathscr{X}}$ restricted to the ball $\{|x| \le c\sqrt{d} + 1\}$ is just the Gaussian measure restricted to the ball (since this ball does not intersect the obstacle), hence satisfies a Poincaré inequality with a constant less than $\frac{1}{2}$. If

$$m = \int_{|x|\le c\sqrt{d}+1} f \, d\mu_{\lambda,\mathscr{X}} / \mu_{\lambda,\mathscr{X}}(|x| \le c\sqrt{d} + 1) ,$$

we have

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) \le \int_D (f - m)^2 d\mu_{\lambda,\mathscr{X}}$$

so that it is enough to control the second moment of $\bar{f} = f - m$.

We write

$$\bar{f} = \chi \, \bar{f} + (1 - \chi) \, \bar{f} = \chi \bar{f} + g$$

where χ is 1-Lipschitz and such that $\mathbb{1}_{|x|\le c\sqrt{d}} \le \chi \le \mathbb{1}_{|x|\le c\sqrt{d}+1}$. g is thus compactly supported in $|x| \ge c\sqrt{d}$ so that we may apply what precedes. In particular

$$\begin{aligned} \int_D \bar{f}^2 \, d\mu_{\lambda,\mathscr{X}} &\le 2 \int_{|x|\le c\sqrt{d}+1} \bar{f}^2 \, d\mu_{\lambda,\mathscr{X}} + 2 \int_D g^2 \, d\mu_{\lambda,\mathscr{X}} \\ &\le \int_{|x|\le c\sqrt{d}+1} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} + 2K \int_D |\nabla g|^2 \, d\mu_{\lambda,\mathscr{X}} \end{aligned}$$

$$\leq \int_{|x|\leq c\sqrt{d}+1} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} + 4K \int_{x\in D, |x|\geq c\sqrt{d}} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}}$$
$$+4K \int_{c\sqrt{d}+1\geq|x|\geq c\sqrt{d}} \bar{f}^2 \, d\mu_{\lambda,\mathscr{X}}$$
$$\leq (1+2K) \int_{|x|\leq c\sqrt{d}+1} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} + 4K \int_{x\in D, |x|\geq c\sqrt{d}} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}}$$
$$\leq (1+6K) \int_D |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} \, .$$

We have thus proved, using (1.6)

Proposition 1.5 *Assume that* $N = 1$. *Let* $h = \frac{b}{3}\sqrt{(d-1)/2}$ *and* $c > 1/(h^3\sqrt{d})$. *Assume that for some* $0 < b < 1$, *we have* $r\sqrt{\lambda} \leq (1-b)\sqrt{(d-1)/2}$ *and that* $|y|\sqrt{\lambda} > 1 + c\sqrt{d} + \sqrt{(d-1)/2}$.

Then the measure $\mu_{\lambda,\mathscr{X}}$ *satisfies a Poincaré inequality* (1.5) *with*

$$C_P(\lambda, B(y,r)) \leq \frac{1}{\lambda}\,(1+6K)\ ,$$

with $K = 2 + h^2 + \frac{1}{2h^2}$.

The main interest of the previous proposition is that it shows that for a single far enough small obstacle the Poincaré constant does not depend on the location of the obstacle. We also have tried to trace a little bit the constants to show that we obtain some tractable explicit upper bound, the final step being to optimize in b (left to the reader).

1.3.6 A General Result for a Single Obstacle with Small Radius

We can gather together all the previous results in the case $N = 1$. For the sake of simplicity the next theorem is not optimal, but readable.

Theorem 1.2 *There exists some universal constant* κ *such that if*

$$r\sqrt{\lambda} \leq \frac{1}{2}\sqrt{(d-1)/2}\,,$$

the measure $\mu_{\lambda,\mathscr{X}}$ *where* $\mathscr{X} = \{y\}$ *is a singleton, satisfies a Poincaré inequality* (1.5) *with*

$$C_P(\lambda, B(y,r)) \leq \frac{\kappa}{\lambda}\,.$$

Proof If d is big enough ($d \geq 33$) we may use Proposition 1.4. If $d \leq 33$ and $|y|\sqrt{\lambda}$ large, we may apply Proposition 1.5 with $b = 1/(2\sqrt{d-1})$. Finally, if $d \leq 33$ and $|y|\sqrt{\lambda}$ is small we may use Proposition 1.2.

Remark 1.2 In comparison with Proposition 1.4, we have spent a rather formidable energy in order to cover the small dimension situation. But the alternate method we have developed for large $|y|$ will be useful in other contexts, in particular for an infinite number of obstacles.

It is also worth noticing that we have used Proposition 1.2 that cannot be extended to more than one obstacle.

1.3.7 The Case of Infinitely Many Obstacles

Now consider the case with more than obstacle. If we look at the localization procedure in Sect. 1.3.5 we see that a key point is to get the value (or a bound) for the Poincaré constant in a neighborhood of the origin. If all obstacle are far enough we can mimic what is done in Sect. 1.3.5. But in general, the n introduced in Assumption 1.1 is not equal to 0, so that we have to look at the Poincaré constant in D_0. Since this set is compact and with a smooth boundary, the finiteness of the Poincaré constant is ensured, for instance by the Down-Meyn-Tweedie theory as we indicate in the introduction.

Unfortunately it is very hard to get some explicit upper bound of this constant depending on all points x_i in $\mathscr{X}$ such that the obstacles $B(x_i, r_i)$ are subsets of D_0. Exactly the same problem occurs in [16] where the value of the Poincaré constant (or the spectral gap) for the parameter ε (using the notations therein) is shown to be quadratic in ε, but with an unknown constant pre-factor.

We can nevertheless mimic what we did in Sect. 1.3.5 replacing the value $1/2$ by the unknown Poincaré constant in D_0. This yields

Theorem 1.3 *For any $1 \leq N \leq +\infty$ (in particular $N = +\infty$), under Assumption 1.1, $\mu_{\lambda,\mathscr{X}}$ satisfies a Poincaré inequality with constant $C_P(\lambda, \mathscr{X}) = \frac{\kappa}{\lambda} < +\infty$ where κ depends on n, d and the structure of the (finite) number of the obstacles that are close to the origin.*

More precisely, with the notations of Proposition 1.5, $\kappa \leq 4K + (2+4K)C_P(n)$ where $C_P(n)$ denotes the Poincaré constant in D_0.

Corollary 1.1 *Ordering the x_i's such that $|x_i| \leq |x_{i+1}|$ for all i, assume that there exists some $0 \leq n < +\infty$ such that $r_i \leq (1-b)\sqrt{(d-1)/2}$ for some $b < 1$ and all $i \geq n$, and that in addition there exists $\varepsilon > 0$ such that for all pair $i \neq j$, $dist(B(x_i, r_i), B(x_j, r_j)) \geq \varepsilon$.*

Then $C_P(\lambda, \mathscr{X}) < +\infty$.

Proof Take $b' = \left(\varepsilon/2\sqrt{(d-1)/2}\right) \wedge b$. The condition on the radii r_i is still satisfied replacing b by b' while $h = \frac{b'}{3}\sqrt{(d-1)/2}$ satisfies $6h \leq \varepsilon$. Hence the balls with radii $r_i + 3h$ are non overlapping and we may apply the previous Theorem.

1.3.8 Others Obstacles Like Hypercubes

Replacing Euclidean balls by others geometries of obstacles requires first to find a Lyapunov function in the neighborhood of each obstacle as in Sect. 1.3.2. We will not discuss this in details here, but only consider the case where we replace the Euclidean ball $B(x_i, r_i)$ by some hypercube, in a nice position.

Namely we consider the x's such that $x = x_i + (z\frac{x_i}{|x_i|} + y_i)$ where y_i belongs to the hyperplane orthogonal to x_i intersected with the $d-1$ l^∞ ball of radius r_i and $z \in [-r_i, r_i]$. In other words we consider hypercubes in d dimensions such that, first the line connecting the origin to the center of mass x_i of the hypercube is orthogonal to some face of the latter, second the hypercube is included in the Euclidean ball $B(x_i, r_i\sqrt{d})$.

In this situation the function W_{x_i} introduced in Sect. 1.3.2 (replacing y by x_i) is still a Lyapunov function with a non-positive normal derivative on the boundary of the hypercube. The reader who is afraid by the singularities of the boundary can "smooth the corners".

The results in Sects. 1.3.6 and 1.3.7 easily extend, but this time with $r_i \leq b$ for some constant b independent of the dimension. Of course we have to assume that all the obstacles are in the nice position described above.

1.4 General Spherical Obstacles Using Stochastic Calculus

As we have seen, provided we are able to find a good Lyapunov function near the obstacles, we are able to control (even if not explicitly) the Poincaré constant in D. The choice we made in the previous section implies a limitation for the radius of the obstacles. What we shall do now is to find a new Lyapunov function near the obstacles. This Lyapunov function will be built by trying to understand how fast the process goes around the obstacles.

Indeed recall the following results on the exponential moments of hitting times (see e.g.[14]).

Proposition 1.6 *Let U be a bounded connected subset with smooth boundary of D and T_U denotes the hitting time of U.*

- *Assume that for some $\theta > 0$ and all $x \in D$, $\mathbb{E}_x\left(e^{\theta\, T_U}\right) < +\infty$. Define $W(x) = \mathbb{E}_x\left(e^{\theta\, T_U}\right)$. Then W belongs to the domain of the generator L of the reflected*

Ornstein-Uhlenbeck process (in particular $\partial W/\partial n = 0$ on ∂D), and satisfies $LW \leq -\theta\, W$ outside of U.

- *For all $x \in D$,*

$$\mathbb{E}_x\left(e^{\theta\, T_U}\right) < +\infty \quad \textit{for all } \theta < \theta(U), \textit{ with } \theta(U) = \frac{\mu_{\lambda,\mathscr{X}}(U)}{16\, C_P(\lambda, \mathscr{X})}\,.$$

Actually, [14] only dealt with diffusion processes, without reflection. But the proof of this Proposition lies on three facts which are still true here: the symmetry of $\mu_{\lambda,\mathscr{X}}$, the existence of a density for the law at time $t > 0$ of the process starting at any x, the results of Proposition 1.4 and Remark 1.6 in [12] which hold true for general Markov processes with a square gradient operator.

Hence provided we can control exponential moments of hitting times, we can build (non explicit) Lyapunov functions.

The discussion below is done for a single obstacle $B(y,r)$. We shall conclude at the end of the section for more than one obstacle.

1.4.1 The Rate of Rotation

To understand how fast the process goes around the obstacle, we introduce a new stochastic process Y_t which is just the reflected Ornstein-Uhlenbeck process in the shell $S = \{r \leq |x-y| \leq r+q\}$ for some positive q, i.e.

$$\begin{cases} dY_t = dW_t \,-\, \lambda\, Y_t\, dt \,+\, (Y_t - y)\, dL_t\,, \\ L_t = \int_0^t \left(\mathbb{1}_{|Y_s-y|=r} - \mathbb{1}_{|Y_s-y|=r+q}\right) dL_s\,. \end{cases} \tag{1.19}$$

Next as usual, we assume that $y = (a,0)$ and write the generic point of the Euclidean space as $x = (x^1, \bar{x})$. Again n denotes the normal vector field $(x^1 - a, \bar{x})$ (pointing either inward or outward), so that, for any nice function g, Ito formula yields

$$g(Y_t) = g(Y_0) + \int_0^t \nabla g(Y_s).dW_s + \int_0^t Lg(Y_s)ds + r\int_0^t \frac{\partial g}{\partial n}(Y_s)\, dL_s\,.$$

Finally we shall look at the process

$$Z_t = \arccos\left(\frac{Y_t^1 - a}{\sqrt{|\bar{Y}_t|^2 + (Y_t^1 - a)^2}}\right) = \varphi(Y_t)\,. \tag{1.20}$$

We can calculate

$$\nabla\varphi(x) = \left(\frac{-|\bar{x}|}{(x^1-a)^2+|\bar{x}|^2}\,,\ \frac{(x^1-a)\,\bar{x}}{|\bar{x}|\,\big((x^1-a)^2+|\bar{x}|^2\big)}\right) \quad \text{so that} \quad \frac{\partial\varphi}{\partial n}(x) = 0\,.$$

Consider

$$M = \{-r-q \le x^1 - a \le -r\,,\ \bar{x} = 0\}\,.$$

If $Y_0 \notin M$, i.e. $Z_0 \neq \pi$, we may apply Ito-Tanaka formula up to time T_M (the first time $Y_.$ hits M) yielding for $t < T_M$,

$$\begin{aligned} Z_t^2 &= Z_0^2 + \int_0^t 2Z_s\,\langle\nabla\varphi(Y_s), dW_s\rangle + \int_0^t |\nabla\varphi(Y_s)|^2\,ds \qquad (1.21)\\ &\quad + \int_0^t \frac{Z_s(2\lambda\,a\,|\bar{Y}_s| + (d-2)\,(Y_s^1-a))}{|\bar{Y}_s|^2+(Y_s^1-a)^2}\,ds\\ &= Z_0^2 + \int_0^t \frac{2Z_s}{\big(|\bar{Y}_s|^2+(Y_s^1-a)^2\big)^{1/2}}\,dB_s\\ &\quad + \int_0^t \frac{1+Z_s(2\lambda\,a\,|\bar{Y}_s| + (d-2)\,(Y_s^1-a))}{|\bar{Y}_s|^2+(Y_s^1-a)^2}\,ds \end{aligned}$$

where $B_.$ is a new standard Brownian motion. We have considered Z^2 instead of Z to kill the local time at 0 of $Z_.$ (since $t < T_M$ the local time of $Z_.$ at π does not appear too).

Introduce the subset

$$K = \{x^1 - a < 0\,,\ |\bar{x}| \le \eta < r\} \cap S\,.$$

Since $M \subset K$ we know that $T_K \le T_M$ so that (1.21) holds for $t \le T_K$. We want to estimate T_K by comparing Z_t with a simpler diffusion process for which estimates are easy to obtain (since they are known).

Set

$$A(t) = \int_0^t \frac{1}{\big(|\bar{Y}_s|^2+(Y_s^1-a)^2\big)}\,ds\,,$$

and $A^{-1}(t)$ the inverse of $A(.)$. Notice that $(t/(r+q)^2) \le A(t) \le (t/r^2)$ so that $r^2 t \le A^{-1}(t) \le (r+q)^2 t$.

Define the time changed process $\tilde{Y}_t = Y_{A^{-1}(t)} = (\tilde{Y}_t^1, \tilde{Y}_t^2)$ and $U_t = Z^2_{A^{-1}(t)}$. Then for $t < A(T_M)$, $U_.$ satisfies

$$U_t = Z_0^2 + \int_0^t 2\sqrt{U_s}\, d\tilde{B}_s + \int_0^t \left(1 + \sqrt{U_s}\,(2\lambda\, a|\tilde{Y}_s^2| + (d-2)\,(\tilde{Y}_s^1 - a))\right) ds\,, \tag{1.22}$$

for some new Brownian motion $\tilde{B}_.$. In order to compare $U_.$ with some CIR process (see Appendix 2) we have to bound the drift term from below.

Remark that for a point $\tilde{y} \in K^c$,

$$|\tilde{y}^2| = \sqrt{(|\tilde{y}^2|)^2 + (\tilde{y}^1 - a)^2}\ \sin(\sqrt{u}) \geq r\ \sin(\sqrt{u}) \geq \frac{\eta\, r}{\pi(r+q)}\sqrt{u}\,.$$

Hence looking separately at the case $\tilde{y}^1 - a > 0$ and $\tilde{y}^1 - a \leq 0$ it follows that the drift term satisfies

$$1 + \sqrt{U_s}\,(2\lambda\, a|\tilde{Y}_s^2| + (d-2)\,(\tilde{Y}_s^1 - a)) \ \geq\ 1 + \left(\frac{2\lambda\, a\,\eta\, r}{\pi(r+q)} - \frac{(d-2)(r+q)(2-\sqrt{\pi})}{\pi}\right) U_s\,.$$

Hence up to time T_K, using standard comparison results for one dimensional diffusions, we know that $U_t \geq V_t$ where

$$dV_t = 2\sqrt{V_t}\, d\tilde{B}_t + (1 + 2\beta V_t)dt$$

i.e. $V_.$ is a generalized squared radial Ornstein-Uhlenbeck process, with

$$2\beta = \frac{2\lambda\, a\,\eta\, r}{\pi(r+q)} - \frac{(d-2)(r+q)(2-\sqrt{\pi})}{\pi} \tag{1.23}$$

provided $\beta \geq 0$.

It follows that $A(T_K)$ is smaller than the first hitting time of π by $V_.$. According to (1.48) with $\delta = 1$, we thus have

$$\mathbb{E}_x(e^{\theta T_K}) < +\infty \quad \text{for all } x \in S \text{ provided} \quad \theta < \frac{\beta}{(r+q)^2}\,.$$

It is thus tempting to define $W(x) = \mathbb{E}_x\left(e^{\theta\, T_K}\right)$, which satisfies $LW = -\theta W$ in $S - K$. This is not yet enough but will be useful (Fig. 1.2).

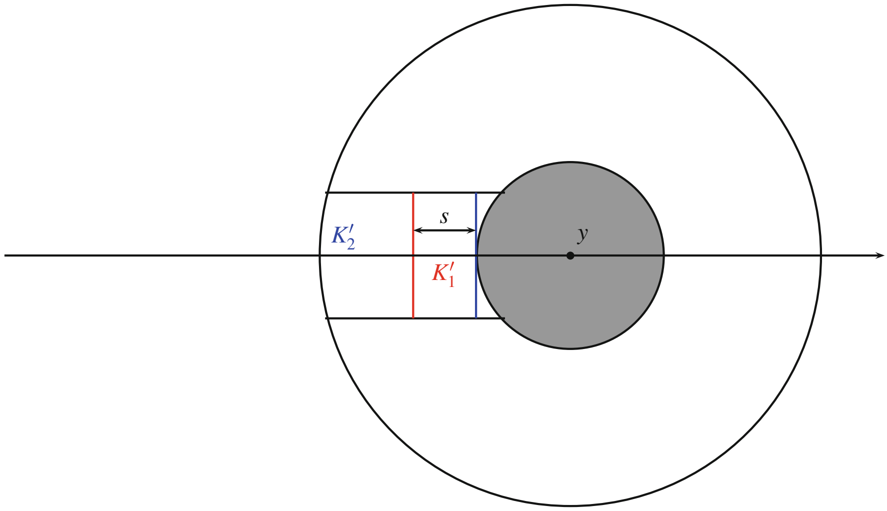


Fig. 1.2 Rotation around the obstacle

1.4.2 The Poincaré Inequality in the Shell S

Using what precedes we shall prove the following first result.

Proposition 1.7 *Let η, s, q be positive numbers such that $\eta + s < r$, $s < q$ and β given by (1.23) is positive. Assume that $a > r + s + \frac{1}{\sqrt{\lambda}}$.*

Then, the (non normalized) restriction of $\mu_{\lambda,r}$ to the shell $S = \{r \leq |x - y| \leq r + q\}$ satisfies a Poincaré inequality

$$\int_S f^2 \, d\mu_{\lambda,r} \leq C_P(\lambda, S) \int_S |\nabla f|^2 \, d\mu_{\lambda,r} + \frac{1}{\mu_{\lambda,r}(S)} \left(\int_S f \, d\mu_{\lambda,r} \right)^2$$

where

$$C_P(\lambda, S) \leq \frac{2(r+q)^2}{\beta} + \frac{2}{\lambda} \left(1 + \frac{(r+q)^2}{\beta \, s^2} \right) \left(\frac{2}{\lambda s^2} + \frac{5}{2} \right).$$

Proof We shall use the results in the previous subsection. Define $W(x) = \mathbb{E}_x \left(e^{\theta T_K} \right)$ for $x \in S$. Then W belongs to the domain of the generator of $Y_.$ (in particular the normal derivative on the shell's boundary vanishes) and satisfies $LW = -\theta W$ in $S - K$.

Consider now

$$K' = \left\{ x^1 - a < 0 \, , \ |\bar{x}| \leq \eta + s < r \right\} \cap S \, .$$

Then as before, using [14] formula (2.14) (in the present framework of our reflected Ornstein-Uhlenbeck process $Y_.$), we have

$$C_P(\lambda, S) \le \frac{2}{\theta} + \left(\frac{2}{\theta\, s^2} + 2\right) C_P(K'). \tag{1.24}$$

It remains to get some bound for $C_P(K')$.

Again we divide K' in two overlapping parts:

$$K'_1 = \left\{-r - q < -r - s < x^1 - a < 0\,,\ |\bar{x}| \le \eta + s < r\right\} \cap S$$

and

$$K'_2 = \left\{x^1 - a < -r\,,\ |\bar{x}| \le \eta + s < r\right\} \cap S\,.$$

Note that K'_2 is convex. Hence the restriction of the Gaussian measure to K'_2 satisfies a Poincaré inequality with constant $1/2\lambda$.

As before it is then sufficient to build some Lyapunov function in K'_1. This time we choose $W(x) = (x^1)^2$. Note that, on one hand, the normal derivative of W on $|\bar{x}| = \eta + s$ is equal to 0, while on the other hand, the (non normalized) inward normal derivative of W on $|x - y| = r$ is equal to $2(x^1 - a)x^1$. The latter is thus negative provided $x^1 > 0$, hence in particular if $a > r + s$.

In addition,

$$LW(x) = 1 - 2\lambda(x^1)^2 \le -\lambda(x^1)^2 \quad \text{in } K'_1 \tag{1.25}$$

as soon as $a > r + s + (1/\sqrt{\lambda})$. Thus, as before we obtain

$$C_P(K') \le \frac{1}{\lambda}\left(\frac{2}{\lambda s^2} + \frac{5}{2}\right).$$

Now let $a' > r' + s' + 1$, $y' = (a', 0)$, $\eta' + s' < r'$. Define $(a, r, s, q, \eta) = \frac{1}{\sqrt{\lambda}}(a', r', c', q', \eta')$ so that $a > r + s + \frac{1}{\sqrt{\lambda}}$. Define $S' = \{r' \le |x - y'| \le r' + q'\}$. The homogeneity property (1.6) is still available in our situation yielding

$$C_P(1, S') = \lambda C_P(\lambda, S)$$

for all λ. β given by (1.23) can thus be written

$$2\beta = \frac{2\, a'\, \eta'\, r'}{\pi(r' + q')} - \frac{(d-2)(r' + q')(2 - \sqrt{\pi})}{\pi\, \sqrt{\lambda}}\,,$$

and goes to $\frac{a'\,\eta'\,r'}{\pi(r'+q')}$ as λ goes to infinity. Hence, letting λ go to infinity in Proposition 1.7 we get

$$C_P(1,S') \leq \frac{2\pi(r'+q')^3}{a'\eta'\,r'} + 2\left(1+\frac{(r'+q')^3}{(r')^2\,a'\,\eta'\,(s')^2}\right)\left(\frac{2}{(s')^2}+\frac{5}{2}\right). \tag{1.26}$$

We have thus obtained

Proposition 1.8 *Let $s < q$, and assume that $a > r+s+1$. Let $0 < \eta < r-s$.*
Then, the (non normalized) restriction of $\mu_{1,r}$ to the shell $S = \{r \leq |x-y| \leq r+q\}$ satisfies a Poincaré inequality

$$\int_S f^2\,d\mu_{1,r} \leq C_P(1,S)\int_S |\nabla f|^2\,d\mu_{1,r} + \frac{1}{\mu_{1,r}(S)}\left(\int_S f\,d\mu_{1,r}\right)^2$$

where $C_P(1,S)$ is given by 1.26 (without ′).

1.4.3 A New Estimate for an Obstacle Which Is Not too Close to the Origin

We may use Proposition 1.8 to build a new Lyapunov function near the obstacle when $\lambda = 1$.

In the situation of the proposition consider $T_{S/2}$ the hitting time of the "half" shell $S' = \{r+(q/2) \leq |x-y| \leq r+q\}$. Then according to Proposition 1.6 we may define $W(x) = \mathbb{E}_x\left(e^{\theta T_{S/2}}\right)$ which satisfies $LW = -\theta W$ for $x \in S - S'$ and $\partial W/\partial n = 0$ on $|x-y| = r$, provided

$$\theta < \frac{1}{8C_P(1,S)}\,\frac{\mu_{1,r}(S')}{\mu_{1,r}(S)}\,. \tag{1.27}$$

Now we can first apply Lemma 1.3 with $2h = q/2$, provided $\theta h^4 > 1$.

It remains to choose all parameters. All conditions are satisfied for instance if

$$\frac{q^4}{4^4}\,\frac{1}{16C_P(1,S)}\,\frac{\mu_{1,r}(S')}{\mu_{1,r}(S)} > 1\,. \tag{1.28}$$

It is not too difficult to be convinced that the ratio of the two measures is uniformly (in r and y) bounded from below, provided $a-r-q > 1$ (1 can be replaced by any positive constant), i.e. provided the origin is far enough from $B(y, r+q)$. Indeed the measure restricted to S is mainly concentrated near the point $(a-r-q, 0)$ which belongs to both S and S'.

Now look at the bound in Proposition 1.8. If r is small (goes to 0), the bound for a given a becomes very bad. Indeed, for $2/s^2$ to be nice, we have to choose s bounded from below, so that q is bounded from below too and since $\eta < r$ the term governed by $1/a\eta$ explodes.

Hence we may choose $r > (1-b)\sqrt{(d-1)/2}$ in order to cover the case which is not covered by Theorem 1.2, or simply $r > \frac{1}{2}$.

Now, we have clearly to choose q as small as possible, but satisfying (1.28). To simplify choose $s = 1/4$ so that $C_P(1, S) \le c$ where c is less than $C(1 + q^3/a)$ for some universal constant C. We see that for (1.28) to be satisfied we only need q to be greater than an universal constant. Hence

Proposition 1.9 *One can find universal constants $m > 0$ and C such that, provided $|y| > r + m$ and $r > \frac{1}{2}$, $C_P(1, B(y, r)) \le C$.*

1.4.4 Finiteness of the Poincaré Constant for an Infinite Number of Spherical Obstacles

Of course we can use the previous construction of a Lyapunov function near the far enough obstacles together with the ideas of Sect. 1.3.7 to cover the case of infinitely many obstacles. To this end, instead of using Lemma 1.3 we should also follow what we have done in Sects. 1.3.3 and 1.3.4, i.e. replace 2λ (= 2 here) by θ defined above in (1.18). But we have to be accurate when using Lyapunov functions near the obstacles, that the enlargements we are using are non overlapping. In particular q and s have to be smaller than the half of the distance between obstacles.

Theorem 1.4 *Let $\mathscr{X} = (x_i)_{1\le i<+\infty}$ a locally finite collection of distinct points, ordered such that $|x_i| \le |x_{i+1}|$ for all i, and $\mathscr{R} = (r_i)_{1\le i<+\infty}$ a collection of non-negative numbers. Assume that there exists $\varepsilon > 0$ with $|x_i - x_j| > r_i + r_j + \varepsilon$ for all $i \neq j$. Define $D = \mathbb{R}^d - \cup_i B(x_i, r_i)$ (for $d \ge 2$) where $B(y, r)$ denotes the Euclidean ball with center y and radius r.*

Then for any $\lambda > 0$, the Gaussian measure $\mu_{\lambda,\mathscr{X}}$ has a finite Poincaré constant and the reflected Ornstein-Uhlenbeck process in D is exponentially ergodic.

Proof Since the conditions are still satisfied when dilating the space we may assume that $\lambda = 1$.

As for the proof of Corollary 1.1 we shall use the Lyapunov functions near the obstacles outside some large enough smooth subset containing the origin to be determined during the proof.

For small obstacles ($r_i \le \frac{1}{2}\sqrt{d-1}$ for instance) we use the Lyapunov function in Sect. 1.3.4. For the large obstacles we use the one introduced in the previous subsection. With the notations of Sect. 1.3.4, and still with $h_i = h$, we obtain

$$A \le \frac{1}{\theta}\left(C + \frac{1}{h^2}B\right) \quad , \quad B \le \frac{1}{2}\left(C + \frac{2}{\alpha\, h}A\right),$$

where $\alpha = \min\{i \text{ large}\,;\; |x_i| - r_i - 2h\}$.

We have to choose h, q, s of order ε (up to well chosen constants), so that for A and B to be controlled by C it is enough that $h^3\,\alpha \geq c(1 + (1/\varepsilon^2))$ for a large enough c.

But it is not difficult to see that $|x_i| - r_i \to +\infty$ as $i \to +\infty$, so that there exists a large enough constant $c > 0$ such that $|x_i| - r_i \geq c(1 + (1/\varepsilon^5))$ and we can conclude.

1.5 Lower Bounds for Non Spherical Obstacles

We obtained in the previous section that for far enough obstacles, the radii of the obstacles do not really increase the value of the Poincaré constant. Hence, roughly speaking, the only radius that really matters is the one of the obstacle containing the origin if such an obstacle exits (of course we did not prove the result in this so general form). But actually this property is strongly linked to the geometry of the obstacle, and we shall see below that replacing spherical obstacles by other geometries will drastically modify the result.

1.5.1 Lower Bounds for Hypercubes via Stochastic Calculus

Replace the ball $\{|x-y| < r\}$ with $y = (a, 0)$ $(a \geq 0)$ by an hypercube, $H_r = \{|x^1 - a| < r\ ,\ |x^j| < r$ for $j \geq 2\}$. As we already said, we may "smooth the corners" for the boundary to be smooth (replacing r by $r + \varepsilon$), so that existence, uniqueness and properties of the reflected process are similar to those we have mentioned for the disc.

Consider the process X_t starting from $x = (a + r, 0)$. Denote by $S(r) = \min_{j\geq 2} S^j(r)$, where $S^j(r)$ is the exit time of $[-r, r]$ by the coordinate $X^j_.$. Up to time $S(r)$, the $X^j_.$'s $(j \geq 2)$ are Ornstein-Uhlenbeck processes, starting at 0, $X^1_.$ is an Ornstein-Uhlenbeck process reflected on $a + r$, starting at $a + r$; and all are independent. Of course $S(r) = T_{U^c(r)}$ where the set $U(r) = \{x^1 \geq a + r\ ;\ \max_{j\geq 2} |x^j| \leq r\}$.

According to Proposition 1.6, if

$$\mathbb{E}_{(a+r,0)}\left(e^{\theta\, S(r)}\right) = +\infty \quad \text{then} \quad C_P(\lambda, \mathscr{X}) \geq \frac{\mu_{\lambda,\mathscr{X}}(U_r^c)}{16\,\theta}\,. \tag{1.29}$$

But according to (1.47) and to the independence of the coordinates of the process, this holds as soon as $\theta > \frac{(d-1)\lambda}{e^{\lambda r^2}-1}$. In particular since $\mu_{\lambda,\mathscr{X}}(U^c(r)) \geq \frac{1}{2}$, we always have

Theorem 1.5 *Let $D = \mathbb{R}^d - H_r$ where H_r is the hypercube described above. Then there exists an universal constant C such that the Poincaré constant in D satisfies $C_P(\lambda, r) \geq \frac{C\, e^{(\lambda r^2)}}{d\lambda}$.*

Recall that we have shown that for small enough obstacles (r of order a dimension free constant) the Poincaré constant is bounded from above by some κ/λ.

What is interesting here is that the lower bound does not depend on the location of y. In particular consider the situation of Theorem 1.4 with an infinite number of hypercubes as obstacles, in the position described in Sect. 1.3.8, i.e. the line joining the origin to the center of mass of each hypercube is orthogonal to some face of the latter. Of course for far enough obstacles the measure of $U_i^c(r_i)$ will still be larger than one half. So *if we allow the existence of a sequence of radii going to infinity the process is no more exponentially ergodic.*

1.5.2 An Isoperimetric Approach for Hypercubes

In this subsection, we present another approach for getting lower bounds. The easiest way to build functions allowing to see the lower bounds we have obtained in the previous subsection, is first to look at indicator of sets, hence isoperimetric bounds.

We define the Cheeger constant $C_C(\lambda, y, r)$ as the smallest constant such that for all subset $A \subset D$ with $\mu_{\lambda,\mathscr{X}}(A) \leq \frac{1}{2}$,

$$C_C(\lambda, y, r)\, \mu^S_{\lambda,\mathscr{X}}(\partial A) \geq \mu_{\lambda,\mathscr{X}}(A)\,. \tag{1.30}$$

Recall that $\mu^S_{\lambda,\mathscr{X}}(\partial A)$ denotes the surface measure of the boundary of A in D defined as

$$\liminf_{h\to 0} \frac{1}{h}\, \mu_{\lambda,\mathscr{X}}(A_h/A)$$

where A_h denotes the Euclidean enlargement of A of size h. The important fact here is that A is considered as a subset of D. In particular, if we denote by ∂S_r the boundary of the square S_r in the plane $\mathbb{R}^2$, $A \cap \partial S_r \subset D$ and so does not belong to the boundary of A in D.

The Cheeger constant is related to the $\mathbb{L}^1$ Poincaré inequality, and it is well known that

$$C_P \leq 4\, C_C^2\,, \tag{1.31}$$

while C_P can be finite but C_C infinite. Hence an upper bound for the Cheeger constant will provide us with an upper bound for the Poincaré constant while a lower bound can only be a hint.

1.5.2.1 Squared Obstacle

For simplicity we shall first assume that $d = 2$, and use the notation in Sect. 1.5.1. Consider for $a > 0$, the subset $A = \{x^1 \geq a + r\,,\ |x^2| \leq r\}$ with boundary $\partial A = \{x^1 \geq a + r\,,\ |x^2| = r\}$.

Recall the basic inequalities, for $0 < b < c \leq +\infty$,

$$\frac{b^2}{1+2b^2}\left(\frac{e^{-b^2}}{b} - \frac{e^{-c^2}}{c}\right) \leq \int_b^c e^{-u^2}\,du \leq \frac{1}{2b}\left(e^{-b^2} - e^{-c^2}\right). \tag{1.32}$$

It follows, for $r\sqrt{\lambda}$ large enough (say larger than one)

$$\begin{aligned}\frac{\mu_{\lambda,\mathscr{X}}(A)}{\mu^S_{\lambda,\mathscr{X}}(\partial A)} &= \frac{\left(\int_{a+r}^{+\infty} e^{-\lambda z^2}\,dz\right)\left(\int_{-r}^{+r} e^{-\lambda u^2}\,du\right)}{2\,e^{-\lambda r^2}\left(\int_{a+r}^{+\infty} e^{-\lambda z^2}\,dz\right)} \\ &\geq \frac{1}{2\sqrt{\lambda}}\,e^{\lambda r^2}\left(1 - \frac{1}{r\sqrt{\lambda}}\,e^{-\lambda r^2}\right),\end{aligned}$$

so that

$$C_C(\lambda, y, r) \geq \frac{1}{2\sqrt{\lambda}}\,e^{\lambda r^2}\left(1 - \frac{1}{r\sqrt{\lambda}}\,e^{-\lambda r^2}\right). \tag{1.33}$$

Note that this lower bound is larger than the one obtained by combining Cheeger's inequality (1.31) and the lower bound for the Poincaré constant obtained in Theorem 1.5, since this combination furnishes an explosion like $e^{\lambda r^2/2}$.

We strongly suspect, though we did not find a rigorous proof, that this set is "almost" the isoperimetric set, in other words that, up to some universal constant, the previous lower bound is also an upper bound for the Cheeger constant. In particular, we believe that this upper bound (hence the upper bound for the Poincaré constant) does not depend on a. Of course, since we know that the isoperimetric constant of the Gaussian measure behaves like $1/\sqrt{\lambda}$, isoperimetric sets for the restriction of the Gaussian measure to D have some (usual) boundary part included in the boundary of the obstacle and our guess reduces to the following statement: if r is large enough, for any subset $B \subset D$ with given Gaussian measure, the standard Gaussian measure of the part of the usual boundary of B that does not intersect ∂D is greater or equal to $C\,e^{-r^2}$ times the measure of the boundary intersecting ∂D.

1.5.2.2 Hypercubes

Of course, what we have just done immediately extends to d dimensions, defining A as $A = \{x^1 \geq a + r\ ,\ |x^i| \leq r \text{ for all } 2 \leq i \leq d\}$ and furnishes exactly the same bound as (1.33) replacing 2 by $2(d-1)$, i.e. in dimension d

$$C_C(\lambda, y, r) \geq \frac{1}{2(d-1)\sqrt{\lambda}}\, e^{\lambda r^2} \left(1 - \frac{1}{r\sqrt{\lambda}}\, e^{-\lambda r^2}\right) . \tag{1.34}$$

In order to get a lower bound for the Poincaré constant, inspired by what precedes, we shall proceed as follows: denote by $A(u)$ the set

$$A(u) = \{x^1 \geq a + r\ ,\ |x^i| \leq u \text{ for all } 2 \leq i \leq d\}\,,$$

and for $r > 1$, choose a Lipschitz function f such that $\mathbb{1}_{A(r-1)} \leq f \leq \mathbb{1}_{A(r)}$, for instance $f(x) = (1 - d(x, A(r-1)))_+$.

If Z_λ denotes the (inverse normalizing) constant in front of the exponential density of the Gaussian kernel (notice that Z_λ goes to 0 as λ goes to infinity), it holds

$$\begin{aligned}\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) &\geq \mu_{\lambda,\mathscr{X}}(A(r-1)) - \big(\mu_{\lambda,\mathscr{X}}(A(r))\big)^2 \\ &\geq Z_\lambda \int_{a+r}^{+\infty} e^{-\lambda u^2} du \left(\left(\int_{-r+1}^{r-1} e^{-\lambda u^2} du\right)^{d-1}\right. \\ &\quad \left. - Z_\lambda \left(\int_{-r}^{r} e^{-\lambda u^2} du\right)^{2(d-1)} \int_{a+r}^{+\infty} e^{-\lambda u^2} du\right),\end{aligned}$$

so that, there exists some universal constant c such that, as soon as $r\sqrt{\lambda} > c$,

$$\mathrm{Var}_\mu(f) \geq \frac{Z_\lambda}{2} \int_{a+r}^{+\infty} e^{-\lambda u^2} du \left(\int_{-r+1}^{r-1} e^{-\lambda u^2} du\right)^{d-1} .$$

At the same time again if $r\sqrt{\lambda} > c$,

$$\begin{aligned}\int |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} &\leq \int \left(\mathbb{1}_{A(r)} - \mathbb{1}_{A(r-1)}\right) d\mu_{\lambda,\mathscr{X}} \\ &\leq Z_\lambda \left(\int_{a+r}^{+\infty} e^{-\lambda u^2} du\right) \frac{e^{-\lambda (r-1)^2}}{(r-1)\lambda}\, (d-1) \left(\int_{-r}^{r} e^{-\lambda u^2} du\right)^{d-2} .\end{aligned}$$

It follows, using homogeneity again, that

$$C_P(\lambda, y, r) \geq \frac{1}{2}\left(\frac{r\sqrt{\lambda}-1}{(d-1)\lambda}\right) e^{(r\sqrt{\lambda}-1)^2} \frac{\left(\int_{-r\sqrt{\lambda}+1}^{r\sqrt{\lambda}-1} e^{-u^2} du\right)^{d-1}}{\left(\int_{-r\sqrt{\lambda}}^{r\sqrt{\lambda}} e^{-u^2} du\right)^{d-2}}$$

$$\geq \left(\frac{r\sqrt{\lambda}-1}{(d-1)\lambda}\right) e^{(r\sqrt{\lambda}-1)^2} \frac{1}{4\sqrt{\pi}} \left(1 - \frac{e^{-(r\sqrt{\lambda}-1)^2}}{r\sqrt{\lambda}-1}\right)^{d-2} . \quad (1.35)$$

Notice that this lower bound is smaller (hence worse) than the one we obtained in Theorem 1.5, and also contain an extra dimension dependent term (the last one). But of course it is much easier to get.

Since 1 is arbitrary, we may replace $r\sqrt{\lambda}-1$ by $r\sqrt{\lambda}-\varepsilon$ for any $0 \leq \varepsilon \leq 1$, the price to pay being some extra factor ε^2 in front of the lower bound for the Poincaré constant.

1.5.3 Back to Spherical Obstacles: Another Lower Bound

It is tempting to develop the same approach in the case of spherical obstacles. First assume $\lambda = 1$.

Introduce for $0 \leq u \leq r$,

$$A(u) = \{x^1 \geq a\,,\ |\bar{x}| \leq u\} \cap D\,.$$

As before we consider, for $\varepsilon < u$, a function $\mathbb{1}_{A(u-\varepsilon)} \leq f \leq \mathbb{1}_{A(u)}$ which is $1/\varepsilon$-Lipschitz. Then

$$\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) \geq \mu_{\lambda,\mathscr{X}}(A(u-\varepsilon)) - \left(\mu_{\lambda,\mathscr{X}}(A(u))\right)^2$$

and

$$\int |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} \leq (1/\varepsilon^2)\,\left(\mu_{\lambda,\mathscr{X}}(A(u)) - \mu_{\lambda,\mathscr{X}}(A(u-\varepsilon))\right)\,,$$

with

$$\mu_{\lambda,\mathscr{X}}(A(u)) = Z_\lambda\,\sigma(S^{d-2}) \int_0^u \left(\int_{a+\sqrt{r^2-s^2}}^{+\infty} e^{-t^2}\,dt\right) s^{d-2}\, e^{-s^2}\, ds\,,$$

and $\sigma(S^{d-2})$ is the Lebesgue measure of the unit sphere. It follows

$$\begin{aligned}(Z_\lambda)^{-1}\int|\nabla f|^2\,d\mu_{\lambda,\mathscr{X}} &\le (\sigma(S^{d-2})/\varepsilon^2)\int_{u-\varepsilon}^{u}\left(\int_{a+\sqrt{r^2-s^2}}^{+\infty} e^{-t^2}\,dt\right)s^{d-2}\,e^{-s^2}\,ds\\ &\le \frac{\sigma(S^{d-2})}{2\varepsilon^2}\int_{u-\varepsilon}^{u}\frac{s^{d-2}}{(a+\sqrt{r^2-s^2})}\,e^{-(a+\sqrt{r^2-s^2})^2}\,e^{-s^2}\,ds\\ &\le \frac{\sigma(S^{d-2})\,u^{d-2}\,e^{-(a^2+r^2)}}{2\varepsilon^2\,(a+\sqrt{r^2-u^2})}\int_{u-\varepsilon}^{u} e^{-2a\,\sqrt{r^2-s^2}}\,ds\,.\end{aligned}$$

To get a precise upper bound for the final integral, we perform the change of variable $v=\sqrt{r^2-s^2}$ so that

$$\begin{aligned}\int_{u-\varepsilon}^{u} e^{-2a\,\sqrt{r^2-s^2}}\,ds &= \int_{\sqrt{r^2-u^2}}^{\sqrt{r^2-(u-\varepsilon)^2}}\frac{v}{\sqrt{r^2-v^2}}\,e^{-2av}\,dv\\ &\le \frac{\sqrt{r^2-(u-\varepsilon)^2}}{2a(u-\varepsilon)}\left(e^{-2a\sqrt{r^2-u^2}}-e^{-2a\sqrt{r^2-(u-\varepsilon)^2}}\right)\\ &\le \frac{\sqrt{r^2-(u-\varepsilon)^2}}{2a(u-\varepsilon)}\,e^{-2a\sqrt{r^2-(u-\varepsilon)^2}}\left(e^{\frac{2a\varepsilon(2u-\varepsilon)}{\sqrt{r^2-(u-\varepsilon)^2}+\sqrt{r^2-u^2}}}-1\right).\end{aligned}$$

Again for $r\ge c$ for some large enough c, and $a+\sqrt{r^2-u^2}\ge 1$, for $u>2\varepsilon$,

$$\begin{aligned}\mathrm{Var}_{\mu_{\lambda,\mathscr{X}}}(f) &\ge \frac{1}{2}\,\mu_{\lambda,\mathscr{X}}(A(u-\varepsilon))\\ &\ge \frac{1}{2}\,Z_\lambda\,\sigma(S^{d-2})\,e^{-(a^2+r^2)}\int_0^{u-\varepsilon}\frac{a+\sqrt{r^2-s^2}}{1+2(a+\sqrt{r^2-s^2})^2}\\ &\quad\times s^{d-2}\,e^{-2a\,\sqrt{r^2-s^2}}\,ds\\ &\ge \frac{1}{2}\,Z_\lambda\,\sigma(S^{d-2})\,e^{-(a^2+r^2)}\int_{u-2\varepsilon}^{u-\varepsilon}\frac{a+\sqrt{r^2-s^2}}{1+2(a+\sqrt{r^2-s^2})^2}\\ &\quad\times s^{d-2}\,e^{-2a\,\sqrt{r^2-s^2}}\,ds\\ &\ge \frac{1}{2}\,Z_\lambda\,\sigma(S^{d-2})\,e^{-(a^2+r^2)}\,\frac{a+\sqrt{r^2-(u-\varepsilon)^2}}{1+2(a+\sqrt{r^2-(u-2\varepsilon)^2})^2}\,(u-2\varepsilon)^{d-2}\\ &\quad\int_{u-2\varepsilon}^{u-\varepsilon} e^{-2a\,\sqrt{r^2-s^2}}\,ds\,.\end{aligned}$$

The last integral is bounded from below by

$$\int_{u-2\varepsilon}^{u-\varepsilon} e^{-2a\sqrt{r^2-s^2}}\, ds \geq \frac{\sqrt{r^2-(u-\varepsilon)^2}}{2a(u-\varepsilon)}\, e^{-2a\sqrt{r^2-(u-\varepsilon)^2}} \times \left(1 - e^{\frac{-2a\varepsilon(2u-3\varepsilon)}{\sqrt{r^2-(u-\varepsilon)^2}+\sqrt{r^2-(u-2\varepsilon)^2}}}\right)$$

We thus deduce

$$C_P(1, B(y,r)) \geq \varepsilon^2\, \frac{(a+\sqrt{r^2-u^2})(a+\sqrt{r^2-(u-\varepsilon)^2})}{1+2(a+\sqrt{r^2-(u-2\varepsilon)^2})^2}\, \frac{(u-\varepsilon)^{d-2}}{u^{d-2}}\, H\,, \tag{1.36}$$

with

$$H = \frac{1 - e^{\frac{-2a\varepsilon(2u-3\varepsilon)}{\sqrt{r^2-(u-\varepsilon)^2}+\sqrt{r^2-(u-2\varepsilon)^2}}}}{e^{\frac{2a\varepsilon(2u-\varepsilon)}{\sqrt{r^2-(u-\varepsilon)^2}+\sqrt{r^2-u^2}}} - 1}\,.$$

For small r (smaller than $c\sqrt{d-1}$ for some small enough c) it is not difficult to show that $C_P(1, B(y,r)) \geq c_d$, and presumably c_d can be chosen independently of d, using again hitting times.

The bound (1.36) is not interesting if $a \gg r$, since in this case H is very small, unless ε is small enough (of order at most r/a), so that the lower bound we obtain goes to 0 with r/a. Hence we shall only look at the case where $a/r \leq C$. Since $2\varepsilon < u < r$, for H to be bounded from below by some universal constant, we see that $au\varepsilon \leq cr$ for some small enough universal constant c, so that we have to choose u and ε of order $\sqrt{r/a}$. It is then not difficult to see that, combined with all what we have done before, this will furnish the following type of lower bound

Proposition 1.10 *There exists a constant C_d such that for all y and r,*

$$C_P(\lambda, B(y,r)) \geq \frac{C_d}{\lambda}\left(1 + \frac{r}{|y| \vee 1}\right).$$

Even for very large r's, the previous method furnishes a dimension dependent bound. Proposition 1.10 is interesting when the obstacle contains the origin, in which case we have a linear dependence in $r/|y|$. Of course, when $y = 0$ we know that the lower bound growths as r^2. Also notice that for large a the previous lower bound is similar to the upper bound we have obtained in the previous section.

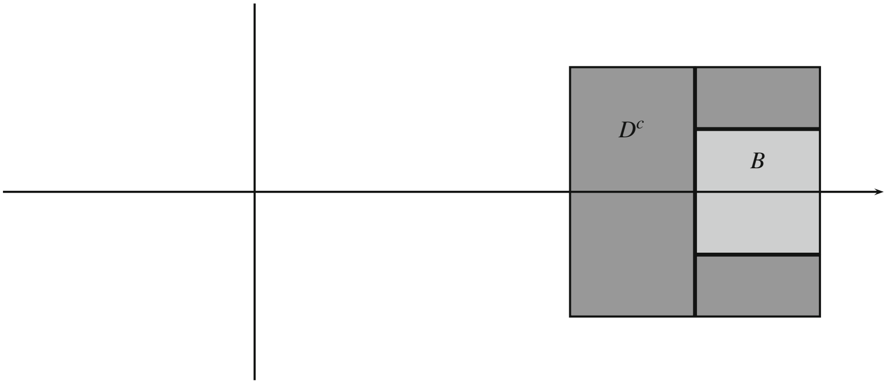


Fig. 1.3 A non convex obstacle D^c in gray. The trap B in lightgray

1.5.4 How to Kill the Poincaré Constant with Far But Small Non Convex Obstacles

In a previous subsection we have seen that an infinity of appropriately oriented squared obstacle with “centers” and radii going to infinity furnishes an infinite Poincaré constant. We shall see here that if we break the convexity of the obstacle, even small obstacles at infinity can kill the Poincaré constant.

For simplicity, we will assume that $d = 2$, and we shall look at $\lambda = 1$ with a non-convex bounded obstacle, namely we consider

$$D^c = \{0 \le y - x^1 \le \alpha\,;\ |x^2| \le \alpha\} \cup \{0 \le x^1 - y \le \alpha\,;\ \frac{\alpha}{2} \le |x^2| \le \alpha\}\,.$$

We simply denote by μ the Gaussian measure restricted to D (Fig. 1.3).

As in the previous subsections we shall introduce some $2/\alpha$-Lipschitz function f such that $\mathbb{1}_A \le f \le \mathbb{1}_B$ with $A = \{0 \le x^1 - y \le \frac{\alpha}{2}\,;\ \frac{\alpha}{2} \ge |x^2|\}$ and $B = \{0 \le x^1 - y \le \alpha\,;\ \frac{\alpha}{2} \ge |x^2|\}$. Hence

$$\mathrm{Var}_\mu(f) \ge \mu(A) - (\mu(B))^2 \quad \text{and} \quad \int |\nabla f|^2\, d\mu \le \frac{4}{\alpha^2}\,(\mu(B) - \mu(A))\,.$$

In addition

$$\mu(A) = Z_1 \left(\int_y^{y+\frac{\alpha}{2}} e^{-u^2}\, du\right) \left(\int_{-\frac{\alpha}{2}}^{\frac{\alpha}{2}} e^{-v^2}\, dv\right),$$

$$\mu(B) = Z_1 \left(\int_y^{y+\alpha} e^{-u^2}\, du\right) \left(\int_{-\frac{\alpha}{2}}^{\frac{\alpha}{2}} e^{-v^2}\, dv\right)$$

so that

$$\frac{\mu(A)}{\mu(B)} \geq \frac{\int_y^{y+\frac{\alpha}{2}} e^{-u^2}\,du}{\int_y^{y+\alpha} e^{-u^2}\,du} \geq \frac{\frac{y^2}{1+2y^2}\left(\frac{e^{-y^2}}{y} - \frac{e^{-(y+\frac{\alpha}{2})^2}}{y+\frac{\alpha}{2}}\right)}{\frac{1}{2y}\left(e^{-y^2} - e^{-(\alpha+y)^2}\right)}$$

$$\geq \frac{2y^2}{1+2y^2}\,\frac{1-e^{-\alpha(y+\frac{\alpha}{4})}}{1-e^{-\alpha(2y+\alpha)}}\,, \tag{1.37}$$

and

$$\frac{\mu(A)}{\mu(B)-\mu(A)} \geq \frac{\int_y^{y+\frac{\alpha}{2}} e^{-u^2}\,du}{\int_{y+\frac{\alpha}{2}}^{y+\alpha} e^{-u^2}\,du} \geq \frac{\frac{y^2}{1+2y^2}\left(\frac{e^{-y^2}}{y} - \frac{e^{-(y+\frac{\alpha}{2})^2}}{y+\frac{\alpha}{2}}\right)}{\frac{1}{2y}\left(e^{-(y+\frac{\alpha}{2})^2} - e^{-(\alpha+y)^2}\right)}$$

$$\geq \frac{2y^2}{1+2y^2}\,\frac{1-e^{-\alpha(y+\frac{\alpha}{4})}}{e^{-\alpha(y+\frac{\alpha}{4})}-e^{-\alpha(2y+\alpha)}}$$

$$\geq \frac{2y^2 e^{\alpha(y+\frac{\alpha}{4})}}{1+2y^2}\,\frac{1-e^{-\alpha(y+\frac{\alpha}{4})}}{1-e^{-\alpha(y+\frac{3\alpha}{4})}}\,. \tag{1.38}$$

$\mu(B)$ goes to 0 as $y \to +\infty$ while there exists some constant c such that $\mu(A) \geq c\,\mu(B)$, provided α is fixed and y large enough (depending on α), in particular as soon as $\alpha y \to +\infty$. As previously we thus have for αy large enough, $\mathrm{Var}_\mu(f) \geq \frac{1}{2}\,\mu(A)$. Gathering all previous results, we thus get $C_P(\mu) \geq \frac{\alpha^2}{8}\,\frac{\mu(A)}{\mu(B)-\mu(A)}$ so that $C_P(\mu)$ explodes (at least) like $e^{\alpha y}$ if $\alpha y \to +\infty$. Hence, even a small non convex obstacle going to infinity, makes the Poincaré constant explode.

More precisely consider an infinite number of such obstacles $(O(y_j, \alpha_j))$ such that one more time the convex face of the obstacle is orthogonal to the line joining the origin to y_j. *If* $\alpha_j \to 0$ *but* $\alpha_j\,|y_j| \to +\infty$, *then the process is not exponentially ergodic.*

Actually it is not difficult to see, though the calculations are a little bit more intricate, that the previous situation is similar to the case of two touching balls as in Fig. 1.4.

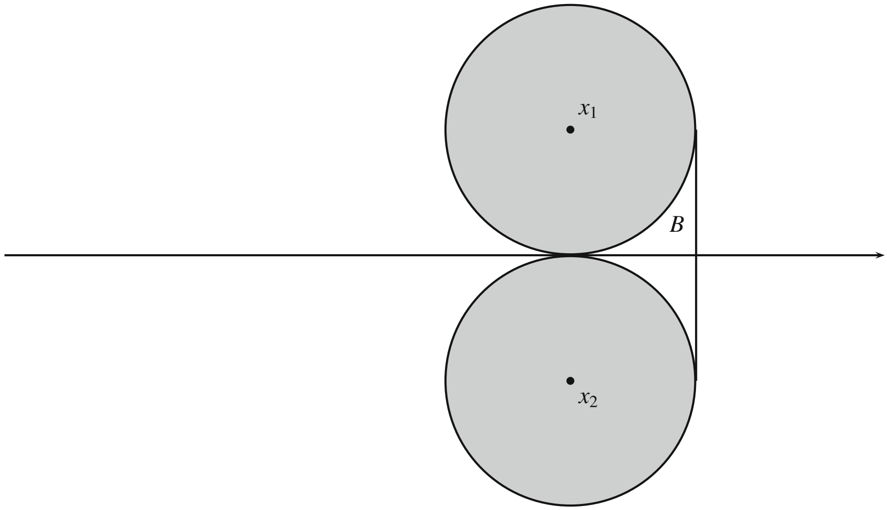


Fig. 1.4 Touching balls

Acknowledgements We want to heartily thank an anonymous referee for an amazing and so accurate work on the paper, correcting many typos, minor and not so minor mistakes. It is a real pleasure nowadays to receive such a report.

Appendix 1: Existence and Uniqueness of the Process

The main (actually unique) result of this section is the following (recall that the notion of solution for a reflected system involves both X and the local times L, see e.g. [9, 24])

Theorem 1.6 *Assume* (1.1)*. Then the system* (1.2) *has a unique (non explosive) strong solution for any allowed starting point* x*. In addition* $\mu_{\lambda,\mathscr{X}}$ *is the unique invariant (actually symmetric) probability measure.*

The remainder of this section is devoted to the proof of this result.

In the sequel we shall denote by L the (formal) infinitesimal generator

$$L = \frac{1}{2}\,\Delta - \lambda \langle x, \nabla \rangle \,, \tag{1.39}$$

whose domain is some extension of the set of smooth functions f compactly supported in $\bar{D}$ such that for all i,

$$\frac{\partial f}{\partial n_i}(y) = 0$$

at any y such that $|y - x_i| = r_i$, where n_i denotes the normal vector field on the sphere of center x_i and radius r_i.

We shall denote by $\mathscr{D}(L)$ this core.

Finite Number of Obstacles

When N is finite, existence of a unique (strong) solution of (1.2) starting from any point (belonging to $\bar{D}$ for (1.2)), up to the explosion time, is standard (see e.g. [9] for references) at least when the boundary of the obstacles is smooth. That is why we have chosen to smooth the hypercubes when looking at this particular situation. The only point is to show that the explosion time is almost surely infinite.

To this end, define

$$d_N = \max_{i=1,\dots,N} |x_i| \quad , \quad r = \max_{i=1,\dots,N} r_i \, , \tag{1.40}$$

and choose a smooth function h_N such that $h_N \geq 1$ everywhere,

$$h_N(x) = 1 \text{ if } |x| < d_N + 2r \quad , \quad h_N(x) = 1 + |x|^2 \text{ if } |x| > d_N + 3r + 1 \, . \tag{1.41}$$

It is enough to remark that $h_N \in \mathscr{D}(L)$ and satisfies

$$L h_N \leq -\lambda \, h_N \quad , \text{ for } |x| > d_L = (d/2\lambda)^{\frac{1}{2}} \vee (d_N + 3r + 1) \, . \tag{1.42}$$

h_N can thus play the role of a Lyapunov function for Hasminskii's non explosion test.

We can thus define for any x in $\bar{D}$ the law $P_t(x, dy)$ of the process at time t, X_t starting from x, as well as a Markov semi-group P_t acting on continuous and bounded functions. It is known that, for all $t > 0$,

$$P_t(x, dy) = p_t(x, y) \, dy$$

where $p_t \in C^\infty(\bar{D})$ (see [9, 10]). Furthermore, the density p_t is everywhere positive. This is a consequence of (1.1) (which implies in particular that D is path connected) and standard tools about the support of the law of the whole process.

$\mu_{\lambda, \mathscr{X}}$ is clearly a symmetric, hence invariant, probability measure. Uniqueness follows from the previous regularity and positivity as usual. Let us denote by q_t the density of the law of X_t w.r.t. $\mu_{\lambda, \mathscr{X}}$ i.e.

$$q_t(x, y) = p_t(x, y) \, \frac{dx}{d\mu_{\lambda, \mathscr{X}}} \, .$$

Application of the Chapman-Kolmogorov formula and standard regularization arguments yield

$$q_{2t}(x, x) = \int q_t(x, y) \, q_t(y, x) \, \mu_{\lambda, \mathscr{X}}(dy) = \int q_t^2(x, y) \, \mu_{\lambda, \mathscr{X}}(dy) \, , \tag{1.43}$$

thanks to symmetry, i.e. $q_t \in \mathbb{L}^2(\mu_{\lambda, \mathscr{X}})$.

Infinite Number of Obstacles

We now consider the case of infinitely many obstacles, still satisfying the non overlapping condition (1.1), for the locally finite collection $\mathscr{X}$. We can thus construct the process up to exit times of an increasing sequence of relatively compact open subsets U_n, each of which containing only a finite number of (closed) obstacles, the remaining (closed) obstacles being included into $(\bar{U_n})^c$. The sequence T_n of exit times of U_n is thus growing to the explosion time, but now it is much more difficult to control this limit.

A standard way is to use Dirichlet forms theory. Namely let us consider

$$\mathscr{E}(f) = \int |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} \tag{1.44}$$

defined for f which are smooth, bounded with bounded derivatives.

Our goal is to show that $\mathscr{E}$ is a conservative local Dirichlet form, so that one can associate to $\mathscr{E}$ a stationary Hunt process $(Y_t)_{t\geq 0}$ which is a non exploding diffusion process. This process coincides with X up to the exit time of U_n for all n, provided X_0 has distribution $\mu_{\lambda,\mathscr{X}}$ (exit time can be equal to 0). But, since $Y_t - Y_0$ is an additive functional of finite energy, it can be decomposed (Lyons-Zheng decomposition) for $0 \leq t \leq T$ into

$$Y_t - Y_0 = M_t + RM_t^T$$

where $M_.$ (resp. $RM_.^T$) is a forward (resp. backward) $\mathbb{L}^2$ martingale with brackets $\langle M \rangle_t = \langle RM^T \rangle_t = t$, hence are Brownian motions. It follows that for any $K > 0$,

$$\begin{aligned} \mathbb{P}\left(\sup_{t\in[0;T]} |Y_t| \geq K\right) &\leq \mathbb{P}\left(\sup_{t\in[0;T]} |Y_t - Y_0| \geq \frac{K}{2} \text{ or } |Y_0| \geq \frac{K}{2}\right) \\ &\leq \mathbb{P}\left(\sup_{t\in[0;T]} |M_t| \geq \frac{K}{4}\right) + \mathbb{P}\left(\sup_{t\in[0;T]} |RM_t^T| \geq \frac{K}{4}\right) \\ &\quad + \mathbb{P}\left(|Y_0| \geq \frac{K}{2}\right) \end{aligned}$$

and Doob's inequality allows us to conclude that the latter upper bound goes to 0 as K goes to infinity. It follows that the supremum of the exit times of the balls $B(0, K)$ is almost surely infinite, hence so does the supremum of the T_n's, implying that Y and X coincide up to any time and that X does not explode, when the initial law is $\mu_{\lambda,\mathscr{X}}$.

Standard arguments (see [19]) imply that there is no explosion starting from quasi every point x (i.e. all x's not belonging to some polar set $\mathscr{N}$, recall that here polar sets coincide with sets of zero capacity), though here we only need that this property holds for $\mu_{\lambda,\mathscr{X}}$ almost all x's, which is an immediate consequence of disintegration of the measure.

Now let x be some point in D, and choose a small ball $B(x,\varepsilon) \subset D$. If $\mathbb{P}_y$ denotes the law of X starting from y as usual, we have for all $z \in B(x,\varepsilon)$,

$$\mathbb{P}_z(\sup_n T_n < +\infty) = \int_{|y-x|=\varepsilon} \mathbb{P}_y(\sup_n T_n < +\infty)\, \eta_z(dy)$$

where η_z denotes the $\mathbb{P}_z$ law of X_τ with τ the exit time of $B(x,\varepsilon)$ (that τ is almost surely finite is well known and actually follows from the arguments below).

Up to the exit time of $B(x,\varepsilon)$, X is just an Ornstein-Uhlenbeck process, so that its law is equivalent to the one of the Brownian motion. For Brownian motion, it is well known that τ is a.s. finite, that the exit measure (starting from z) is simply the harmonic measure (related to z) on the sphere $S(x,\varepsilon)$, hence is equivalent to the surface measure σ_x. Thus the same properties hold true for our Ornstein-Uhlenbeck process.

It follows that η_z is equivalent to the surface measure σ_x on the sphere $S(x,\varepsilon)$, so that η_z and η_x are equivalent.

(One can see e.g. [11] theorem 4.18 for much more sophisticated situations).

Choose $z \notin \mathscr{N}$. The previous formula shows that for η_z almost all $y \in S(x,\varepsilon)$, $\mathbb{P}_y(\sup_n T_n < +\infty) = 0$, so that the same holds η_x almost surely and finally $\mathbb{P}_x(\sup_n T_n < +\infty) = 0$, showing that no explosion occurs starting from any point.

It remains to show that $\mathscr{E}$ is a conservative and local Dirichlet form. To this end introduce the truncated form

$$\mathscr{E}_n(f) = \frac{1}{\mu_{\lambda,\mathscr{X}}(U_n)} \int_{U_n} |\nabla f|^2 \, d\mu_{\lambda,\mathscr{X}} \tag{1.45}$$

corresponding to the reflected O-U process in U_n with hard obstacles. It is not difficult to see that we can build the open sets U_n in such a way that ∂U_n is smooth. It thus follows that $\mathscr{E}_n$ is a conservative and local Dirichlet form, to which is associated a non-exploding process X^n. The same reasoning as before shows that we can start from any point $x \in U_n$. We use the superscript n for the Markov law corresponding to $\mathscr{E}_n$

Let τ_K be the exit time from the ball $B(0,K)$ and let n_K be such that for $n \geq n_K$, $B(0,K) \subset U_n$. All processes X^n $(n \geq n_K)$, starting from $x \in B(0,K)$, coincide up to time τ_K (and coincide with the possibly exploding X). Now choose some initial measure $\nu(dy) = u(y)dy$ where u is bounded and has compact support included in $B(0,R)$. Then ν is absolutely continuous with respect to $\mu^n_{\lambda,r}$ and one can find some constant $C(K,\nu)$ such that

$$\| \frac{d\nu}{d\mu^n_{\lambda,\mathscr{X}}} \|_\infty \leq C(K,\nu) \quad \text{for all } n \geq n_K .$$

For any $T > 0$, it yields, using the Lyons-Zheng decomposition as before

$$\begin{aligned}
\mathbb{P}_\nu\left(\sup_{t\in[0;T]} |X^n_t| \geq K\right) &\leq C(K,\nu)\,\mathbb{P}_{\mu^n_{\lambda,r}}\left(\sup_{t\in[0;T]} |X^n_t - X^n_0| \geq \frac{K}{2} \text{ or } |X^n_0| \geq \frac{K}{2}\right)\\
&\leq C(K,\nu)\left(\mathbb{P}_{\mu^n_{\lambda,r}}\left(\sup_{t\in[0;T]} |M^n_t| \geq \frac{K}{4}\right)\right.\\
&\quad \left. + \mathbb{P}_{\mu^n_{\lambda,r}}\left(\sup_{t\in[0;T]} |RM^{T,n}_t| \geq \frac{K}{4}\right)\right)\\
&\quad + C(K,\nu)\,\mathbb{P}_{\mu^n_{\lambda,r}}\left(|X^n_0| \geq \frac{K}{2}\right)\\
&\leq C(K,\nu)\left(C_1\, e^{-C_2 K^2/T} + \mu^n_{\lambda,\mathscr{X}}(B^c(0,K/2))\right)\\
&\leq C(K,\nu)\left(C_1\, e^{-C_2 K^2/T} + \frac{\mu_{\lambda,\mathscr{X}}(B^c(0,K/2))}{\mu_{\lambda,\mathscr{X}}(U_n)}\right)
\end{aligned}$$

for well chosen universal constants C_1, C_2. It immediately follows that $\mathbb{P}_\nu(\tau_K \leq T)$ (here we consider the process X) goes to 0 as K goes to $+\infty$, so that the process starting from ν does not explode. This is of course sufficient for our purpose, since conservativeness follows by choosing a sequence ν_j converging to $\mu_{\lambda,\mathscr{X}}$.

Remark 1.3 Once the non explosion is proven, standard arguments show that the process is Feller. Hence compact sets are closed petite sets in the terminology of [17, 18]. We refer to the latter reference for a complete discussion.

Appendix 2: Useful Estimates for Exponential Moments of Hitting Times

In this section we shall recall some estimates of exponential moments of hitting times for some special linear processes. Denotes by $S(r)$ the first exit time of the symmetric interval $[-r, r]$ for a one dimensional process.

For the linear Brownian motion it is well known, (see [27] Exercise 3.10) that

$$E_0\left(e^{\theta\, S(r)}\right) = \frac{1}{\cos(r\sqrt{2\theta})} < +\infty$$

if and only if

$$\theta \leq \frac{\pi^2}{8\,r^2}\,.$$

Surprisingly enough (at least for us) a precise description of the Laplace transform of $S(r)$ for the O-U process is very recent: it was first obtained in [21]. A simpler proof is contained in [22] Theorem 3.1. The result reads as follows

Theorem 1.7 (See [21, 22]) *If $S(r)$ denotes the exit time from $[-r, r]$ of a linear O-U process with drift $-\lambda x$ ($\lambda > 0$), then for $\theta \geq 0$,*

$$E_0\left(e^{-\theta\, S(r)}\right) = \frac{1}{{}_1F_1\left(\frac{\theta}{2\lambda}\,,\, \frac{1}{2}\,,\, \lambda\, r^2\right)}\,,$$

where ${}_1F_1$ denotes the confluent hypergeometric function.

The function ${}_1F_1$ is also denoted by Φ (in [21] for instance) or by M in [1] (where it is called Kummer function) and is defined by

$$ {}_1F_1(a, b, z) = \sum_{k=0}^{+\infty} \frac{(a)_k}{(b)_k}\, \frac{z^k}{k!} \quad \text{where} \quad (a)_k = a(a+1)\ldots(a+k-1)\,,\ (a)_0 = 1\,. \tag{1.46}$$

In our case, $b = \frac{1}{2}$, so that ${}_1F_1$ is an analytic function, as a function of both z and θ. It follows that $\theta \mapsto E_0\left(e^{-\theta\, S(r)}\right)$ can be extended, by analytic continuation, to $\theta < 0$ as long as λr^2 is not a zero of ${}_1F_1(\frac{\theta}{2\lambda}\,,\, \frac{1}{2}\,,\, .)$.

The zeros of the confluent hypergeometric function are difficult to study. Here we are looking for the first negative real zero. For $-1 < a < 0$, $b > 0$, it is known (and easy to see) that there exists only one such zero, denoted here by u. Indeed ${}_1F_1(a, b, 0) = 1$ and all terms in the expansion (1.46) are negative for $z > 0$ except the first one, implying that the function is decaying to $-\infty$ as $z \to +\infty$. However, an exact or an approximate expression for u are unknown (see the partial results of Slater in [1, 28], or in [20]). Our situation however is simpler than the general one, and we shall obtain a rough but sufficient bound.

First, comparing with the Brownian motion, we know that for all $\lambda > 0$ we must have

$$\frac{-\theta}{\lambda} \leq \frac{\pi^2}{8(r\sqrt{\lambda})^2}\,.$$

So, if $\lambda\, r^2 > \pi^2/8$ and $-\theta/2\lambda \geq 1/2$, the Laplace transform (or the exponential moment) is infinite. We may thus assume that $-\theta/2\lambda < 1/2$.

Hence, for ${}_1F_1\left(\frac{\theta}{2\lambda}\,,\, \frac{1}{2}\,,\, \lambda\, r^2\right)$ to be negative it is enough that

$$\begin{aligned} 1 &< \frac{-\theta}{\lambda}\left((\lambda\, r^2) + \sum_{k=2}^{+\infty} \frac{(1+\frac{\theta}{2\lambda})(2+\frac{\theta}{2\lambda})\ldots(k-1+\frac{\theta}{2\lambda})}{(1+\frac{1}{2})(2+\frac{1}{2})\ldots(k-1+\frac{1}{2})}\, \frac{(\lambda\, r^2)^k}{k!}\right) \\ &< \frac{-\theta}{\lambda}\left(\sum_{k=1}^{+\infty} \frac{(\lambda\, r^2)^k}{k!}\right)\,, \end{aligned}$$

i.e.

$$\text{as soon as}\quad \beta = -\theta > \frac{\lambda}{e^{\lambda r^2} - 1} \quad \text{then} \quad \mathbb{E}_0\left(e^{\beta S(r)}\right) = +\infty . \tag{1.47}$$

So there is a drastically different behavior between both processes.

Finally we shall also need estimates for a general CIR process or generalized squared radial Ornstein-Uhlenbeck process, i.e. the solution of

$$dU_t = 2\sqrt{U_t} dB_t + (\delta + 2\beta\, U_t)\, dt$$

when $\beta > 0$ and $\delta > 0$. According to [21] Theorem 3, for $\theta > 0$,

$$E_0\left(e^{-\theta\, S(u)}\right) = \frac{e^{\beta\, u}}{{}_1F_1\left(\frac{(\theta+\beta\delta)}{2\beta}\, ,\, \frac{\delta}{2}\, ,\, \beta\, u\right)} . \tag{1.48}$$

It follows that for $0 < \theta < \beta\, \delta$, $E_0\left(e^{\theta\, S(u)}\right) < +\infty$.

Appendix 3: The Case $N = 1$: Another Estimate for a General y Using Decomposition of Variance

A very usual method to deal with dimension controls is the decomposition of variance. This method can be used here in order to transfer the results of Proposition 1.1 to a non centered obstacle. Though the results are non optimal in many directions, the method contains some interesting features.

In this section for simplicity we will first assume that $\lambda = 1$, and second that $d \geq 3$. Recall that we are looking here at the case of an unique spherical obstacle $B(y, r)$, so that we simply denote by $\mu_{d,r}$ the restricted Gaussian measure $\mu_{\lambda,\mathscr{X}}$. Since we will use an induction procedure on the dimension d we explicitly make it appear in the notation.

Using rotation invariance we may also assume that $y = (a, 0)$ for some $a \in \mathbb{R}^+$, 0 being the null vector of $\mathbb{R}^{d-1}$. So, writing $x = (u, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{d-1}$,

$$\mu_{d,r}(du, d\bar{x}) = \nu^0_{d-1,R(u)}(d\bar{x})\, \mu_1(du)\, ,$$

where $\nu^0_{d-1,R(u)}(d\bar{x})$ is the $d - 1$ dimensional Gaussian measure restricted to $B^c(0, R(u))$ as in Sect. 1.2.1 with $R(u) = \sqrt{\left(\left(r^2 - (u - a)^2\right)_+\right)}$ and μ_1 is the first marginal of $\mu_{d,r}$ given by

$$\mu_1(du) = \frac{\gamma_{d-1}(B^c(0, R(u)))}{\gamma_d(B^c(y, r))}\ \gamma_1(du)\,,$$

γ_n denoting the n dimensional Gaussian measure $c_n\, e^{-|x|^2}\, dx$.

The standard decomposition of variance tells us that for a nice f,

$$\mathrm{Var}_{\mu_{d,r}}(f) = \int \left(\mathrm{Var}_{\nu^0_{d-1,R(u)}}(f)\right)\, \mu_1(du) + \mathrm{Var}_{\mu_1}(\bar{f})\,, \tag{1.49}$$

where

$$\bar{f}(u) = \int\, f(u, \bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x})\,.$$

According to Proposition 1.1, on one hand, it holds for all u,

$$\mathrm{Var}_{\nu^0_{d-1,R(u)}}(f) \le \left(1 + \frac{(r^2 - (u-a)^2)_+}{d-1}\right) \int\, |\nabla_{\bar{x}} f|^2\, d\nu^0_{d-1,R(u)}\,, \tag{1.50}$$

so that

$$\int \left(\mathrm{Var}_{\nu^0_{d-1,R(u)}}(f)\right)\, \mu_1(du) \le \left(1 + \frac{r^2}{d-1}\right) \int\, |\nabla_{\bar{x}} f|^2\, d\mu_{d,r}\,. \tag{1.51}$$

On the other hand, μ_1 is a logarithmically bounded perturbation of γ_1 hence satisfies some Poincaré inequality so that

$$\mathrm{Var}_{\mu_1}(\bar{f}) \le C_1 \int \left|\frac{d\bar{f}}{du}\right|^2\, d\mu_1\,. \tag{1.52}$$

So we have first to get a correct bound for C_1, second to understand what $\frac{d\bar{f}}{du}$ is.

A Bound for C_1

Since μ_1 is defined on the real line, upper and lower bounds for C_1 may be obtained by using Muckenhoupt bounds (see [3] Theorem 6.2.2). Unfortunately we were not able to obtain the corresponding explicit expression in our situation as μ_1 is not sufficiently explicitly given to use Muckenhoupt criterion. So we shall give various upper bounds using other tools.

The usual Holley-Stroock perturbation argument combined with the Poincaré inequality for γ_1 imply that

$$C_1 \le \frac{1}{2}\,\frac{\sup_u\{\gamma_{d-1}(B^c(0,R(u)))\}}{\inf_u\{\gamma_{d-1}(B^c(0,R(u)))\}} \le \frac{1}{2}\,\frac{\int_0^{+\infty}\rho^{d-2}\,e^{-\rho^2}\,d\rho}{\int_r^{+\infty}\rho^{d-2}\,e^{-\rho^2}\,d\rho}$$
$$= \frac{1}{2}\left(1 + \frac{\int_0^r \rho^{d-2}\,e^{-\rho^2}\,d\rho}{\int_r^{+\infty}\rho^{d-2}\,e^{-\rho^2}\,d\rho}\right). \tag{1.53}$$

Using the first inequality and the usual lower bound for the denominator, it follows that

$$\text{for all } r>0, \quad C_1 \le \pi^{(d-2)/2}\,\frac{e^{r^2}}{r^{d-3}}\,.$$

The function $\rho \mapsto \rho^{d-2}\,e^{-\rho^2}$ increases up to its maximal value which is attained for $\rho^2=(d-2)/2$ and then decreases to 0. It follows, using the second form of the inequality (1.53) that

- if $r \le \sqrt{\frac{d-2}{2}}$ we have $C_1 \le \frac{1}{2}+r^2$, while
- if $r \ge \sqrt{\frac{d-2}{2}}$ we have

$$C_1 \le \frac{1}{2} + \left(\frac{d-2}{2}\right)^{\frac{d-2}{2}}\, e^{-\frac{d-2}{2}}\,\frac{e^{r^2}}{r^{d-4}}\,.$$

These bounds are quite bad for large r's but do not depend on y.

Why is it bad? First for $a=0$ (corresponding to the situation of Sect. 1.2.1) we know that $C_1 \le 1+\frac{r^2}{d}$ according to Proposition 1.1 applied to functions depending on x_1. Actually the calculations we have done in the proof of Proposition 1.1, are unchanged for $f(z)=z_1$, so that it is immediately seen that $C_1 \ge \max(\frac{1}{2},\frac{r^2}{d})$.

Intuitively the case $a=0$ is the worst one, though we have no proof of this. We can nevertheless give some hints.

The natural generator associated to μ_1 is

$$L_1 = \frac{d^2}{du^2} - \left(u - \frac{d}{du}\log(\gamma_{d-1}(B^c(0,R(u))))\right)\frac{d}{du}$$
$$= \frac{d^2}{du^2} - u\,\frac{d}{du} + \frac{(u-a)\,(R(u))^{d-3}\,e^{-R^2(u)}}{\int_{R(u)}^{+\infty}\rho^{d-2}\,e^{-\rho^2}\,d\rho}\,\mathbb{1}_{|u-a|\le r}\,\frac{d}{du}\,.$$

The additional drift term behaves badly for $a \le u \le a+r$, since in this case it is larger than $-u$, while for $u \le a$ it is smaller. In stochastic terms it means that one can compare the induced process with the Ornstein-Uhlenbeck process except possibly for $a \leq u \leq a+r$. In analytic terms let us look for a Lyapunov function for L_1. As for the O-U generator the simplest one is $g(u) = u \mapsto u^2$ for which

$$L_1 g \leq 2 - 4u^2 + 4u(u-a)\, \mathbb{1}_{a \leq u \leq a+r}\,.$$

Remember that $a \geq r$ so that $-au \leq -\frac{1}{2}\,u^2$. It follows

$$\text{provided } a \geq r, \quad L_1 g \leq 2 - 2g\,. \tag{1.54}$$

For $|u| \geq \sqrt{2}$ we then have $L_1 g(u) \leq -g(u)$, so that g is a Lyapunov function outside the interval $[-\sqrt{2}, \sqrt{2}]$ and the restriction of μ_1 to this interval coincides (up to the constants) with the Gaussian law γ_1 hence satisfies a Poincaré inequality with constant $\frac{1}{2}$ on this interval. According to the results in [5] we recalled in the previous section, we thus have that C_1 is bounded above by some universal constant c.

We may gather our results

Lemma 1.6 *The following upper bound holds for C_1:*

(1) (small obstacle) if $r \leq \sqrt{\frac{d-2}{2}}$ we have $C_1 \leq \frac{1}{2} + r^2$,
(2) (far obstacle) if $|y| > r + \sqrt{2}$, $C_1 \leq c$ for some universal constant c,
(3) (centered obstacle) if $y = 0$, $C_1 \leq 1 + \frac{r^2}{d}$,
(4) in all other cases, there exists $c(d)$ such that $C_1 \leq c(d)\,\frac{e^{r^2}}{r^{d-3}}$.

We conjecture that actually $C_1 \leq C(1+r^2)$ for some universal constant C.

Remark 1.4 In a recent work [25], the authors obtain a much better upper bound in case (4) (in fact a constant) when the origin belongs to the boundary of the ball and $d = 3$. ◇

Controlling $\frac{d\bar{f}}{du}$

It remains to understand what $\frac{d\bar{f}}{du}$ is and to compute the integral of its square against μ_1.

Recall that

$$\bar{f}(u) = \int f(u, \bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x})\,.$$

Hence

$$\bar{f}(u) = \mathbb{1}_{|u-a|>r} \int f(u,\bar{x})\, \nu^0_{d-1,0}(d\bar{x})$$
$$+ \mathbb{1}_{|u-a|\le r} \int_{\mathbb{S}^{d-2}} \int_{R(u)}^{+\infty} f(u,\rho\,\theta)\, \frac{\rho^{d-2}\, e^{-\rho^2}}{c(d)\,\gamma_{d-1}(B^c(0,R(u)))}\, d\rho\, d\theta\,,$$

where $d\theta$ is the non-normalized surface measure on the unit sphere $\mathbb{S}^{d-2}$ and $c(d)$ the normalization constant for the Gaussian measure. Hence, for $|u-a| \neq r$ we have

$$\frac{d}{du}\bar{f}(u) = \int \frac{\partial f}{\partial x_1}(u,\bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x})$$
$$- \mathbb{1}_{|u-a|\le r} \int f(u,\bar{x})\, \mathbb{1}_{|\bar{x}|>R(u)}\, \frac{\frac{d}{du}\,(\gamma_{d-1}(B^c(0,R(u))))}{\gamma^2_{d-1}(B^c(0,R(u)))}\, \gamma_{d-1}(d\bar{x})$$
$$- \mathbb{1}_{|u-a|\le r}\, \frac{R'(u)\, R^{d-2}(u)\, e^{-R^2(u)}}{c(d)\,\gamma_{d-1}(B^c(0,R(u)))} \int_{\mathbb{S}^{d-2}} f(u,R(u)\,\theta)\, d\theta\,.$$

Notice that if f only depends on u, $\bar{f} = f$ so that

$$\frac{d}{du}\bar{f}(u) = \frac{\partial f}{\partial x_1}(u) = \int \frac{\partial f}{\partial x_1}(u)\, \nu^0_{d-1,R(u)}(d\bar{x})\,,$$

and thus the sum of the two remaining terms is equal to 0. Hence in computing the sum of the two last terms, we may replace f by $f - \int f(u,\bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x})$ or if one prefers, we may assume that the latter $\int f(u,\bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x})$ vanishes. Observe that this change will not affect the gradient in the $\bar{x}$ direction.

Assuming this, the second term becomes

$$- \mathbb{1}_{|u-a|\le r}\, \frac{\frac{d}{du}\,(\gamma_{d-1}(B^c(0,R(u))))}{\gamma_{d-1}(B^c(0,R(u)))} \int f(u,\bar{x})\, \nu^0_{d-1,R(u)}(d\bar{x}) \;=\; 0\,.$$

We thus have using a scale change

$$\frac{d\bar{f}}{du} = \int \frac{\partial f}{\partial x_1}(u,\bar{x}) \nu^0_{d-1,R(u)}(d\bar{x})$$
$$- \mathbb{1}_{|u-a|\le r}\, \frac{R'(u)\, e^{-R^2(u)}}{c(d)\,\gamma_{d-1}(B^c(0,R(u)))} \int_{\mathbb{S}^{d-2}(R(u))} f(u,\,\theta)\, d\theta\,.$$

Our goal is to control the last term using the gradient of f. One good way to do it is to use the Green-Riemann formula, in a well adapted form. Indeed, let V be a vector field written as

$$V(\bar{x}) = -\frac{\varphi(|\bar{x}|)}{|\bar{x}|^{d-1}}\,\bar{x} \quad \text{where } \varphi(R(u)) = R^{d-2}(u)\,. \tag{1.55}$$

This choice is motivated by the fact that the divergence, $\nabla.(\bar{x}/|\bar{x}|^{d-1}) = 0$ on the whole $\mathbb{R}^{d-1} - \{0\}$.

Of course in what follows we may assume that $R(u) > 0$, so that all calculations make sense. The Green-Riemann formula tells us that, denoting $g_u(\bar{x}) = f(u, \bar{x})$, for some well chosen φ

$$\begin{aligned}\int_{\mathbb{S}^{d-2}(R(u))} f(u,\,\theta)\,d\theta &= \int_{\mathbb{S}^{d-2}(R(u))} g_u\,\langle V, (-\bar{x}/|\bar{x}|)\rangle\,d\theta \\ &= \int \mathbb{1}_{|\bar{x}|\geq R(u)}\,\nabla.(g_u V)(\bar{x})\,d\bar{x} \\ &= -\int \mathbb{1}_{|\bar{x}|\geq R(u)}\,\langle \nabla g_u(\bar{x}), (\bar{x}/|\bar{x}|^{d-1})\rangle\,\varphi(|\bar{x}|)\,d\bar{x} \\ &\quad - \int \mathbb{1}_{|\bar{x}|\geq R(u)}\,g_u(\bar{x})\,(\varphi'(|\bar{x}|)/|\bar{x}|^{d-2})\,d\bar{x}\,.\end{aligned}$$

Now we choose $\varphi(s) = R^{d-2}(u)\,e^{R^2(u)}\,e^{-s^2}$ and recall that $R'(u) = -((u - a)/R(u))\,\mathbb{1}_{|u-a|\leq r}$. We have finally obtained

$$\begin{aligned}&\mathbb{1}_{|u-a|\leq r}\,\frac{R'(u)\,e^{-R^2(u)}}{c(d)\,\gamma_{d-1}(B^c(0,R(u)))}\int_{\mathbb{S}^{d-2}(R(u))} f(u,\,\theta)\,d\theta \\ &= \mathbb{1}_{|u-a|\leq r}\,(u-a)\,R^{d-3}(u)\int \langle \nabla_{\bar{x}} f(u,\bar{x}), (\bar{x}/|\bar{x}|^{d-1})\rangle\,\nu^0_{d-1,R(u)}(d\bar{x}) \\ &\quad - \mathbb{1}_{|u-a|\leq r}\,(u-a)\,R^{d-3}(u)\,2\int (f(u,\bar{x})/|\bar{x}|^{d-3})\,\nu^0_{d-1,R(u)}(d\bar{x})\,.\end{aligned}$$

We will bound the above quantities for each fixed u. To control the first one we use Cauchy-Schwarz inequality, while for the second one we use first Cauchy-Schwarz inequality yielding a term containing $\int f^2(u,\bar{x})\,\nu^0_{d-1,R(u)}(d\bar{x})$ and then the Poincaré inequality for $\nu^0_{d-1,R(u)}$ we obtained in Proposition 1.1, since $\int f(u,\bar{x})\,\nu^0_{d-1,R(u)}(d\bar{x}) = 0$. This yields

$$\begin{aligned}\left|\frac{d\bar{f}}{du}\right| &\leq \int \left|\frac{\partial f}{\partial x_1}\right|\,\nu^0_{d-1,R(u)}(d\bar{x}) \\ &\quad + 2\left(\int |\nabla_{\bar{x}} f|^2\,\nu^0_{d-1,R(u)}(d\bar{x})\right)^{\frac{1}{2}}\,(A_1(u) + 2\,A_2(u))\end{aligned}$$

where

$$A_1^2(u) = |u-a|^2 \, \mathbb{1}_{|u-a|\leq r} \, R^{2d-6}(u) \, \left(\int |\bar{x}|^{4-2d} \, \nu^0_{d-1,R(u)}(d\bar{x}) \right) \, ,$$

and

$$A_2^2(u) = |u-a|^2 \, \mathbb{1}_{|u-a|\leq r} \, R^{2d-6}(u) \, \left(1 + \frac{R^2(u)}{d-1} \right) \left(\int |\bar{x}|^{6-2d} \, \nu^0_{d-1,R(u)}(d\bar{x}) \right) \, .$$

It follows

$$\int \left| \frac{d\bar{f}}{du} \right|^2 d\mu_1 \leq 2 \int \left| \frac{\partial f}{\partial x_1} \right|^2 \mu_{d,r}(du, d\bar{x}) + 4 \sup_u \, (A_1(u) + 2\, A_2(u))^2 \int |\nabla_{\bar{x}} f|^2 \, \mu_{d,r}(du, d\bar{x}) \, .$$

It remains to study the final supremum.

Recalling that $\nu^0_{d-1,R(u)}$ is the (normalized) Gaussian measure restricted to $|\bar{x}| \geq R(u)$, we see that what we have to do is to get upper bounds for quantities like

$$\int_R^{+\infty} \rho^{-k} \, e^{-\rho^2} \, d\rho$$

for $k = d-2$ or $k = d-4$, and a lower bound for

$$\int_R^{+\infty} \rho^{d-2} \, e^{-\rho^2} \, d\rho \, .$$

It is easily seen that for $k \in \mathbb{N}$,

$$\int_R^{+\infty} \rho^{-k} \, e^{-\rho^2} \, d\rho \leq \frac{1}{R^{k+1}} \int_R^{+\infty} \rho \, e^{-\rho^2} \, d\rho \leq \frac{e^{-R^2}}{2\, R^{1+k}} \quad \text{if } R \geq 1, \tag{1.56}$$

$$\int_R^{+\infty} \rho^{-k} \, e^{-\rho^2} \, d\rho \leq \frac{1}{2e} + \frac{1}{(k-1)R^{k-1}} \quad \text{if } R \leq 1, k \geq 2, \tag{1.57}$$

$$\int_R^{+\infty} \rho^{-1} \, e^{-\rho^2} \, d\rho \leq \frac{1}{2e} + \ln(1/R) \quad \text{if } \ R \leq 1, \tag{1.58}$$

$$\int_R^{+\infty} e^{-\rho^2} \, d\rho \leq 1 + \frac{1}{2e} \quad \text{if } \ R \leq 1, \tag{1.59}$$

$$\int_R^{+\infty} \rho \, e^{-\rho^2} \, d\rho \leq \frac{1}{2} \quad \text{if } R \leq 1, \tag{1.60}$$

$$\int_R^{+\infty} \rho^{d-2}\, e^{-\rho^2}\, d\rho \geq \frac{R^{d-3}}{2\, e^{R^2}} \quad \text{if } R \geq 1, \tag{1.61}$$

$$\int_R^{+\infty} \rho^{d-2}\, e^{-\rho^2}\, d\rho \geq \frac{1}{2\, e} \quad \text{if } R \leq 1. \tag{1.62}$$

Applying the previous bounds first for $k = d-2$ (corresponding to A_1) and $k = d-4$ (corresponding to A_2), we obtain first for $R(u) \geq 1$,

$$A_1^2(u) \leq r^2\, R^{-2}(u) \;\leq\; r^2\,,$$

$$A_2^2(u) \leq r^2 \left(1 + \frac{R^2(u)}{d-1}\right) \;\leq\; r^2 \left(1 + \frac{r^2}{d-1}\right)$$

while for $R(u) \leq 1$,

$$A_1^2(u) \leq (1+2e)\, r^2\,,$$

$$A_2^2(u) \leq 2\,(1+2e)\, r^2\,,$$

the latter bounds being obtained after discussing according to the dimension $d = 3, 4, 5$ or larger than 6.

Gathering all these results together we may state

Theorem 1.8 *Assume $d \geq 3$. There exists a function $C(r, d)$ such that, for all $y \in \mathbb{R}^d$,*

$$C_P(1, y, r) \;\leq\; C(r, d)\,.$$

Furthermore, there exists some universal constant c such that

$$C(r, d) \leq C_1(r)\, C_2(r)\,,$$

$C_1(r)$ being given in Lemma 1.6 and $C_2(r)$ satisfying

1. *if $r \leq \sqrt{(d-1)/2}$, $C_2(r) \leq c\,(1 + r^2)$,*
2. *if $r \geq 1$, $C_2(r) \leq c\, r^2 \left(1 + \frac{r^2}{d-1}\right)$.*

Remark 1.5 The previous theorem is interesting as it shows that when $N = 1$, the Poincaré constant is bounded uniformly in y and it furnishes some tractable bounds.

The method suffers nevertheless two defaults. First it does not work for $d = 2$, in which case the conditioned measure does no more satisfy a Poincaré inequality. More important for our purpose, the method does not extend to more than one obstacle, unless the obstacles have a particular location.

References

1. M. Abramowitz, I.A. Stegun (eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Books on Intermediate and Advanced Mathematics (Dover Publications Inc., New York, 1972)
2. S.B. Alexander, I.D. Berg, R.L. Bishop, The Riemannian obstacle problem. Illinois J. Math. **31**, 167–184 (1987)
3. C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*. Panoramas et Synthèses, vol. 10 (Société Mathématique de France, Paris, 2000)
4. A. Athreya, T. Kolba, J.C. Mattingly, Propagating Lyapunov functions to prove noise–induced stabilization. Electron. J. Probab. **17**, 1–38 (2012)
5. D. Bakry, F. Barthe, P. Cattiaux, A. Guillin, A simple proof of the Poincaré inequality for a large class of probability measures. Electron. J. Probab. **13**, 60–66 (2008)
6. D. Bakry, P. Cattiaux, A. Guillin, Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. J. Funct. Anal. **254**, 727–759 (2008)
7. J. Birrell, D.P. Herzog, J. Wehr, Transition from ergodic to explosive behavior in a family of stochastic differential equations. Stoch. Process. Appl. **122**(5), 1519–1539 (2012)
8. S.G. Bobkov, Spectral gap and concentration for some spherically symmetric probability measures, in *Geometric Aspects of Functional Analysis, Israel Seminar 2000–2001*. Lecture Notes in Mathematics, vol. 1807 (Springer, Berlin, 2003), pp. 37–43
9. P. Cattiaux, Hypoellipticité et hypoellipticité partielle pour les diffusions avec une condition frontière. Ann. Inst. Henri Poincaré. Prob. Stat. **22**, 67–112 (1986)
10. P. Cattiaux, Régularité au bord pour les densités et les densités conditionnelles d'une diffusion réfléchie hypoelliptique. Stochast. Stoch. Rep. **20**, 309–340 (1987)
11. P. Cattiaux, Calcul stochastique et opérateurs dégénérés du second ordre: II Problème de Dirichlet. Bull. Sci. Math. **115**, 81–122 (1991)
12. P. Cattiaux, A. Guillin, Deviation bounds for additive functionals of Markov processes. ESAIM Probab. Stat. **12**, 12–29 (2008)
13. P. Cattiaux, A. Guillin, F.Y. Wang, L. Wu, Lyapunov conditions for super Poincaré inequalities. J. Funct. Anal. **256**(6), 1821–1841 (2009)
14. P. Cattiaux, A. Guillin, P.A. Zitt, Poincaré inequalities and hitting times. Ann. Inst. Henri Poincaré. Prob. Stat. **49**(1), 95–118 (2013)
15. P. Cattiaux, M. Fradon, A. Kulik, S. Roelly, Long time behavior of stochastic hard balls systems. Bernoulli **22**(2), 681–710 (2016)
16. P. Diaconis, G. Lebeau, L. Michel, Geometric analysis for the Metropolis algorithm on Lipschitz domains. Invent. Math. **185**(2), 239–281 (2011)
17. R. Douc, G. Fort, E. Moulines, P. Soulier, Practical drift conditions for subgeometric rates of convergence. Ann. Appl. Probab. **14**(3), 1353–1377 (2004)
18. N. Down, S.P. Meyn, R.L. Tweedie, Exponential and uniform ergodicity of Markov processes. Ann. Prob. **23**(4), 1671–1691 (1995)
19. M. Fukushima, Y. Oshima, M. Takeda, *Dirichlet Forms and Symmetric Markov Processes.* Studies in Mathematics (De Gruyter, Berlin, 1994)
20. L. Gatteschi, New inequalities for the zeros of confluent hypergeometric functions, in *Asymptotic and Computational Analysis (Winnipeg, MB, 1989)*. Lecture Notes in Pure and Applied Mathematics, vol. 124 (Dekker, New York, 1990), pp. 175–192
21. A. Göing-Jaeschke, M. Yor, A survey and some generalizations of Bessel processes. Bernoulli **9**(2), 313–349 (2003)
22. P. Graczyk, T. Jakubowski, Exit times and Poisson kernels of the Ornstein-Uhlenbeck diffusion. Stoch. Models **24**(2), 314–337 (2008)
23. T. Hargé, Diffraction pour l'équation de la chaleur. Duke Math. J. **73**(3), 713–744 (1994)
24. N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North Holland, Amsterdam, 1981)

25. A.M. Kulik, T.D. Tymoshkevych, Lift zonoid order and functional inequalities. Teor. Imovir. Mat. Stat. **89**, 75–90 (2013). Reprinted in Theory Probab. Math. Stat. **89**, 83–99 (2014)
26. E.H. Lieb, R. Seiringer, J. Yngvason, Poincaré inequalities in punctured domains. Ann. Math. **158**, 1067–1080 (2003)
27. D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion* (Springer, Berlin, 1991)
28. L.J. Slater, The real zeros of the confluent hypergeometric function. Proc. Camb. Philos. Soc. **52**, 626–635 (1956)

Chapter 2
A Probabilistic Look at Conservative Growth-Fragmentation Equations

Florian Bouguet

Abstract In this note, we consider general growth-fragmentation equations from a probabilistic point of view. Using Foster-Lyapunov techniques, we study the recurrence of the associated Markov process depending on the growth and fragmentation rates. We prove the existence and uniqueness of its stationary distribution, and we are able to derive precise bounds for its tails in the neighborhoods of both 0 and $+\infty$. This study is systematically compared to the results obtained so far in the literature for this class of integro-differential equations.

Keywords Growth-fragmentation · Markov process · Stationary measure · Tail of distribution · Foster-Lyapunov criterion.

2.1 Introduction

In this work, we consider the growth and fragmentation of a population of microorganisms (typically, bacteria or cells) through a structured quantity x which rules the division. For instance, one can consider x to be the size of a bacterium. The bacteria grow and, from time to time, split into two daughters. This behavior leads to an integro-differential equation, which can also model numerous phenomena involving fragmentation, like polymerization, network congestions or neurosciences. In the context of a dividing population, we refer to [22, Chapter 4] for background and biological motivations, and to [11, 21] for motivations in determining the eigenelements of the equation, which correspond to the Malthusian parameter of the population (see [22]). Regardless, if we denote by $u(t, x)$ the concentration of individuals of size x at time t, such dynamics lead to the following *growth-*

F. Bouguet (✉)
Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, France

CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, France

Inria, Villers-lès-Nancy, France
e-mail: florian.bouguet@ens-rennes.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_2

fragmentation equation:

$$\partial_t u(t,x) + \partial_x[\tau(x)u(t,x)] + \beta(x)u(t,x) = 2\int_x^\infty \beta(y)\kappa(x,y)u(t,y)dy, \tag{2.1}$$

for $x, t > 0$, where τ and β are the respective growth rate and fragmentation rate of the population, and κ is the fragmentation kernel (here we adopt the notation of [7]). Not that because of the factor 2 in the right-hand side of (2.1), the mass of the total particle system increases with time, so that this equation is not conservative.

The evolution of this population, or rather its probabilistic counterpart, has also been widely studied for particular growth and division rates. In the context of network congestions, this is known as the TCP window size process, which received a lot of attention recently (see [1, 4, 8, 18]). Let us provide the probabilistic interpretation of this mechanism. Consider a bacterium of size X, which grows at rate τ and randomly splits at rate β following a kernel κ, as before. We shall denote by $Q(x,dy) := x\kappa(yx,x)dy$ to deal with the relative size of the daughters compared to the mother's, so that

$$\int_0^x f(y)\kappa(y,x)dy = \int_0^1 f(xy)Q(x,dy).$$

We shall naturally assume that, for any $x > 0$,

$$\int_0^x \kappa(y,x)dy = \int_0^1 Q(x,dy) = 1.$$

If we dismiss one of the two daughters and carry on the study only with the other one, the growth and fragmentation of the population can also be modeled by a *piecewise deterministic Markov process* (PDMP) $(X_t)_{t\geq 0}$ with càdlàg trajectories a.s. The dynamics of X are ruled by its infinitesimal generator, defined for any function f in its domain $\mathcal{D}(\mathcal{L})$:

$$\mathcal{L}f(x) := \tau(x)f'(x) + \beta(x)\int_0^1 [f(xy) - f(x)]Q(x,dy). \tag{2.2}$$

We shall call X a *cell process* (not to be confused with a growth-fragmentation process, see Remark 2.2). It is a Feller process, and we denote by $(P_t)_{t\geq 0}$ its semigroup (for reminders about Feller processes or PDMPs, see [9, 14]). If we denote by $\mu_t = \mathscr{L}(X_t)$ the probability law of X_t, the Kolmogorov's forward equation $\partial_t(P_t f) = \mathcal{L}P_t f$ is the weak formulation of

$$\partial_t \mu_t = \mathcal{L}'\mu_t, \tag{2.3}$$

where $\mathcal{L}'$ is the adjoint operator of $\mathcal{L}$ in $L^2(\mathbb{L})$ where $\mathbb{L}$ stands for the Lebesgue measure. Now, if μ_t admits a density $u(t, \cdot)$ with respect to $\mathbb{L}$, then (2.3) writes

$$\partial_t u(t,x) = \mathcal{L}' u(t,x) = -\partial_x[\tau(x)u(t,x)] - \beta(x)u(t,x) + \int_x^\infty \beta(y)\kappa(x,y)u(t,y)dy. \tag{2.4}$$

Note that (2.4) is the conservative version of (2.1), since for any $t \geq 0$, $\int_0^\infty u(t,x)dx = 1$, which comes from the fact that there is only one bacterium at a time.

Remark 2.1 (Link with Biology) Working with the probabilistic version of the problem allows us not to require the absolute continuity of μ_t nor $Q(x, \cdot)$. This is useful since many biological models set $Q(x, \cdot) = \delta_{1/2}$ (equal mitosis) or $Q(x, \cdot) = \mathscr{U}([0, 1])$ (uniform mitosis). Note that biological models usually assume that $\int_0^1 yQ(x, dy) = 1/2$, so that the mass of the mother is conserved after the fragmentation, which is automatically satisfied for a symmetric division, but we do not require this hypothesis in our study. We stress that it is possible to study both daughters with a structure of random tree, as in [3, 5, 12], the latter also drawing a bridge between the stochastic and deterministic viewpoints.

In the articles [2, 7, 11], the authors investigate the behavior of the first eigenvalue and eigenfunction of (2.1), with a focus on the dependence on the growth rate τ and the division rate β. Although it has been previously done for specific rates (e.g. [15]), they work in the setting of general functions τ and β. The aim of the present paper is to provide a probabilistic counterpart to the aforementioned articles, by studying the Markov process $(X_t)_{t\geq 0}$ generated by (2.2), and to explain the assumptions for the well-posedness of the problem. We provide a probabilistic justification to the links between the growth and fragmentation rates, with the help of the renowned Foster-Lyapunov criterion. We shall also study the tails of distribution of the stationary measure of the process when it exists. We will see that, although the assumptions are similar, there is a difference between the tails of the stationary distribution in the conservative case and in the non-conservative case.

Remark 2.2 (Cell Processes and Growth-Fragmentation Processes) The name cell process comes from the paper [5], where the author provides a general construction for the so-called growth-fragmentation processes, with the structure of branching processes. This construction allows to study the family of all bacteria (or cells) alive at time t. Let us stress that, in [5], the process is allowed to divide on a dense set of times; the setting of PDMPs does not capture such a phenomenon, but does not require the process X to converge a.s. at infinity. The construction of a growth-fragmentation process is linked to the study of the non-conservative growth-fragmentation equation (2.1), whereas our construction of the cell process enables us to study the behavior of its conservative version (2.4). In Sect. 2.2, we shall see that there is no major differences for the well-posedness of the equation in the conservative and the non-conservative settings. However, the tails of the stationary distribution are rather different in the two frameworks, so the results of Sect. 2.3 are to be compared to the computations of [2] when $\lambda = 0$. The Malthusian parameter λ being the exponential growth rate of the mass of the population, it is clear that it is null in the conservative case.

The rest of this paper is organized as follows: in Sect. 2.2, we study the Harris recurrence of X as well as the existence and uniqueness of its stationary distribution π, and we compare our conditions to those of [7]. In Sect. 2.3, we study the moments of π, we derive precise upper bounds for its tails of distribution in the neighborhoods of both 0 and $+\infty$ and we compare our conditions to those of [2].

2.2 Balance Between Growth and Fragmentation

To investigate the assumptions used in [7], we turn to the study of the Markov process generated by (2.2). More precisely, we will provide a justification to the balance between τ and β with the help of a Foster-Lyapunov criterion. Note that we shall not require the fragmentation kernel $Q(x, dy)$ to admit a density with respect to the Lebesgue measure $\mathbb{L}(dy)$. Moreover, in order to be as general as possible, we do not stick to the biological framework and thus do not assume that $\int_0^1 Q(x, dy) = 1/2$, which will be (technically) replaced by Assumption 2.2.i) below.

We start by stating general assumptions on the growth and fragmentation rates.

Assumption 2.1 (Behavior of τ and β) *Assume that:*

i) The functions β and τ are continuous, and τ is locally Lipschitz.
ii) For any $x > 0$, $\beta(x), \tau(x) > 0$.
iii) There exist constants $\gamma_0, \gamma_\infty, \nu_0, \nu_\infty$ and $\beta_0, \beta_\infty, \tau_0, \tau_\infty > 0$ such that

$$\beta(x) \underset{x\to 0}{\sim} \beta_0 x^{\gamma_0}, \quad \beta(x) \underset{x\to\infty}{\sim} \beta_\infty x^{\gamma_\infty}, \quad \tau(x) \underset{x\to 0}{\sim} \tau_0 x^{\nu_0}, \quad \tau(x) \underset{x\to\infty}{\sim} \tau_\infty x^{\nu_\infty}.$$

Note that, if τ and β satisfy Assumption 2.1, then Assumptions (2.18) and (2.19) in [7] are fulfilled (by taking $\mu = |\gamma_\infty|$ or $\mu = |\nu_\infty|$, and $r_0 = |\nu_0|$ therein).

The following assumption concerns the expected behavior of the fragmentation, and is easy to check in most cases, especially if $Q(x, \cdot)$ does not depend on x. For any $a \in \mathbb{R}$, we define the moment of order a of $Q(x, \cdot)$ by

$$M_x(a) := \int_0^1 y^a Q(x, dy), \quad M(a) := \sup_{x>0} M_x(a).$$

Assumption 2.2 (Moments of Q) *Assume that:*

i) There exists $a > 0$ such that $M(a) < 1$.
ii) There exists $b > 0$ such that $M(-b) < +\infty$.
iii) For any $x > 0$, $Q(x, \{1\}) = 0$.

Note that, in particular, Assumption 2.2.i) and ii) imply that, for any $x > 0$, $Q(x, \{1\}) < 1$ and $Q(x, \{0\}) = 0$. Assumption 2.2.iii) means that there are no *phantom jumps*, i.e. divisions of the bacteria without loss of mass. It is easy to deal with a process with phantom jumps with the following thinning technique: if X is generated by (2.2) and Q admits the decomposition

$$Q(x, dy) = Q(x, \{1\})\delta_1 + (1 - Q(x, \{1\}))Q'(x, dy),$$

then notice that (2.2) writes

$$\mathcal{L}f(x) = \tau(x)f'(x) + \beta'(x)\int_0^1 [f(xy) - f(x)]Q'(x, dy),$$

with $\beta'(x) = (1 - Q(x, \{1\})\beta(x)$ and $Q'(x, \{1\}) = 0$.

Let us make another assumption, concerning the balance between the growth rate and the fragmentation rate in the neighborhoods of 0 and $+\infty$, which is fundamental to obtain an interesting Markov process.

Assumption 2.3 (Balance of β and τ) *Assume that*

$$\gamma_0 > \nu_0 - 1, \quad \gamma_\infty > \nu_\infty - 1.$$

Let us mention that Assumptions 2.1.iii) and 2.3, could be replaced by integrability conditions in the neighborhoods of 0 or $+\infty$, see Assumptions (2.21) and (2.22) in [7]. However, we make those hypotheses for the sake of simplicity, and for easier comparisons of our results to [2, 7].

Remark 2.3 (The Critical Case) This remark concerns the whole paper, and may be omitted at first reading. Throughout Sect. 2.2, we can weaken Assumption 2.3 with the following:

i) Either

$$\gamma_0 > \nu_0 - 1, \quad \text{or} \quad \gamma_0 = \nu_0 - 1 \text{ and } \frac{b}{M(-b) - 1} < \frac{\beta_0}{\tau_0}. \tag{2.5}$$

ii) Either

$$\gamma_\infty > \nu_\infty - 1 \quad \text{or} \quad \gamma_\infty = \nu_\infty - 1 \text{ and } \frac{a}{1 - M(a)} < \frac{\tau_\infty}{\beta_\infty}. \tag{2.6}$$

Indeed, a careful reading of the proof of Theorem 2.4 shows that computations are similar, and the only change lies in the coefficients in (2.13) and (2.14), which are still negative under (2.5) and (2.6). This corresponds to the critical case of the growth-fragmentation equations (see for instance [6, 10]).

However, the behavior of the tail of the stationary distribution changes radically in the critical case. As a consequence, Sect. 2.3 is written in the framework of Assumption 2.3 only. Indeed, it is crucial to be able to choose a as large as possible (which is ensured in Assumption 2.8), so that π admits moments of any order. This is not possible under (2.6), since then

$$\lim_{a\to+\infty} \frac{a}{1-M(a)} = +\infty,$$

so that the Foster-Lyapunov criterion does not apply and we expect the stationary measure to have heavy tails.

Define V as a smooth, convex function on $(0,\infty)$ such that

$$V(x) = \begin{cases} x^{-b} & \text{if } x \in (0,1], \\ x^{a} & \text{if } x \in [2,+\infty), \end{cases} \tag{2.7}$$

where a and b satisfy Assumption 2.2. We can now state the main result of this article.

Theorem 2.4 (Behavior of the Cell Process) *Let X be a PDMP generated by* (2.2)*. If Assumptions 2.1–2.3 are in force, then X is irreducible, Harris recurrent and aperiodic, compact sets are petite for X, and the process possesses a unique (up to a multiplicative constant) stationary measure π.*

Moreover, if

$$b \geq \nu_0 - 1, \quad a \geq -\gamma_\infty,$$

then X is positive Harris recurrent and π is a probability measure.

Furthermore, if

$$\nu_0 \leq 1, \quad \gamma_\infty \geq 0,$$

then X is exponentially ergodic in $(1+V)$*-norm.*

Remark 2.4 (Link with the Conditions of [7]) We highlight the equivalence of Assumption 2.3 and [7, Eq. (2.4) and (2.5)]. The condition [7, Eq. (2.6)] writes in our context

$$\int_0^u Q(x,dy) \leq \min\left(1, Cu^{\bar{\gamma}}\right),$$

which is implied by Assumption 2.2.ii) together with the condition $b \geq \nu_0 - 1$, as soon as $Q(x,\cdot)$ admits a density with respect to $\mathbb{L}$. Let us also mention that counterexamples for the existence of the stationary measure are provided in [11], when β is constant and τ is affine.

Before proving Theorem 2.4, let us shortly present the Foster-Lyapunov criterion, which is the main tool for our proof (the interested reader may find deeper insights in [19] or [20]). The idea is to find a so-called Lyapunov function V controlling the excursions of X out of petite sets. Recall that a set $K \subseteq \mathbb{R}_+$ is petite if there exists a probability distribution $\mathscr{A}$ over $\mathbb{R}_+$ and some non-trivial positive measure ν over $\mathbb{R}_+$ such that, for any $x \in K$, $\int_0^\infty \delta_x P_t \mathscr{A}(dt) \geq \nu$. We produce here three criteria, adapted from [20, Theorems 3.2, 4.2 and 6.1], which provide stronger and stronger results. Recall that, for some norm-like function V, we define the V-norm of a probability measure μ by

$$\|\mu\|_V := \sup_{|f| \leq V} |\mu(f)| = \sup_{|f| \leq V} \left| \int f d\mu \right|.$$

Theorem 2.5 (Foster-Lyapunov Criterion) *Let X be a Markov process with càdlàg trajectories a.s. Let $V \geq 1$ be a continuous norm-like real-valued function. Assume that compact sets of $(0, +\infty)$ are petite for X.*

i) If there exist a compact set K and a positive constant α' such that

$$\mathcal{L}V \leq \alpha' \mathbb{1}_K,$$

then X is Harris recurrent and possesses a unique (up to a multiplicative constant) stationary measure π.

ii) Moreover, if there exist a function $f \geq 1$ and a positive constant α such that

$$\mathcal{L}V \leq -\alpha f + \alpha' \mathbb{1}_K,$$

then X is positive Harris recurrent, π is a probability measure and $\pi(f) < +\infty$.

iii) Moreover, if $f \geq V$, then X is exponentially ergodic and there exist $C, v > 0$ such that

$$\|\mu_t - \pi\|_{1+V} \leq C(1 + \mu_0(V))e^{-vt}.$$

Note that the exponential rate v provided in Theorem 2.5 is not explicit; if one wants to obtain quantitative speeds of convergence, it is often useful to turn to ad hoc coupling methods (see [4] for instance). Also, note that Assumption 2.2 is sufficient but not necessary to derive ergodicity from a Foster-Lyapunov criterion, since we only need the limits in (2.13) and (2.14) to be negative. Namely, we only ask the fragmentation kernel $Q(x, \cdot)$ to be not too close to 0 and 1, uniformly over x.

Remark 2.5 (Construction of V) If we are able to prove a Foster-Lyapunov criterion with a norm-like function V, we want to choose V as explosive as possible (i.e. such that $V(x)$ goes quickly to $+\infty$ when $x \to 0$ or $x \to +\infty$) to obtain better bounds for the tail of π, since $\pi(V)$ is finite: this is the purpose of Sect. 2.3. If we define V with (2.7), this choice brings us to choose a and b as large as possible in Assumption 2.2. However, the larger a and b, the slower the convergence (because of the term $\mu_0(V)$), so there is a balance to find here.

For many particular cell processes, it is possible to build a Lyapunov function of the form $x \mapsto e^{\theta x}$, so that π admits exponential moments up to θ. We shall use a similar function in Sect. 2.3 to obtain bounds for the tails of the stationary distribution.

Proof (of Theorem 2.4) We denote by φ_z the unique maximal solution of $\partial_t y(t) = \tau(y(t))$ with initial condition z, and let $a, b > 0$ be as in Assumption 2.2. Firstly, we prove that compact sets are petite for $(X_t)_{t\geq 0}$. Let $z_2 > z_1 > z_0 > 0$ and $z \in [z_0, z_1]$. Since $\tau > 0$ on $[z_0, z_2]$, the function φ_z is a diffeomorphism from $[0, \varphi_z^{-1}(z_2)]$ to $[z, z_2]$; let $t = \varphi_{z_0}^{-1}(z_2)$ be the maximum time for the flow to reach z_2 from $[z_0, z_1]$. Denote by X^z the process generated by (2.2) such that $\mathscr{L}(X_0) = \delta_z$, and T_n^z the epoch of its nth jump. Let $\mathscr{A} = \mathscr{U}([0, t])$. For any $x \in [z_1, z_2]$, we have

$$\begin{aligned}
\int_0^\infty \mathbb{P}(X_s^z \leq x)\mathscr{A}(ds) &\geq \frac{1}{t}\int_0^t \mathbb{P}(X_s^z \leq x | T_1^z > \varphi_z^{-1}(z_2))\mathbb{P}(T_1^z > \varphi_z^{-1}(z_2))ds \\
&\geq \frac{\mathbb{P}(T_1^z > \varphi_z^{-1}(z_2))}{t}\int_0^t \mathbb{P}(\varphi_z(s) \leq x)ds \\
&\geq \frac{\mathbb{P}(T_1^z > \varphi_z^{-1}(z_2))}{t}\int_0^{\varphi_z^{-1}(x)} ds \\
&\geq \frac{\mathbb{P}(T_1^z > \varphi_z^{-1}(z_2))}{t}\int_z^x (\varphi_z^{-1})'(u)du. \qquad (2.8)
\end{aligned}$$

Since β and τ are bounded on $[z_0, z_2]$, the following inequalities hold:

$$\begin{aligned}
\mathbb{P}(T_1^z > \varphi_z^{-1}(z_2)) &= \exp\left(-\int_0^{\varphi_z^{-1}(z_2)} \beta(\varphi_z(s))ds\right) \\
&= \exp\left(-\int_z^{z_2} \beta(u)(\varphi_z^{-1})'(u)du\right) \\
&\geq \exp\left(-(z_2 - z_0)\sup_{[z_0,z_2]}\left(\beta(\varphi_z^{-1})'\right)\right) \\
&\geq \exp\left(-(z_2 - z_0)\left(\sup_{[z_0,z_2]}\beta\right)\left(\inf_{[z_0,z_2]}\tau\right)^{-1}\right),
\end{aligned}$$

since $\sup_{[z_0,z_2]}(\varphi_z^{-1})' = \left(\inf_{[z_0,z_2]}\tau\right)^{-1}$. Hence, there exists a constant C such that, (2.8) writes, for $x \in [z_1, z_2]$,

$$\int_0^\infty \mathbb{P}(X_s^z \leq x)\mathscr{A}(ds) \geq C(x - z_1),$$

which is also

$$\int_0^\infty \delta_z P_s \mathscr{A}(ds) \geq C\mathbb{L}_{[z_1,z_2]},$$

where $\mathbb{L}_K$ is the Lebesgue measure restricted to a Borelian set K. Hence, by definition, $[z_0, z_1]$ is a petite set for the process X.

Now, let us show that the process (X_t) is $\mathbb{L}_{(0,\infty)}$-irreducible with similar arguments. Let $z_1 > z_0 > 0$ and $z > 0$. If $z \leq z_0$,

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{z_0 \leq X_t^z \leq z_1\}} dt\right] &\geq \mathbb{P}(T_1^z > \varphi_z^{-1}(z_1))\mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{z_0 \leq X_t^z \leq z_1\}} dt \,\middle|\, T_1^z > \varphi_z^{-1}(z_1)\right] \\ &\geq \exp\left(-(z_1 - z_0)\left(\sup_{[z_0,z_1]} \beta\right)\left(\inf_{[z_0,z_1]} \tau\right)^{-1}\right)\varphi_{z_0}^{-1}(z_1).\end{aligned} \tag{2.9}$$

If $z > z_0$, for any $t_0 > 0$ and $n \in \mathbb{N}$, the process X^z has a positive probability of jumping n times before time t_0. Recall that $\int_0^1 y^a Q(x, dy) \leq M(a) < 1$. For any $n > (\log(z) - \log(z_0)) \log(M(a)^{-1})^{-1}$, let $0 < \varepsilon < z_0^a - (zM(a)^n)^a$. By continuity of $(x, t) \mapsto \varphi_x(t)$, there exists $t_0 > 0$ small enough such that, $\forall (x, t) \in [0, z] \times [0, t_0]$,

$$\varphi_x(t)^a \leq x^a + \frac{\varepsilon}{n+1}, \quad \mathbb{E}[(X_{t_0}^z)^a | T_n^z \leq t_0] \leq (zM(a)^n)^a + \varepsilon < z_0^a.$$

Then, using Markov's inequality

$$\mathbb{P}(X_{t_0}^z \leq z_0 | T_n^z \leq t_0 < T_{n+1}^z) \geq 1 - \frac{\mathbb{E}[(X_{t_0}^z)^a | T_n^z \leq t_0 < T_{n+1}^z]}{z_0^a} > 0.$$

Then, $\mathbb{P}(X_{t_0}^z \leq z_0) > 0$ for any t_0 small enough, and, using (2.9)

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{z_0 \leq X_t^z \leq z_1\}} dt\right] &\geq \mathbb{E}\left[\int_{t_0}^\infty \mathbb{1}_{\{z_0 \leq X_t^z \leq z_1\}} dt \,\middle|\, X_{t_0}^z \leq z_0\right]\mathbb{P}(X_{t_0}^z \leq z_0) \\ &\geq \exp\left(-(z_1 - z_0)\left(\sup_{[z_0,z_1]} \beta\right)\left(\inf_{[z_0,z_1]} \tau\right)^{-1}\right) \\ &\quad \times \varphi_{z_0}^{-1}(z_1)\mathbb{P}(X_{t_0}^z \leq z_0) \\ &> 0.\end{aligned}$$

Aperiodicity is easily proven with similar arguments.

We turn to the proof of the Lyapunov condition. For $x \geq 2$, $V(x) = x^a$ and

$$\begin{aligned}
\mathcal{L}V(x) &= a\frac{\tau(x)}{x}V(x) + \beta(x)\int_0^1 V(xy)Q(x,dy) - \beta(x)V(x) \\
&\leq \left(a\frac{\tau(x)}{x} - \beta(x)\right)V(x) + \beta(x)\int_0^{1/x}(xy)^{-b}Q(x,dy) \\
&\quad + \beta(x)\int_{1/x}^{2/x} 2^a Q(x,dy) + \beta(x)\int_{2/x}^{1}(xy)^a Q(x,dy) \\
&\leq \left(a\frac{\tau(x)}{x} - \beta(x)\right)V(x) + \beta(x)\left(x^{-b}M_x(-b) + 2^a + x^a M_x(a)\right) \\
&\leq \left(a\frac{\tau(x)}{x} - \beta(x)\left(1 - M_x(a) - \frac{M_x(-b)}{x^b V(x)} - \frac{2^a}{V(x)}\right)\right)V(x). \qquad (2.10)
\end{aligned}$$

For $x \leq 1$, $V(x) = x^{-b}$ and

$$\mathcal{L}V(x) = \left(-b\frac{\tau(x)}{x} + \beta(x)(M_x(-b) - 1)\right)V(x). \qquad (2.11)$$

Combining $\gamma_\infty > \nu_\infty - 1$ with Assumption 2.2.i), for x large enough we have

$$\begin{aligned}
&a\frac{\tau(x)}{x} - \beta(x)\left(1 - M_x(a) - \frac{M_x(-b)}{x^b V(x)} - \frac{2^a}{xV(x)}\right) \\
&\leq a\frac{\tau(x)}{x} - \beta(x)\left(1 - M(a) + o(1)\right) \leq 0.
\end{aligned}$$

Likewise, combining $\gamma_0 > \nu_0 - 1$ with Assumption 2.2.ii),

$$-b\frac{\tau(x)}{x} + \beta(x)(M_x(-b) - 1) \leq -b\frac{\tau(x)}{x} + \beta(x)\left(M(-b) - 1\right) \leq 0$$

for x close enough to 0. Then, Theorem 2.5.i) entails that X is Harris recurrent, thus admits a unique stationary measure (see for instance [17]).

Note that (2.10) writes

$$\mathcal{L}V(x) \leq -\beta_\infty(1 - M(a) + o(1))x^{a+\gamma_\infty},$$

so that, if we can choose $a \geq -\gamma_\infty$, then

$$\lim_{x\to\infty} -\beta_\infty(1 - M(a) + o(1))x^{a+\gamma_\infty} < 0.$$

Likewise, (2.11) writes

$$\mathcal{L}V(x) \le -(b\tau_0 + o(1))x^{\nu_0-1-b}, \tag{2.12}$$

so, if $b \ge \nu_0 - 1$, we get

$$\lim_{x\to 0} -(b\tau_0 + o(1))x^{\nu_0-1-b} < 0.$$

Then, there exist positive constants A, α, α'

$$\mathcal{L}V \le -\alpha f + \alpha' \mathbb{1}_{[1/A,A]},$$

where $f \ge 1$ is a smooth function, such that $f(x) = x^{\nu_0-1-b}$ for x close to 0, and $f(x) = x^{a+\gamma_\infty}$ for x large enough. Then, Theorem 2.5.ii) ensures positive Harris recurrence for X.

Now, if we assume $\gamma_\infty \ge 0$ and $\nu_0 \le 1$ in addition, then there exists $\alpha > 0$ such that

$$\begin{aligned}&\lim_{x\to+\infty} a\frac{\tau(x)}{x} - \beta(x)\left(1 - M_x(a) - \frac{M_x(-b)}{x^b V(x)} - \frac{2^a}{xV(x)}\right)\\ &\le \lim_{x\to+\infty} a\frac{\tau(x)}{x} - \beta(x)\left(1 - M(a) + o(1)\right) \le -\alpha,\end{aligned} \tag{2.13}$$

and

$$\lim_{x\to 0} -b\frac{\tau(x)}{x} + \beta(x)(M_x(-b) - 1) \le \lim_{x\to 0} -b\frac{\tau(x)}{x} + \beta(x)\left(M(-b) - 1\right) \le -\alpha. \tag{2.14}$$

Combining (2.10) and (2.11) with (2.13) and (2.14) respectively, and since V is bounded on $[1, 2]$, there exist positive constants A, α' such that

$$\mathcal{L}V \le -\alpha V + \alpha' \mathbb{1}_{[1/A,A]}.$$

The function V is thus a Lyapunov function, for which Theorem 2.5.iii) entails exponential ergodicity for X.

2.3 Tails of the Stationary Distribution

In this section, we use, and reinforce when necessary, the results of Theorem 2.4 to study the asymptotic behavior of the tails of distribution of the stationary measure π. We will naturally divide this section into two parts, to study the behavior of $\pi(dx)$

as $x \to 0$ and as $x \to +\infty$. Hence, throughout this section, we shall assume that X satisfies Assumptions 2.1–2.3. The key point is to use the fact that $\pi(f) < +\infty$ provided in the second part of Theorem 2.5. We recall that $\mathbb{L}$ stands for the Lebesgue measure on $\mathbb{R}$.

In order to compare our results to those of [2, Theorem 1.8], we consider the same framework and make the following assumption:

Assumption 2.6 (Density of Q and π) *Assume that:*

i) For any $x > 0$, $Q(x, \cdot) \ll \mathbb{L}$ and $Q(x, dy) = q(y)dy$, and there exist constants $q_0, q_1 \geq 0$ and $\mu_0, \mu_1 > -1$ such that

$$q(x) \underset{x\to 0}{=} q_0 x^{\mu_0} + o(x^{\mu_0}), \quad q(x) \underset{x\to 1}{=} q_1(1-x)^{\mu_1} + o((1-x)^{\mu_1}).$$

ii) $\pi \ll \mathbb{L}$ and $\pi(dx) = G(x)dx$, and there exist constants $G_0, G_\infty, \widetilde{G}_\infty > 0$ and $\alpha_0, \alpha_\infty, \widetilde{\alpha}_\infty \in \mathbb{R}$ such that

$$G(x) \underset{x\to 0}{\sim} G_0 x^{\alpha_0}, \quad G(x) \underset{x\to +\infty}{\sim} G_\infty x^{\alpha_\infty} \exp\left(-\widetilde{G}_\infty x^{\widetilde{\alpha}_\infty}\right).$$

We do not require the coefficients q_0, q_1 to be (strictly) positive, so that this assumption can also cover the case $Q(x, dy) = \delta_r(dy)$ for $0 < r < 1$, which is widely used for modeling physical or biological situations. For the sake of simplicity, the hypotheses concerning the density of π (resp. Q) in the neighborhood of both 0 and $+\infty$ (resp. 0 and 1) are gathered in Assumption 2.6, but it is clear that we only need either the assumption on the left behavior or on the right behavior to precise the fractional moments or the exponential moments of the stationary distribution. In the same spirit, we could weaken $Q(x, \cdot) \ll \mathbb{L}$ into $Q(x, \cdot)$ admitting a density with respect to $\mathbb{L}$ only in the neighborhoods of 0 and 1, bounded above by q.

Remark 2.6 (Absolute Continuity of π) At first glance, Assumption 2.6.ii) may seem disconcerting since π is unknown; this is the very goal of this section to study its moments. However, for some models it is possible to prove the absolute continuity of π, or even get a non-tractable formula for its density (see e.g. [4, 13, 16] for the particular case of the TCP window size process). In such cases, the question of existence of its moments is still not trivial. Still, Assumption 2.6.ii) is stated only to make easier comparisons with the estimates obtained with deterministic methods, and is not needed for the important results of the present paper. However, we stress that Assumption 2.6.i) is in a way more fundamental, since it implies directly Assumption 2.8, which is needed to study the behavior of π in the neighborhood of $+\infty$ (see Proposition 2.9).

Theorem 2.7 (Negative Moments of π) *Let X be the PDMP generated by* (2.2). *If Assumptions 2.1–2.3 hold, and if*

$$b \geq \nu_0 - 1, \tag{2.15}$$

then

$$\int_0^1 x^{\nu_0-1-b}\pi(dx) < +\infty.$$

Moreover, if Assumption 2.6 holds and $\mu_0 + 2 - \nu_0 > 0$*, then*

$$\alpha_0 \geq \mu_0 + 1 - \nu_0.$$

Proof The first part of the theorem is a straightforward consequence of (2.12).

Combining Assumption 2.2.ii) with Assumption 2.6.i), we naturally have to take $b < \mu_0 + 1$. Thus, for any $\varepsilon \in (0, \mu_0 + 1)$, we take $b = \mu_0 + 1 - \varepsilon$. Define V with (2.7) as before, so that, for $x \leq 1$, $V(x) = x^{-\mu_0-1+\varepsilon}$ and

$$\mathcal{L}V(x) \leq \left(-b\frac{\tau(x)}{x} + \beta(x)(M(-b) - 1)\right) V(x) \underset{x\to 0}{\sim} -b\tau_0 x^{\nu_0-\mu_0-2+\varepsilon}.$$

Applying Theorem 2.5.ii) with $f(x) = x^{\nu_0-\mu_0-2+\varepsilon}$, which tends to $+\infty$ when $\mu_0 + 1 - \nu_0 > -1$, we have $\pi(f) < +\infty$ so

$$\int_0^1 x^{\alpha_0+\nu_0-\mu_0-2+\varepsilon} dx < +\infty, \quad \alpha_0 > 1 + \mu_0 - \nu_0 + \varepsilon,$$

for any $\varepsilon > 0$. Thus $\alpha_0 \geq \mu_0 + 1 - \nu_0$.

Now, we turn to the study of the tail of distribution of $\pi(dx)$ as $x \to +\infty$. Since choosing a polynomial function as a Lyapunov function can only provide the existence of moments for π, we need to introduce a more coercive function to study in detail the behavior of its tail of distribution and get the existence of exponential moments. We begin with the following assumption.

Assumption 2.8 (Uniform Asymptotic Bound of the Fragmentation) *Let* $\theta = \gamma_\infty + 1 - \nu_\infty$*. Assume there exists* $0 < C < 1$ *such that, for any* $\varepsilon > 0$ *and* $0 < \eta < \beta_\infty(\theta\tau_\infty)^{-1}$*, there exists* $x_0 > 0$ *such that*

$$\sup_{x \geq x_0} \int_0^1 y^{-\varepsilon} \exp\left(\eta x^\theta (y^\theta - 1)\right) Q(x, dy) < 1 - C.$$

It is easy to understand this assumption if

$$\widetilde{V}(x) = x^{-\varepsilon} e^{\eta x^\theta} \tag{2.16}$$

and if $\mathscr{L}(Y^{(x)}) = Q(x, .)$; then, Assumption 2.8 rewrites

$$\sup_{x \geq x_0} \frac{\mathbb{E}[\widetilde{V}(xY^{(x)})]}{\widetilde{V}(x)} \leq 1 - C.$$

Once again, this is asking the fragmentation kernel to be not too close to 1. As we will see, this is quite natural when Q has a regular behavior around 0 and 1.

Proposition 2.9 *Assumption 2.8 holds for any $C \in (0, 1)$ whenever Assumption 2.6.i) holds.*

Proof Define $\widetilde{V}$ as in (2.16) for $x \geq 1$, larger than 1, and increasing and smooth on $\mathbb{R}$. For any (large) $x > 0$, for any (small) $\delta > 0$,

$$\begin{aligned}
\mathbb{E}[\widetilde{V}(xY^{(x)})] &= \mathbb{E}[\widetilde{V}(xY^{(x)})|Y^{(x)} \leq 1 - \delta]\mathbb{P}(Y^{(x)} \leq 1 - \delta) \\
&\quad + \mathbb{E}\left[\widetilde{V}(xY^{(x)})\mathbb{1}_{\{Y^{(x)} > 1 - \delta\}}\right] \\
&\leq \widetilde{V}((1 - \delta)x) + \mathbb{E}\left[\widetilde{V}(xY^{(x)})\mathbb{1}_{\{Y^{(x)} > 1 - \delta\}}\right].
\end{aligned} \tag{2.17}$$

It is clear that, for $\delta < 1$,

$$\lim_{x \to +\infty} \frac{\widetilde{V}((1 - \delta)x)}{\widetilde{V}(x)} = \lim_{x \to +\infty} (1 - \delta)^{-\varepsilon} \exp\left(-\eta(1 - (1 - \delta)^{\theta})x^{\theta}\right) = 0.$$

On the other hand, using Hölder's inequality with $q > \max(1, -1/\mu_1)$ and $p^{-1} + q^{-1} = 1$, as well as a Taylor expansion, there exists some constant $C_\delta \geq 1$ such that

$$\begin{aligned}
\int_{1-\delta}^{1} \widetilde{V}(xy) q_1 (1 - y)^{\mu_1} dy &= q_1 \widetilde{V}(x) \int_0^{\delta} \exp\left(\eta x^{\theta}((1 - y)^{\theta} - 1)\right) y^{\mu_1}(1 - y)^{-\varepsilon} dy \\
&\leq \frac{q_1}{(1 - \delta)^{\epsilon}} \left[\int_0^{\delta} y^{q\mu_1} dy\right]^{1/q} \\
&\quad \times \widetilde{V}(x) \left[\int_0^{\delta} \exp\left(\eta p x^{\theta}((1 - y)^{\theta} - 1)\right) dy\right]^{1/p} \\
&\leq C_\delta \widetilde{V}(x) \left[\int_0^{\delta} \exp\left(-\eta p \theta x^{\theta} y\right) dy\right]^{1/p} \\
&\leq C_\delta \widetilde{V}(x) \left[\frac{1 - \exp\left(-\eta p \theta x^{\theta} \delta\right)}{\eta p \theta x^{\theta}}\right]^{1/p}.
\end{aligned}$$

The term $(\eta p\theta x^\theta)^{-1}(1-\exp\left(-\eta p\theta x^\theta\delta\right))$ converges to 0 as $x\to+\infty$, so that, for any $C\in(0,1)$, there exists $x_0>0$ such that, for any $x\geq x_0$,

$$\left[\frac{1-\exp\left(-\eta p\theta x^\theta\delta\right)}{\eta p\theta x^\theta}\right]^{1/p}\leq\frac{1-C}{2C_\delta},\qquad\frac{\widetilde{V}((1-\delta)x)}{\widetilde{V}(x)}\leq\frac{1-C}{2}.$$

Plugging these bounds into (2.17) achieves the proof.

Now, we can characterize the weight of the asymptotic tail of π and recover [2, Theorem 1.7].

Theorem 2.10 (Exponential Moments of π) *Let X be the PDMP generated by* (2.2)*. If Assumptions 2.1–2.3 and 2.8 hold, then*

$$\int_1^{+\infty}x^{\nu_\infty-1-\varepsilon}\exp\left(\eta x^\theta\right)\pi(dx)<+\infty,\quad\theta=\gamma_\infty+1-\nu_\infty,\quad\eta=\frac{C\beta_\infty}{\theta\tau_\infty},\quad\varepsilon>0.$$

Moreover, if Assumption 2.6 is also in force, then either:

- $\widetilde{\alpha}_\infty>\gamma_\infty+1-\nu_\infty$;
- $\widetilde{\alpha}_\infty=\gamma_\infty+1-\nu_\infty$ *and* $\widetilde{G}_\infty>C\beta_\infty((\gamma_\infty+1-\nu_\infty)\tau_\infty)^{-1}$;
- $\widetilde{\alpha}_\infty=\gamma_\infty+1-\nu_\infty$, $\widetilde{G}_\infty=C\beta_\infty((\gamma_\infty+1-\nu_\infty)\tau_\infty)^{-1}$ *and* $\alpha_\infty\geq-\nu_\infty$.

Remark 2.7 (Link with the Estimates of [2]) Note that the Assumption 2.3 and (2.15) correspond to the assumptions required for [2, Theorem 1.8] to hold, with the correspondence

$$\mu_0\leftrightarrow\mu-1,\quad\nu_0\leftrightarrow\alpha_0,\quad\mu_0+2-\nu_0>0\leftrightarrow\mu+1-\alpha_0>0.$$

Actually, the authors also assume this strict inequality to prove the existence of the stationary distribution, which we relax here, and we need it only in Theorem 2.15 to provide a lower bound for α_0, which rules the tail of the stationary distribution in the neighborhood of 0. By using Lyapunov methods, there is no hope in providing an upper bound for α_0, but we can see that this inequality is optimal by comparing it to [2, Theorem 1.8] so that, in fact,

$$\alpha_0=\mu_0+1-\nu_0.$$

If $\nu_0>1$, we do not recover the same equivalents for the distribution of G around 0. This is linked to the fact that there is a phase transition in the non-conservative equation at $\nu_0=1$, since the tail of G relies deeply on the function

$$\Lambda(x)=\int_1^x\frac{\lambda+\beta(y)}{\tau(y)}dy,$$

where λ is the Malthusian parameter of the equation, which is the growth of the profiles of the integro-differential equation. However, we deal here with the conservative case, for which this parameter is null. The bounds that we provide are indeed consistent with the computations of the proof of [2, Theorem 1.8] in the case $\lambda = 0$.

Concerning the estimates as $x \to +\infty$, as mentioned above, we can not recover upper bounds, and then sharp estimates, for $\alpha_\infty, G_\infty, \widetilde{\alpha}_\infty$ with Foster-Lyapunov methods. From the proof of Theorem 2.10, it is clear that the parameters η and θ are optimal if one wants to apply Theorem 2.5. Under second-order-type assumptions like [2, Hypothesis 1.5], it is clear that

$$\Lambda(x) = \int_1^x \frac{\beta(y)}{\tau(y)} dy \underset{x\to+\infty}{\sim} \frac{\beta_\infty}{\tau_\infty(\gamma_\infty + 1 - \nu_\infty)} x^{\gamma_\infty+1-\nu_\infty}.$$

This explains the precise value of η, but we pay the price of having slightly less general hypotheses about Q than [2] with a factor C arising from Assumption 2.8, which leads to have no disjunction of cases for α_∞. Also, since we deal with the case $\lambda = 0$, the equivalent of the function Λ is different from the aforementioned paper when $\gamma_\infty < 0$, so that $\max\{\gamma_\infty, 0\}$ does not appear in our computations.

Proof (of Theorem 2.10) Let $\widetilde{V}$ be as in (2.16), that is

$$\widetilde{V}(x) = x^{-\varepsilon} e^{\eta x^\theta},$$

with η, θ given in Theorem 2.10. Then, following the computations of the proof of Theorem 2.4, we get, for $x > x_0$,

$$\begin{aligned}
\mathcal{L}\widetilde{V}(x) &\le \left(\eta\theta\tau_\infty x^{\theta-1+\nu_\infty} - C\beta_\infty x^{\gamma_\infty} - \varepsilon\tau_\infty x^{\nu_\infty-1}\right)(1+o(1))\widetilde{V}(x) \\
&\le -\varepsilon\tau_\infty x^{\nu_\infty-1}(1+o(1))\widetilde{V}(x) \\
&\le -\frac{\varepsilon\tau_\infty}{2} x^{\nu_\infty-1}\widetilde{V}(x).
\end{aligned}$$

Using Theorem 2.5.ii) with $f(x) = x^{\nu_\infty-1}\widetilde{V}(x)$, the last inequality ensures that

$$\int_1^{+\infty} f(x)\pi(dx) < +\infty.$$

Now, in the setting of Assumption 2.6, the following holds:

$$\int_1^{+\infty} f(x)\pi(dx) < +\infty \iff \int_1^{+\infty} x^{\nu_\infty-1-\varepsilon+\alpha_\infty} \exp\left(\eta x^\theta - \widetilde{G}_\infty x^{\widetilde{\alpha}_\infty}\right) dx < +\infty. \tag{2.18}$$

It is clear then that the disjunction of cases of Theorem 2.10 is the only way for the integral on the right-hand side of (2.18) to be finite.

Acknowledgements The author wants to thank Pierre Gabriel for fruitful discussions about growth-fragmentation equations, as well as Eva Löcherbach, Florent Malrieu and Jean-Christophe Breton for their precious help and comments. The referee is also warmly thanked for his constructive remarks. This work was financially supported by the ANR PIECE (ANR-12-JS01-0006-01), and the Centre Henri Lebesgue (programme "Investissements d'avenir" ANR-11-LABX-0020-01).

References

1. R. Azaïs, J.-B. Bardet, A. Génadot, N. Krell, P.-A. Zitt, Piecewise deterministic Markov process—recent results, in *Journées MAS 2012. ESAIM Proceedings*, vol. 44 (EDP Sciences, Les Ulis, 2014), pp. 276–290
2. D. Balagué, J.A. Cañizo, P. Gabriel, Fine asymptotics of profiles and relaxation to equilibrium for growth-fragmentation equations with variable drift rates. Kinet. Relat. Models **6**(2), 219–243 (2013)
3. V. Bansaye, J.-F. Delmas, L. Marsalle, V.C. Tran, Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. Ann. Appl. Probab. **21**(6), 2263–2314 (2011)
4. J.-B. Bardet, A. Christen, A. Guillin, F. Malrieu, P.-A. Zitt, Total variation estimates for the TCP process. Electron. J. Probab. **18**(10), 1–21 (2013)
5. J. Bertoin, Markovian growth-fragmentation processes. HAL e-prints (2015)
6. J. Bertoin, A.R. Watson, Probabilistic aspects of critical growth-fragmentation equations. Adv. Appl. Probab. **48**(A), 37–61 (2016)
7. V. Calvez, M. Doumic, P. Gabriel, Self-similarity in a general aggregation-fragmentation problem. Application to fitness analysis. J. Math. Pures Appl. (9) **98**(1), 1–27 (2012)
8. D. Chafaï, F. Malrieu, K. Paroux, On the long time behavior of the TCP window size process. Stoch. Process. Appl. **120**(8), 1518–1534 (2010)
9. M. Davis, *Markov Models & Optimization*, vol. 49 (CRC Press, Boca Raton, 1993)
10. M. Doumic, M. Escobedo, Time asymptotics for a critical case in fragmentation and growth-fragmentation equations. Kinet. Relat. Models **9**(2), 251–297 (2016)
11. M. Doumic Jauffret, P. Gabriel, Eigenelements of a general aggregation-fragmentation model. Math. Models Methods Appl. Sci. **20**(5), 757–783 (2010)
12. M. Doumic, M. Hoffmann, N. Krell, L. Robert, Statistical estimation of a growth-fragmentation model observed on a genealogical tree. Bernoulli **21**(3), 1760–1799 (2015)
13. V. Dumas, F. Guillemin, P. Robert, A Markovian analysis of additive-increase multiplicative-decrease algorithms. Adv. Appl. Probab. **34**(1), 85–111 (2002)
14. S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (Wiley, New York, 1986)
15. M.L. Greer, P. van den Driessche, L. Wang, G.F. Webb, Effects of general incidence and polymer joining on nucleated polymerization in a model of prion proliferation. SIAM J. Appl. Math. **68**(1), 154–170 (2007)
16. I. Grigorescu, M. Kang, Recurrence and ergodicity for a continuous AIMD model (2009). http://www.math.miami.edu/~igrigore/pp/b_alpha_0.pdf
17. H. Kaspi, A. Mandelbaum, On Harris recurrence in continuous time. Math. Oper. Res. **19**(1), 211–222 (1994)
18. A.H. Löpker, J.S.H. van Leeuwaarden, Transient moments of the TCP window size process. J. Appl. Probab. **45**(1), 163–175 (2008)
19. S.P. Meyn, R.L. Tweedie, *Markov Chains and Stochastic Stability*. Communications and Control Engineering Series (Springer, London, 1993)
20. S.P. Meyn, R.L. Tweedie, Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. Adv. Appl. Probab. **25**(3), 518–548 (1993)

21. P. Michel, Existence of a solution to the cell division eigenproblem. Math. Models Methods Appl. Sci. **16**(7, Suppl.), 1125–1153 (2006)
22. B. Perthame, *Transport Equations in Biology*. Frontiers in Mathematics (Birkhäuser Verlag, Basel, 2007)

Chapter 3
Iterated Proportional Fitting Procedure and Infinite Products of Stochastic Matrices

J. Brossard and C. Leuridan

Abstract The iterative proportional fitting procedure (IPFP), introduced in 1937 by Kruithof, aims to adjust the elements of an array to satisfy specified row and column sums. Thus, given a rectangular non-negative matrix X_0 and two positive marginals a and b, the algorithm generates a sequence of matrices $(X_n)_{n\geq 0}$ starting at X_0, supposed to converge to a biproportional fitting, that is, to a matrix Y whose marginals are a and b and of the form $Y = D_1 X_0 D_2$, for some diagonal matrices D_1 and D_2 with positive diagonal entries.

When a biproportional fitting does exist, it is unique and the sequence $(X_n)_{n\geq 0}$ converges to it at an at least geometric rate. More generally, when there exists some matrix with marginal a and b and with support included in the support of X_0, the sequence $(X_n)_{n\geq 0}$ converges to the unique matrix whose marginals are a and b and which can be written as a limit of matrices of the form $D_1 X_0 D_2$.

In the opposite case (when there exists no matrix with marginals a and b whose support is included in the support of X_0), the sequence $(X_n)_{n\geq 0}$ diverges but both subsequences $(X_{2n})_{n\geq 0}$ and $(X_{2n+1})_{n\geq 0}$ converge.

In the present paper, we use a new method to prove again these results and determine the two limit-points in the case of divergence. Our proof relies on a new convergence theorem for backward infinite products $\cdots M_2 M_1$ of stochastic matrices M_n, with diagonal entries $M_n(i,i)$ bounded away from 0 and with bounded ratios $M_n(j,i)/M_n(i,j)$. This theorem generalizes Lorenz' stabilization theorem. We also provide an alternative proof of Touric and Nedić's theorem on backward infinite products of doubly-stochastic matrices, with diagonal entries bounded away from 0. In both situations, we improve slightly the conclusion, since we establish not only the convergence of the sequence $(M_n \cdots M_1)_{n\geq 0}$, but also its finite variation.

J. Brossard · C. Leuridan (✉)
Institut Fourier, Université Grenoble Alpes, Grenoble Cedex, France
e-mail: Jean_Brossard@orange.fr; Christophe.Leuridan@univ-grenoble-alpes.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_3

Keywords Infinite products of stochastic matrices · Contingency matrices · Distributions with given marginals · Iterative proportional fitting · Relative entropy · I-divergence

Subject Classifications 15B51, 62H17, 62B10, 68W40

3.1 Introduction

3.1.1 The Iterative Proportional Fitting Procedure

The aim of the iterative proportional fitting procedure is to find a non-negative matrix with given row and columns sums and having zero entries at some given places. Fix two integers $p \geq 2$, $q \geq 2$ (namely the sizes of the matrices to be considered) and two vectors $a = (a_1, \dots, a_p)$, $b = (b_1, \dots, b_q)$ with positive components such that $a_1 + \cdots + a_p = b_1 + \cdots + b_q = 1$ (namely the target marginals). We assume that the common value of the sums $a_1 + \cdots + a_p$ and $b_1 + \cdots + b_q$ is 1 for convenience only, to enable probabilistic interpretations, but this is not a true restriction.

We introduce the following notations for any $p \times q$ real matrix X:

$$X(i,+) = \sum_{j=1}^{q} X(i,j), \quad X(+,j) = \sum_{i=1}^{p} X(i,j), \quad X(+,+) = \sum_{i=1}^{p} \sum_{j=1}^{q} X(i,j),$$

and we set $R_i(X) = X(i,+)/a_i$, $C_j(X) = X(+,j)/b_j$.

The IPFP has been introduced in 1937 by Kruithof (*method of the double factors* in third appendix of [9]) to estimate telephone traffic between central stations. This procedure starts from a $p \times q$ non-negative matrix X_0 such that the sum of the entries on each row or column is positive (so X_0 has at least one positive entry on each row or column) and works as follows.

- For each $i \in [\![1, p]\!]$, divide the row i of X_0 by the positive number $R_i(X_0)$. This yields a matrix X_1 satisfying the same assumptions as X_0 and having the desired row-marginals.
- For each $j \in [\![1, q]\!]$, divide the column j of X_1 by the positive number $C_j(X_1)$. This yields a matrix X_2 satisfying the same assumptions as X_0 and having the desired column-marginals.
- Repeat the operations above starting from X_2 to get X_3, X_4, and so on.

Let $\mathscr{M}_{p,q}(\mathbf{R}_+)$ be the set of all $p \times q$ matrices with non-negative entries, and

$$\Gamma_0 := \{X \in \mathscr{M}_{p,q}(\mathbf{R}_+) : \forall i \in [\![1,p]\!],\ X(i,+) > 0,\ \forall j \in [\![1,q]\!],\ X(+,j) > 0\},$$

$$\Gamma_1 := \{X \in \Gamma_0 : X(+,+) = 1\}$$

$$\Gamma_R := \Gamma(a,*) = \{X \in \Gamma_0 : \forall i \in [\![1,p]\!],\ X(i,+) = a_i\},$$

$$\Gamma_C := \Gamma(*,b) = \{X \in \Gamma_0 : \forall j \in [\![1,q]\!],\ X(+,j) = b_j\},$$

$$\Gamma := \Gamma(a,b) = \Gamma_R \cap \Gamma_C.$$

For every integer $m \geq 1$, denote by Δ_m the set of all $m \times m$ diagonal matrices with positive diagonal entries.

The IPFP consists in applying alternatively the transformations $T_R : \Gamma_0 \to \Gamma_R$ and $T_C : \Gamma_0 \to \Gamma_C$ defined by

$$T_R(X)(i,j) = R_i(X)^{-1} X(i,j) \text{ and } T_C(X)(i,j) = C_j(X)^{-1} X(i,j).$$

The homogeneity of the map T_R shows that replacing X_0 with $X_0(+,+)^{-1}X_0$ does not change the matrices X_n for $n \geq 1$, so there is no restriction to assume that $X_0 \in \Gamma_1$.

Note that Γ_R and Γ_C are subsets of Γ_1 and are closed subsets of $\mathscr{M}_{p,q}(\mathbf{R}_+)$. Therefore, if $(X_n)_{n\geq 0}$ converges, its limit belongs to the set Γ. Furthermore, by construction, the matrices X_n belong to the set

$$\begin{aligned}\Delta_p X_0 \Delta_q &= \{D_1 X_0 D_2 : D_1 \in \Delta_p, D_2 \in \Delta_q\} \\ &= \{(\alpha_i X_0(i,j)\beta_j) : (\alpha_1,\ldots,\alpha_p) \in (\mathbf{R}_+^*)^p,\ (\beta_1,\ldots,\beta_q) \in (\mathbf{R}_+^*)^q\}.\end{aligned}$$

According to the terminology used by Pretzel, we will say that X_0 and X_n are diagonally equivalent. In particular, the matrices X_n have by construction the same support, where the support of a matrix $X \in \mathscr{M}_{p,q}(\mathbf{R}_+)$ is defined by

$$\text{Supp}(X) = \{(i,j) \in [\![1,p]\!] \times [\![1,q]\!] : X(i,j) > 0\}.$$

For every limit point L of $(X_n)_{n\geq 0}$, we get $\text{Supp}(L) \subset \text{Supp}(X_0)$ and this inclusion may be strict. In particular, if $(X_n)_{n\geq 0}$ converges, its limit belongs to the set

$$\Gamma(X_0) := \Gamma(a,b,X_0) = \{S \in \Gamma : \text{Supp}(S) \subset \text{Supp}(X_0)\}.$$

When the set $\Gamma(X_0)$ is empty, the sequence $(X_n)_{n\geq 0}$ cannot converge, and no precise behavior was established until 2013, when Gietl and Reffel showed that both subsequences $(X_{2n})_{n\geq 0}$ and $(X_{2n+1})_{n\geq 0}$ converge [7].

In the opposite case, when $\Gamma(a, b)$ contains some matrix with support included in X_0, various proofs of the convergence of $(X_n)_{n\geq 0}$ are known (Bacharach [2] in 1965, Bregman [3] in 1967, Sinkhorn [13] in 1967, Csiszár [5] in 1975, Pretzel [11] in 1980 and others (see [4] and [12] to get an exhaustive review). Moreover, the limit can be described using some probabilistic tools that we introduce now.

3.1.2 Probabilistic Interpretations and Tools

At many places, we shall identify a, b and matrices X in Γ_1 with the probability measures on $[\![1, p]\!]$, $[\![1, q]\!]$ and $[\![1, p]\!] \times [\![1, q]\!]$ given by $a(\{i\}) = a_i$, $b(\{j\}) = b_j$ and $X(\{(i, j)\} = X(i, j)$. Through this identification, the set Γ_1 can be seen as the set of all probability measures on $[\![1, p]\!] \times [\![1, q]\!]$ whose marginals charge every point; the set Γ can be seen as the set of all probability measures on $[\![1, p]\!] \times [\![1, q]\!]$ having marginals a and b. This set is non-empty since it contains the probability $a \otimes b$.

The I-divergence, or Kullback-Leibler divergence, also called relative entropy, plays a key role in the study of the iterative proportional fitting algorithm. For every X and Y in Γ_1, the relative entropy of Y with respect to X is

$$D(Y||X) = \sum_{(i,j)\in \text{Supp}(X)} Y(i, j) \ln \frac{Y(i, j)}{X(i, j)} \text{ if Supp}(Y) \subset \text{Supp}(X),$$

and $D(Y||X) = +\infty$ otherwise, with the convention $0 \ln 0 = 0$. Although D is not a distance, the quantity $D(Y||X)$ measures in some sense how much Y is far from X since $D(Y||X) \geq 0$, with equality if and only if $Y = X$.

In 1968, Ireland and Kullback [8] gave an incomplete proof of the convergence of $(X_n)_{n\geq 0}$ when X_0 is positive, relying on the properties of the I-divergence. Yet, the I-divergence can be used to prove the convergence when the set $\Gamma(X_0)$ is non-empty, and to determine the limit. The maps T_R and T_C can be viewed has I-projections on Γ_R and Γ_C in the sense that for every $X \in \Gamma_0$, $T_R(X)$ (respectively $T_C(X)$) is the only matrix achieving the least upper bound of $D(Y||X)$ over all Y in Γ_R (respectively Γ_C).

In 1975, Csiszár established (theorem 3.2 in [5]) that, given a finite collection of linear sets $\mathscr{E}_1, \dots, \mathscr{E}_k$ of probability distributions on a finite set, and a distribution R such that $\mathscr{E}_1 \cap \dots \cap \mathscr{E}_k$ contains some probability distribution which is absolutely continuous with regard to R, the sequence obtained from R by applying cyclically the I-projections on $\mathscr{E}_1, \dots, \mathscr{E}_k$ converges to the I-projection of R on $\mathscr{E}_1 \cap \dots \cap \mathscr{E}_k$. This result applies to our context (the finite set is $[\![1, p]\!] \times [\![1, q]\!]$, the linear sets $\mathscr{E}_1$ and $\mathscr{E}_2$ are Γ_R and Γ_C) and shows that if the set $\Gamma = \Gamma_R \cap \Gamma_C$ contains some matrix with support included in X_0, then $(X_n)_{n\geq 0}$ converges to the I-projection of X_0 on Γ.

3.2 Old and New Results

Since the behavior of the sequence $(X_n)_{n\geq 0}$ depends only on the existence or the non-existence of elements of Γ with support equal to or included in $\mathrm{Supp}(X_0)$, we state a criterion which determines in which case we are.

Consider two subsets A of $[\![1,p]\!]$ and B of $[\![1,q]\!]$ such that X_0 is null on $A \times B$. Note $A^c = [\![1,p]\!] \setminus A$ and $B^c = [\![1,q]\!] \setminus B$. Then for every $S \in \Gamma(X_0)$,

$$a(A) = \sum_{i\in A} a_i = \sum_{(i,j)\in A\times B^c} S(i,j) \leq \sum_{(i,j)\in[1,p]\times B^c} S(i,j) = \sum_{j\in B^c} b_j = b(B^c).$$

If $a(A) = b(B^c)$, S must be null on $A^c \times B^c$. If $a(A) > b(B^c)$, we get a contradiction, so $\Gamma(X_0)$ is empty.

Actually, these causes of the non-existence of elements of Γ with support equal to or included in $\mathrm{Supp}(X_0)$ provide necessary and sufficient conditions. We give two criterions, the first one was already stated by Bacharach [2]. Pukelsheim gave a different formulation of these conditions in theorems 2 and 3 of [12]. We use a different method, relying on the theory of linear system of inequalities, and give a more precise statement, namely item 3 of critical case below.

Theorem 3.1 (Criteria to Distinguish the Cases)

1. *(Case of incompatibility) The set $\Gamma(X_0)$ is empty if and only if there exist two subsets $A \subset [\![1,p]\!]$ and $B \subset [\![1,q]\!]$ such that X_0 is null on $A \times B$ and $a(A) > b(B^c)$.*
2. *(Critical case) Assume now that $\Gamma(X_0)$ is not empty. Then*

   a. *There exists a matrix $S_0 \in \Gamma(X_0)$ whose support contains the support of every matrix in $\Gamma(X_0)$.*

   b. *The support of S_0 is strictly contained in* $\mathrm{Supp}(X_0)$ *if and only if there exist two non-empty subsets A of $[\![1,p]\!]$ and B of $[\![1,q]\!]$ such that $a(A) = b(B^c)$ and X_0 is null on $A \times B$ but has a positive entry on $A^c \times B^c$.*

   c. *More precisely, the support of S_0 is the complement in* $\mathrm{Supp}(X_0)$ *of the union of all products $A^c \times B^c$ over all non-empty subsets $A \times B$ of $[\![1,p]\!] \times [\![1,q]\!]$ such that X_0 is null on $A \times B$ and $a(A) = b(B^c)$.*

Note that the assumption that X_0 has at least a positive entry on each row or column prevents A and B from being full when X_0 is null on $A \times B$. The additional condition $a(A) > b(B^c)$ (or $a(A) = b(B^c)$) is equivalent to the condition $a(A) + b(B) > 1$ (or $a(A) + b(B) = 1$), so rows and column play a symmetric role.

The condition $a(A) > b(B^c)$ and the positivity of all components of a and b also prevent A and B from being empty. We will call *cause of incompatibility* any (non-empty) block $A \times B \subset [\![1,p]\!] \times [\![1,q]\!]$ such that X_0 is null on $A \times B$ and $a(A) > b(B^c)$. If the set $\Gamma(X_0)$ is non-empty, we will call *cause of criticality* any non-empty block $A \times B \subset [\![1,p]\!] \times [\![1,q]\!]$ such that X_0 is null on $A \times B$ and $a(A) = b(B^c)$.

Given a convergent sequence $(x_n)_{n\geq 0}$ of vectors in some normed vector space $(E, ||\cdot||)$, with limit $\ell \in E$, we will say that *the rate of convergence is geometric* (respectively *at least geometric*) if $0 < \lim_n ||x_n - x_\infty||^{1/n} < 1$ (respectively $\limsup_n ||x_n - x_\infty||^{1/n} < 1$).

We now describe the asymptotic behavior of sequence $(X_n)_{n\geq 0}$ in each case. The first case is already well-known.

Theorem 3.2 (Case of Fast Convergence) *Assume that Γ contains some matrix with same support as X_0. Then*

1. *The sequences $(R_i(X_{2n})_{n\geq 0})$ and $(C_j(X_{2n+1})_{n\geq 0})$ converge to* 1 *at an at least geometric rate.*
2. *The sequence $(X_n)_{n\geq 0}$ converges to some matrix X_∞ which has the same support as X_0. The rate of convergence is at least geometric.*
3. *The limit X_∞ is the only matrix in $\Gamma \cap \Delta_p X_0 \Delta_q$ (in particular X_0 and X_∞ are diagonally equivalent). It is also the unique matrix achieving the minimum of $D(Y||X_0)$ over all $Y \in \Gamma(X_0)$.*

For example, if $p = q = 2$, $a_1 = b_1 = 2/3$, $a_2 = b_2 = 1/3$, and

$$X_0 = \frac{1}{4}\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then for every $n \geq 1$, X_n or $X_n^\top$ is equal to

$$\frac{1}{3(2^n - 1)}\begin{pmatrix} 2^n & 2^n - 2 \\ 2^n - 1 & 0 \end{pmatrix},$$

depending on whether n is odd or even. The limit is

$$X_\infty = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

The second case is also well-known, except the fact that the quantities $R_i(X_{2n}) - 1$ and $C_j(X_{2n+1}) - 1$ are $o(n^{-1/2})$.

Theorem 3.3 (Case of Slow Convergence) *Assume that Γ contains some matrix with support included in* Supp(X_0) *but contains no matrix with support equal to* Supp(X_0). *Then*

1. *The series*

$$\sum_{n\geq 0}(R_i(X_{2n}) - 1)^2 \text{ and } \sum_{n\geq 0}(C_j(X_{2n+1}) - 1)^2$$

converge.

2. *The sequences* $(\sqrt{n}(R_i(X_n) - 1))_{n \geq 0}$ *and* $(\sqrt{n}(C_j(X_n) - 1))_{n \geq 0}$ *converge to* 0. *In particular, the sequences* $(R_i(X_n))_{n \geq 0}$ *and* $(C_j(X_n))_{n \geq 0}$ *converge to* 1.
3. *The sequence* $(X_n)_{n \geq 0}$ *converges to some matrix* X_∞ *whose support contains the support of every matrix in* $\Gamma(X_0)$.
4. *The limit* X_∞ *is the unique matrix achieving the minimum of* $D(Y||X_0)$ *over all* $Y \in \Gamma(X_0)$.
5. *If* $(i, j) \in \text{Supp}(X_0) \setminus \text{Supp}(X_\infty)$*, the infinite product* $R_i(X_0)C_j(X_1)R_i(X_2) C_j(X_3)\cdots$ *is infinite.*

Actually the assumption that Γ contains no matrix with support equal to $\text{Supp}(X_0)$ can be removed; but when this assumption fails, Theorem 3.2 applies, so Theorem 3.3 brings nothing new. When this assumption holds, the last conclusion of Theorem 3.3 shows that the rate of convergence cannot be in $o(n^{-1-\varepsilon})$ for any $\varepsilon > 0$.

However, a rate of convergence in $\Theta(n^{-1})$ is possible, and we do not know whether other rates of slow convergence may occur. For example, consider $p = q = 2$, $a_1 = a_2 = b_1 = b_2 = 1/2$, and

$$X_0 = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then for every $n \geq 1$, X_n or $X_n^\top$ is equal to

$$\frac{1}{2n+2}\begin{pmatrix} 1 & n \\ n+1 & 0 \end{pmatrix},$$

depending on whether n is odd or even. The limit is

$$X_\infty = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

When Γ contains no matrix with support included in $\text{Supp}(X_0)$, we already know by Gietl and Reffel's theorem [7] that both sequences $(X_{2n})_{n \geq 0}$ and $(X_{2n+1})_{n \geq 0}$ converge. The convergence may be slow, so Aas gives in [1] an algorithm to fasten the convergence. Aas' algorithm finds and exploits the block structure associated to the inconsistent problem of finding a non-negative matrix whose marginals are a and b and whose support is contained in $\text{Supp}(X_0)$. The next two theorems give a complete description of the two limit points and how to find them.

Theorem 3.4 (Case of Divergence) *Assume that* Γ *contains no matrix with support included in* $\text{Supp}(X_0)$.

Then there exist some positive integer $r \leq \min(p, q)$*, some partitions* $\{I_1, \dots, I_r\}$ *of* $[\![1, p]\!]$ *and* $\{J_1, \dots, J_r\}$ *of* $[\![1, q]\!]$ *such that:*

1. $(R_i(X_{2n}))_{n\geq 0}$ *converges to* $\lambda_k = b(J_k)/a(I_k)$ *whenever* $i \in I_k$*;*
2. $(C_j(X_{2n+1}))_{n\geq 0}$ *converges* λ_k^{-1} *whenever* $j \in J_k$*;*
3. $X_n(i, j) = 0$ *for every* $n \geq 0$ *whenever* $i \in I_k$ *and* $j \in J_{k'}$ *with* $k < k'$*;*
4. $X_n(i, j) \to 0$ *as* $n \to +\infty$ *at a geometric rate whenever* $i \in I_k$ *and* $j \in J_{k'}$ *with* $k > k'$*;*
5. *The sequence* $(X_{2n})_{n\geq 0}$ *converges to the unique matrix achieving the minimum of* $D(Y||X_0)$ *over all* $Y \in \Gamma(a', b, X_0)$*, where* $a'_i/a_i = \lambda_k$ *whenever* $i \in I_k$*;*
6. *The sequence* $(X_{2n+1})_{n\geq 0}$ *converges to the unique matrix achieving the minimum of* $D(Y||X_0)$ *over all* $Y \in \Gamma(a, b', X_0)$*, where* $b'_j/b_j = \lambda_k^{-1}$ *whenever* $j \in J_k$*;*
7. *For every* $k \in [\![1, r]\!]$*,* $a'(I_k) = b(J_k)$ *and* $a(I_k) = b'(J_k)$*. Moreover, the support of any matrix in* $\Gamma(a', b, X_0) \cup \Gamma(a, b', X_0)$ *is contained in* $I_1 \times J_1 \cup \dots \cup I_r \times J_r$*.*
8. *Let* $D_1 = \mathrm{Diag}(a'_1/a_1, \dots, a'_p/a_p)$ *and* $D_2 = \mathrm{Diag}(b_1/b'_1, \dots, b_q/b'_q)$*. Then for every* $S \in \Gamma(a, b', X_0)$*,* $D_1 S = S D_2 \in \Gamma(a', b, X_0)$*, and all matrices in* $\Gamma(a', b, X_0)$ *can be written in this way.*

For example, if $p = q = 2$, $a_1 = b_1 = 1/3$, $a_2 = b_2 = 2/3$, and

$$X_0 = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then for every $n \geq 1$, X_n or $X_n^\top$ is equal to

$$\frac{1}{3(3 \times 2^{n-1} - 1)}\begin{pmatrix} 1 & 3 \times 2^{n-1} - 2 \\ 2(3 \times 2^{n-1} - 1) & 0 \end{pmatrix},$$

depending on whether n is odd or even. We get $a'_1 = b'_1 = 2/3$ and $a'_2 = b'_2 = 1/3$ since the two limit points are

$$\frac{1}{3}\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \text{ and } \frac{1}{3}\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

The symmetry in Theorem 3.4 shows that the limit points of $(X_n)_{n\geq 0}$ would be the same if we would applying T_C first instead of T_R.

Actually, the assumption that $\Gamma(X_0)$ is empty can be removed and it is not used in the proof of Theorem 3.4. Indeed, when $\Gamma(X_0)$ is non-empty, the conclusions still hold with $r = 1$ and $\lambda_1 = 1$, but Theorem 3.4 brings nothing new in this case.

Theorem 3.4 does not indicate what the partitions $\{I_1, \dots, I_r\}$ and $\{J_1, \dots, J_r\}$ are. Actually the integer r, the partitions $\{I_1, \dots, I_r\}$ and $\{J_1, \dots, J_r\}$ depend only on a, b and on the support of X_0. and can be determined recursively as follows. This gives a complete description of the two limit points mentioned in Theorem 3.4.

Theorem 3.5 (Determining the Partitions in Case of Divergence) *Keep the assumption and the notations of Theorem 3.4. Fix $k \in [\![1, r]\!]$, set $P = [\![1, p]\!] \setminus (I_1 \cup \ldots \cup I_{k-1})$, $Q = [\![1, q]\!] \setminus (J_1 \cup \ldots \cup J_{k-1})$, and consider the restricted problem associated to the marginals $a(\cdot|P) = (a_i/a(P))_{i\in P}$, $b(\cdot|Q) = (b_j/b(Q))_{j\in Q}$ and to the initial condition $(X_0(i, j))_{(i,j)\in P\times Q}$.*

If $k = r$, this restricted problem admits some solution.

If $k < r$, the set $A_k \times B_k := I_k \times (Q \setminus J_k)$ is a cause of incompatibility of this restricted problem. More precisely, among all causes of incompatibility $A \times B$ maximizing the ratio $a(A)/b(Q \setminus B)$, it is the one which maximizes the set A and minimizes the set B.

Note that if a cause of incompatibility $A \times B$ maximizes the ratio $a(A)/b(B^c)$, then it is maximal for the inclusion order. We now give an example to illustrate how Theorem 3.5 enables us to determine the partitions $\{I_1, \ldots, I_r\}$ and $\{J_1, \ldots, J_r\}$. In the following array, the $*$ and 0 indicate the positive and the null entries in X_0; the last column and row indicate the target sums on each row and column.

$$
\begin{array}{cccccc||c}
* & * & 0 & 0 & 0 & 0 & 0.25 \\
0 & * & * & 0 & 0 & 0 & 0.25 \\
0 & * & * & * & 0 & 0 & 0.25 \\
* & * & * & * & 0 & * & 0.15 \\
* & 0 & * & * & * & * & 0.10 \\
\hline\hline
0.05 & 0.05 & 0.1 & 0.2 & 0.2 & 0.4 & 1
\end{array}
$$

We indicate below in underlined boldface characters some blocks $A \times B$ of zeroes which are causes of incompatibility, and the corresponding ratios $a(A)/b(B^c)$.

$$
\begin{array}{l}
A = \{1, 2, 3, 4\} \\
B = \{5\} \\
\dfrac{a(A)}{b(B^c)} = \dfrac{0.9}{0.8}
\end{array}
\qquad
\begin{array}{cccccc||c}
* & * & 0 & 0 & \underline{\mathbf{0}} & 0 & 0.25 \\
0 & * & * & 0 & \underline{\mathbf{0}} & 0 & 0.25 \\
0 & * & * & * & \underline{\mathbf{0}} & 0 & 0.25 \\
* & * & * & * & \underline{\mathbf{0}} & * & 0.15 \\
* & 0 & * & * & * & * & 0.10 \\
\hline\hline
0.05 & 0.05 & 0.1 & 0.2 & 0.2 & 0.4 & 1
\end{array}
$$

$$
\begin{array}{l}
A = \{2, 3\} \\
B = \{1, 5, 6\} \\
\dfrac{a(A)}{b(B^c)} = \dfrac{0.5}{0.35}
\end{array}
\qquad
\begin{array}{cccccc||c}
* & * & 0 & 0 & 0 & 0 & 0.25 \\
\underline{\mathbf{0}} & * & * & 0 & \underline{\mathbf{0}} & \underline{\mathbf{0}} & 0.25 \\
\underline{\mathbf{0}} & * & * & * & \underline{\mathbf{0}} & \underline{\mathbf{0}} & 0.25 \\
* & * & * & * & 0 & * & 0.15 \\
* & 0 & * & * & * & * & 0.10 \\
\hline\hline
0.05 & 0.05 & 0.1 & 0.2 & 0.2 & 0.4 & 1
\end{array}
$$

$$
\begin{array}{l}
A = \{1\} \\
B = \{3,4,5,6\} \\
\dfrac{a(A)}{b(B^c)} = \dfrac{0.25}{0.1}
\end{array}
\qquad
\begin{array}{cccccc||c}
* & * & \underline{\mathbf{0}} & \underline{\mathbf{0}} & \underline{\mathbf{0}} & \underline{\mathbf{0}} & 0.25 \\
0 & * & * & 0 & 0 & 0 & 0.25 \\
0 & * & * & * & 0 & 0 & 0.25 \\
* & * & * & * & 0 & * & 0.15 \\
* & 0 & * & * & * & * & 0.10 \\
\hline\hline
0.05 & 0.05 & 0.1 & 0.2 & 0.2 & 0.4 & 1
\end{array}
$$

$$
\begin{array}{l}
A = \{1,2\} \\
B = \{4,5,6\} \\
\dfrac{a(A)}{b(B^c)} = \dfrac{0.5}{0.2}
\end{array}
\qquad
\begin{array}{cccccc||c}
* & * & 0 & \underline{\mathbf{0}} & \underline{\mathbf{0}} & \underline{\mathbf{0}} & 0.25 \\
0 & * & * & \underline{\mathbf{0}} & \underline{\mathbf{0}} & \underline{\mathbf{0}} & 0.25 \\
0 & * & * & * & 0 & 0 & 0.25 \\
* & * & * & * & 0 & * & 0.15 \\
* & 0 & * & * & * & * & 0.10 \\
\hline\hline
0.05 & 0.05 & 0.1 & 0.2 & 0.2 & 0.4 & 1
\end{array}
$$

One checks that the last two blocks are those which maximize the ratio $a(A)/b(B^c)$. Among these two blocks, the latter has a bigger A and a smaller B, so it is $A_1 \times B_1$. Therefore, $I_1 = \{1,2\}$ and $J_1 = \{1,2,3\}$, and we look at the restricted problem associated to the marginals $a(\cdot|I_1^c)$, $b(\cdot|J_1^c)$ and to the initial condition $(X_0(i,j))_{(i,j)\in I_1^c \times J_1^c}$. The dots below indicate the removed rows and columns.

$$
\begin{array}{c|ccc||c}
\cdots & \cdot & \cdot & \cdot & \cdot \\
\cdots & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdots & * & 0 & 0 & 0.5 \\
\cdots & * & 0 & * & 0.3 \\
\cdots & * & * & * & 0.2 \\
\hline\hline
\cdots & 0.25 & 0.25 & 0.5 & 1
\end{array}
$$

Two causes of incompatibility have to be considered, namely $\{3,4\} \times \{5\}$ and $\{3\} \times \{5,6\}$. The latter maximizes the ratio $a(A)/b(J_1^c \setminus B)$, so it is $A_2 \times B_2$. Therefore, $I_2 = \{3\}$ and $J_2 = \{4\}$, and we look at the restricted problem below.

$$
\begin{array}{c|cc||c}
\cdots\cdot & \cdot & \cdot & \cdot \\
\cdots\cdot & \cdot & \cdot & \cdot \\
\cdots\cdot & \cdot & \cdot & \cdot \\
\hline
\cdots\cdot & 0 & * & 0.6 \\
\cdots\cdot & * & * & 0.4 \\
\hline\hline
\cdots\cdot & 0.33 & 0.67 & 1
\end{array}
$$

This time, there is no cause of incompatibility, so $r = 3$, and the sets $I_3 = \{4,5\}$, $J_3 = \{5,6\}$ contain all the remaining indices. We indicate below the block structure defined by the partitions $\{I_1, I_2, I_3\}$ and $\{J_1, J_2, J_3\}$ (for readability, our example was chosen in such a way that each block is made of consecutive indices). By Theorem 3.4, the limit of the sequence $(X_{2n})_{n\geq 0}$ admits marginals a' and b, its support is included in Supp(X_0) and also in $(I_1 \times J_1) \cup (I_2 \times J_2) \cup (I_3 \times J_3)$, namely it solves the problem below.

*	*	0	0	0	0	0.1
0	*	*	0	0	0	0.1
0	0	0	*	0	0	0.2
0	0	0	0	0	*	0.36
0	0	0	0	*	*	0.24
0.05	0.05	0.1	0.2	0.2	0.4	1

We observe that each cause of incompatibility $A \times B$ related to the block structure, namely $I_1 \times (J_2 \cup J_3)$ or $(I_1 \cup I_2) \times J_3$, becomes a cause of criticality with regard to the margins a' and b, namely $a'(A) = b(B^c)$, so $\lim_n X_{2n}$ has zeroes on $A^c \times B^c$. This statement fails for the "weaker" cause of incompatibility $\{1, 2, 3, 4\} \times \{5\}$. Yet, it still holds for the cause of incompatibility $\{1\} \times \{3, 4, 5, 6\}$; this rare situation occurs because there were two causes of incompatibility maximizing the ratio $a(A)/b(B^c)$, namely $\{1\} \times \{3, 4, 5, 6\}$ and $\{1, 2\} \times \{4, 5, 6\}$. Hence $\lim_n X_{2n}$ has another additional zero at position (2, 2). We add dashlines below to make visible this refinement of the block structure. Here, no minimization of I-divergence is required to get the limit of $(X_{2n})_{n\geq 0}$ since the set $\Gamma(a', b, X_0)$ contains only one matrix, namely

0.05	0.05	0	0	0	0	0.1
0	0	0.1	0	0	0	0.1
0	0	0	.2	0	0	0.2
0	0	0	0	0	0.36	0.36
0	0	0	0	0.2	0.04	0.24
0.05	0.05	0.1	0.2	0.2	0.4	1

.

We note that the convergence of $X_{2n}(2, 2)$ to 0 is slow since

$$\lim_{n\to+\infty} \frac{X_{2n+2}(2,2)}{X_{2n}(2,2)} = \lim_{n\to+\infty} \frac{1}{R_2(X_{2n})C_2(X_{2n+1})} = \frac{1}{\lambda_1 \lambda_1^{-1}} = 1.$$

This is a typical situation in which the following observation is useful.

Theorem 3.6 *The matrices* $\lim_n X_{2n}$ *and* $\lim_n X_{2n+1}$ *have the same support* Σ *and* Σ *is the union of the supports of all matrices in* $\Gamma(a', b, X_0) \cup \Gamma(a, b', X_0)$. *Moreover, if* X'_0 *denotes the matrix obtained from* X_0 *by setting to* 0 *all entries outside* Σ, *then the limit points provided by the IPFP starting from* X_0 *and from* X'_0 *coincide.*

Aas mentions this fact (proposition 1 in [1]) as a result of Pretzel (last part of theorem 1 in [11]), although Pretzel considers only the case where the set $\Gamma(a, b, X_0)$ is not empty. We prove Theorem 3.6 by adapting Pretzel's proof.

The set Σ can be determined by applying Theorem 3.1 (critical case, item (c)) to the marginals a' and b. The interest of Theorem 3.6 is that starting from X'_0 ensures an at least geometric rate of convergence, since Theorem 3.2 applies when one performs the IPFP on the marginals a' and b (or a and b') and the initial matrix X'_0.

That is why Aas investigates the inherent block structure. Actually, the splitting considered by Aas is finer that the splitting provided by our Theorems 3.4 and 3.5; in the example above, Aas would split I_1 into $\{1, 2\}$ and $\{3\}$, and J_1 into $\{1\}$ and $\{2\}$. Up to this distinction, most of the statements provided by our Theorems 3.4 and 3.5 are explicitly or implicitly present in Aas' paper, which focuses on an algorithmic point of view.

The convergence of the sequences $(X_{2n})_{n\geq 0}$ and $(X_{2n+1})_{n\geq 0}$ was already established by Gietl and Reffel [7] with the help of I-divergence. Our proof is completely different (although I-divergence helps us to determine the limit points). Our first step is to prove the convergence of the sequences $(R_i(X_{2n}))_{n\geq 0}$ and $(C_j(X_{2n+1}))_{n\geq 0}$ by exploiting recursion relations involving stochastic matrices. The proof relies on the next general result on infinite products of stochastic matrices. Theorem 3.7 below will only be used to prove Theorems 3.4 and 3.5, so apart from Lemma 3.1, Sect. 3.3 can be skipped if the reader is only interested by the new proof of Theorems 3.1–3.3.

A sequence $(x_n)_{n\geq 0}$ of vectors in some normed vector space $(E, ||\cdot||)$ will be said to have a *finite variation* if the series $\sum_n ||x_{n+1} - x_n||$ converges. The finite variation implies the convergence when $(E, ||\cdot||)$ is a Banach space, in particular when E has a finite dimension.

Theorem 3.7 *Let $(M_n)_{n\geq 1}$ be some sequence of $d \times d$ stochastic matrices. Assume that there exists some constants $\gamma > 0$, and $\rho \geq 1$ such that for every $n \geq 1$ and i, j in $[\![1, d]\!]$, $M_n(i, i) \geq \gamma$ and $M_n(i, j) \leq \rho M_n(j, i)$. Then the sequence $(M_n \cdots M_1)_{n\geq 1}$ has a finite variation, so it converges to some stochastic matrix L. Moreover, the series $\sum_n M_n(i, j)$ and $\sum_n M_n(j, i)$ converge whenever the rows of L with indexes i and j are different.*

An important literature deals with infinite products of stochastic matrices, with various motivations: study of inhomogeneous Markov chains, of opinion dynamics... See for example [14]. Backward infinite products converge much more often than forward infinite products. Many theorems involve the ergodic coefficients of stochastic matrices. For a $d \times d$ stochastic matrix M, the ergodic coefficient is

$$\tau(M) = \min_{1\leq i,i'\leq d} \sum_{j=1}^{d} \min(M(i, j), M(i', j)) \in [0, 1].$$

The difference $1 - \tau(M)$ is the maximal total variation distance between the lines of M seen as probabilities on $[\![1, d]\!]$. These theorems do not apply in our context.

To our knowledge, Theorem 3.7 is new. The closest statements we found in the literature are Lorenz' stabilization theorem (theorem 2 of [10]) and a theorem of Touri and Nedić on infinite product of bistochastic matrices (theorem 7 of [17], relying on theorem 6 of [16]). The method we use to prove Theorem 3.7 is different from theirs.

On the one hand, Theorem 3.7 provides a stronger conclusion (namely finite variation and not only convergence) and has weaker assumptions than Lorenz' stabilization theorem. Indeed, Lorenz assumes that each M_n has a positive diagonal and a symmetric support, and that the entries of all matrices M_n are bounded below by some $\delta > 0$; this entails the assumptions of our Theorem 3.7, with $\gamma = \delta$ and $\rho = \delta^{-1}$.

On the other hand, Lorenz' stabilization theorem gives more precisions on the limit $L = \lim_{n\to+\infty} M_n \dots M_1$. In particular, if the support of M_n does not depends on n, then Lorenz shows that by applying a same permutation on the rows and on the columns of L, one gets a block-diagonal matrix in which each diagonal block is a consensus matrix, namely a stochastic matrix whose rows are all the same. This additional conclusion does not hold anymore under our weaker assumptions. For example, for every $r \in [-1, 1]$, consider the stochastic matrix

$$M(r) = \frac{1}{2}\begin{pmatrix} 1+r & 1-r \\ 1-r & 1+r \end{pmatrix}.$$

One checks that for every r_1 and r_2 in $[-1, 1]$, $M(r_2)M(r_1) = M(r_2 r_1)$. Let $(r_n)_{n\geq 1}$ be any sequence of numbers in $]0, 1]$ whose infinite product converges to some $\ell > 0$. Then our assumptions hold with $\gamma = 1/2$ and $\rho = 1$ and the matrices $M(r_n)$ have the same support. Yet, the limit of the products $M(r_n)\cdots M(r_1)$, namely $M(\ell)$, has only positive coefficients and is not a consensus matrix.

Note also that, given an arbitrary sequence $(M_n)_{n\geq 1}$ of stochastic matrices, assuming only that the diagonal entries are bounded away from 0 does not ensure the convergence of the infinite product $\cdots M_2 M_1$. Indeed, consider the triangular stochastic matrices

$$T_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

A recursion shows that for every $n \geq 1$ and $(\varepsilon_1, \dots, \varepsilon_n) \in \{0, 1\}^n$,

$$T_{\varepsilon_n} \cdots T_{\varepsilon_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ r & 1 - 2^{-n} - r & 2^{-n} \end{pmatrix}, \text{ where } r = \sum_{k=1}^{n} \frac{\varepsilon_k}{2^{n+1-k}}.$$

Hence, one sees that the infinite product $\cdots T_0 T_1 T_0 T_1 T_0 T_1$ diverges.

Yet, for a sequence of *doubly-stochastic* matrices, it is sufficient to assume that the diagonal entries are bounded away from 0. This result was proved by Touri and Nedić (theorem 5 of [15] or theorem 7 of [17], relying on theorem 6 of [16]). We provide a simpler proof and a slight improvement, showing that the sequence $(M_n \dots M_1)_{n \geq 1}$ not only converges but also has a finite variation.

Theorem 3.8 *Let $(M_n)_{n\geq 1}$ be some sequence of $d \times d$ doubly-stochastic matrices. Assume that there exists some constant $\gamma > 0$ such that for every $n \geq 1$ and i in $[\![1, d]\!]$, $M_n(i,i) \geq \gamma$. Then the sequence $(M_n \dots M_1)_{n\geq 1}$ has a finite variation, so it converges to some stochastic matrix L. Moreover, the series $\sum_n M_n(i,j)$ and $\sum_n M_n(j,i)$ converge whenever the rows of L with indexes i and j are different.*

Our proof relies on the following fact: for each column vector $V \in \mathbf{R}^d$, set

$$D(V) = \sum_{1\leq i,j\leq d} |V(i) - V(j)| \text{ and } ||V||_1 = \sum_{1\leq i\leq d} |V(i)|.$$

We will call *dispersion of* V the quantity $D(V)$. Then, under the assumptions of Theorem 3.8, the inequality

$$\gamma ||M_{n+1} \cdots M_1 V - M_n \cdots M_1 V||_1 \leq D(M_n \cdots M_1 V) - D(M_{n+1} \cdots M_1 V).$$

holds for every $n \geq 0$.

3.3 Infinite Products of Stochastic Matrices

3.3.1 Proof of Theorem 3.7

We begin with an elementary lemma.

Lemma 3.1 *Let M be any $m \times n$ stochastic matrix and $V \in \mathbf{R}^n$ be a column vector. Denote by $\underline{M}$ the smallest entry of M, by $\underline{V}$, $\overline{V}$ and* $\mathrm{diam}(V) = \overline{V} - \underline{V}$ *the smallest entry, the largest entry and the diameter of V. Then*

$$\underline{MV} \geq (1 - \underline{M})\, \underline{V} + \underline{M}\, \overline{V} \geq \underline{V},$$

$$\overline{MV} \leq \underline{M}\, \underline{V} + (1 - \underline{M})\, \overline{V} \leq \overline{V},$$

so

$$\mathrm{diam}(MV) \leq (1 - 2\underline{M})\, \mathrm{diam}(V).$$

Proof Call $M(i, j)$ the entries of M and $V(1), \ldots, V(n)$ the entries of V. Let j_1 and j_2 be indexes such that $V(j_1) = \underline{V}$ and $V(j_2) = \overline{V}$. Then for every $i \in [\![1, m]\!]$,

$$\begin{aligned}
(MV)_i &= \sum_{j \neq j_2} M(i, j)V(j) + M(i, j_2)V(j_2) \\
&\geq \sum_{j \neq j_2} M(i, j)\underline{V} + M(i, j_2)\overline{V} \\
&= \underline{V} + M(i, j_2)(\overline{V} - \underline{V}) \\
&\geq \underline{V} + \underline{M}\ (\overline{V} - \underline{V}) \\
&\geq \underline{V}.
\end{aligned}$$

The first inequality follows. Applying it to $-V$ yields the second inequality.

The interesting case is when $n \geq 2$, so $\underline{M} \leq 1/2$ and $1 - 2\underline{M} \geq 0$. Yet, the lemma and the proof above still apply when $n = 1$, since $\overline{MV} = \underline{MV} = \overline{V} = \underline{V}$ in this case.

We now restrict ourselves to square stochastic matrices. To every column vector $V \in \mathbf{R}^d$, we associate the column vector $V^{\uparrow} \in \mathbf{R}^d$ obtained by ordering the components in non-decreasing order. In particular $V^{\uparrow}(1) = \underline{V}$ and $V^{\uparrow}(d) = \overline{V}$.

In the next lemmas and corollary, we establish inequalities that will play a key role in the proof of Theorem 3.7.

Lemma 3.2 *Let M be some $d \times d$ stochastic matrix with diagonal entries bounded below by some constant $\gamma > 0$, and $V \in \boldsymbol{R}^d$ be a column vector with components in increasing order $V(1) \leq \ldots \leq V(d)$. Let σ be a permutation of $[\![1, d]\!]$ such that $(MV)(\sigma(1)) \leq \ldots \leq (MV)(\sigma(d))$. For every $i \in [\![1, d]\!]$, set*

$$A_i = \sum_{j=1}^{i-1} M(\sigma(i), j)\ [V(i) - V(j)],\ \ B_i = \sum_{j=i+1}^{d} M(\sigma(i), j)\ [V(j) - V(i)],$$

with the natural conventions $A_1 = B_d = 0$. The following statements hold.

1. *For every $i \in [\![1, d]\!]$, $(MV)^{\uparrow}(i) - V^{\uparrow}(i) = B_i - A_i$.*
2. *All the terms in the sums defining A_i and B_i are non-negative.*
3. *$B_i \geq M(\sigma(i), \sigma(i))\ [V(\sigma(i)) - V(i)] \geq \gamma\ [V(\sigma(i)) - V(i)]$ whenever $i < \sigma(i)$.*
4. *Let $a < b$ in $[\![1, d]\!]$. If the orbit $O(a)$ of a associated to the permutation σ contains some integer at least equal to b, then*

$$V(b) - V(a) \leq \gamma^{-1} \sum_{i \in O(a) \cap [\![1, b-1]\!]} B_i \boldsymbol{1}_{[i < \sigma(i)]} \leq \gamma^{-1} \sum_{i \in O(a) \cap [\![1, b-1]\!]} B_i .$$

5. *One has*

$$\sum_{i=1}^{d} \left|V(\sigma(i)) - V(i)\right| \leq 2\gamma^{-1} \sum_{i=1}^{d} B_i \mathbf{1}_{[i<\sigma(i)]} \leq 2\gamma^{-1} \sum_{i=1}^{d-1} B_i.$$

Proof By assumption,

$$(MV)^{\uparrow}(i) - V^{\uparrow}(i) = (MV)(\sigma(i)) - V(i) = \sum_{j=1}^{d} M(\sigma(i), j)\,[V(j) - V(i)] = -A_i + B_i,$$

which yields the first item. The next two items follow directly from the assumptions $V(1) \leq \ldots \leq V(d)$ and $M(j, j) \geq \gamma$ for every $j \in [\![1, d]\!]$.

Under the assumptions of item 4, the integer $n(a, b) = \min\{n \geq 1 : \sigma^n(a) \geq b\}$ is well-defined and

$$\begin{aligned} V(b) - V(a) &\leq V(\sigma^{n(a,b)}(a)) - V(a) \\ &\leq \sum_{k=0}^{n(a,b)-1} [V(\sigma^{k+1}(a)) - V(\sigma^k(a))]\mathbf{1}_{\sigma^k(a)<\sigma^{k+1}(a)} \\ &\leq \gamma^{-1} \sum_{k=0}^{n(a,b)-1} B_{\sigma^k(a)}\mathbf{1}_{\sigma^k(a)<\sigma^{k+1}(a)} \end{aligned}$$

by item 3. Item 4 follows.

Since the sum of $V(\sigma(i)) - V(i)$ over all $i \in [\![1, d]\!]$ is null and since $V(1) \leq \ldots \leq V(d)$, one has

$$\sum_{i=1}^{d} \left|V(\sigma(i)) - V(i)\right| = 2\sum_{i=1}^{d} \big(V(\sigma(i)) - V(i)\big)\mathbf{1}_{[i<\sigma(i)]} \leq 2\gamma^{-1} \sum_{i=1}^{d-1} B_i \mathbf{1}_{[i<\sigma(i)]},$$

by item 3. The proof is complete.

We denote by $||\cdot||_1$ the norm on $\mathbf{R}^d$ defined as the sum of the absolute values of the components.

Lemma 3.3 *Keep the assumptions and the notations of Lemma 3.2. Assume that there exists some constant $\rho \geq 1$ such that for every i, j in $[\![1, d]\!]$, $M(i, j) \leq \rho M(j, i)$. Set $C = d(d-1)\max(\gamma^{-1}, \rho)$. Then the following statements hold*

1. *For every $i \in [\![1, d]\!]$, $A_i \leq (i-1)\max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{i-1})$.*
2. *For every $m \in [\![1, d]\!]$, $B_1 + \ldots + B_m \leq (1 + C + \cdots + C^{m-1})||(MV)^{\uparrow} - V^{\uparrow}||_1$.*

Proof Given $i \in [\![2, d]\!]$ and $j \in [\![1, i-1]\!]$, let us check that

$$A_{i,j} := M(\sigma(i), j)\,[V(i) - V(j)] \leq \max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{i-1}).$$

We distinguish two cases.

- If the orbit of some $k \in [\![1, j]\!]$ contains some integer at least equal to i, then inequality 4 of Lemma 3.2 applied with $(a, b) = (k, i)$ yields

$$A_{i,j} \leq V(i) - V(j) \leq V(i) - V(k) \leq \gamma^{-1} \sum_{z \in O(k) \cap [\![1, i-1]\!]} B_z.$$

- Otherwise, the orbit of every element of $[\![1, j]\!]$ is contained in $[\![1, i-1]\!]$, so the orbit of every element of $[\![i, d]\!]$ is contained in $[\![j+1, d]\!]$. In particular, the orbits $O(\sigma(i)) = O(i)$ and $O(j)$ are disjoint. Applying inequality 3 and inequality 4 of Lemma 3.2, once with $(a, b) = (\sigma(i), i)$, once with $(a, b) = (j, \sigma^{-1}(j))$ yields

$$\begin{aligned}
A_{i,j} &= M(\sigma(i), j)\,[V(i) - V(\sigma(i))] + M(\sigma(i), j)\,[V(\sigma(i)) - V(\sigma^{-1}(j))] \\
&\quad + M(\sigma(i), j)\,[V(\sigma^{-1}(j) - V(j)] \\
&\leq \mathbf{1}_{\sigma(i)<i}\,[V(i) - V(\sigma(i))] + \mathbf{1}_{\sigma^{-1}(j)<\sigma(i)}\,\rho M(j, \sigma(i))[V(\sigma(i)) - V(\sigma^{-1}(j))] \\
&\quad + \mathbf{1}_{j<\sigma^{-1}(j)}\,[V(\sigma^{-1}(j)) - V(j)] \\
&\leq \gamma^{-1} \sum_{z \in O(i) \cap [\![1, i-1]\!]} B_z + \rho B_{\sigma^{-1}(j)} + \gamma^{-1} \sum_{z \in O(j) \cap [\![1, \sigma^{-1}(j)-1]\!]} B_z \\
&\leq \max(\gamma^{-1}, \rho) \sum_{z \in [\![1, i-1]\!]} B_z.
\end{aligned}$$

In both cases, we have got the inequality $A_{i,j} \leq \max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{i-1})$. Summing over all $j \in [\![1, i-1]\!]$ yields item 1.

Let $m \in [\![1, d]\!]$. Equality $(MV)^{\uparrow}(i) - V^{\uparrow}(i) = B_i - A_i$ and item 1 yield

$$\begin{aligned}
\sum_{i=1}^{m} B_i &\leq \sum_{i=1}^{m} |(MV)^{\uparrow}(i) - V^{\uparrow}(i)| + \sum_{i=1}^{m} A_i \\
&\leq ||(MV)^{\uparrow} - V^{\uparrow}||_1 + \sum_{i=1}^{m} (i-1) \max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{i-1}) \\
&\leq ||(MV)^{\uparrow} - V^{\uparrow}||_1 + \sum_{i=1}^{m} (d-1) \max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{m-1}) \\
&\leq ||(MV)^{\uparrow} - V^{\uparrow}||_1 + C \sum_{i=1}^{m-1} B_i
\end{aligned}$$

The particular case where $m = 1$ yields $B_1 \leq ||(MV)^\uparrow - V^\uparrow||_1$. Item 2 follows by induction.

Corollary 3.1 *Let M be some $d \times d$ stochastic matrix with diagonal entries bounded below by some $\gamma > 0$ Assume that there exists some constant $\rho \geq 1$ such that for every $n \geq 1$ and i, j in $[\![1, d]\!]$, $M(i, j) \leq \rho M(j, i)$. Set $C = d(d-1)\max(\gamma^{-1}, \rho)$. Then for any column $V \in \boldsymbol{R}^d$ the following statements hold.*

1. *Fix $m \in [\![1, d]\!]$ and $s \geq 1 + (m-1)\max(\gamma^{-1}, \rho)$.*

$$\sum_{i=1}^{m} s^{-i}(MV)^\uparrow(i) \geq \sum_{i=1}^{m} s^{-i} V^\uparrow(i).$$

2. $||MV - V||_1 \leq (2 + C + \cdots + C^{d-2})||(MV)^\uparrow - V^\uparrow||_1.$

Proof By applying a same permutation to the components of V, to the rows and to the columns of M, one may assume that $V(1) \leq \ldots \leq V(d)$. Let σ be a permutation of $[\![1, d]\!]$ such that $(MV)(\sigma(1)) \leq \ldots \leq (MV)(\sigma(d))$. Then Lemmas 3.2 and 3.3 apply.

For every $i \in [\![1, m]\!]$,

$$(MV)^\uparrow(i) - V^\uparrow(i) = B_i - A_i \geq B_i - (m-1)\max(\gamma^{-1}, \rho)(B_1 + \ldots + B_{i-1}).$$

Summing over i yields

$$\begin{aligned}
\sum_{i=1}^{m} s^{-i}((MV)^\uparrow(i) - V^\uparrow(i)) &\geq \sum_{j=1}^{m} s^{-j} B_j - \sum_{i=2}^{m}\sum_{j=1}^{i-1} s^{-i}(m-1)\max(\gamma^{-1}, \rho) B_j \\
&= \sum_{j=1}^{m}\Big(s^{-j} - (m-1)\max(\gamma^{-1}, \rho)\sum_{i=j+1}^{m} s^{-j}\Big) B_j \\
&\geq \sum_{j=1}^{m}\Big(s^{-j} - (m-1)\max(\gamma^{-1}, \rho)\frac{s^{-(j+1)}}{1 - s^{-1}}\Big) B_j \\
&= \sum_{j=1}^{m} \frac{s^{-j}}{s-1}\left[s - 1 - (m-1)\max(\gamma^{-1}, \rho)\right] B_j \\
&\geq 0.
\end{aligned}$$

Furthermore, for every $i \in [\![1, d]\!]$,

$$\left|(MV)(\sigma(i)) - V(\sigma(i))\right| - \left|(MV)(\sigma(i)) - V(i)\right| \leq \left|V(\sigma(i)) - V(i)\right|.$$

Summing over i and using the last statements of Lemma 3.2 and Corollary 3.3 yield

$$\begin{aligned} ||MV - V||_1 - ||(MV)^{\uparrow} - V^{\uparrow}||_1 &\leq \sum_{i=1}^{d} \left|V(\sigma(i)) - V(i)\right| \\ &\leq 2\gamma^{-1} \sum_{i=1}^{d-1} B_i \\ &\leq (1 + C + \cdots + C^{d-2})||(MV)^{\uparrow} - V^{\uparrow}||_1 \end{aligned}$$

The proof is complete.

We now derive the last step of the proof of Theorem 3.7. Indeed, applying the next corollary to each vector of the canonical basis on $\mathbf{R}^d$ yields Theorem 3.7.

Corollary 3.2 *Let $(M_n)_{n\geq 1}$ be a sequence of $d \times d$ stochastic matrices. Assume that there exists some constants $\gamma > 0$, and $\rho \geq 1$ such that for every $n \geq 1$ and i, j in $[\![1, d]\!]$, $M_n(i, i) \geq \gamma$ and $M_n(i, j) \leq \rho M_n(j, i)$. For every column vector $V \in \boldsymbol{R}^d$, the sequence of vectors $(V_n)_{n\geq 0} := (M_n \dots M_1 V)_{n\geq 0}$ has a finite variation, so it converges. Moreover, the series $\sum_n M_n(i, j)$ and $\sum_n M_n(j, i)$ converge whenever the two sequences $(V_n(i))_{n\geq 0}$ and $(V_n(j))_{n\geq 0}$ have a different limit.*

Proof Fix $s \geq 1 + (d-1)\max(\gamma^{-1}, \rho)$. For each n, one can apply Corollary 3.1 to the matrix M_{n+1} and to the vector V_n.

If $m \in [\![1, d]\!]$, the sequence $(s^{-1}V_n^{\uparrow}(1) + \cdots + s^{-m}V_n^{\uparrow}(m))_{n\geq 0}$ is non-decreasing by Corollary 3.1 (first part) and bounded above by $s^{-1}V^{\uparrow}(d) + \cdots + s^{-m}V^{\uparrow}(d)$, thanks to Lemma 3.1, so it has a finite variation and converges. By difference, each sequence $(V_n^{\uparrow}(i))_{n\geq 0}$ has a finite variation. The convergence of the series $\sum_n ||V_{n+1}^{\uparrow} - V_n^{\uparrow}||_1$ follows, and also $\sum_n ||V_{n+1} - V_n||_1$ by Corollary 3.1 (second part).

Call $\lambda_1 < \ldots < \lambda_r$ the distinct values of $\lim_{n\to\infty} V_n(i)$ for $i \in [\![1, d]\!]$. For each $k \in [\![1, r]\!]$, set $I_k = \{i \in [\![1, d]\!] : \lim_{n\to\infty} V_n(i) = \lambda_k\}$, $J_k = I_1 \cup \ldots \cup I_k$ and call m_k the size of J_k. Fix $\varepsilon > 0$ such that $2\varepsilon < \gamma \min(\lambda_2 - \lambda_1, \ldots, \lambda_r - \lambda_{r-1})$, so that the intervals $[\lambda_k - \varepsilon, \lambda_k + \varepsilon]$ are pairwise disjoint. Then one can find some non-negative integer N, such that $V_n(i) \in [\lambda_k - \varepsilon, \lambda_k + \varepsilon]$ for every $n \geq N$, $k \in [\![1, r]\!]$, and $i \in I_k$.

Given $k \in [\![1, r-1]\!]$ and $n \geq N$, we show below that

$$\sum_{i=1}^{m_k} s^{-i}[V_{n+1}^{\uparrow}(i) - V_n^{\uparrow}(i)] \geq (\lambda_{k+1} - \lambda_k - 2\varepsilon)s^{-m_k} \sum_{i \in J_k} \sum_{j \in J_k^c} M_{n+1}(i, j).$$

This inequality together with the convergence of the sequence $(V_n^{\uparrow})_{n\geq 0}$ and the inequalities $M_n(j,i) \leq \rho M_n(i,j)$ will yield the convergence of the series $\sum_n M_n(i,j)$ and $\sum_n M_n(j,i)$ for every $(i,j) \in J_k \times J_k^c$.

Fix $n \geq N$ and a permutation σ of $[\![1, d]\!]$ such that $V_{n+1}(\sigma(1)) \leq \ldots \leq V_{n+1}(\sigma(d))$. Note that $\sigma([\![1, m_k]\!]) = J_k$.

The column vector U_n defined by $U_n(j) = \min(V_n(j), \lambda_k + \varepsilon)$ has the same m_k least components as V_n (corresponding to the indexes $j \in J_k$), so $U_n^{\uparrow}$ have the same m_k first components as $V_n^{\uparrow}$. Furthermore, $V_n(j) - U_n(j) \geq \lambda_{k+1} - \lambda_k - 2\varepsilon$ for every $j \in J_k^c$. Hence, for every $i \in [\![1, d]\!]$,

$$\begin{aligned}
V_{n+1}^{\uparrow}(i) - (M_{n+1}U_n)(\sigma(i)) &= V_{n+1}(\sigma(i)) - (M_{n+1}U_n)(\sigma(i)) \\
&= (M_{n+1}V_n - M_{n+1}U_n))(\sigma(i)) \\
&= \sum_{j=1}^{d} M_{n+1}(\sigma(i), j)\ (V_n(j) - U_n(j)) \\
&\geq (\lambda_{k+1} - \lambda_k - 2\varepsilon) \sum_{j \in J_k^c} M_{n+1}(\sigma(i), j).
\end{aligned}$$

Fix $s \geq 1 + (m_k - 1)\max(\gamma^{-1}, \rho)$. Then

$$\begin{aligned}
\sum_{i=1}^{m_k} s^{-i}\big(V_{n+1}^{\uparrow}(i) - (M_{n+1}U_n)(\sigma(i))\big) &\geq (\lambda_{k+1} - \lambda_k - 2\varepsilon)\ \sum_{i=1}^{m_k} s^{-i} \sum_{j \in J_k^c} M_{n+1}(\sigma(i), j) \\
&\geq (\lambda_{k+1} - \lambda_k - 2\varepsilon)\ s^{-m_k} \sum_{i=1}^{m_k} \sum_{j \in J_k^c} M_{n+1}(\sigma(i), j) \\
&= (\lambda_{k+1} - \lambda_k - 2\varepsilon)\ s^{-m_k} \sum_{i \in J_k} \sum_{j \in J_k^c} M_{n+1}(i, j).
\end{aligned}$$

But the rearrangement inequality[1] and the first part of Corollary 3.1 yield

$$\begin{aligned}\sum_{i=1}^{m_k} s^{-i}(M_{n+1}U_n)(\sigma(i)) &\geq \sum_{i=1}^{m_k} s^{-i}(M_{n+1}U_n)^{\uparrow}(i) \\ &\geq \sum_{i=1}^{m_k} s^{-i}U_n^{\uparrow}(i) \\ &= \sum_{i=1}^{m_k} s^{-i}V_n^{\uparrow}(i).\end{aligned}$$

We get the desired inequality by additioning the last two inequalities.

The proof is complete.

3.3.2 *Proof of Theorem 3.8*

The proof we give is simpler than the proof of Theorem 3.7, although some arguments are very similar. We begin with the key lemma.

Lemma 3.4 *Let M be some $d \times d$ doubly-stochastic matrix with diagonal entries bounded below by some constant $\gamma > 0$, and $V \in \boldsymbol{R}^d$ be any column vector. Call* dispersion of V *the quantity*

$$D(V) = \sum_{1 \leq i,j \leq d} |V(i) - V(j)|.$$

Then $D(V) - D(MV) \geq \gamma ||MV - V||_1$.

Proof On the one hand, for every i and j in $[\![1, d]\!]$,

$$\begin{aligned}(MV)(i) - (MV)(j) &= \sum_{1 \leq k \leq d} M(i,k)V(k) - \sum_{1 \leq l \leq d} M(j,l)V(l) \\ &= \sum_{1 \leq k,l \leq d} M(i,k)M(j,l)(V(k) - V(l)),\end{aligned}$$

[1]Namely $\sum_{i=1}^{d} U^{\uparrow}(d+1-i)V^{\uparrow}(i) \leq \sum_{i=1}^{d} U(i)V(i) \leq \sum_{i=1}^{d} U^{\uparrow}(i)V^{\uparrow}(i)$ for every U and V in $\mathbf{R}^d$.

so

$$D(MV) = \sum_{1\le i,j\le d} \Big| \sum_{1\le k,l\le d} M(i,k)M(j,l)(V(k)-V(l)) \Big|.$$

On the other hand

$$\begin{aligned} D(V) &= \sum_{1\le k,l\le d} |V(k)-V(l)| \\ &= \sum_{1\le i,j\le d} \sum_{1\le k,l\le d} M(i,k)M(j,l)|V(k)-V(l)|. \end{aligned}$$

By difference, $D(V) - D(MV)$ is the sum over all i and j in $[\![1,d]\!]$ of the non negative quantities

$$\Delta(i,j) = \sum_{1\le k,l\le d} M(i,k)M(j,l)|V(k)-V(l)| - \Big| \sum_{1\le k,l\le d} M(i,k)M(j,l)(V(k)-V(l)) \Big|.$$

Thus

$$D(V) - D(MV) \ge \sum_{1\le i\le d} \Delta(i,i).$$

But for every $i \in [\![1,d]\!]$,

$$\begin{aligned} \Delta(i,i) &= \sum_{1\le k,l\le d} M(i,k)M(i,l)|V(k)-V(l)| - 0 \\ &\ge \sum_{1\le k\le d} M(i,k)M(i,i)|V(k)-V(i)| \\ &\ge \gamma \sum_{1\le k\le d} M(i,k)|V(k)-V(i)| \\ &\ge \gamma \Big| \sum_{1\le k\le d} M(i,k)(V(k)-V(i)) \Big| \\ &= \gamma\ |(MV)(i) - V(i)|. \end{aligned}$$

The result follows.

We now derive the last step of the proof of Theorem 3.8. Indeed, applying the next corollary to each vector of the canonical basis on $\mathbf{R}^d$ yields Theorem 3.8.

Corollary 3.3 *Let $(M_n)_{n\geq 1}$ be any sequence of $d \times d$ bistochastic matrices with diagonal entries bounded below by some $\gamma > 0$. For every column vector $V \in \boldsymbol{R}^d$, the sequence $(V_n)_{n\geq 0} := (M_n \dots M_1 V)_{n\geq 0}$ has a finite variation, so it converges. Moreover, the series $\sum_n M_n(i, j)$ and $\sum_n M_n(j, i)$ converge whenever the two sequences $(V_n(i))_{n\geq 0}$ and $(V_n(j))_{n\geq 0}$ have a different limit.*

Proof Lemma 3.4 yields $\gamma||V_{n+1} - V_n||_1 \leq D(V_n) - D(V_{n+1})$ for every $n \geq 0$. In particular, the sequence $((D(V_n))_{n\geq 0}$ is non-increasing and bounded below by 0, so it converges. The convergence of the series $\sum_n ||V_{n+1} - V_n||_1$ and the convergence of the sequence $(V_n)_{n\geq 0}$ follow.

Call $\lambda_1 < \dots < \lambda_r$ the distinct values of $\lim_{n\to\infty} V_n(i)$ for $i \in [\![1, d]\!]$. For each $k \in [\![1, r]\!]$, set $I_k = \{i \in [\![1, d]\!] : \lim_{n\to\infty} V_n(i) = \lambda_k\}$, and $J_k = I_1 \cup \dots \cup I_k$.

The proof of the convergence of the series $\sum_n M_{n+1}(i, j)$ for every $(i, j) \in J_k \times J_k^c$. works like the proof of Corollary 3.2, with r replaced by 1, thanks to Lemma 3.5 stated below, so the rearrangement inequality becomes an equality.

Using the equality

$$\sum_{(i,j)\in J_k^c\times J_k} M_{n+1}(i, j) = |J_k| - \sum_{(i,j)\in J_k\times J_k} M_{n+1}(i, j) = \sum_{(i,j)\in J_k\times J_k^c} M_{n+1}(i, j),$$

we derive the convergence of the series $\sum_n M_{n+1}(i, j)$ for every $(i, j) \in J_k^c \times J_k$. The proof is complete.

Lemma 3.5 *Let M be some $d \times d$ doubly-stochastic matrix. Then for every column vector $V \in \boldsymbol{R}^d$ and $m \in [\![1, d]\!]$*

$$\sum_{i=1}^{m} (MV)^{\uparrow}(i) \geq \sum_{i=1}^{m} V^{\uparrow}(i).$$

Proof By applying a same permutation to the columns of M and to the components of V, one may assume that $V(1) \leq \dots \leq V(d)$. By applying a permutation to the rows of M, one may assume also that $(MV)(1) \leq \dots \leq (MV)(d)$. Since M is doubly-stochastic, the real numbers

$$S(j) = \sum_{i=1}^{m} M(i, j) \text{ for } j \in [\![1, d]\!]$$

are in $[0, 1]$ and add up to m. Moreover

$$\sum_{i=1}^{m} (MV)(i) = \sum_{j=1}^{d} S(j)V(j).$$

Hence

$$\begin{aligned}\sum_{i=1}^{m}(MV)(i)-\sum_{j=1}^{m}V(j) &= \sum_{j=m+1}^{d}S(j)V(j)+\sum_{j=1}^{m}(S(j)-1)V(j)\\ &\geq \sum_{j=m+1}^{d}S(j)V(m)+\sum_{j=1}^{m}(S(j)-1)V(m)\\ &= 0.\end{aligned}$$

We are done.

3.4 Proof of Theorem 3.1

3.4.1 Condition for the Non-existence of a Solution with Support Included in Supp(X_0)

We assume that Γ contains no matrix with support included in Supp(X_0), namely that the system

$$\begin{cases} \forall i \in [\![1,p]\!],\ X(i,+)=a_i \\ \forall j \in [\![1,q]\!],\ X(+,j)=b_j \\ \forall (i,j) \in [\![1,p]\!]\times[\![1,q]\!],\ X(i,j)\geq 0 \\ \forall (i,j) \in \mathrm{Supp}(X_0)^c,\ X(i,j)=0 \end{cases}$$

is inconsistent.

This system can be seen as a system of linear inequalities of the form $\ell(X) \leq c$ (where ℓ is some linear form and c some constant) by splitting each equality $\ell(X) = c$ into the two inequalities $\ell(X) \leq c$ and $\ell(X) \geq c$, and by transforming each inequality $\ell(X) \geq c$ into the equivalent inequality $-\ell(X) \leq -c$. But theorem 4.2.3 in [18] (a consequence of Farkas' or Fourier's lemma) states that a system of linear inequalities of the form $\ell(X) \leq c$ is inconsistent if and only if some linear combination with non-negative weights of the linear inequalities yields the inequality $0 \leq -1$.

Given such a linear combination, call $\alpha_{i,+}$, $\alpha_{i,-}\beta_{j,+}$, $\beta_{j,-}$, $\gamma_{i,j,+}$, $\gamma_{i,j,-}$ the weights associated to the inequalities $X(i,+) \leq a_i$, $-X(i,+) \leq -a_i$, $X(+,j) \leq b_j$, $-X(+,j) \leq -b_j$, $X(i,j) \leq 0$, $-X(i,j) \leq 0$. When $(i,j) \in \mathrm{Supp}(X_0)$, the inequality $X(i,j) \leq 0$ does not appear in the system, so we set $\gamma_{i,j,+} = 0$. Then the real numbers $\alpha_i := \alpha_{i,+} - \alpha_{i,-}$, $\beta_j := \beta_{j,+} - \beta_{j,-}$, $\gamma_{i,j} := \gamma_{i,j,+} - \gamma_{i,j,-}$ satisfy the following conditions:

- for every $(i, j) \in [\![1, p]\!] \times [\![1, q]\!]$, $\alpha_i + \beta_j + \gamma_{i,j} = 0$,
- $\sum_{i=1}^{p} \alpha_i a_i + \sum_{j=1}^{q} \beta_j b_j = -1$,
- $\gamma_{i,j} \leq 0$ whenever $(i, j) \in \mathrm{Supp}(X_0)$.

Let U and V be two random variables with respective laws

$$\sum_{i=1}^{p} a_i \delta_{\alpha_i} \text{ and } \sum_{j=1}^{q} b_j \delta_{-\beta_j}.$$

Then

$$\int_{\mathbf{R}} \big(P[U > t] - P[V > t]\big)\mathrm{d}t = \mathbf{E}[U] - \mathbf{E}[V] = \sum_{i=1}^{p} \alpha_i a_i + \sum_{j=1}^{q} \beta_j b_j = -1 < 0,$$

so there exists some real number t such that $P[U > t] - P[V > t] < 0$. Consider the sets $A = \{i \in [\![1, p]\!] : \alpha_i \leq t\}$ and $B = \{j \in [\![1, q]\!] : \beta_j < -t\}$. Then for every $(i, j) \in A \times B$, $-\gamma_{i,j} = \alpha_i + \beta_j < 0$, so $(i, j) \notin \mathrm{Supp}(X_0)$. In other words, X_0 is null on $A \times B$. Moreover,

$$\begin{aligned} a(A) - b(B^c) &= \sum_{i \in A} a_i - \sum_{j \in B^c} b_j = P[U \leq t] - P[-V \geq -t] \\ &= P[V > t] - P[U > t] > 0. \end{aligned}$$

Hence the block $A \times B$ is a cause of incompatibility. The proof is complete.

3.4.2 *Condition for the Existence of Additional Zeroes Shared by Every Solution in $\Gamma(X_0)$*

We now assume that Γ contains some matrix with support included in $\mathrm{Supp}(X_0)$.

Using the convexity of $\Gamma(X_0)$, one can construct a matrix $S_0 \in \Gamma(X_0)$ whose support contains the support of every matrix in $\Gamma(X_0)$. This yields item (a).

We now prove items (b) and (c). The observations made in the introduction show that $\mathrm{Supp}(S_0) \subset \mathrm{Supp}(X_0)$ and that S_0 is null on $A^c \times B^c$ whenever $A \times B$ is a non-empty subset of $[\![1, p]\!] \times [\![1, q]\!]$ such that X_0 is null on $A \times B$ and $a(A) = b(B^c)$. This yields the 'if' part of item (b) and one inclusion in item (c).

To prove the 'only if' part of item (b) and the reverse inclusion in item (c), fix $(i_0, j_0) \in \mathrm{Supp}(X_0) \setminus \mathrm{Supp}(S_0)$. Then for every $p \times q$ matrix X with real entries,

$$\left.\begin{array}{r}\forall i \in [\![1, p]\!],\ X(i,+) = a_i \\ \forall j \in [\![1, q]\!],\ X(+,j) = b_j \\ \forall (i,j) \in [\![1, p]\!] \times [\![1, q]\!],\ X(i,j) \geq 0 \\ \forall (i,j) \in \mathrm{Supp}(X_0)^c,\ X(i,j) = 0\end{array}\right\} \Longrightarrow X(i_0, j_0) \leq 0.$$

The system in the left-hand side of the implication is consistent since $\Gamma(X_0)$ is non-empty by assumption. As before, the system at the left-hand side of the implication can be seen as a system of linear inequalities of the form $\ell(X) \leq c$.

We now use theorem 4.2.7 in [18] (a consequence of Farkas' or Fourier's lemma) which states that any linear inequation which is a consequence of some consistent system of linear inequalities of the form $\ell(X) \leq c$ can be deduced from the system and from the inequality $0 \leq 1$ by linear combinations with non-negative weights.

Given such a linear combination, call $\alpha_{i,+}, \alpha_{i,-} \beta_{j,+}, \beta_{j,-}, \gamma_{i,j,+}, \gamma_{i,j,-}$ and η the weights associated to the inequalities $X(i,+) \leq a_i$, $-X(i,+) \leq -a_i$, $X(+,j) \leq b_j$, $-X(+,j) \leq -b_j$, $X(i,j) \leq 0$, $-X(i,j) \leq 0$, and $0 \leq 1$. When $(i,j) \in \mathrm{Supp}(X_0)$, the inequality $X(i,j) \leq 0$ does not appear in the system, so we set $\gamma_{i,j,+} = 0$. Then the real numbers $\alpha_i := \alpha_{i,+} - \alpha_{i,-}$, $\beta_j := \beta_{j,+} - \beta_{j,-}$, $\gamma_{i,j} := \gamma_{i,j,+} - \gamma_{i,j,-}$ satisfy the following conditions.

- for every $(i,j) \in [\![1, p]\!] \times [\![1, q]\!]$, $\alpha_i + \beta_j + \gamma_{i,j} = \delta_{i,i_0}\delta_{j,j_0}$,
- $\displaystyle\sum_{i=1}^{p} \alpha_i a_i + \sum_{j=1}^{q} \beta_j b_j + \eta = 0.$
- $\gamma_{i,j} \leq 0$ whenever $(i,j) \in \mathrm{Supp}(X_0)$,
- $\eta \geq 0$

Let U and V be two random variables with respective laws

$$\sum_{i=1}^{p} a_i \delta_{\alpha_i} \text{ and } \sum_{j=1}^{q} b_j \delta_{-\beta_j}.$$

Then

$$\int_{\mathbf{R}} \big(P[U > t] - P[V > t]\big)\,\mathrm{d}t = \mathbf{E}[U] - \mathbf{E}[V] = \sum_{i=1}^{p} \alpha_i a_i + \sum_{j=1}^{q} \beta_j b_j = -\eta \leq 0.$$

Set $u_0 = \min(\alpha_1, \ldots, \alpha_p) = \operatorname{ess\,inf} U$ and $v_0 = \max(-\beta_1, \ldots, -\beta_q) = \operatorname{ess\,sup} V$. Then $P[U > t] - P[V > t] \geq 0$ when $t < u_0$ or $t \geq v_0$. Hence one can find $t \in [u_0, v_0[$ such that $P[U > t] - P[V > t] \leq 0$. Consider the sets $A = \{i \in [\![1, p]\!] : \alpha_i \leq t\}$ and $B = \{j \in [\![1, q]\!] : \beta_j < -t\}$. For every $(i,j) \in A \times B$, $\gamma_{i,j} = \delta_{i,i_0}\delta_{j,j_0} - \alpha_i - \beta_j > 0$, so $(i,j) \notin \mathrm{Supp}(X_0)$. In other words, X_0 is null on

$A \times B$. Moreover,

$$a(A) = \sum_{i \in A} a_i = P[U \leq t] \geq P[U = u_0] > 0,$$

$$b(B) = \sum_{j \in B} b_j = P[-V < -t] \geq P[V = v_0] > 0,$$

so A and B are non-empty and

$$a(A) - b(B^c) = P[U \leq t] - P[V \leq t] = P[V > t] - P[U > t] \geq 0.$$

The last inequality is necessarily an equality, since otherwise $A \times B$ would be a cause of incompatibility. Hence $a(A) = b(B^c)$, so $A \times B$ is cause of criticality. This proves the 'only if' part of item (b).

We also know that $(i_0, j_0) \notin A \times B$ since $X_0(i_0, j_0) > 0$. If we knew that $(i_0, j_0) \in A^c \times B^c$, we would get the reverse inclusion in item (c). Unfortunately, this statement may fail with the choice of t made above.

Assume, that $(i_0, j_0) \notin A^c \times B^c$. Since $a(A) = b(B^c)$, the linear system defining $\Gamma(a, b, X_0)$ can be split into three independent consistent subsystems, namely

$$\forall (i, j) \in (A \times B) \cup (A^c \times B^c), X(i, j) = 0$$

and the two systems

$$\begin{cases} \forall i \in I, \ X(i, J) = a_i \\ \forall j \in J, \ X(I, j) = b_j \\ \forall (i, j) \in I \times J, \ X(i, j) \geq 0 \\ \forall (i, j) \in (I \times J) \setminus \text{Supp}(X_0), \ X(i, j) = 0 \end{cases}$$

where the block $I \times J$ is either $A \times B^c$ or $A^c \times B$.

If $(i_0, j_0) \notin A^c \times B^c$, then (i_0, j_0) belongs to one of these two blocks, say $I_1 \times J_1$. Since $a(I_1) = b(J_1)$ and since the equality $X(i_0, j_0) = 0$ is a consequence of the consistent subsystem above with $I \times J = I_1 \times J_1$, one can apply the proof of item (b) to the marginals $a(\cdot|I_1)$ and $b(\cdot|J_1)$ and the restriction of X_0 on $I_1 \times J_1$. This yields a subset $A_1 \times B_1$ of $I_1 \times J_1$ such that X_0 is null on $A_1 \times B_1$, $a(A_1) = b(B_1^c)$, and $(i_0, j_0) \notin A_1 \times B_1$.

If $(i_0, j_0) \in A_1^c \times B_1^c$, we are done. Otherwise, (i_0, j_0) belongs to one of the two blocks $A_1 \times B_1^c$ or $A_1^c \times B_1$, say $I_2 \times J_2$, and the recursive construction goes on. This construction necessarily stops after a finite number of steps and produces a block $A' \times B'$ on which X_0 is null, such that $a(A') = b(B'^c)$ and $(i_0, j_0) \in A'^c \times B'^c$. Item (c) follows.

3.5 Tools and Preliminary Results

3.5.1 Results on the Quantities $R_i(X_{2n})$ and $C_j(X_{2n+1})$

Lemma 3.6 *Let $X \in \Gamma_1$. Then*

$$\sum_{i=1}^{q} a_i R_i(X) = \sum_{i,j} X(i,j) = \sum_j b_j = 1$$

and for every $j \in [\![1, q]\!]$,

$$C_j(T_R(X)) = \sum_{i=1}^{p} \frac{X(i,j)}{b_j} R_i(X)^{-1}.$$

When $X \in \Gamma_C$, this equality expresses $C_j(T_R(X))$ as a weighted (arithmetic) mean of the quantities $R_i(X)^{-1}$, with weights $X(i,j)/b_j$.

For each $X \in \Gamma_1$, call $R(X)$ the column vector with components $R_1(X), \dots, R_p(X)$ and $C(X)$ the column vector with components $C_1(X), \dots, C_q(X)$. Set

$$\underline{R}(X) = \min_i R_i(X),\ \overline{R}(X) = \max_i R_i(X),\ \underline{C}(X) = \min_j C_j(X),\ \overline{C}(X) = \max_i C_j(X).$$

Corollary 3.4 *The intervals*

$$[\overline{C}(X_1)^{-1}, \underline{C}(X_1)^{-1}],\ [\underline{R}(X_2), \overline{R}(X_2)],\ [\overline{C}(X_3)^{-1}, \underline{C}(X_3)^{-1}],\ [\underline{R}(X_4), \overline{R}(X_4)], \cdots$$

contain 1 *and form a non-increasing sequence.*

In Lemma 3.6, one can invert the roles of the lines and the columns. Given $X \in \Gamma_C$, the matrix $T_R(X)$ is in Γ_R so the quantities $R_i(T_C(T_R(X)))$ can be written as weighted (arithmetic) means of the $C_j(T_R(X))^{-1}$. But the $C_j(T_R(X))$ can be written as a weighted (arithmetic) means of the quantities $R_k(X)^{-1}$. Putting things together, one gets weighted arithmetic means of weighted harmonic means. Next lemma shows how to transform these into weighted arithmetic means by modifying the weights.

Lemma 3.7 *Let $X \in \Gamma_C$. Then $R(T_C(T_R(X))) = P(X)R(X)$, where $P(X)$ is the $p \times p$ matrix given by*

$$P(X)(i,k) = \sum_{j=1}^{q} \frac{T_R(X)(i,j)T_R(X)(k,j)}{a_i b_j C_j(T_R(X))}.$$

The matrix $P(X)$ is stochastic. Moreover it satisfies for every i and k in $[\![1,p]\!]$,

$$P(X)(i,i) \geq \frac{\underline{a}}{\overline{b}\,\overline{C}(T_R(X))q}$$

and

$$P(X)(k,i) \leq \frac{\overline{a}}{\underline{a}} P(X)(i,k).$$

Proof For every $i \in [\![1,p]\!]$,

$$R_i(T_C(T_R(X))) = \sum_{j=1}^{q} \frac{T_R(X)(i,j)}{a_i} \frac{1}{C_j(T_R(X))}.$$

But the assumption $X \in \Gamma_C$ yields

$$1 = \frac{1}{b_j} \sum_{k=1}^{p} X(k,j) = \frac{1}{b_j} \sum_{k=1}^{p} T_R(X)(k,j) R_k(X).$$

Hence,

$$R_i(T_C(T_R(X))) = \sum_{j=1}^{q} \frac{T_R(X)(i,j)}{a_i} \sum_{k=1}^{p} \frac{T_R(X)(k,j)}{b_j C_j(T_R(X))} R_k(X).$$

These equalities can be written as $R(T_C(T_R(X))) = P(X)R(X)$, where $P(X)$ is the $p \times p$ matrix whose entries are given in the statement of Lemma 3.7. By construction, the entries of $P(X)$ are non-negative and for every $i \in [\![1,p]\!]$,

$$\sum_{k=1}^{p} P(X)(i,k) = \sum_{j=1}^{q} \frac{T_R(X)(i,j)}{a_i b_j C_j(T_R(X))} \sum_{k=1}^{p} T_R(X)(k,j) = \sum_{j=1}^{q} \frac{T_R(X)(i,j)}{a_i} = 1.$$

Moreover, since

$$\sum_{j=1}^{q} T_R(X)(i,j)^2 \geq \frac{1}{q} \Big(\sum_{j=1}^{q} T_R(X)(i,j) \Big)^2 = \frac{a_i^2}{q}$$

we have

$$\begin{aligned} P(X)(i,i) &\geq \frac{1}{a_i\,\overline{b}\,\overline{C}(T_R(X))}\sum_{j=1}^{q} T_R(X)(i,j)^2 \\ &= \frac{a_i}{\overline{b}\,\overline{C}(T_R(X))\,q} \\ &\geq \frac{\underline{a}}{\overline{b}\,\overline{C}(T_R(X))\,q}. \end{aligned}$$

The last inequality to be proved follows directly from the symmetry of the matrix $(a_i P(X)(i,k))_{1\leq i,k\leq p}$.

3.5.2 A Function Associated to Each Element of Γ_1

Definition 3.1 For every X and S in Γ_1, we set

$$F_S(X) = \prod_{(i,j)\in[\![1,p]\!]\times[\![1,q]\!]} X(i,j)^{S(i,j)},$$

with the convention $0^0 = 1$.

We note that $0 \leq F_S(X) \leq 1$, and that $F_S(X) > 0$ if and only if $\mathrm{Supp}(S) \subset \mathrm{Supp}(X)$.

Lemma 3.8 *Let $S \in \Gamma_1$. For every $X \in \Gamma_1$, $F_S(X) \leq F_S(S)$, with equality if and only if $X = S$. Moreover, if* $\mathrm{Supp}(S) \subset \mathrm{Supp}(X)$, *then $D(S||X) = \ln(F_S(S)/F_S(X))$.*

Proof Assume that $\mathrm{Supp}(S) \subset \mathrm{Supp}(X)$. The definition of F_S and the arithmetic-geometric inequality yield

$$\frac{F_S(S)}{F_S(X)} = \prod_{i,j}\Big(\frac{X(i,j}{S(i,j)}\Big)^{S(i,j)} \leq \sum_{i,j} S(i,j)\Big(\frac{X(i,j}{S(i,j)}\Big) = \sum_{i,j} X(i,j) = 1,$$

with equality if and only if $X(i,j) = S(i,j)$ for every $(i,j) \in \mathrm{Supp}(S)$. The result follows.

Lemma 3.9 *Let $X \in \Gamma_1$.*

- *For every $S \in \Gamma_R$ such that* $\mathrm{Supp}(S) \subset \mathrm{Supp}(X)$, *one has $F_S(X) \leq F_S(T_R(X))$, and the ratio $F_S(X)/F_S(T_R(X))$ does not depend on S.*
- *For every $S \in \Gamma_C$ such that* $\mathrm{Supp}(S) \subset \mathrm{Supp}(X)$, *one has $F_S(X) \leq F_S(T_C(X))$, and the ratio $F_S(X)/F_S(T_C(X))$ does not depend on S.*

Proof Let $S \in \Gamma_R$. For every $(i, j) \in \mathrm{Supp}(X)$, $X(i, j)/(T_R(X))(i, j) = R_i(X)$ so the arithmetic-geometric mean inequality yields

$$\frac{F_S(X)}{F_S(T_R(X))} = \prod_{i,j} R_i(X)^{S_{i,j}} = \prod_i R_i(X)^{a_i} \leq \sum_i a_i R_i(X) = \sum_{i,j} X(i, j) = 1.$$

The first statement follows. The second statement is proved in the same way.

Corollary 3.5 *Assume that $\Gamma(X_0)$ is not empty. Then:*

1. *for every $S \in \Gamma(X_0)$, the sequence $(F_S(X_n))_{n\geq 0}$ is non-decreasing and bounded above, so it converges;*
2. *for every (i, j) in the union of the supports* Supp(S) *over all $S \in \Gamma(X_0)$, the sequence $(X_n(i, j))_{n\geq 1}$ is bounded away from* 0.

Proof Lemmas 3.8 and 3.9 yield the first item. Given $S \in \Gamma(X_0)$ and $(i, j) \in \mathrm{Supp}(S)$, we get for every $n \geq 0$, $X_n(i, j)^{S(i,j)} \geq F_S(X_n) \geq F_S(X_0) > 0$. The second item follows.

The first item of Corollary 3.5 will yield the first items of Theorem 3.3. The second item of Corollary 3.5 is crucial to establish the geometric rate of convergence in Theorem 3.2 and the convergence in Theorem 3.3.

3.6 Proof of Theorem 3.2

In this section, we assume that Γ contains some matrix having the same support as X_0, and we establish the convergences with at least geometric rate stated in Theorem 3.2. The main tools are Lemma 3.1 and the second item of Corollary 3.5. Corollary 3.5 shows that the non-zero entries of all matrices X_n are bounded below by some positive real number γ. Therefore, the non-zero entries of all matrices $X_n X_n^\top$ are bounded below by γ^2. These matrix have the same support as $X_0 X_0^\top$.

By Lemma 3.7, for every $n \geq 1$, $R(X_{2n+2}) = P(X_{2n})R(X_{2n})$, where $P(X_{2n})$ is a stochastic matrix given by

$$P(X_{2n})(i, i') = \sum_{j=1}^{q} \frac{X_{2n+1}(i, j)X_{2n+1}(i', j)}{a_i b_j C_j(X_{2n+1})}.$$

These matrices have also the same support as $X_0 X_0^\top$. Moreover, by Lemma 3.7 and Corollary 3.4,

$$P(X_{2n})(i, i') \geq \frac{1}{\overline{a}\,\overline{b}\,\overline{C}(X_3)}(X_{2n+1}X_{2n+1}^\top)(i, i'),$$

so the non-zero entries of $P(X_{2n})$ are bounded below by $\gamma^2/(\overline{a}\,\overline{b}\,\overline{C}(X_3)) > 0$.

We now define a binary relation on the set $[\![1, p]\!]$ by

$$i \mathscr{R} i' \iff X_0 X_0^\top(i, i') > 0 \iff \exists j \in [\![1, q]\!],\ X_0(i, j) X_0(i', j) > 0.$$

The matrix $X_0 X_0^\top$ is symmetric with positive diagonal (since on each line, X_0 has at least a positive entry), so the relation $\mathscr{R}$ is symmetric and reflexive. Call $I_1, \dots, I_r$ the connected components of the graph G associated to $\mathscr{R}$, and d the maximum of their diameters. For each k, set $J_k = \{j \in [\![1, q]\!] : \exists i \in I_k : X_0(i, j) > 0\}$.

Lemma 3.10 *The sets $J_1, \dots, J_k$ form a partition of $[\![1, q]\!]$ and the support of X_0 is contained in $I_1 \times J_1 \cup \dots \cup I_r \times J_r$. Therefore, the support of $X_0 X_0^\top$ is contained in $I_1 \times I_1 \cup \dots \cup I_r \times I_r$, so one can get a block-diagonal matrix by permuting suitably the lines of X_0.*

Proof By assumption, the sum of the entries of X_0 on any row or any column is positive. Given $k \in [\![1, r]\!]$ and $i \in I_k$, there exists $j \in [\![1, q]\!]$ such that $X_0(i, j) > 0$, so J_k is not empty.

Fix now $j \in [1, q]$. There exists $i \in [\![1, p]\!]$ such that $X_0(i, j) > 0$. Such an i belongs to some connected component I_k, so j belongs J_k. If j also belongs to $J_{k'}$, then $X_0(i', j) > 0$ for some $i' \in I_{k'}$, so $X_0 X_0^\top(i, i') \geq X_0(i, j) X_0(i', j) > 0$, hence i and i' belong to the connected component of G, so $k' = k$.

The other statements follow.

Lemma 3.11 *For $n \geq 1$, set $P_{2n} = P(X_{2n})$ and $M_n = P_{2n+2d-2} \cdots P_{2n+2} P_{2n}$. Call c the minimum of all positive entries of all matrices P_{2n}. Then $c > 0$, and for every $n \geq 1$, $i \in I_k$ and $i' \in I_{k'}$.*

$$M_n(i, i') \geq c^d \quad \text{if } k = k'.$$
$$M_n(i, i') = 0 \quad \text{if } k \neq k'.$$

Proof The positivity of c has already be proved at the beginning of the present section. Moreover, $M_n(i, i')$ is the sum of the products

$$P_{2n+2d-2}(i, i_1) P_{2n+2d-4}(i_1, i_2) \cdots P_{2n+2}(i_{d-2}, i_{d-1}) P_{2n}(i_{d-1}, i').$$

over all $(i_1, \dots, i_{d-1})$ in $[\![1, p]\!]^{d-1}$.

If $k \neq k'$, all these products are 0, since no path can connect i and i' in the graph G, so $M_n(i, i') = 0$.

If $k = k'$, one can find a path $i = i_0, \dots, i_\ell = i'$ in the graph G, with length $\ell \leq d$. Setting $i_{\ell+1} = \dots = i_d$ if $\ell < d$, we get

$$P_{2n+2d-2}(i, i_1) P_{2n+2d-4}(i_1, i_2) \cdots P_{2n+2}(i_{d-2}, i_{d-1}) P_{2n}(i_{d-1}, i') \geq c^d.$$

The result follows.

Keep the notations of the last lemma. Then $R(X_{2n+2d}) = M_n R(X_{2n})$ for every $n \geq 1$. For each $k \in [\![1, r]\!]$, Lemma 3.1 applied to the submatrix $(M_n(i, i'))_{i,i' \in I_k}$ and the vector $R_{I_k}(X_{2n}) = (R_i(X_{2n}))_{i \in I_k}$ yields

$$\operatorname{diam}(R_{I_k}(X_{2n+2d})) \leq (1 - c^d) \operatorname{diam}(R_{I_k}(X_{2n})).$$

But Lemma 3.1 applied to the submatrix $(P(X_{2n})(i, i'))_{i,i' \in I_k}$ shows that the intervals

$$\Big[\min_{i \in I_k} R_i(X_{2n}), \max_{i \in I_k} R_i(X_{2n})\Big]$$

indexed by $n \geq 1$ form a non-increasing sequence. Hence, each sequence $(R_i(X_{2n}))$ tends to a limit which does not depend on $i \in I_k$, and the speed of convergence it at least geometric.

Call λ_k this limit. By Lemma 3.10, we have for every $n \geq 1$,

$$\sum_{i \in I_k} a_i R_i(X_{2n}) = \sum_{(i,j) \in I_k \times J_k} X_{2n}(i, j) = \sum_{j \in J_k} X_{2n}(+, j) = \sum_{j \in J_k} b_j$$

Passing to the limit yields

$$\lambda_k \sum_{i \in I_k} a_i = \sum_{j \in J_k} b_j,$$

whereas the assumption that Γ contains some matrix S having the same support as X_0 yields

$$\sum_{i \in I_k} a_i = \sum_{(i,j) \in I_k \times J_k} S(i, j) = \sum_{j \in J_k} b_j.$$

Thus $\lambda_k = 1$.

We have proved that each sequence $(R_i(X_{2n}))_{n \geq 0}$ tends to 1 with at least geometric rate. The same arguments work for the sequences $(C_j(X_{2n+1}))_{n \geq 0}$. Therefore, each infinite product $R_i(X_0)R_i(X_2)\cdots$ or $C_j(X_1)C_j(X_3)\cdots$ converges at an at least geometric rate. The convergence of the sequence $(X_n)_{n \geq 0}$ with at a least geometric rate follows.

Moreover, call α_i and β_j the inverses of the infinite products $R_i(X_0)R_i(X_2)\cdots$ and $C_j(X_1)C_j(X_3)\cdots$ and X_∞ the limit of $(X_n)_{n \geq 0}$. Then $X_\infty(i, j) = \alpha_i \beta_j X_0(i, j)$, so X_∞ belongs to the set $\Delta_p X_0 \Delta_q$. As noted in the introduction, we have also $X_\infty \in \Gamma$. It remains to prove that X_∞ is the only matrix in $\Gamma \cap \Delta_p X_0 \Delta_q$ and the only matrix which achieves the least upper bound of $D(Y||X_0)$ over all $Y \in \Gamma(X_0)$.

Let E_{X_0} be the vector space of all matrices in $\mathscr{M}_{p,q}(\mathbf{R})$ which are null on $\mathrm{Supp}(X_0)^c$ (which can be identified canonically with $\mathbf{R}^{\mathrm{Supp}(X_0)}$), and $E^+_{X_0}$ be the convex subset of all non-negative matrices in E_{X_0}. The subset $E^{+*}_{X_0}$ of all matrices in E_{X_0} which are positive on $\mathrm{Supp}(X_0)^c$, is open in E_{X_0}, dense in $E^+_{X_0}$ and contains X_∞. Consider the map f_{X_0} from $E^+_{X_0}$ to $\mathbf{R}$ defined by

$$f_{X_0}(Y) = \sum_{(i,j)\in \mathrm{Supp}(X_0)} Y(i,j) \ln \frac{Y(i,j)}{X_0(i,j)},$$

with the convention $t \ln t = 0$. This map is strictly convex since the map $t \mapsto t \ln t$ from $\mathbf{R}_+$ to $\mathbf{R}$ is. Its differential at any point $Y \in E^{+*}_{X_0}$ is given by

$$\mathrm{d}f_{X_0}(Y)(H) = \sum_{(i,j)\in \mathrm{Supp}(X_0)} \Big(\ln \frac{Y(i,j)}{X_0(i,j)} + 1 \Big) H(i,j).$$

Now, if Y_0 is any matrix in $\Gamma \cap \Delta_p X_0 \Delta_q$ (including the matrix X_∞), the quantities $\ln(Y_0(i,j)/X_0(i,j))$ can be written $\lambda_i + \mu_j$. Thus for every matrix $H \in E(X_0)$ with null row-sums and column-sums, $\mathrm{d}f_{X_0}(Y_0)(H) = 0$, hence the restriction of f_{X_0} to $\Gamma(X_0)$ has a strict global minimum at Y_0. The proof is complete.

The Case of Positive Matrices

The proof of the convergence at an at least geometric rate can be notably simplified when X_0 has only positive entries. In this case, Fienberg [6] used geometric arguments to prove the convergence of the iterated proportional fitting procedure at an at least geometric rate. We sketch another proof using the observation made by Fienberg that the ratios

$$\frac{X_n(i,j)X_n(i',j')}{X_n(i,j')X_n(i',j)}$$

are independent of n, since they are preserved by the transformations T_R and T_C. Call κ the least of these positive constants. Using Corollary 3.4, one checks that the average of the entries of X_n on each row or column remains bounded below by some constant $\gamma > 0$. Thus for every location (i,j) and $n \geq 1$, one can find two indexes $i' \in [\![1,p]\!]$ and $j' \in [\![1,q]\!]$ such that $X_n(i',j) \geq \gamma$ and $X_n(i,j') \geq \gamma$, so

$$X_n(i,j) \geq X_n(i,j)X_n(i',j') \geq \kappa X_n(i,j')X_n(i',j) \geq \kappa\gamma^2.$$

This shows that the entries of the matrices X_n remain bounded away from 0, so the ratios $X_n(i,j)/b_j$ and $X_n(i,j)/a_i$ are bounded below by some constant $c > 0$

independent of $n \geq 1$, i and j. Set

$$\rho_n = \frac{\overline{R}(X_n)}{\underline{R}(X_n)} \text{ if } n \text{ is even, } \rho_n = \frac{\overline{C}(X_n)}{\underline{C}(X_n)} \text{ if } n \text{ is odd.}$$

For every $n \geq 1$, the matrix $(X_{2n}(i,j)/b_j)_{1 \leq i \leq p, 1 \leq j \leq q}$ is stochastic, so the equalities

$$C_j(X_{2n+1}) = \sum_{i=1}^{p} (X_{2n}(i,j)/b_j) R_i(X_{2n+1})^{-1}$$

and Lemma 3.1 yields $\underline{C}(X_{2n+1}) \geq \overline{R}(X_{2n})^{-1}$ and

$$\overline{C}(X_{2n+1}) - \underline{C}(X_{2n+1}) \leq (1-2c)(\underline{R}(X_{2n})^{-1} - \overline{R}(X_{2n})^{-1}).$$

Thus,

$$\rho_{2n+1} - 1 = \frac{\overline{C}(X_{2n+1}) - \underline{C}(X_{2n+1})}{\underline{C}(X^{(2n+1)})} \leq (1-2c)(\rho_{2n} - 1).$$

We prove the inequality $\rho_{2n} - 1 \leq (1-2c)(\rho_{2n-1} - 1)$ in the same way. Hence $\rho_n \to 1$ at an at least geometric rate. The result follows by Corollary 3.4.

3.7 Proof of Theorem 3.3

We now assume that Γ contains some matrix with support included in $\mathrm{Supp}(X_0)$.

3.7.1 Asymptotic Behavior of the Sequences $(R(X_n))$ and $(C(X_n))$

The first item of Corollary 3.5 yields the convergence of the infinite product

$$\prod_i R_i(X_0)^{a_i} \times \prod_j C_j(X_1)^{b_j} \times \prod_i R_i(X_2)^{a_i} \times \prod_j C_j(X_3)^{b_j} \times \cdots .$$

Set $g(t) = t - 1 - \ln t$ for every $t > 0$. Using the equalities

$$\forall n \geq 1, \sum_{i=1}^{p} a_i R_i(X_{2n}) = \sum_{j=1}^{q} b_j C_j(X_{2n-1}) = 1,$$

we derive the convergence of the series

$$\sum_i a_i g(R_i(X_0)) + \sum_j b_j g(C_j(X_1)) + \sum_i a_i g(R_i(X_2)) + \sum_j b_j g(C_j(X_3)) + \cdots .$$

But g is null at 1, positive everywhere else, and tends to infinity at 0+ and at $+\infty$. By positivity of the a_i and b_j, we get the convergence of all series

$$\sum_{n\geq 0} g(R_i(X_{2n})) \text{ and } \sum_{n\geq 0} g(C_j(X_{2n+1}))$$

and therefore the convergence of all sequences $(R_i(X_{2n}))_{n\geq 0}$ and to $(C_j(X_{2n+1}))_{n\geq 0}$ towards 1. Hence, the series $\sum_n (R_i(X_{2n}) - 1)^2$ and $\sum_{n\geq 0}(C_j(X_{2n+1}) - 1)^2$ converge since $g(t) \sim (t-1)^2/2$ as $t \to 1$.

We now use a quantity introduced by Bregman [3] and called L_1-error by Pukelsheim [12]. For each $X \in \Gamma_1$, set

$$\begin{aligned} e(X) &= \sum_{i=1}^{p} \Big| \sum_{j=1}^{q} X(i,j) - a_i \Big| + \sum_{j=1}^{q} \Big| \sum_{i=1}^{p} X(i,j) - b_j \Big| \\ &= \sum_{i=1}^{p} a_i |R_i(X) - 1| + \sum_{j=1}^{q} b_j |C_j(X) - 1|. \end{aligned}$$

The convexity of the square function yields

$$\frac{e(X)^2}{2} \leq \sum_{i=1}^{p} a_i (R_i(X) - 1)^2 + \sum_{j=1}^{q} b_j (C_j(X) - 1)^2.$$

Thus the series $\sum_n e(X_n)^2$ converges. But the sequence $(e(X_n))_{n\geq 1}$ is non-increasing (the proof of this fact is recalled below). Therefore, for every $n \geq 1$,

$$0 \leq \frac{n}{2} e(X_n)^2 \leq \sum_{n/2\leq k\leq n} e(X_k)^2.$$

Convergences $ne(X_n)^2 \to 0$, $\sqrt{n}(R_i(X_n) - 1) \to 0$ and $\sqrt{n}(C_j(X_n) - 1) \to 0$ follow.

To check the monotonicity of $(e(X_n))_{n\geq 1}$, note that $T_R(X) \in \Gamma_R$ for every $X \in \Gamma_C$, so

$$\begin{aligned} e(T_R(X)) &= \sum_{j=1}^{q} \Big| \sum_{i=1}^{p} X(i,j)\, R_i(X)^{-1} - b_j \Big| \\ &= \sum_{j=1}^{q} \Big| \sum_{i=1}^{p} X(i,j)\, (R_i(X)^{-1} - 1) \Big| \\ &\leq \sum_{i=1}^{p} \sum_{j=1}^{q} X(i,j)\, |R_i(X)^{-1} - 1| \\ &= \sum_{i=1}^{p} a_i R_i(X)\, |R_i(X)^{-1} - 1| \\ &= e(X). \end{aligned}$$

In the same way, $e(T_C(Y)) \leq e(Y)$ for every $Y \in \Gamma_R$.

3.7.2 Convergence and Limit of (X_n)

Since $\Gamma(X_0)$ is not empty, we can fix a matrix $S_0 \in \Gamma(X_0)$ whose support is maximum, like in Theorem 3.1, critical case, item (a).

Let L be a limit point of the sequence $(X_n)_{n\geq 0}$, so L is the limit of some subsequence $(X_{\varphi(n)})_{n\geq 0}$. As noted in the introduction, $\mathrm{Supp}(L) \subset \mathrm{Supp}(X_0)$. But for every $i \in [\![1,p]\!]$ and $j \in [\![1,q]\!]$, $R_i(L) = \lim R_i(X_{\varphi(n)}) = 1$ and $C_j(L) = \lim C_j(X_{\varphi(n)}) = 1$. Hence $L \in \Gamma(X_0)$. Corollary 3.5 yields the inclusion $\mathrm{Supp}(S_0) \subset \mathrm{Supp}(L)$ hence for every $S \in \Gamma(X_0)$, $\mathrm{Supp}(S) \subset \mathrm{Supp}(S_0) \subset \mathrm{Supp}(L) \subset \mathrm{Supp}(X_0)$, so the quantities $F_S(X_0)$ and $F_S(L)$ are positive.

By Lemma 3.9, the ratios $F_S(X_{\varphi(n)})/F_S(X_0)$ do not depend on $S \in \Gamma(X_0)$, so by continuity of F_S, the ratio $F_S(L)/F_S(X_0)$ does not depend on $S \in \Gamma(X_0)$. But by Lemma 3.8,

$$\ln \frac{F_S(L)}{F_S(X_0)} = \ln \frac{F_S(S)}{F_S(X_0)} - \ln \frac{F_S(S)}{F_S(L)} = D(S||X_0) - D(S||L),$$

and $D(S||L) \geq 0$ with equality if and only if $S = L$. Therefore, L is the only element achieving the greatest lower bound of $D(S||X_0)$ over all $S \in \Gamma(X_0)$.

$$L = \arg\min_{S\in\Gamma(X_0)} D(S||X_0).$$

We have proved the unicity of the limit point of the sequence $(X_n)_{n\geq 0}$. By compactness of $\Gamma(X_0)$, the convergence follows.

Remark 3.1 Actually, one has $\mathrm{Supp}(S_0) = \mathrm{Supp}(L)$. Indeed, Theorem 3.3 shows that for every $(i, j) \in \mathrm{Supp}(X_0) \setminus \mathrm{Supp}(S_0)$, $X_n(i, j) \to 0$ as $n \to +\infty$. This fact could be retrieved by using the same arguments as in the proof of Theorem 3.1 to show that on the set $\Gamma(X_0)$, the linear form $X \mapsto X(i_0, j_0)$ coincides with some linear combination of the affine forms $X \mapsto R_i(X) - 1$, $i \in [\![1, p]\!]$, and $X \mapsto C_j(X) - 1$, $j \in [\![1, q]\!]$.

3.8 Proof of Theorems 3.4–3.6

We recall that neither proof below uses the assumption that $\Gamma(X_0)$ is empty.

3.8.1 Proof of Theorem 3.4

3.8.1.1 Convergence of the Sequences $(R(X_{2n}))$ and $(C(X_{2n+1}))$

By Lemma 3.7 and Corollary 3.4, we have for every $n \geq 1$, $R(X_{2n+2}) = P(X_{2n})R(X_{2n})$, where $P(X_{2n})$ is a stochastic matrix such that for every i and k in $[\![1, p]\!]$,

$$P(X_{2n})(i,i) \geq \frac{\underline{a}}{\overline{b}\,\overline{C}(X_{2n+1})q} \geq \frac{\underline{a}\,\underline{R}(X_2)}{\overline{b}q}$$

and

$$P(X_{2n})(k,i) \leq (\overline{a}/\underline{a})P(X_{2n})(i,k).$$

The sequence $(P(X_{2n}))_{n\geq 0}$ satisfies the assumption of Corollary 3.2 and Theorem 3.7, so any one of these two results ensures the convergence of the sequence $(R(X_{2n}))_{n\geq 0}$. By Corollary 3.4, the entries of these vectors stay in the interval $[\underline{R}(X_2), \overline{R}(X_2)]$, so the limit of each entry is positive. The same arguments show that the sequence $(C(X_{2n+1}))_{n\geq 0}$ also converges to some vector with positive entries.

3.8.1.2 Relations Between the Components of the Limits, and Block Structure

Call $\lambda_1 < \ldots < \lambda_r$ the different values of the limits of the sequences $(R_i(X_{2n}))_{n\geq 0}$, and $\mu_1 > \ldots > \mu_s$ the different values of the limits of the sequences

$(C_j(X_{2n+1}))_{n\geq 0}$. The values of these limits will be precised later. Consider the sets

$$I_k = \{i \in [\![1, p]\!] : \lim R_i(X_{2n}) = \lambda_k\} \text{ for } k \in [\![1, r]\!],$$

$$J_l = \{j \in [\![1, q]\!] : \lim C_j(X_{2n+1}) = \mu_l\} \text{ for } l \in [\![1, s]\!].$$

When $(i, j) \in I_k \times J_l$, the sequence $(R_i(X_{2n})C_j(X_{2n+1}))_{n\geq 0}$ converges to $\lambda_k \mu_l$. If $\lambda_k \mu_l > 1$, this entails the convergence to 0 of the sequence $(X_n(i, j))_{n\geq 0}$ with a geometric rate; and if $\lambda_k \mu_l < 1$, this entails the nullity of all $X_n(i, j)$ (otherwise the sequence $(X_n(i, j))_{n\geq 0}$ would go to infinity). But for all $n \geq 1$, $R_i(X_{2n}) = 1$ and $C_j(X_{2n+1}) = 1$, so at least one entry of the matrices X_n on each line or column does not converge to 0. This forces the equalities $s = r$ and $\mu_k = \lambda_k^{-1}$ for every $k \in [\![1, r]\!]$.

3.8.1.3 Convergence of the Sequences (X_{2n}) and (X_{2n+1})

Let L be any limit point of the sequence $(X_{2n})_{n\geq 0}$, so L is the limit of some subsequence $(X_{2\varphi(n)})_{n\geq 0}$. By definition of a', $R_i(L) = \lim R_i(X_{2\varphi(n)}) = a'_i/a_i$ for every $i \in [\![1, p]\!]$. Moreover, $\text{Supp}(L) \subset \text{Supp}(X_0)$, so L belongs to $\Gamma(a', b, X_0)$.

Like in Sect. 3.7.2, we check that the quantity

$$D(S||X_0) - D(S||L) = \ln(F_S(L)/F_S(X_0))$$

does not depend on $S \in \Gamma(a', b, X_0)$, so L is the unique matrix achieving the greatest lower bound of $D(S||X_0)$ over all $S \in \Gamma(a', b, X_0)$. The convergence of $(X_{2n})_{n\geq 0}$ follows by compactness of $\Gamma(a', b, X_0)$.

By Lemma 3.9, the ratios $F_S(X_{2\varphi(n)})/F_S(X_0)$ do not depend on $S \in \Gamma(X_0)$, so by continuity of F_S, the ratio $F_S(L)/F_S(X_0)$ The same arguments show that the sequence $(X_{2n+1})_{n\geq 0}$ converges to the unique matrix achieving the greatest lower bound of $D(S||X_0)$ over all $S \in \Gamma(a, b', X_0)$.

3.8.1.4 Formula for λ_k

We know that the sequence $(X_n(i, j))_{n\geq 0}$ converges to 0 whenever $i \in I_k$ and $j \in J_l$ with $k \neq l$. Thus the support of $L = \lim X_{2n}$ is contained in $I_1 \times J_1 \cup \ldots \cup I_r \times J_r$. But L belongs to $\Gamma(a', b, X_0)$, so for every $k \in [\![1, r]\!]$

$$\lambda_k a(I_k) = \sum_{i \in I_k} a'_i = \sum_{(i,j) \in I_k \times J_k} L(i, j) = \sum_{j \in J_k} b_j = b(J_k).$$

3.8.1.5 Properties of Matrices in $\Gamma(a', b, X_0)$ and $\Gamma(a', b, X_0)$

Let $S \in \Gamma(a, b', X_0)$.

Let $k \in [\![1, r-1]\!]$, $A_k = I_1 \cup \cdots \cup I_k$ and $B_k = J_{k+1} \cup \cdots \cup I_r$. We already know that X_0 is null on $A_k \times B_k$, so S is also null on this set. Moreover, for every $l \in [\![1, r]\!]$,

$$a(I_l) = \lambda_l^{-1} b(J_l) = \sum_{j \in J_l} \lambda_l^{-1} b_j = \sum_{j \in J_l} b'_j = b'(J_l).$$

Summation over all $l \in [\![1, k]\!]$ yields $a(A_k) = b'(B_k^c)$. Hence by Theorem 3.1 (critical case), S is null on the set $A_k^c \times B_k^c = (I_{k+1} \cup \cdots \cup I_r) \times (J_1 \cup \cdots \cup J_k)$.

This shows that the support of S is included in $I_1 \times J_1 \cup \cdots \cup I_r \times J_r$. This block structure and the equalities $a'_i/a_i = b_j/b'_j = \lambda_k$ whenever $(i, j) \in I_k \times J_k$ yield the equality $D_1 S = S D_2$. This matrix has the same support as S. Moreover, its i-th row is a'_i/a_i times the i-th row of S, so its i-th row-sum is $a'_i/a_i \times a_i = a'_i$; in the same way its j-th column is b_j/b'_j times the j-th column of S, so its j-th column sum is $b_j/b'_j \times b'_j = b_j$. As symmetric conclusions hold for every matrix in $\Gamma(a', b, X_0)$, the proof is complete.

3.8.2 *Proof of Theorem 3.5*

Fix $S \in \Gamma(a', b, X_0)$ (we know by Theorem 3.4 that this set is not empty). Let $k \in [\![1, r]\!]$, $P = [\![1, p]\!] \setminus (I_1 \cup \ldots \cup I_{k-1})$, $Q = [\![1, q]\!] \setminus (J_1 \cup \ldots \cup J_{k-1})$.

If $k = r$, then $P = I_r$ and $Q = J_r$. As $a'_i = a_i b(J_r)/a(I_r)$ for every $i \in I_r$, we have $a'(I_r) = b(J_r)$ and $a'_i/a'(I_r) = a_i b(J_r)/a(I_r)$ for every $i \in I_r$. Therefore, the matrix $(S(i, j)/a'(I_r))$ is a solution of the restricted problem associated to the marginals $a(\cdot|P) = (a_i/a(I_r))_{i \in P}$, $b(\cdot|Q) = (b_j/b(Q))_{j \in Q}$ and to the initial condition $(X_0(i, j))_{(i,j) \in P \times Q}$.

If $k < r$, then $P = I_k \cup \ldots \cup I_r$ and $Q = J_k \cup \ldots \cup J_r$. Let $A_k = I_k$ and $B_k = J_{k+1} \cup \ldots \cup J_r$. By Theorem 3.4, the matrix X_0 is null on product $A_k \times B_k$. Moreover, the inequalities $\lambda_1 < \ldots < \lambda_r$ and $a(I_l) > 0$ for every $l \in [\![1, r]\!]$ yield

$$b(Q) = \sum_{l=k}^{r} b(J_l) = \sum_{l=k}^{r} \lambda_l a(I_l) > \lambda_k \sum_{l=k}^{r} a(I_l) = \lambda_k a(P),$$

so

$$\frac{a(A_k|P)}{b(Q \setminus B_k|Q)} = \frac{a(I_k)/a(P)}{b(J_k)/b(Q)} = \lambda_k^{-1} \frac{b(Q)}{a(P)} > 1.$$

Hence $A_k \times B_k$ is a cause of incompatibility of the restricted problem associated to the marginals $a(\cdot|P) = (a_i/a(P))_{i \in P}$, $b(\cdot|Q) = (b_j/b(Q))_{j \in Q}$ and to the initial condition $(X_0(i,j))_{(i,j)\in P \times Q}$.

Now, assume that X_0 is null on some subset $A \times B$ of $P \times Q$. Then S is also null on $A \times B$, so for every $l \in [\![k, r]\!]$,

$$\lambda_k a(A \cap I_l) \leq \lambda_l a(A \cap I_l) = a'(A \cap I_l) = S((A \cap I_l) \times ((Q \setminus B) \cap J_l)) \leq b((Q \setminus B) \cap J_l).$$

Summing this inequalities over all $l \in [\![k, r]\!]$ yields $\lambda_k a(A) \leq b(Q \setminus B)$, so

$$\frac{a(A)}{b(Q \setminus B)} \leq \lambda_k^{-1} = \frac{a(I_k)}{b(I_k)} = \frac{a(A_k)}{b(Q \setminus B_k)}.$$

Moreover, if equality holds in the last inequality, then for every $l \in [\![k, r]\!]$,

$$\lambda_k a(A \cap I_l) = \lambda_l a(A \cap I_l) = b((Q \setminus B) \cap J_l).$$

This yields $A \cap I_l = (Q \setminus B) \cap J_l = \emptyset$ for every $l \in [\![k+1, r]\!]$, thus $A \subset I_k = A_k$ and $Q \setminus B \subset I_k$, namely $B \supset B_k$. The proof is complete.

3.8.3 Proof of Theorem 3.6

The proof relies the next two lemmas, from Pretzel, relying on notion of diagonal equivalence. We provide proofs to keep the paper self-contained. The first one differs from Pretzel's original proof. Recall that two matrices X and Y in $\mathscr{M} p, q(\mathbf{R}_+)$ are said to be diagonally equivalent if there exists $D' \in \Delta_p$ and $D'' \in \Delta_q$ such that $Y = D'XD''$. In particular, X and Y must have the same support to be diagonally equivalent.

Lemma 3.12 (Property 1 of [11]) *Let X and Y be in $\mathscr{M} p, q(\boldsymbol{R}_+)$. If X and Y are diagonally equivalent and have the same marginals then $X = Y$.*

Proof By assumption, $Y = D'XD''$ for some $D' \in \Delta_p$ and $D'' \in \Delta_q$. Call $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_q$ the diagonal entries of D' and D''. For every $(i, j) \in [\![1, p]\!] \times [\![1, q]\!]$, $Y(i, j) = \alpha_i \beta_j X(i, j)$, so

$$D(Y||X) = \sum_{i,j} Y(i,j)(\ln \alpha_i + \ln \beta_j) = \sum_i Y(i,+) \ln \alpha_i + \sum_j Y(+,j) \ln \beta_j.$$

In the same way,

$$D(X||Y) = \sum_{i,j} X(i,j)(\ln(\alpha_i^{-1}) + \ln(\beta_j^{-1})) = -\sum_i X(i,+) \ln \alpha_i - \sum_j X(+,j) \ln \beta_j.$$

Since X and Y have the same marginals, the non-negative quantities $D(Y|X)$ and $D(X|Y)$ are opposite, so they are null. Hence $X = Y$.

Lemma 3.13 (Lemma 2 of [11]) *Let X and Y be in $\mathscr{M}p, q(\boldsymbol{R}_+)$. If Y has the same support as X and is the limit of a some sequence $(Y_n)_{n\geq 0}$ of matrices which are diagonally equivalent to X, then X and Y are diagonally equivalent.*

Proof For each $n \geq 0$, one can find some positive real numbers $\alpha_n(1), \dots, \alpha_n(p)$ and $\beta_n(1), \dots, \beta_n(q)$ such that $Y_n(i,j) = \alpha_n(i)\beta_n(j)X(i,j)$ for every $(i,j) \in [\![1,p]\!] \times [\![1,q]\!]$. By assumption, the sequence $(\alpha_n(i)\beta_n(j))_{n\geq 0}$ converges to a positive number whenever $(i,j) \in \mathrm{Supp}(X)$.

In a way similar to the beginning of the proof of Theorem 3.4, we define a non-oriented graph G on $[\![1,p]\!]$ as follows: (i,i') is an edge if and only if there exists some $j \in [\![1,q]\!]$ such that $X(i,j)X(i',j) > 0$. Then the sequence $(\alpha_n(i)/\alpha_n(i'))_{n\geq 0}$ converges whenever i and i' belong to a same connected component of G.

Call $I_1, \dots, I_r$ the connected components of the graph G. For each $k \in [\![1,r]\!]$ choose $i_k \in I_k$ and set $J_k = \{j \in [\![1,q]\!] : \exists i \in I_k : X(i,j) > 0\}$. Then the sets $J_1, \dots, J_k$ form a partition of $[\![1,q]\!]$ and the support of X is contained in $I_1 \times J_1 \cup \dots \cup I_r \times J_r$.

For every $n \geq 0$, set $\alpha'_n(i) = \alpha_n(i)/\alpha_n(i_k)$ whenever $i \in I_k$ and $\beta'_n(j) = \beta_n(j)\alpha_n(i_k)$ whenever $j \in J_k$. Then $Y_n(i,j) = \alpha'_n(i)\beta'_n(j)X(i,j)$ for every $(i,j) \in [\![1,p]\!] \times \in [\![1,q]\!]$. Since all sequences $(\alpha'_n(i))_{n\geq 0}$ and $(\beta'_n(j))_{n\geq 0}$ converge to a positive limit, we deduce that X and Y are diagonally equivalent.

We now prove Theorem 3.6.

Set $L_{\mathrm{even}} = \lim_n X_{2n}$ and $L_{\mathrm{odd}} = \lim_n X_{2n+1}$. Letting n go to infinity in the equality $X_{2n+1} = T_R(X_{2n})$ yields $L_{\mathrm{odd}} = D_1^{-1}L_{\mathrm{even}}(i,j)$, where $D_1 = \mathrm{Diag}(a'_1/a_1, \dots, a'_p/a_p)$. We deduce that L_{even} and L_{odd} have the same support Σ.

By Theorem 3.4, L_{even} is the only matrix achieving the minimum of $D(Y||X_0)$ over all $Y \in \Gamma(a', b, X_0)$. Thus, by Theorem 3.3, L_{even} is also the limit of the sequence provided by the IPFP performed with the marginals a' and b and the initial matrix X_0, so Σ is the maximum of the supports of all matrices in $\Gamma(a', b, X_0)$.

By construction, for each $n \geq 0$, one can find $D'_n \in \Delta_p$ and $D''_n \in \Delta_q$ such that $X_{2n} = D'_n X_0 D''_n$. Since D'_n and D''_n are diagonal, $(D'_n X'_0 D''_n)(i,j) = X_{2n}(i,j)$ whenever $(i,j) \in \Sigma$ and $(D'_n X'_0 D''_n)(i,j) = 0$ whenever $(i,j) \in \Sigma^c$. Hence $\lim_n D'_n X'_0 D''_n = L_{\mathrm{even}}$ since $\Sigma = \mathrm{Supp}(L_{\mathrm{even}})$. But X'_0 and L_{even} have the same support, so they are diagonally equivalent by Lemma 3.13.

Call $(X'_n)_{n\geq 0}$ (respectively $(X''_n)_{n\geq 1}$) the sequence provided by the IPFP performed on the marginals a, b (respectively a', b) and the initial matrix X'_0. Equivalently, one could also start from $X'_0(+,+)^{-1}X'_0$ to have an initial matrix in Γ_1.

Since $L_{\text{even}} \in \Gamma(a', b)$ and $\text{Supp}(L_{\text{even}}) = \text{Supp}(X'_0)$, Theorem 3.2 applies, so the limit $L'' = \lim_n X''_n$ exists, L'' belongs to $\Gamma(a', b)$ and L'' is diagonally equivalent to X'_0 and therefore to L_{even}. By Lemma 3.12, we get $L'' = L_{\text{even}}$. Since all matrices X'_n and X''_n have the same support Σ, contained in $I_1 \times J_1 \cup \cdots \cup I_1 \times J_1$ by Theorem 3.4, a recursion shows that for every $n \geq 0$, $X'_{2n} = X''_{2n}$ and $X'_{2n+1} = D_1^{-1} X''_{2n+1}$. Hence $\lim_n X'_{2n} = L'' = L_{\text{even}}$.

A similar proof works for L_{odd} and the set $\Gamma(a, b', X_0)$.

Acknowledgements We thank A. Coquio, D. Piau, G. Geenens, F. Pukelsheim and the referee for their careful reading and their useful remarks.

References

1. E. Aas, Limit points of the iterative scaling procedure. Ann. Oper. Res. **215**(1), 15–23 (2014)
2. M. Bacharach, Estimating nonnegative matrices from marginal data. Int. Econ. Rev. **6**(3), 294–310 (1965)
3. L.M. Bregman, Proof of the convergence of Sheleikhovskii's method for a problem with transportation constraints. USSR Comput. Math. Math. Phys. **7**(1), 191–204 (1967)
4. J.B. Brown, P.J. Chase, A.O. Pittenger, Order independence and factor convergence in iterative scaling. Linear Algebra Appl. **190**, 1–39 (1993)
5. I. Csiszár, I-divergence geometry of probability distributions and minimization problems. Ann. Probab. **3**(1), 146–158 (1975)
6. S. Fienberg, An iterative procedure for estimation in contingency tables. Ann. Math. Stat. **41**(3), 907–917 (1970)
7. C. Gietl, F. Reffel, Accumulation points of the iterative scaling procedure. Metrika **73**(1), 783–798 (2013)
8. C.T. Ireland, S. Kullback, Contingency tables with given marginals. Biometrika **55**(1), 179–188 (1968)
9. J. Kruithof, Telefoonverkeers rekening (Calculation of telephone traffic). De Ingenieur 52, pp. E15–E25 (1937). Partial English translation: Krupp, R.S. http://wwwhome.ewi.utwente.nl/~ptdeboer/misc/kruithof-1937-translation.html
10. J. Lorenz, A stabilization theorem for dynamics of continuous opinions. Phys. A Stat. Mech. Appl. **355**(1), 217–223 (2005)
11. O. Pretzel, Convergence of the iterative scaling procedure for non-negative matrices. J. Lond. Math. Soc. **21**, 379–384 (1980)
12. F. Pukelsheim, Biproportional matrix scaling and the iterative proportional fitting procedure. Ann. Oper. Res. **215**, 269–283 (2014)
13. R. Sinkhorn, Diagonal equivalence to matrices with prescribed row and column sums. Am. Math. Mon. **74**(4), 402–405 (1967)
14. B. Touri, Product of random stochastic matrices and distributed averaging. Springer Thesis (2012)
15. B. Touri, A. Nedic, Alternative characterization of ergodicity for doubly stochastic chains, in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, FL (2011), pp. 5371–5376
16. B. Touri, A. Nedić, On ergodicity, infinite flow and consensus in random models. IEEE Trans. Autom. Control **56**(7), 1593–1605 (2011)
17. B. Touri, A. Nedić, On backward product of stochastic matrices. Automatica **48**(8), 1477–1488 (2012)
18. R. Webster, *Convexity* (Oxford University Press, Oxford, 1994)

Chapter 4
Limiting Eigenvectors of Outliers for Spiked Information-Plus-Noise Type Matrices

Mireille Capitaine

Abstract We consider an Information-Plus-Noise type matrix where the Information matrix is a spiked matrix. When some eigenvalues of the random matrix separate from the bulk, we study how the corresponding eigenvectors project onto those of the spikes. Note that, in an Appendix, we present alternative versions of the earlier results of Bai and Silverstein (Random Matrices Theory Appl 1(1):1150004, 44, 2012) ("noeigenvalue outside the support of the deterministic equivalent measure") and Capitaine (Indiana Univ Math J 63(6):1875–1910, 2014) ("exact separation phenomenon") where we remove some technical assumptions that were difficult to handle.

Keywords Random matrices · Spiked information-plus-noise type matrices · Eigenvalues · Eigenvectors · Outliers · Deterministic equivalent measure · Exact separation phenomenon

4.1 Introduction

In this paper, we consider the so-called Information-Plus-Noise type model

$$M_N = \Sigma_N \Sigma_N^* \quad \text{where} \quad \Sigma_N = \sigma \frac{X_N}{\sqrt{N}} + A_N,$$

defined as follows.

- $n = n(N), n \leq N, c_N = n/N \to_{N \to +\infty} c \in]0; 1]$.
- $\sigma \in]0; +\infty[$.
- $X_N = [X_{ij}]_{1 \leq i \leq n; 1 \leq j \leq N}$ where $\{X_{ij}, i \in \mathbb{N}, j \in \mathbb{N}\}$ is an infinite set of complex random variables such that $\{\Re(X_{ij}), \Im(X_{ij}), i \in \mathbb{N}, j \in \mathbb{N}\}$ are independent

M. Capitaine (✉)
CNRS, Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse Cedex, France
e-mail: mireille.capitaine@math.univ-toulouse.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_4

centered random variables with variance $1/2$ and satisfy

1. There exists $K > 0$ and a random variable Z with finite fourth moment for which there exists $x_0 > 0$ and an integer number $n_0 > 0$ such that, for any $x > x_0$ and any integer numbers $n_1, n_2 > n_0$, we have

$$\frac{1}{n_1 n_2} \sum_{i \le n_1, j \le n_2} P\left(|X_{ij}| > x\right) \le K P\left(|Z| > x\right). \tag{4.1}$$

2.

$$\sup_{(i,j)\in\mathbb{N}^2} \mathbb{E}(|X_{ij}|^3) < +\infty. \tag{4.2}$$

- Let ν be a compactly supported probability measure on $\mathbb{R}$ whose support has a finite number of connected components. Let $\Theta = \{\theta_1; \ldots; \theta_J\}$ where $\theta_1 > \ldots > \theta_J \ge 0$ are J fixed real numbers independent of N which are outside the support of ν. Let $k_1, \ldots, k_J$ be fixed integer numbers independent of N and $r = \sum_{j=1}^J k_j$. Let $\beta_j(N) \ge 0$, $r+1 \le j \le n$, be such that $\frac{1}{n}\sum_{j=r+1}^n \delta_{\beta_j(N)}$ weakly converges to ν and

$$\max_{r+1 \le j \le n} \text{dist}(\beta_j(N), \text{supp}(\nu)) \underset{N\to\infty}{\longrightarrow} 0 \tag{4.3}$$

where $\text{supp}(\nu)$ denotes the support of ν.
Let $\alpha_j(N)$, $j = 1, \ldots, J$, be real nonnegative numbers such that

$$\lim_{N\to+\infty} \alpha_j(N) = \theta_j.$$

Let A_N be a $n \times N$ deterministic matrix such that, for each $j = 1, \ldots, J$, $\alpha_j(N)$ is an eigenvalue of $A_N A_N^*$ with multiplicity k_j, and the other eigenvalues of $A_N A_N^*$ are the $\beta_j(N)$, $r+1 \le j \le n$. Note that the empirical spectral measure of $A_N A_N^*$ weakly converges to ν.

Remark 4.1 Note that assumption such as (4.1) appears in [14]. It obviously holds if the X_{ij}'s are identically distributed with finite fourth moment.

For any Hermitian $n \times n$ matrix Y, denote by $\text{spect}(Y)$ its spectrum, by

$$\lambda_1(Y) \ge \ldots \ge \lambda_n(Y)$$

the ordered eigenvalues of Y and by μ_Y the empirical spectral measure of Y:

$$\mu_Y := \frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i(Y)}.$$

For a probability measure τ on $\mathbb{R}$, denote by g_τ its Stieltjes transform defined for $z \in \mathbb{C} \setminus \mathbb{R}$ by

$$g_\tau(z) = \int_{\mathbb{R}} \frac{d\tau(x)}{z - x}.$$

When the X_{ij}'s are identically distributed, Dozier and Silverstein established in [15] that almost surely the empirical spectral measure μ_{M_N} of M_N converges weakly towards a nonrandom distribution $\mu_{\sigma,\nu,c}$ which is characterized in terms of its Stieltjes transform which satisfies the following equation: for any $z \in \mathbb{C}^+$,

$$g_{\mu_{\sigma,\nu,c}}(z) = \int \frac{1}{(1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))z - \frac{t}{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)} - \sigma^2(1-c)} d\nu(t). \tag{4.4}$$

This result of convergence was extended to independent but non identically distributed random variables by Xie in [30]. (Note that, in [21], the authors investigated the case where σ is replaced by a bounded sequence of real numbers.) In [11], the author carries on with the study of the support of the limiting spectral measure previously investigated in [16] and later in [25, 28] and obtains that there is a one-to-one relationship between the complement of the limiting support and some subset in the complement of the support of ν which is defined in (4.6) below.

Proposition 4.1 *Define differentiable functions $\omega_{\sigma,\nu,c}$ and $\Phi_{\sigma,\nu,c}$ on respectively $\mathbb{R} \setminus supp(\mu_{\sigma,\nu,c})$ and $\mathbb{R} \setminus supp(\nu)$ by setting*

$$\omega_{\sigma,\nu,c} : \begin{array}{l} \mathbb{R} \setminus supp(\mu_{\sigma,\nu,c}) \to \mathbb{R} \\ x \mapsto x(1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(x))^2 - \sigma^2(1-c)(1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(x)) \end{array} \tag{4.5}$$

and

$$\Phi_{\sigma,\nu,c} : \begin{array}{l} \mathbb{R} \setminus supp(\nu) \to \mathbb{R} \\ x \mapsto x(1 + c\sigma^2 g_\nu(x))^2 + \sigma^2(1-c)(1 + c\sigma^2 g_\nu(x)) \end{array}.$$

Set

$$\mathcal{E}_{\sigma,\nu,c} := \left\{ x \in \mathbb{R} \setminus supp(\nu), \Phi'_{\sigma,\nu,c}(x) > 0, g_\nu(x) > -\frac{1}{\sigma^2 c} \right\}. \tag{4.6}$$

$\omega_{\sigma,\nu,c}$ is an increasing analytic diffeomorphism with positive derivative from $\mathbb{R} \setminus supp(\mu_{\sigma,\nu,c})$ to $\mathcal{E}_{\sigma,\nu,c}$, with inverse $\Phi_{\sigma,\nu,c}$.

Moreover, extending previous results in [25] and [8] involving the Gaussian case and finite rank perturbations, [11] establishes a one-to-one correspondence between the θ_i's that belong to the set $\mathcal{E}_{\sigma,\nu,c}$ (counting multiplicity) and the outliers in the spectrum of M_N. More precisely, setting

$$\Theta_{\sigma,\nu,c} = \left\{\theta \in \Theta, \Phi'_{\sigma,\nu,c}(\theta) > 0, g_\nu(\theta) > -\frac{1}{\sigma^2 c}\right\}, \tag{4.7}$$

and

$$\mathscr{S} = \text{ supp } (\mu_{\sigma,\nu,c}) \cup \left\{\Phi_{\sigma,\nu,c}(\theta), \theta \in \Theta_{\sigma,\nu,c}\right\}, \tag{4.8}$$

we have the following results.

Theorem 4.1 ([11]) *For any $\epsilon > 0$,*

$$\mathbb{P}[\textit{for all large } N, \text{spect}(M_N) \subset \{x \in \mathbb{R}, \textit{dist}(x, \mathscr{S}) \leq \epsilon\}] = 1.$$

Theorem 4.2 ([11]) *Let θ_j be in $\Theta_{\sigma,\nu,c}$ and denote by $n_{j-1} + 1, \ldots, n_{j-1} + k_j$ the descending ranks of $\alpha_j(N)$ among the eigenvalues of $A_N A_N^*$. Then the k_j eigenvalues $(\lambda_{n_{j-1}+i}(M_N),\ 1 \leq i \leq k_j)$ converge almost surely outside the support of $\mu_{\sigma,\nu,c}$ towards $\rho_{\theta_j} := \Phi_{\sigma,\nu,c}(\theta_j)$. Moreover, these eigenvalues asymptotically separate from the rest of the spectrum since (with the conventions that $\lambda_0(M_N) = +\infty$ and $\lambda_{N+1}(M_N) = -\infty$) there exists $\delta_0 > 0$ such that almost surely for all large N,*

$$\lambda_{n_{j-1}}(M_N) > \rho_{\theta_j} + \delta_0 \textit{ and } \lambda_{n_{j-1}+k_j+1}(M_N) < \rho_{\theta_j} - \delta_0. \tag{4.9}$$

Remark 4.2 Note that Theorems 4.1 and 4.2 were established in [11] for A_N as (4.14) below and with $\mathscr{S} \cup \{0\}$ instead of $\mathscr{S}$ but they hold true as stated above and in the more general framework of this paper. Indeed, these extensions can be obtained sticking to the proof of the corresponding results in [11] but using the new versions of [3] and of the exact separation phenomenon of [11] which are presented in the Appendix 1 of the present paper.

The aim of this paper is to study how the eigenvectors corresponding to the outliers of M_N project onto those corresponding to the spikes θ_i's. Note that there are some pioneering results investigating the eigenvectors corresponding to the outliers of finite rank perturbations of classical random matricial models: [27] in the real Gaussian sample covariance matrix setting, and [7, 8] dealing with finite rank additive or multiplicative perturbations of unitarily invariant matrices. For a general perturbation, dealing with sample covariance matrices, Péché and Ledoit [23] introduced a tool to study the average behaviour of the eigenvectors but it seems that this did not allow them to focus on the eigenvectors associated with the eigenvalues that separate from the bulk. It turns out that further studies [6, 10] point out that the angle between the eigenvectors of the outliers of the deformed model and the eigenvectors associated to the corresponding original spikes is determined by Biane-Voiculescu's subordination function. For the model investigated in this paper, such a free interpretation holds but we choose not to develop this free probabilistic point of view in this paper and we refer the reader to the paper [13]. Here is the main result of the paper.

Theorem 4.3 *Let θ_j be in $\Theta_{\sigma,\nu,c}$ (defined in (4.7)) and denote by $n_{j-1}+1, \ldots, n_{j-1}+k_j$ the descending ranks of $\alpha_j(N)$ among the eigenvalues of $A_N A_N^*$. Let $\xi(j)$ be a normalized eigenvector of M_N relative to one of the eigenvalues $(\lambda_{n_{j-1}+q}(M_N), 1 \leq q \leq k_j)$. Denote by $\| \cdot \|_2$ the Euclidean norm on $\mathbb{C}^n$. Then, almost surely*

(i) $\lim_{N \to +\infty} \left\| P_{Ker\,(\alpha_j(N) I_N - A_N A_N^*)} \xi(j) \right\|_2^2 = \tau(\theta_j)$
where

$$\tau(\theta_j) = \frac{1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(\rho_{\theta_j})}{\omega'_{\sigma,\nu,c}(\rho_{\theta_j})} = \frac{\Phi'_{\sigma,\nu,c}(\theta_j)}{1 + \sigma^2 c g_\nu(\theta_j)} \tag{4.10}$$

(ii) for any θ_i in $\Theta_{\sigma,\nu,c} \setminus \{\theta_j\}$,

$$\lim_{N \to +\infty} \left\| P_{Ker\,(\alpha_i(N) I_N - A_N A_N^*)} \xi(j) \right\|_2 = 0.$$

The sketch of the proof of Theorem 4.3 follows the analysis of [10] as explained in Sect. 4.2. In Sect. 4.3, we prove a universal result allowing to reduce the study to estimating expectations of Gaussian resolvent entries carried on Sect. 4.4. In Sect. 4.5, we explain how to deduce Theorem 4.3 from the previous Sections. In an Appendix 1, we present alternative versions on the one hand of the result in [3] about the lack of eigenvalues outside the support of the deterministic equivalent measure, and, on the other hand, of the result in [11] about the exact separation phenomenon. These new versions deal with random variables whose imaginary and real parts are independent but remove the technical assumptions ((1.10) and "$b_1 > 0$" in Theorem 1.1 in [3] and "$\omega_{\sigma,\nu,c}(b) > 0$" in Theorem 1.2 in [11]). This allows us to claim that Theorem 4.2 holds in our context (see Remark 4.2). Finally, we present, in Appendix 2, some technical lemmas that are used throughout the paper.

4.2 Sketch of the Proof

Throughout the paper, for any $m \times p$ matrix B, $(m, p) \in \mathbb{N}^2$, we will denote by $\|B\|$ the largest singular value of B, and by $\|B\|_2 = \{Tr(BB^*)\}^{\frac{1}{2}}$ its Hilbert-Schmidt norm.

The proof of Theorem 4.3 follows the analysis in two steps of [10].

Step A First, we shall prove that, for any orthonormal system $(\xi_1, \cdots, \xi_{k_j})$ of eigenvectors associated to the k_j eigenvalues $\lambda_{n_{j-1}+q}(M_N)$, $1 \leq q \leq k_j$, the following convergence holds almost surely: $\forall l = 1, \dots, J$,

$$\sum_{p=1}^{k_j} \left\| P_{\ker(\alpha_l(N) I_N - A_N A_N^*)} \xi_p \right\|_2^2 \to_{N \to +\infty} \frac{k_j \delta_{jl} (1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(\rho_{\theta_j}))}{\omega'_{\sigma,\nu,c}(\rho_{\theta_j})}. \tag{4.11}$$

Note that for any smooth functions h and f on $\mathbb{R}$, if $v_1, \dots, v_n$ are eigenvectors associated to $\lambda_1(A_N A_N^*), \dots, \lambda_n(A_N A_N^*)$ and $w_1, \dots, w_n$ are eigenvectors associated to $\lambda_1(M_N), \dots, \lambda_n(M_N)$, one can easily check that

$$\mathrm{Tr}\left[h(M_N) f(A_N A_N^*)\right] = \sum_{m,p=1}^{n} h(\lambda_p(M_N)) f(\lambda_m(A_N A_N^*)) |\langle v_m, w_p \rangle|^2. \tag{4.12}$$

Thus, since $\alpha_l(N)$ on one hand and the k_j eigenvalues of M_N in $(\rho_{\theta_j} - \varepsilon, \rho_{\theta_j} + \varepsilon[)$ (for ϵ small enough) on the other hand, asymptotically separate from the rest of the spectrum of respectively $A_N A_N^*$ and M_N, a fit choice of h and f will allow the study of the restrictive sum $\sum_{p=1}^{k_j} \left\| P_{\ker(\alpha_l(N) I_N - A_N A_N^*)} \xi_p \right\|_2^2$. Therefore proving (4.11) is reduced to the study of the asymptotic behaviour of $\mathrm{Tr}\left[h(M_N) f(A_N A_N^*)\right]$ for some functions f and h respectively concentrated on a neighborhood of θ_l and ρ_{θ_j}.

Step B In the second, and final, step, we shall use a perturbation argument identical to the one used in [10] to reduce the problem to the case of a spike with multiplicity one, case that follows trivially from Step A.

Step B closely follows the lines of [10] whereas Step A requires substantial work. We first reduce the investigations to the mean Gaussian case by proving the following.

Proposition 4.2 *Let X_N as defined in Sect. 4.1. Let $\mathcal{G}_N = [\mathcal{G}_{ij}]_{1 \le i \le n, 1 \le j \le N}$ be a $n \times N$ random matrix with i.i.d. standard complex normal entries. Let h be a function in $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ with compact support, and Γ_N be a $n \times n$ Hermitian matrix such that*

$$\sup_{n,N} \|\Gamma_N\| < \infty \text{ and } \sup_{n,N} \mathrm{rank}(\Gamma_N) < \infty. \tag{4.13}$$

Then almost surely,

$$\mathrm{Tr}\left(h\left(\left(\sigma \frac{X_N}{\sqrt{N}} + A_N\right)\left(\sigma \frac{X_N}{\sqrt{N}} + A_N\right)^*\right) \Gamma_N\right)$$
$$- \mathbb{E}\left(\mathrm{Tr}\left[h\left(\left(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N\right)\left(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N\right)^*\right) \Gamma_N\right]\right) \to_{N \to +\infty} 0.$$

The asymptotic behaviour of $\mathbb{E}\Big(\mathrm{Tr}\Big[h\Big(\Big(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N\Big)\Big(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N\Big)^*\Big) f(A_N A_N^*)\Big]\Big)$ can be deduced, by using the bi-unitarily invariance of the distribution of $\mathcal{G}_N$, from the following Proposition 4.3 and Lemma 4.18.

Proposition 4.3 *Let* $\mathcal{G}_N = [\mathcal{G}_{ij}]_{1 \le i \le n, 1 \le j \le N}$ *be a* $n \times N$ *random matrix with i.i.d. complex standard normal entries. Assume that* A_N *is such that*

$$A_N = \begin{pmatrix} d_1(N) & & & (0) \\ & & (0) & \\ & \ddots & & (0) \\ (0) & & & \\ & & d_n(N) & (0) \end{pmatrix} \tag{4.14}$$

where $n = n(N)$, $n \le N$, $c_N = n/N \to_{N \to +\infty} c \in]0; 1]$, *for* $i = 1, \ldots, n$, $d_i(N) \in \mathbb{C}$, $\sup_N \max_{i=1,\ldots,n} |d_i(N)| < +\infty$ *and* $\frac{1}{n}\sum_{i=1}^{n} \delta_{|d_i(N)|^2}$ *weakly converges to a compactly supported probability measure* ν *on* $\mathbb{R}$ *when* N *goes to infinity. Define for all* $z \in \mathbb{C} \setminus \mathbb{R}$,

$$G_N^{\mathcal{G}}(z) = \left(zI - \left(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N \right) \left(\sigma \frac{\mathcal{G}_N}{\sqrt{N}} + A_N \right)^* \right)^{-1}.$$

Define for any $q = 1, \ldots, n$,

$$\gamma_q(N) = (A_N A_N^*)_{qq} = |d_q(N)|^2. \tag{4.15}$$

There is a polynomial P *with nonnegative coefficients, a sequence* $(u_N)_N$ *of nonnegative real numbers converging to zero when* N *goes to infinity and some nonnegative real number* l, *such that for any* (p, q) *in* $\{1, \ldots, n\}^2$, *for all* $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\mathbb{E}\left(\left(G_N^{\mathcal{G}}(z) \right)_{pq} \right) = \frac{1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)}{\omega_{\sigma,\nu,c}(z) - \gamma_q(N)} \delta_{pq} + \Delta_{p,q,N}(z), \tag{4.16}$$

with

$$\left| \Delta_{p,q,N}(z) \right| \le (1 + |z|)^l P(|\Im z|^{-1}) u_N.$$

4.3 Proof of Proposition 4.2

In the following, we will denote by $o_C(1)$ any deterministic sequence of positive real numbers depending on the parameter C and converging for each fixed C to zero when N goes to infinity. The aim of this section is to prove Proposition 4.2.
Define for any $C > 0$,

$$\begin{aligned} Y_{ij}^C &= \Re X_{ij} \mathbb{1}_{|\Re X_{ij}| \le C} - \mathbb{E}\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}| \le C} \right) \\ &\quad + \mathrm{i} \left\{ \Im X_{ij} \mathbb{1}_{|\Im X_{ij}| \le C} - \mathbb{E}\left(\Im X_{ij} \mathbb{1}_{|\Im X_{ij}| \le C} \right) \right\}. \end{aligned} \tag{4.17}$$

Set

$$\theta^* = \sup_{(i,j)\in\mathbb{N}^2} \mathbb{E}(|X_{ij}|^3) < +\infty.$$

We have

$$\begin{aligned}\mathbb{E}\left(|X_{ij} - Y_{ij}^C|^2\right) &= \mathbb{E}\left(|\Re X_{ij}|^2 \mathbb{1}_{|\Re X_{ij}|>C}\right) + \mathbb{E}\left(|\Im X_{ij}|^2 \mathbb{1}_{|\Im X_{ij}|>C}\right) \\ &\quad - \left\{\mathbb{E}\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}|>C}\right)\right\}^2 - \left\{\mathbb{E}\left(\Im X_{ij} \mathbb{1}_{|\Im X_{ij}|>C}\right)\right\}^2 \\ &\leq \frac{\mathbb{E}\left(|\Re X_{ij}|^3\right) + \mathbb{E}\left(|\Im X_{ij}|^3\right)}{C}\end{aligned}$$

so that

$$\sup_{i\geq 1, j\geq 1} \mathbb{E}\left(|X_{ij} - Y_{ij}^C|^2\right) \leq \frac{2\theta^*}{C}.$$

Note that

$$\begin{aligned}1 - 2\mathbb{E}\left(|\Re Y_{ij}^C|^2\right) &= 1 - 2\mathbb{E}\left\{\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}|\leq C} - \mathbb{E}\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}|\leq C}\right)\right)^2\right\} \\ &= 2\left[\frac{1}{2} - \mathbb{E}\left(|\Re X_{ij}|^2 \mathbb{1}_{|\Re X_{ij}|\leq C}\right)\right] + 2\left\{\mathbb{E}\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}|\leq C}\right)\right\}^2 \\ &= 2\mathbb{E}\left(|\Re X_{ij}|^2 \mathbb{1}_{|\Re X_{ij}|>C}\right) + 2\left\{\mathbb{E}\left(\Re X_{ij} \mathbb{1}_{|\Re X_{ij}|>C}\right)\right\}^2,\end{aligned}$$

so that

$$\sup_{i\geq 1, j\geq 1} |1 - 2\mathbb{E}\left(|\Re Y_{ij}^C|^2\right)| \leq \frac{4\theta^*}{C}.$$

Similarly

$$\sup_{i\geq 1, j\geq 1} |1 - 2\mathbb{E}\left(|\Im Y_{ij}^C|^2\right)| \leq \frac{4\theta^*}{C}.$$

Let us assume that $C > 8\theta^*$. Then, we have

$$\mathbb{E}\left(|\Re Y_{ij}^C|^2\right) > \frac{1}{4} \text{ and } \mathbb{E}\left(|\Im Y_{ij}^C|^2\right) > \frac{1}{4}.$$

Define for any $C > 8\theta^*$, $X^C = (X_{ij}^C)_{1\leq i\leq n;1\leq j\leq N}$, where for any $1 \leq i \leq n$, $1 \leq j \leq N$,

$$X_{ij}^C = \frac{\Re Y_{ij}^C}{\sqrt{2\mathbb{E}\left(|\Re Y_{ij}^C|^2\right)}} + \mathrm{i}\frac{\Im Y_{ij}^C}{\sqrt{2\mathbb{E}\left(|\Im Y_{ij}^C|^2\right)}}. \tag{4.18}$$

Let $\mathscr{G} = [\mathscr{G}_{ij}]_{1\leq i\leq n,1\leq j\leq N}$ be a $n\times N$ random matrix with i.i.d. standard complex normal entries, independent from X_N, and define for any $\alpha > 0$,

$$X^{\alpha,C} = \frac{X^C + \alpha\mathscr{G}}{\sqrt{1+\alpha^2}}.$$

Now, for any $n \times N$ matrix B, let us introduce the $(N+n) \times (N+n)$ matrix

$$\mathscr{M}_{N+n}(B) = \begin{pmatrix} 0_{n\times n} & B + A_N \\ B^* + A_N^* & 0_{N\times N} \end{pmatrix}.$$

Define for any $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\tilde{G}(z) = \left(zI_{N+n} - \mathscr{M}_{N+n}\left(\sigma\frac{X_N}{\sqrt{N}}\right)\right)^{-1},$$

and

$$\tilde{G}^{\alpha,C}(z) = \left(zI_{N+n} - \mathscr{M}_{N+n}\left(\sigma\frac{X^{\alpha,C}}{\sqrt{N}}\right)\right)^{-1}.$$

Denote by $\mathbb{U}(n+N)$ the set of unitary $(n+N)\times(n+N)$ matrices. We first establish the following approximation result.

Lemma 4.1 *There exist some positive deterministic functions u and v on $[0,+\infty[$ such that $\lim_{C\to+\infty} u(C) = 0$ and $\lim_{\alpha\to 0} v(\alpha) = 0$, and a polynomial P with nonnegative coefficients such that for any α and $C > 8\theta^*$, we have that*

- *almost surely, for all large N,*

$$\sup_{U\in\mathbb{U}(n+N)} \sup_{(i,j)\in\{1,\ldots,n+N\}^2} \sup_{z\in\mathbb{C}\setminus\mathbb{R}} |\Im z|^2 \left|(U^*\tilde{G}^{\alpha,C}(z)U)_{ij} - (U^*\tilde{G}(z)U)_{ij}\right|$$
$$\leq u(C) + v(\alpha), \tag{4.19}$$

- *for all large N,*

$$\sup_{U\in\mathbb{U}(n+N)}\sup_{(i,j)\in\{1,\ldots,n+N\}^2}\sup_{z\in\mathbb{C}\setminus\mathbb{R}}\frac{1}{P(|\Im z|^{-1})}$$
$$\times\left|\mathbb{E}\left((U^*\tilde{G}^{\alpha,C}(z)U)_{ij}-(U^*\tilde{G}(z)U)_{ij}\right)\right|$$
$$\leq u(C)+v(\alpha)+o_C(1). \tag{4.20}$$

Proof Note that

$$X_{ij}^C-Y_{ij}^C=\Re X_{ij}^C\left(1-\sqrt{2}\mathbb{E}\left(|\Re Y_{ij}^C|^2\right)^{1/2}\right)+\mathrm{i}\Im X_{ij}^C\left(1-\sqrt{2}\mathbb{E}\left(|\Im Y_{ij}^C|^2\right)^{1/2}\right)$$
$$=\Re X_{ij}^C\frac{1-2\mathbb{E}\left(|\Re Y_{ij}^C|^2\right)}{1+\sqrt{2}\mathbb{E}\left(|\Re Y_{ij}^C|^2\right)^{1/2}}+\mathrm{i}\Im X_{ij}^C\frac{1-2\mathbb{E}\left(|\Im Y_{ij}^C|^2\right)}{1+\sqrt{2}\mathbb{E}\left(|\Im Y_{ij}^C|^2\right)^{1/2}}.$$

Then,

$$\left\{\sup_{(i,j)\in\mathbb{N}^2}\mathbb{E}\left(|X_{ij}^C-Y_{ij}^C|^2\right)\right\}^{1/2}\leq\frac{4\theta^*}{C},\ \text{and}\ \sup_{(i,j)\in\mathbb{N}^2}\mathbb{E}\left(|X_{ij}^C-Y_{ij}^C|^3\right)<\infty.$$

It is straightforward to see, using Lemma 4.17, that for any unitary $(n+N)\times(n+N)$ matrix U,

$$\left|(U^*\tilde{G}^{\alpha,C}(z)U)_{ij}-(U^*\tilde{G}(z)U)_{ij}\right|$$
$$\leq\frac{\sigma}{|\Im z|^2}\left\|\frac{X_N-X^{\alpha,C}}{\sqrt{N}}\right\|$$
$$\leq\frac{\sigma}{|\Im z|^2}\left\{\left\|\frac{X_N-Y^C}{\sqrt{N}}\right\|+\left\|\frac{X^C-Y^C}{\sqrt{N}}\right\|\right.$$
$$\left.+\left(1-\frac{1}{\sqrt{1+\alpha^2}}\right)\left\|\frac{X^C}{\sqrt{N}}\right\|+\alpha\left\|\frac{\mathcal{G}}{\sqrt{N}}\right\|\right\}. \tag{4.21}$$

From Bai-Yin's theorem (Theorem 5.8 in [2]), we have

$$\left\|\frac{\mathcal{G}}{\sqrt{N}}\right\|=2+o(1).$$

Applying Remark 4.3 to the $(n+N)\times(n+N)$ matrix $\tilde{B}=\begin{pmatrix}0_{n\times n} & B\\ B^* & 0_{N\times N}\end{pmatrix}$ for $B\in\{X_N-Y^C, X^C-Y^C, X^C\}$ (see also Appendix B of [14]), we have that almost surely

$$\limsup_{N\to+\infty}\left\|\frac{X^C}{\sqrt{N}}\right\| \leq 2\sqrt{2},\ \limsup_{N\to+\infty}\left\|\frac{X^C - Y^C}{\sqrt{N}}\right\| \leq \frac{8\sqrt{2}\theta^*}{C},$$

and

$$\limsup_{N\to+\infty}\left\|\frac{X_N - Y^C}{\sqrt{N}}\right\| \leq 4\sqrt{\frac{\theta^*}{C}}.$$

Then, (4.19) readily follows.

Let us introduce

$$\Omega_{N,C} = \left\{\left\|\frac{\mathscr{G}}{\sqrt{N}}\right\| \leq 4, \left\|\frac{X^C}{\sqrt{N}}\right\| \leq 4, \left\|\frac{X_N - Y^C}{\sqrt{N}}\right\| \leq 8\sqrt{\frac{\theta^*}{C}}, \left\|\frac{X^C - Y^C}{\sqrt{N}}\right\| \leq \frac{16\theta^*}{C}\right\}.$$

Using (4.21), we have

$$\begin{aligned}
&\left|\mathbb{E}\left((U^*\tilde{G}^{\alpha,C}(z)U)_{ij} - (U^*\tilde{G}(z)U)_{ij}\right)\right| \\
&\quad\leq \frac{4\sigma}{|\Im z|^2}\left[2\sqrt{\frac{\theta^*}{C}} + \frac{4\theta^*}{C} + \alpha + \left(1 - \frac{1}{\sqrt{1+\alpha^2}}\right)\right] \\
&\qquad + \frac{2}{|\Im z|}\mathbb{P}(\Omega_{N,C}^c).
\end{aligned}$$

Thus (4.20) follows.

Now, Lemmas 4.18, 4.1 and 4.19 readily yields the following approximation lemma.

Lemma 4.2 *Let h be in $\mathscr{C}^\infty(\mathbb{R},\mathbb{R})$ with compact support and $\tilde{\Gamma}_N$ be a $(n+N)\times(n+N)$ Hermitian matrix such that*

$$\sup_{n,N}\|\tilde{\Gamma}_N\| < \infty \text{ and } \sup_{n,N}\operatorname{rank}(\tilde{\Gamma}_N) < \infty. \tag{4.22}$$

Then, there exist some deterministic functions on $[0,+\infty[$, u and v, such that $\lim_{C\to+\infty} u(C) = 0$ and $\lim_{\alpha\to 0} v(\alpha) = 0$, such that for all $C > 0$, $\alpha > 0$, we have almost surely for all large N,

$$\left|\operatorname{Tr}\left[h\left((\mathscr{M}_{N+n}\left(\frac{X^{\alpha,C}}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right] - \operatorname{Tr}\left[h\left((\mathscr{M}_{N+n}\left(\frac{X_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right| \leq a_{C,\alpha}^{(1)}, \tag{4.23}$$

and for all large N,

$$\left|\mathbb{E}\mathrm{Tr}\left[h\left((\mathscr{M}_{N+n}\left(\frac{X^{\alpha,C}}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]-\mathbb{E}\mathrm{Tr}\left[h\left((\mathscr{M}_{N+n}\left(\frac{X_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right|\leq a^{(2)}_{C,\alpha,N}, \tag{4.24}$$

where

$$a^{(1)}_{C,\alpha}=u(C)+v(\alpha),\ a^{(2)}_{C,\alpha,N}=u(C)+v(\alpha)+o_C(1).$$

Note that the distributions of the independent random variables $\Re(X^{\alpha,C}_{ij})$, $\Im(X^{\alpha,C}_{ij})$ are all a convolution of a centred Gaussian distribution with some variance v_α, with some law with bounded support in a ball of some radius $R_{C,\alpha}$; thus, according to Lemma 4.20, they satisfy a Poincaré inequality with some common constant $C_{PI}(C,\alpha)$ and therefore so does their product (see Appendix 2). An important consequence of the Poincaré inequality is the following concentration result.

Lemma 4.3 *Lemma 4.4.3 and Exercise 4.4.5 in [1] or Chapter 3 in [24]. There exists $K_1>0$ and $K_2>0$ such that for any probability measure $\mathbb{P}$ on $\mathbb{R}^{\mathbb{M}}$ which satisfies a Poincaré inequality with constant C_{PI}, and for any Lipschitz function F on $\mathbb{R}^M$ with Lipschitz constant $|F|_{Lip}$, we have*

$$\forall\epsilon>0,\ \mathbb{P}\left(|F-\mathbb{E}_{\mathbb{P}}(F)|>\epsilon\right)\leq K_1\exp\left(-\frac{\epsilon}{K_2\sqrt{C_{PI}}|F|_{Lip}}\right).$$

In order to apply Lemma 4.3, we need the following preliminary lemmas.

Lemma 4.4 (See Lemma 8.2 [10]) *Let f be a real $C_{\mathscr{L}}$-Lipschitz function on $\mathbb{R}$. Then its extension on the $N\times N$ Hermitian matrices is $C_{\mathscr{L}}$-Lipschitz with respect to the Hilbert-Schmidt norm.*

Lemma 4.5 *Let $\tilde{\Gamma}_N$ be a $(n+N)\times(n+N)$ matrix and h be a real Lipschitz function on $\mathbb{R}$. For any $n\times N$ matrix B,*

$$\left\{(\Re B(i,j),\ \Im B(i,j))_{1\leq i\leq n,1\leq j\leq N}\right\}\mapsto Tr\left[h\left((\mathscr{M}_{N+n}\left(B\right)\right)\tilde{\Gamma}_N\right]$$

is Lipschitz with constant bounded by $\sqrt{2}\left\|\tilde{\Gamma}_N\right\|_2\|h\|_{Lip}$.

Proof

$$\begin{aligned}&\left|\mathrm{Tr}\left[h(\mathscr{M}_{N+p}(B))\tilde{\Gamma}_N\right]-\mathrm{Tr}\left[h(\mathscr{M}_{N+p}(B'))\tilde{\Gamma}_N\right]\right|\\&\leq\left\|\tilde{\Gamma}_N\right\|_2\left\|h(\mathscr{M}_{N+p}(B))-h(\mathscr{M}_{N+p}(B'))\right\|_2\\&\leq\left\|\tilde{\Gamma}_N\right\|_2\|h\|_{Lip}\left\|\mathscr{M}_{N+p}(B)-\mathscr{M}_{N+p}(B')\right\|_2\end{aligned}\tag{4.25}$$

where we used Lemma 4.4 in the last line. Now,

$$\left\| \mathscr{M}_{N+p}(B) - \mathscr{M}_{N+p}(B^{'}) \right\|_2^2 = 2 \left\| B - B^{'} \right\|_2^2 . \tag{4.26}$$

Lemma 4.5 readily follows from (4.25) and (4.26).

Lemma 4.6 *Let $\tilde{\Gamma}_N$ be a $(n+N)\times(n+N)$ matrix such that $\sup_{N,n} \left\| \tilde{\Gamma}_N \right\|_2 \leq K$. Let h be a real Lipschitz function on $\mathbb{R}$. $F_N = \mathrm{Tr}\left[h\left(\mathscr{M}_{N+p}\left(\frac{X^{\alpha,C}}{\sqrt{N}} \right) \right) \tilde{\Gamma}_N \right]$ satisfies the following concentration inequality*

$$\forall \epsilon > 0,\ \mathbb{P}\left(|F_N - \mathbb{E}(F_N)| > \epsilon \right) \leq K_1 \exp\left(- \frac{\epsilon \sqrt{N}}{K_2(\alpha, C) K \|h\|_{Lip}} \right),$$

for some positive real numbers K_1 and $K_2(\alpha, C)$.

Proof Lemma 4.6 follows from Lemmas 4.5 and 4.3 and basic facts on Poincaré inequality recalled at the end of Appendix 2.

By Borel-Cantelli's Lemma, we readily deduce from the above Lemma the following

Lemma 4.7 *Let $\tilde{\Gamma}_N$ be a $(n+N)\times(n+N)$ matrix such that $\sup_{N,n} \left\| \tilde{\Gamma}_N \right\|_2 \leq K$. Let h be a real $\mathscr{C}^1$- function with compact support on $\mathbb{R}$.*

$$\mathrm{Tr}\left[h\left(\mathscr{M}_{N+p}\left(\sigma \frac{X^{\alpha,C}}{\sqrt{N}} \right) \right) \tilde{\Gamma}_N \right] - \mathbb{E}\left[\mathrm{Tr}\left[h\left(\mathscr{M}_{N+p}\left(\sigma \frac{X^{\alpha,C}}{\sqrt{N}} \right) \right) \tilde{\Gamma}_N \right] \right]$$
$$\xrightarrow{a.s}_{N\to+\infty} 0. \tag{4.27}$$

Now, we will establish a comparison result with the Gaussian case for the mean values by using the following lemma (which is an extension of Lemma 4.10 below to the non-Gaussian case) as initiated by Khorunzhy et al. [22] in Random Matrix Theory.

Lemma 4.8 *Let ξ be a real-valued random variable such that $\mathbb{E}(|\xi|^{p+2}) < \infty$. Let ϕ be a function from $\mathbb{R}$ to $\mathbb{C}$ such that the first $p+1$ derivatives are continuous and bounded. Then,*

$$\mathbb{E}(\xi\phi(\xi)) = \sum_{a=0}^{p} \frac{\kappa_{a+1}}{a!} \mathbb{E}(\phi^{(a)}(\xi)) + \epsilon, \tag{4.28}$$

where κ_a are the cumulants of ξ, $|\epsilon| \leq K \sup_t |\phi^{(p+1)}(t)| \mathbb{E}(|\xi|^{p+2})$, K only depends on p.

Lemma 4.9 *Let $\mathcal{G}_N = [\mathcal{G}_{ij}]_{1\le i\le n, 1\le j\le N}$ be a $n \times N$ random matrix with i.i.d. complex $N(0,1)$ Gaussian entries. Define*

$$\tilde{G}^{\mathcal{G}}(z) = \left(zI_{N+n} - \mathcal{M}_{N+n}\left(\sigma \frac{\mathcal{G}_N}{\sqrt{N}}\right)\right)^{-1}$$

for any $z \in \mathbb{C} \setminus \mathbb{R}$. There exists a polynomial P with nonnegative coefficients such that for all large N, for any $(i,j) \in \{1,\ldots,n+N\}^2$, for any $z \in \mathbb{C}\setminus\mathbb{R}$, for any unitary $(n+N)\times(n+N)$ matrix U,

$$\left|\mathbb{E}\left[(U^*\tilde{G}^{\mathcal{G}}(z)U)_{ij}\right] - \mathbb{E}\left[(U^*\tilde{G}(z)U)_{ij}\right]\right| \le \frac{1}{\sqrt{N}} P(|\Im z|^{-1}). \tag{4.29}$$

Moreover, for any $(N+n)\times(N+n)$ matrix $\tilde{\Gamma}_N$ such that

$$\sup_{n,N} \|\tilde{\Gamma}_N\| < \infty \text{ and } \sup_{n,N} \operatorname{rank}(\tilde{\Gamma}_N) < \infty, \tag{4.30}$$

and any function h in $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ with compact support, there exists some constant $K > 0$ such that, for any large N,

$$\left|\mathbb{E}\left[\operatorname{Tr}\left[h\left(\mathcal{M}_{N+n}\left(\sigma\frac{X_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right] - \mathbb{E}\left[\operatorname{Tr}\left[h\left(\mathcal{M}_{N+n}\left(\sigma\frac{\mathcal{G}_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right]\right| \le \frac{K}{\sqrt{N}}. \tag{4.31}$$

Proof We follow the approach of [26] chapters 18 and 19 consisting in introducing an interpolation matrix $X_N(\alpha) = \cos\alpha X_N + \sin\alpha \mathcal{G}_N$ for any α in $[0; \frac{\pi}{2}]$ and the corresponding resolvent matrix $\tilde{G}(\alpha, z) = \left(zI_{N+n} - \mathcal{M}_{N+n}\left(\sigma\frac{X_N(\alpha)}{\sqrt{N}}\right)\right)^{-1}$ for any $z \in \mathbb{C}\setminus\mathbb{R}$. We have, for any $(s,t) \in \{1,\ldots,n+N\}^2$,

$$\mathbb{E}\tilde{G}^{\mathcal{G}}_{st}(z) - \mathbb{E}\tilde{G}_{st}(z) = \int_0^{\frac{\pi}{2}} \mathbb{E}\left(\frac{\partial}{\partial\alpha}\tilde{G}_{st}(\alpha,z)\right) d\alpha$$

with

$$\begin{aligned}\frac{\partial}{\partial\alpha}\tilde{G}_{st}(\alpha,z) = \frac{\sigma}{2\sqrt{N}} \sum_{l=1}^{n}\sum_{k=n+1}^{n+N} &\left\{\left[\tilde{G}_{sl}(\alpha,z)\tilde{G}_{kt}(\alpha,z) + \tilde{G}_{sk}(\alpha,z)\tilde{G}_{lt}(\alpha,z)\right]\right.\\ &\times\left[-\sin\alpha\Re X_{l(k-n)} + \cos\alpha\Re\mathcal{G}_{l(k-n)}\right]\\ &+i\left[\tilde{G}_{sl}(\alpha,z)\tilde{G}_{kt}(\alpha,z) - \tilde{G}_{sk}(\alpha,z)\tilde{G}_{lt}(\alpha,z)\right]\\ &\left.\times\left[-\sin\alpha\Im X_{l(k-n)} + \cos\alpha\Im\mathcal{G}_{l(k-n)}\right]\right\}.\end{aligned}$$

Now, for any $l = 1, \ldots, n$ and $k = n+1, \ldots, n+N$, using Lemma 4.8 for $p = 1$ and for each random variable ξ in the set $\{\Re X_{l(k-n)}, \Re \mathcal{G}_{l(k-n)}, \Im X_{l(k-n)}, \Im \mathcal{G}_{l(k-n)}\}$, and for each ϕ in the set

$$\left\{(U^*\tilde{G}(\alpha, z))_{ip}(\tilde{G}(\alpha, z)U)_{qj};\, (p, q) = (l, k) \text{ or } (k, l), (i, j) \in \{1, \ldots, n+N\}^2\right\},$$

one can easily see that there exists some constant $K > 0$ such that

$$\left|\mathbb{E}(U^*\tilde{G}^{\mathcal{G}}(z)U)_{ij} - \mathbb{E}(U^*\tilde{G}(z)U)_{ij}\right| \leq \frac{K}{N^{3/2}} \sup_{Y \in \mathcal{H}_{n+N}(\mathbb{C})} \sup_{V \in \mathbb{U}(n+N)} S_V(Y)$$

where $\mathcal{H}_{n+N}(\mathbb{C})$ denotes the set of $(n+N) \times (n+N)$ Hermitian matrices and $S_V(Y)$ is a sum of a finite number independent of N and n of terms of the form

$$\sum_{l=1}^{n} \sum_{k=n+1}^{n+N} \left|\left(U^*R(Y)\right)_{ip_1} (R(Y))_{p_2p_3} (R(Y))_{p_4p_5} (R(Y)U)_{p_6j}\right| \tag{4.32}$$

with $R(Y) = (zI_{N+n} - Y)^{-1}$ and $\{p_1, \ldots, p_6\}$ contains exactly three k and three l. When $(p_1, p_6) = (k, l)$ or (l, k), then, using Lemma 4.17,

$$\begin{aligned}
&\sum_{l=1}^{n} \sum_{k=n+1}^{n+N} \left|\left(U^*R(Y)\right)_{ip_1} (R(Y))_{p_2p_3} (R(Y))_{p_4p_5} (R(Y)U)_{p_6j}\right| \\
&\leq \frac{1}{|\Im z|^2} \sum_{k,l=1}^{n+N} \left|\left(U^*R(Y)\right)_{il} (R(Y)U)_{kj}\right| \\
&\leq \frac{(N+n)}{|\Im z|^2} \left(\sum_{l=1}^{n+N} \left|\left(U^*R(Y)\right)_{il}\right|^2\right)^{1/2} \left(\sum_{k=1}^{n+N} \left|(R(Y)U)_{kj}\right|^2\right)^{1/2} \\
&= \frac{(N+n)}{|\Im z|^2} \left(\left(U^*R(Y)R(Y)^*U\right)_{ii}\right)^{1/2} \left(\left(U^*R(Y)^*R(Y)U\right)_{jj}\right)^{1/2} \\
&\leq \frac{(N+n)}{|\Im z|^4}
\end{aligned}$$

When $p_1 = p_6 = k$ or l, then, using Lemma 4.17,

$$\begin{aligned}
&\sum_{l=1}^{n} \sum_{k=n+1}^{n+N} \left|\left(U^*R(Y)\right)_{ip_1} (R(Y))_{p_2p_3} (R(Y))_{p_4p_5} (R(Y)U)_{p_6j}\right| \\
&\leq \frac{N+n}{|\Im z|^2} \sum_{l=1}^{n+N} \left|\left(U^*R(Y)\right)_{il} (R(Y)U)_{lj}\right|
\end{aligned}$$

$$\leq \frac{(N+n)}{|\Im z|^2}\left(\sum_{l=1}^{n+N}\left|\left(U^*R(Y)\right)_{il}\right|^2\right)^{1/2}\left(\sum_{l=1}^{n+N}\left|(R(Y)U)_{lj}\right|^2\right)^{1/2}$$

$$= \frac{(N+n)}{|\Im z|^2}\left(\left(U^*R(Y)R(Y)^*U\right)_{ii}\right)^{1/2}\left(\left(U^*R(Y)^*R(Y)U\right)_{jj}\right)^{1/2}$$

$$\leq \frac{(N+n)}{|\Im z|^4}$$

(4.29) readily follows.

Then by Lemma 4.19, there exists some constant $K > 0$ such that, for any N and n, for any $(i, j) \in \{1, \ldots, n+N\}^2$, any unitary $(n+N) \times (n+N)$ matrix U,

$$\limsup_{y \to 0^+}\left|\int\left[\mathbb{E}(U^*\tilde{G}(t+\mathrm{i}y)U)_{ij} - \mathbb{E}(U^*\tilde{G}^{\mathscr{G}}(t+\mathrm{i}y)U)_{ij}\right]h(t)dt\right| \leq \frac{K}{\sqrt{N}}. \tag{4.33}$$

Thus, using (4.97) and (4.30), we can deduce (4.31) from (4.33).

The above comparison lemmas allow us to establish the following convergence result.

Proposition 4.4 *Let h be a function in $\mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ with compact support and let $\tilde{\Gamma}_N$ be a $(n+N)\times(n+N)$ matrix such that $\sup_{n,N} \operatorname{rank}(\tilde{\Gamma}_N) < \infty$ and $\sup_{n,N} \|\tilde{\Gamma}_N\| < \infty$. Then we have that almost surely*

$$Tr\left[h\left((\mathscr{M}_{N+n}\left(\sigma\frac{X_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right] - \mathbb{E}\left[Tr\left[h\left((\mathscr{M}_{N+n}\left(\sigma\frac{\mathscr{G}_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right] \longrightarrow_{N\to+\infty} 0. \tag{4.34}$$

Proof Lemmas 4.2, 4.7 and 4.9 readily yield that there exist some positive deterministic functions u and v on $[0, +\infty[$ with $\lim_{C\to+\infty} u(C) = 0$ and $\lim_{\alpha\to 0} v(\alpha) = 0$, such that for any $C > 0$ and any $\alpha > 0$, almost surely

$$\limsup_{N\to+\infty}\left|Tr\left[h\left((\mathscr{M}_{N+n}\left(\sigma\frac{X_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right] - \mathbb{E}\left[Tr\left[h\left((\mathscr{M}_{N+n}\left(\sigma\frac{\mathscr{G}_N}{\sqrt{N}}\right)\right)\tilde{\Gamma}_N\right]\right]\right| \leq u(C) + v(\alpha).$$

The result follows by letting α go to zero and C go to infinity.

Now, note that, for any $N \times n$ matrix B, for any continuous real function h on $\mathbb{R}$, and any $n \times n$ Hermitian matrix Γ_N, we have

$$Tr\left(h\left((B+A_N)(B+A_N)^*\right)\Gamma_N\right) = Tr\left[\tilde{h}\left(\mathscr{M}_{N+n}(B)\right)\tilde{\Gamma}_N\right]$$

where $\tilde{h}(x) = h(x^2)$ and $\tilde{\Gamma}_N = \begin{pmatrix} \Gamma_N & (0) \\ (0) & (0) \end{pmatrix}$. Thus, Proposition 4.4 readily yields Proposition 4.2.

4.4 Proof of Proposition 4.3

The aim of this section is to prove Proposition 4.3 which deals with Gaussian random variables.Therefore we assume here that A_N is as (4.14) and set $\gamma_q(N) = (A_N A_N^*)_{qq}$. In this section, we let X stand for $\mathcal{G}_N$, A stands for A_N, G denotes the resolvent of $M_N = \Sigma\Sigma^*$ where $\Sigma = \sigma\frac{\mathcal{G}_N}{\sqrt{N}} + A_N$ and g_N denotes the mean of the Stieltjes transform of the spectral measure of M_N, that is

$$g_N(z) = \mathbb{E}\left(\frac{1}{n} Tr G(z)\right),\ z \in \mathbb{C} \setminus \mathbb{R}.$$

4.4.1 Matricial Master Equation

To obtain Eq. (4.35) below, we will use many ideas from [17]. The following Gaussian integration by part formula is the key tool in our approach.

Lemma 4.10 (Lemma 2.4.5 [1]) *Let ξ be a real centered Gaussian random variable with variance 1. Let Φ be a differentiable function with polynomial growth of Φ and Φ'. Then,*

$$\mathbb{E}\left(\xi\Phi(\xi)\right) = \mathbb{E}\left(\Phi'(\xi)\right).$$

Proposition 4.5 *Let z be in $\mathbb{C} \setminus \mathbb{R}$. We have for any (p, q) in $\{1, \ldots, n\}^2$,*

$$\mathbb{E}\left(G_{pq}(z)\right)\Bigg\{z(1-\sigma^2 c_N g_N(z)) - \frac{\gamma_q(N)}{1-\sigma^2 c_N g_N(z)} - \sigma^2(1-c_N) + \frac{\sigma^2}{N}\sum_{p=1}^{n}\nabla_{pp}(z)\Bigg\} = \delta_{pq} + \nabla_{pq}(z), \tag{4.35}$$

where

$$\nabla_{pq} = \frac{1}{1-\sigma^2 c_N g_N}\left\{\frac{\sigma^2}{N}\frac{\mathbb{E}\left(G_{pq}\right)}{1-\sigma^2 c_N g_N}\Delta_3 + \Delta_2(p,q) + \Delta_1(p,q).\right\}, \tag{4.36}$$

$$\Delta_1(p,q) = \sigma^2 \mathbb{E}\left\{\left[\frac{1}{N}TrG - \mathbb{E}\left(\frac{1}{N}TrG\right)\right](G\Sigma\Sigma^*)_{pq}\right\}, \tag{4.37}$$

$$\Delta_2(p,q) = \frac{\sigma^2}{N}\mathbb{E}\left\{Tr(GA\Sigma^*)\left[G_{pq} - \mathbb{E}\left(G_{pq}\right)\right]\right\}, \tag{4.38}$$

$$\Delta_3 = \sigma^2 \mathbb{E}\left\{\left[\frac{1}{N}TrG - \mathbb{E}\left(\frac{1}{N}TrG\right)\right]Tr(\Sigma^* GA)\right\}. \tag{4.39}$$

Proof Using Lemma 4.10 with $\xi = \Re X_{ij}$ or $\xi = \Im X_{ij}$ and $\Phi = G_{pi}\overline{\Sigma_{qj}}$, we obtain that for any j, q, p,

$$\mathbb{E}\left[\left(G\frac{\sigma X}{\sqrt{N}}\right)_{pj}\overline{\Sigma_{qj}}\right] = \sum_{i=1}^{n}\mathbb{E}\left[G_{pi}\frac{\sigma X_{ij}}{\sqrt{N}}\overline{\Sigma_{qj}}\right] \tag{4.40}$$

$$= \frac{\sigma^2}{N}\sum_{i=1}^{n}\mathbb{E}\left[(G\Sigma)_{pj}\, G_{ii}\overline{\Sigma_{qj}}\right] + \frac{\sigma^2}{N}\mathbb{E}(G_{pq}) \tag{4.41}$$

$$= \frac{\sigma^2}{N}\mathbb{E}\left[(TrG)\,(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right] + \frac{\sigma^2}{N}\mathbb{E}(G_{pq}). \tag{4.42}$$

On the other hand, we have

$$\mathbb{E}\left[(GA)_{pj}\,\overline{\Sigma_{qj}}\right] = \mathbb{E}\left[(GA)_{pj}\,\overline{A_{qj}}\right] + \sum_{i=1}^{n}\mathbb{E}\left[G_{pi}A_{ij}\frac{\sigma\overline{X_{qj}}}{\sqrt{N}}\right] \tag{4.43}$$

$$= \mathbb{E}\left[(GA)_{pj}\,\overline{A_{qj}}\right] + \frac{\sigma^2}{N}\mathbb{E}\left[G_{pq}\left(\Sigma^* GA\right)_{jj}\right] \tag{4.44}$$

where we applied Lemma 4.10 with $\xi = \Re X_{qj}$ or $\xi = \Im X_{qj}$ and $\Psi = G_{pi}A_{ij}$. Summing (4.42) and (4.44) yields

$$\mathbb{E}\left[(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right] = \frac{\sigma^2}{N}\mathbb{E}(G_{pq}) + \frac{\sigma^2}{N}\mathbb{E}\left[(TrG)\,(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right] \tag{4.45}$$

$$+\frac{\sigma^2}{N}\mathbb{E}\left[G_{pq}\left(\Sigma^* GA\right)_{jj}\right] + \mathbb{E}\left[(GA)_{pj}\,\overline{A_{qj}}\right]. \tag{4.46}$$

Define

$$\Delta_1(j) = \frac{\sigma^2}{N}\mathbb{E}\left[(TrG)\,(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right] - \frac{\sigma^2}{N}\mathbb{E}\left[TrG\right]\mathbb{E}\left[(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right].$$

From (4.46), we can deduce that

$$\mathbb{E}\left[(G\Sigma)_{pj}\,\overline{\Sigma_{qj}}\right] = \frac{1}{1-\sigma^2 c_N g_N}\left\{\frac{\sigma^2}{N}\mathbb{E}(G_{pq}) + \frac{\sigma^2}{N}\mathbb{E}\left[G_{pq}\left(\Sigma^* GA\right)_{jj}\right]\right.$$
$$\left.+\mathbb{E}\left[(GA)_{pj}\,\overline{A_{qj}}\right] + \Delta_1(j)\right\}.$$

Then, summing over j, we obtain that

$$\mathbb{E}\left[\left(G\Sigma\Sigma^*\right)_{pq}\right] = \frac{1}{1-\sigma^2 c_N g_N}\left\{\sigma^2\mathbb{E}(G_{pq}) + \frac{\sigma^2}{N}\mathbb{E}\left[G_{pq}Tr\left(\Sigma^* GA\right)\right]\right.$$
$$\left.+\mathbb{E}\left[\left(GAA^*\right)_{pq}\right] + \Delta_1(p,q)\right\}, \tag{4.47}$$

where $\Delta_1(p,q)$ is defined by (4.37). Applying Lemma 4.10 with $\xi = \Re X_{ij}$ or $\Im X_{ij}$ and $\Psi = (GA)_{ij}$, we obtain that

$$\mathbb{E}\left[Tr\left(\frac{\sigma X^*}{\sqrt{N}}GA\right)\right] = \frac{\sigma^2}{N}\mathbb{E}\left[TrG\;Tr\left(\Sigma^* GA\right)\right].$$

Thus,

$$\mathbb{E}\left[Tr\left(\Sigma^* GA\right)\right] = \mathbb{E}\left[Tr\left(A^* GA\right)\right] + \sigma^2 c_N g_N \mathbb{E}\left[Tr\left(\Sigma^* GA\right)\right] + \Delta_3,$$

where Δ_3 is defined by (4.39) and then

$$\mathbb{E}\left[Tr\left(\Sigma^* GA\right)\right] = \frac{1}{1-\sigma^2 c_N g_N}\left\{\mathbb{E}\left[Tr\left(GAA^*\right)\right] + \Delta_3\right\}. \tag{4.48}$$

(4.48) and (4.38) imply that

$$\frac{\sigma^2}{N}\mathbb{E}\left[G_{pq}Tr\left(\Sigma^* GA\right)\right] = \frac{\sigma^2}{N}\frac{\mathbb{E}(G_{pq})}{1-\sigma^2 c_N g_N}\left\{\mathbb{E}\left[Tr\left(GAA^*\right)\right] + \Delta_3\right\} + \Delta_2(p,q), \tag{4.49}$$

where $\Delta_2(p,q)$ is defined by (4.38). We can deduce from (4.47) and (4.49) that

$$\begin{aligned}
&\mathbb{E}\left[\left(G\Sigma\Sigma^*\right)_{pq}\right] \\
&= \frac{1}{1-\sigma^2 c_N g_N}\left\{\sigma^2\mathbb{E}(G_{pq}) + \mathbb{E}\left[\left(GAA^*\right)_{pq}\right]\right. \\
&\quad + \frac{\sigma^2}{N}\frac{\mathbb{E}\left[G_{pq}\right]}{1-\sigma^2 c_N g_N}\mathbb{E}\left[Tr\left(GAA^*\right)\right] \\
&\quad \left. + \frac{\sigma^2}{N}\frac{\mathbb{E}(G_{pq})}{1-\sigma^2 c_N g_N}\Delta_3 + \Delta_1(p,q) + \Delta_2(p,q)\right\}. \qquad (4.50)
\end{aligned}$$

Using the resolvent identity and (4.50), we obtain that

$$\begin{aligned}
z\mathbb{E}\left(G_{pq}\right) = \frac{1}{1-\sigma^2 c_N g_N}&\left\{\sigma^2\mathbb{E}(G_{pq}) + \mathbb{E}\left[\left(GAA^*\right)_{pq}\right]\right. \\
&\left. + \frac{\sigma^2}{N}\frac{\mathbb{E}\left[G_{pq}\right]}{1-\sigma^2 c_N g_N}\mathbb{E}\left[Tr\left(GAA^*\right)\right]\right\} + \delta_{pq} + \nabla_{pq} \qquad (4.51)
\end{aligned}$$

where ∇_{pq} is defined by (4.36). Taking $p = q$ in (4.51), summing over p and dividing by n, we obtain that

$$\begin{aligned}
zg_N &= \frac{\sigma^2 g_N}{1-\sigma^2 c_N g_N} + \frac{Tr\left[\mathbb{E}(G)AA^*\right]}{n(1-\sigma^2 c_N g_N)} \qquad (4.52) \\
&\quad + \frac{\sigma^2 g_N Tr\left[\mathbb{E}(G)AA^*\right]}{N(1-\sigma^2 c_N g_N)^2} + 1 + \frac{1}{n}\sum_{p=1}^{n}\nabla_{pp}. \qquad (4.53)
\end{aligned}$$

It readily follows that

$$\frac{Tr\left[\mathbb{E}(G)AA^*\right]}{n(1-\sigma^2 c_N g_N)}\left(\frac{\sigma^2 c_N g_N}{(1-\sigma^2 c_N g_N)} + 1\right) = \left(z - \frac{\sigma^2}{(1-\sigma^2 c_N g_N)}\right)g_N - 1 - \frac{1}{n}\sum_{p=1}^{n}\nabla_{pp}.$$

Therefore

$$\frac{Tr\left[\mathbb{E}(G)AA^*\right]}{n(1-\sigma^2 c_N g_N)} = zg_N(1-\sigma^2 c_N g_N) - (1-\sigma^2 c_N g_N) - \sigma^2 g_N - (1-\sigma^2 c_N g_N)\frac{1}{n}\sum_{p=1}^{n}\nabla_{pp}. \qquad (4.54)$$

(4.54) and (4.51) yield

$$\mathbb{E}(G_{pq}) \times \left\{ z(1-\sigma^2 c_N g_N) - \frac{\gamma_q}{1-\sigma^2 c_N g_N} - \sigma^2(1-c_N) + \frac{\sigma^2}{N}\sum_{p=1}^{n} \nabla_{pp} \right\}$$
$$= \delta_{pq} + \nabla_{pq}.$$

Proposition 4.5 follows.

4.4.2 Variance Estimates

In this section, when we state that some quantity $\Delta_N(z)$, $z \in \mathbb{C} \setminus \mathbb{R}$, is equal to $O(\frac{1}{N^p})$, this means precisely that there exist some polynomial P with nonnegative coefficients and some positive real number l which are all independent of N such that for any $z \in \mathbb{C} \setminus \mathbb{R}$,

$$|\Delta_N(z)| \leq \frac{(|z|+1)^l P(|\Im z|^{-1})}{N^p}.$$

We present now the different estimates on the variance. They rely on the following Gaussian Poincaré inequality (see Appendix 2). Let $Z_1, \ldots, Z_q$ be q real independent centered Gaussian variables with variance σ^2. For any $\mathscr{C}^1$ function $f : \mathbb{R}^q \to \mathbb{C}$ such that f and $\mathrm{grad} f$ are in $L^2(\mathscr{N}(0, \sigma^2 I_q))$, we have

$$\mathbf{V}\left\{f(Z_1, \ldots, Z_q)\right\} \leq \sigma^2 \mathbb{E}\left(\|(\mathrm{grad} f)(Z_1, \ldots, Z_q)\|_2^2 \right), \tag{4.55}$$

denoting for any random variable a by $\mathbf{V}(a)$ its variance $\mathbb{E}(|a - \mathbb{E}(a)|^2)$. Thus, $(Z_1, \ldots, Z_q)$ satisfies a Poincaré inequality with constant $C_{PI} = \sigma^2$.

The following preliminary result will be useful to these estimates.

Lemma 4.11 *There exists $K > 0$ such for all N,*

$$\mathbb{E}\left(\lambda_1\left(\frac{XX^*}{N}\right)\right) \leq K.$$

Proof According to Lemma 7.2 in [19], we have for any $t \in]0; N/2]$,

$$\mathbb{E}\left[\mathrm{Tr}\left(\exp t \frac{XX^*}{N}\right)\right] \leq n \exp\left((\sqrt{c_N}+1)^2 t + \frac{1}{N}(c_N+1)t^2 \right).$$

By the Chebychev's inequality, we have

$$\exp\left(t\mathbb{E}\left(\lambda_1\left(\frac{XX^*}{N}\right)\right)\right) \leq \mathbb{E}\left(\exp t\lambda_1\left(\frac{XX^*}{N}\right)\right)$$
$$\leq \mathbb{E}\left[\mathrm{Tr}\left(\exp t\frac{XX^*}{N}\right)\right]$$
$$\leq n\exp\left((\sqrt{c_N}+1)^2t+\frac{1}{N}(c_N+1)t^2\right).$$

It follows that

$$\mathbb{E}\left(\lambda_1\left(\frac{XX^*}{N}\right)\right) \leq \frac{1}{t}\log n+(\sqrt{c_N}+1)^2+\frac{1}{N}(c_N+1)t.$$

The result follows by optimizing in t.

Lemma 4.12 *There exists $C>0$ such that for all large N, for all $z\in\mathbb{C}\setminus\mathbb{R}$,*

$$\mathbb{E}\left(\left|\frac{1}{n}\mathrm{Tr}G-\mathbb{E}(\frac{1}{n}\mathrm{Tr}G)\right|^2\right) \leq \frac{C}{N^2|\Im z|^4}, \tag{4.56}$$

$$\forall (p,q)\in\{1,\ldots,n\}^2,\ \mathbb{E}\left(|G_{pq}-\mathbb{E}(G_{pq})|^2\right) \leq \frac{C}{N|\Im z|^4}, \tag{4.57}$$

$$\mathbb{E}\left(|\mathrm{Tr}\Sigma^*GA-\mathbb{E}(\mathrm{Tr}\Sigma^*GA)|^2\right) \leq \frac{C(1+|z|)^2}{|\Im z|^4}. \tag{4.58}$$

Proof Let us define $\Psi:\mathbb{R}^{2(n\times N)}\to M_{n\times N}(\mathbb{C})$ by

$$\Psi:\ \{x_{ij},y_{ij},i=1,\ldots,n,\ j=1,\ldots,N\}\to\sum_{i=1,\ldots,n}\ \sum_{j=1,\ldots,N}\left(x_{ij}+\mathrm{i}y_{ij}\right)e_{ij},$$

where e_{ij} stands for the $n\times N$ matrix such that for any (p,q) in $\{1,\ldots,n\}\times\{1,\ldots,N\}$, $(e_{ij})_{pq}=\delta_{ip}\delta_{jq}$. Let F be a smooth complex function on $M_{n\times N}(\mathbb{C})$ and define the complex function f on $\mathbb{R}^{2(n\times N)}$ by setting $f=F\circ\Psi$. Then,

$$\|\mathrm{grad} f(u)\|_2=\sup_{V\in M_{n\times N}(\mathbb{C}),TrVV^*=1}\left|\frac{d}{dt}F(\Psi(u)+tV)_{|t=0}\right|.$$

Now, $X=\Psi(\Re(X_{ij}),\Im(X_{ij}),1\leq i\leq n,1\leq j\leq N)$ where the distribution of the random variable $(\Re(X_{ij}),\Im(X_{ij}),1\leq i\leq n,1\leq j\leq N)$ is $\mathcal{N}(0,\frac{1}{2}I_{2nN})$.

Hence consider $F:\ H\to\frac{1}{n}\mathrm{Tr}\left(zI_n-\left(\sigma\frac{H}{\sqrt{N}}+A\right)\left(\sigma\frac{H}{\sqrt{N}}+A\right)^*\right)^{-1}$.

Let $V \in M_{n\times N}(\mathbb{C})$ such that $TrVV^* = 1$.

$$\frac{d}{dt}F(X+tV)|_{t=0}$$
$$= \frac{1}{n}\left\{\mathrm{Tr}\left(G\sigma\frac{V}{\sqrt{N}}\left(\sigma\frac{X}{\sqrt{N}}+A\right)^* G\right)+\mathrm{Tr}\left(G\left(\sigma\frac{X}{\sqrt{N}}+A\right)\sigma\frac{V^*}{\sqrt{N}}G\right)\right\}.$$

Moreover using Cauchy-Schwartz's inequality and Lemma 4.17, we have

$$\left|\frac{1}{n}\mathrm{Tr}\left(G\sigma\frac{V}{\sqrt{N}}\left(\sigma\frac{X}{\sqrt{N}}+A\right)^* G\right)\right|$$
$$\leq \frac{\sigma}{n}(TrVV^*)^{\frac{1}{2}}\left[\frac{1}{N}Tr(\left(\sigma\frac{X}{\sqrt{N}}+A\right)\left(\sigma\frac{X}{\sqrt{N}}+A\right)^* G^2(G^*)^2)\right]^{\frac{1}{2}}$$
$$\leq \frac{\sigma}{\sqrt{N}\sqrt{n}|\Im z|^2}\left[\lambda_1\left(\left(\sigma\frac{X}{\sqrt{N}}+A\right)\left(\sigma\frac{X}{\sqrt{N}}+A\right)^*\right)\right]^{\frac{1}{2}}.$$

We get obviously the same bound for $|\frac{1}{n}\mathrm{Tr}\left(G\left(\sigma\frac{X}{\sqrt{N}}+A\right)\sigma\frac{V^*}{\sqrt{N}}G\right)|$. Thus

$$\mathbb{E}\left(\|\mathrm{grad} f\left(\Re(X_{ij}),\Im(X_{ij}),1\leq i\leq n,1\leq j\leq N\right)\|_2^2\right)$$
$$\leq \frac{4\sigma^2}{|\Im z|^4 Nn}\mathbb{E}\left[\lambda_1\left(\left(\sigma\frac{X}{\sqrt{N}}+A\right)\left(\sigma\frac{X}{\sqrt{N}}+A\right)^*\right)\right]. \tag{4.59}$$

(4.56) readily follows from (4.55), (4.59), Theorem A.8 in [2], Lemma 4.11 and the fact that $\|A_N\|$ is uniformly bounded. Similarly, considering

$$F:\ H\to \mathrm{Tr}\left[\left(zI_N-\left(\sigma\frac{H}{\sqrt{N}}+A\right)\left(\sigma\frac{H}{\sqrt{N}}+A\right)^*\right)^{-1}E_{qp}\right],$$

where E_{qp} is the $n\times n$ matrix such that $(E_{qp})_{ij}=\delta_{qi}\delta_{pj}$, we can obtain that, for any $V\in M_{n\times N}(\mathbb{C})$ such that $\mathrm{Tr}VV^*=1$,

$$\left|\frac{d}{dt}F(X+tV)_{|_{t=0}}\right|$$
$$\leq \frac{\sigma}{\sqrt{N}}\left\{\left((GG^*)_{pp}\left(G^*\Sigma\Sigma^*G\right)_{qq}\right)^{1/2}+\left((G^*G)_{qq}\left(G\Sigma\Sigma^*G^*\right)_{pp}\right)^{1/2}\right\}.$$

Thus, one can get (4.57) in the same way. Finally, considering

$$F:\ H\to \mathrm{Tr}\left[\left(\sigma\frac{H}{\sqrt{N}}+A\right)^*\left(zI_N-\left(\sigma\frac{H}{\sqrt{N}}+A\right)\left(\sigma\frac{H}{\sqrt{N}}+A\right)^*\right)^{-1}A\right],$$

we can obtain that, for any $V \in M_{n\times N}(\mathbb{C})$ such that $\mathrm{Tr} VV^* = 1$,

$$\left|\frac{d}{dt}F(X+tV)_{|t=0}\right| \le \sigma \left\{ \left(\frac{1}{N}\mathrm{Tr}\Sigma^* GA\Sigma^* GG^*\Sigma A^* G^*\Sigma\right)^{1/2} + \left(\frac{1}{N}\mathrm{Tr}GA\Sigma^* G\Sigma\Sigma^* G^*\Sigma A^* G^*\right)^{1/2} + \left(\frac{1}{N}\mathrm{Tr}GAA^*G^*\right)^{1/2} \right\}$$

Using Lemma 4.17 (i), Theorem A.8 in [2], Lemma 4.11, the identity $\Sigma\Sigma^* G = G\Sigma\Sigma^* = -I + zG$, and the fact that $\|A_N\|$ is uniformly bounded, the same analysis allows to prove (4.58).

Corollary 4.1 *Let $\Delta_1(p,q)$, $\Delta_2(p,q)$, $(p,q) \in \{1,\dots,n\}^2$, and Δ_3 be as defined in Proposition 4.5. Then there exist a polynomial P with nonnegative coefficients and a nonnegative real number l such that, for all large N, for any $z \in \mathbb{C}\setminus\mathbb{R}$,*

$$\Delta_3(z) \le \frac{P(|\Im z|^{-1})(1+|z|)^l}{N}, \tag{4.60}$$

and for all $(p,q) \in \{1,\dots,n\}^2$,

$$\Delta_1(p,q)(z) \le \frac{P(|\Im z|^{-1})(1+|z|)^l}{N}, \tag{4.61}$$

$$\Delta_2(p,q)(z) \le \frac{P(|\Im z|^{-1})(1+|z|)^l}{N\sqrt{N}}. \tag{4.62}$$

Proof Using the identity

$$GM_N = -I + zG,$$

(4.61) readily follows from Cauchy-Schwartz inequality, Lemma 4.17 and (4.56). (4.62) and (4.60) readily follows from Cauchy-Schwartz inequality and Lemma 4.12.

4.4.3 Estimates of Resolvent Entries

In order to deduce Proposition 4.3 from Proposition 4.5 and Corollary 4.1, we need the two following Lemmas 4.13 and 4.14.

Lemma 4.13 *For all $z \in \mathbb{C} \setminus \mathbb{R}$,*

$$\frac{1}{\left|1-\sigma^2 c_N g_N(z)\right|} \leq \frac{|z|}{|\Im z|}, \tag{4.63}$$

$$\frac{1}{\left|1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)\right|} \leq \frac{|z|}{|\Im z|}. \tag{4.64}$$

Proof Since μ_{M_N} is supported by $[0, +\infty[$, (4.63) readily follows from

$$\begin{aligned}\frac{1}{\left|1-\sigma^2 c_N g_N(z)\right|} &= \frac{|z|}{\left|z-\sigma^2 c_N z g_N(z)\right|} \\ &\leq \frac{|z|}{\left|\Im(z-\sigma^2 c_N z g_N(z))\right|} = \frac{|z|}{|\Im z|\left(1+\sigma^2 c_N \mathbb{E}\int \frac{t}{|z-t|^2} d\mu_{M_N}(t)\right)}.\end{aligned}$$

(4.64) may be proved similarly.

Corollary 4.1 and Lemma 4.13 yield that, there is a polynomial Q with nonnegative coefficients, a sequence b_N of nonnegative real numbers converging to zero when N goes to infinity and some nonnegative integer number l, such that for any p, q in $\{1, \ldots, n\}$, for all $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\nabla_{pq} \leq (1+|z|)^l Q(|\Im z|^{-1}) b_N, \tag{4.65}$$

where ∇_{pq} was defined by (4.36).

Lemma 4.14 *There is a sequence v_N of nonnegative real numbers converging to zero when N goes to infinity such that for all $z \in \mathbb{C} \setminus \mathbb{R}$,*

$$\left|g_N(z) - g_{\mu_{\sigma,\nu,c}}(z)\right| \leq \left\{\frac{|z|^2+2}{|\Im z|^2} + \frac{1}{|\Im z|}\right\} v_N. \tag{4.66}$$

Proof First note that it is sufficient to prove (4.66) for $z \in \mathbb{C}^+ := \{z \in \mathbb{C}; \Im z > 0\}$ since $g_N(\bar{z}) - g_{\mu_{\sigma,\nu,c}}(\bar{z}) = \overline{g_N(z) - g_{\mu_{\sigma,\nu,c}}(z)}$. Fix $\epsilon > 0$. According to Theorem A.8 and Theorem 5.11 in [2], and the assumption on A_N, we can choose $K > \max\{2/\varepsilon; x, x \in \operatorname{supp}(\mu_{\sigma,\nu,c})\}$ large enough such that $\mathbb{P}(\|M_N\| > K)$ goes to zero as N goes to infinity. Let us write

$$g_N(z) = \mathbb{E}\left(\frac{1}{n}\mathrm{Tr} G_N(z) \mathbb{1}_{\|M_N\| \leq K}\right) + \mathbb{E}\left(\frac{1}{n}\mathrm{Tr} G_N(z) \mathbb{1}_{\|M_N\| > K}\right). \tag{4.67}$$

For any $z \in \mathbb{C}^+$ such that $|z| > 2K$, we have

$$\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr} G_N(z) \mathbb{1}_{\|M_N\| \leq K}\right)\right| \leq \frac{1}{K} \leq \frac{\epsilon}{2} \text{ and } \left|g_{\mu_{\sigma,\nu,c}}(z)\right| \leq \frac{1}{K} \leq \frac{\epsilon}{2}.$$

Thus, $\forall z \in \mathbb{C}^+$, such that $|z| > 2K$, we can deduce that

$$\begin{aligned}
&\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \frac{(\Im z)^2}{|z|^2+2} \\
&\le \left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \\
&\le \varepsilon.
\end{aligned} \tag{4.68}$$

Now, it is clear that $\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N\mathbb{1}_{\|M_N\|\le K}\right)$ is a sequence of locally bounded holomorphic functions on $\mathbb{C}^+$ which converges towards $g_{\mu_{\sigma,\nu,c}}$. Hence, by Vitali's Theorem, $\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N\mathbb{1}_{\|M_N\|\le K}\right)$ converges uniformly towards $g_{\mu_{\sigma,\nu,c}}$ on each compact subset of $\mathbb{C}^+$. Thus, there exists $N(\epsilon) > 0$, such that for any $N \ge N(\epsilon)$, for any $z \in \mathbb{C}^+$, such that $|z| \le 2K$ and $\Im z \ge \varepsilon$,

$$\begin{aligned}
&\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \frac{(\Im z)^2}{|z|^2+2} \\
&\le \left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \\
&\le \varepsilon.
\end{aligned} \tag{4.69}$$

Finally, for any $z \in \mathbb{C}^+$, such that $\Im z \in]0; \varepsilon[$, we have

$$\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \frac{(\Im z)^2}{|z|^2+2} \le \frac{2}{\Im z}\frac{(\Im z)^2}{|z|^2+2} \le \Im z \le \varepsilon. \tag{4.70}$$

It readily follows from (4.68), (4.69) and (4.70) that for $N \ge N(\epsilon)$,

$$\sup_{z\in\mathbb{C}^+}\left\{\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \frac{(\Im z)^2}{|z|^2+2}\right\} \le \varepsilon$$

Moreover, for $N \ge N'(\epsilon) \ge N(\epsilon)$, $\mathbb{P}(\|M_N\| > K) \le \varepsilon$. Therefore, for $N \ge N'(\epsilon)$, we have for any $z \in \mathbb{C}^+$,

$$\begin{aligned}
&\left|g_N(z) - g_{\mu_{\sigma,\nu,c}}(z)\right| \\
&\le \frac{|z|^2+2}{|\Im z|^2}\sup_{z\in\mathbb{C}^+}\left\{\left|\mathbb{E}\left(\frac{1}{n}\mathrm{Tr}G_N(z)\mathbb{1}_{\|M_N\|\le K}\right) - g_{\mu_{\sigma,\nu,c}}(z)\right| \frac{(\Im z)^2}{|z|^2+2}\right\} \\
&\quad + \frac{1}{\Im z}\mathbb{P}(\|M_N\| > K) \\
&\le \varepsilon\left\{\frac{|z|^2+2}{|\Im z|^2} + \frac{1}{\Im z}\right\}.
\end{aligned} \tag{4.71}$$

Thus, the proof is complete by setting

$$v_N = \sup_{z\in\mathbb{C}^+} \left\{ \left| g_N(z) - g_{\mu_{\sigma,\nu,c}}(z) \right| \left(\frac{|z|^2+2}{|\Im z|^2} + \frac{1}{\Im z} \right)^{-1} \right\}.$$

Now set

$$\tau_N = (1-\sigma^2 c_N g_N(z))z - \frac{\gamma_q(N)}{1-\sigma^2 c_N g_N(z)} - \sigma^2(1-c_N)$$

and

$$\tilde{\tau}_N = (1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))z - \frac{\gamma_q(N)}{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)} - \sigma^2(1-c). \tag{4.72}$$

Lemmas 4.13 and 4.14 yield that there is a polynomial R with nonnegative coefficients, a sequence w_N of nonnegative real numbers converging to zero when N goes to infinity and some nonnegative real number l, such that for all $z \in \mathbb{C} \setminus \mathbb{R}$,

$$|\tau_N - \tilde{\tau}_N| \leq (1+|z|)^l R(|\Im z|^{-1}) w_N. \tag{4.73}$$

Now, one can easily see that,

$$\left| \Im \left\{ (1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))z - \frac{\gamma_q(N)}{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)} - \sigma^2(1-c) \right\} \right| \geq |\Im z|, \tag{4.74}$$

so that

$$\left| \frac{1}{\tilde{\tau}_N} \right| \leq \frac{1}{|\Im z|}. \tag{4.75}$$

Note that

$$\frac{1}{\tilde{\tau}_N} = \frac{(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))}{\omega_{\sigma,\nu,c}(z) - \gamma_q(N)}. \tag{4.76}$$

Then, (4.16) readily follows from Proposition 4.5, (4.65), (4.73), (4.75), (4.76), and (ii) Lemma 4.17. The proof of Proposition 4.3 is complete.

4.5 Proof of Theorem 4.3

We follow the two steps presented in Sect. 4.2.

Step A We first prove (4.11).

Let $\eta > 0$ small enough and N large enough such that for any $l = 1, \dots, J$, $\alpha_l(N) \in [\theta_l - \eta, \theta_l + \eta]$ and $[\theta_l - 2\eta, \theta_l + 2\eta]$ contains no other element of the spectrum of $A_N A_N^*$ than $\alpha_l(N)$. For any $l = 1, \dots, J$, choose $f_{\eta,l}$ in $\mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ with support in $[\theta_l - 2\eta, \theta_l + 2\eta]$ such that $f_{\eta,l}(x) = 1$ for any $x \in [\theta_l - \eta, \theta_l + \eta]$ and $0 \le f_{\eta,l} \le 1$. Let $0 < \epsilon < \delta_0$ where δ_0 is introduced in Theorem 4.2. Choose $h_{\varepsilon,j}$ in $\mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ with support in $[\rho_{\theta_j} - \varepsilon, \rho_{\theta_j} + \varepsilon]$ such that $h_{\varepsilon,j} \equiv 1$ on $[\rho_{\theta_j} - \varepsilon/2, \rho_{\theta_j} + \varepsilon/2]$ and $0 \le h_{\varepsilon,j} \le 1$.

Almost surely for all large N, M_N has k_j eigenvalues in $]\rho_{\theta_j} - \varepsilon/2, \rho_{\theta_j} + \varepsilon/2[$. According to Theorem 4.2, denoting by $(\xi_1, \cdots, \xi_{k_j})$ an orthonormal system of eigenvectors associated to the k_j eigenvalues of M_N in $(\rho_{\theta_j} - \varepsilon/2, \rho_{\theta_j} + \varepsilon/2)$, it readily follows from (4.12) that almost surely for all large N,

$$\sum_{n=1}^{k_j} \left\| P_{\ker(\alpha_l(N) I_n - A_N A_N^*)} \xi_n \right\|^2 = \operatorname{Tr}\left[h_{\varepsilon,j}(M_N) f_{\eta,l}(A_N A_N^*) \right].$$

Applying Proposition 4.2 with $\Gamma_N = f_{\eta,l}(A_N A_N^*)$ and $K = k_l$, the problem of establishing (4.11) is reduced to prove that

$$\mathbb{E}\left(\operatorname{Tr}\left[h_{\varepsilon,j}\left(\left(\sigma \frac{\mathscr{G}_N}{\sqrt{N}} + A_N \right) \left(\sigma \frac{\mathscr{G}_N}{\sqrt{N}} + A_N \right)^* \right) f_{\eta,l}(A_N A_N^*) \right] \right)$$
$$\to_{N \to +\infty} \frac{k_j \delta_{jl} (1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(\rho_{\theta_j}))}{\omega'_{\sigma,\nu,c}(\rho_{\theta_j})}. \tag{4.77}$$

Using a Singular Value Decomposition of A_N and the biunitarily invariance of the distribution of $\mathscr{G}_N$, we can assume that A_N is as (4.14) and such that for any $j = 1, \dots, J$,

$$(A_N A_N^*)_{ii} = \alpha_j(N) \quad \text{for } i = k_1 + \ldots + k_{j-1} + l, l = 1, \dots, k_j.$$

Now, according to Lemma 4.18,

$$\mathbb{E}\left(\operatorname{Tr}\left[h_{\varepsilon,j}\left(\left(\sigma \frac{\mathscr{G}_N}{\sqrt{N}} + A_N \right) \left(\sigma \frac{\mathscr{G}_N}{\sqrt{N}} + A_N \right)^* \right) f_{\eta,l}(A_N A_N^*) \right] \right)$$
$$= - \lim_{y \to 0^+} \frac{1}{\pi} \int \Im \mathbb{E} \operatorname{Tr}\left[G_N^{\mathscr{G}}(t + iy) f_{\eta,l}(A_N A_N^*) \right] h_{\varepsilon,j}(t) dt,$$

with, for all large N,

$$\mathbb{E}\mathrm{Tr}\left[G_N^{\mathscr{G}}(t+\mathrm{i}y)f_{\eta,l}(A_NA_N^*)\right] = \sum_{k=k_1+\cdot+k_{l-1}+1}^{k_1+\cdot+k_l} f_{\eta,l}(\alpha_l(N))\mathbb{E}[G_N^{\mathscr{G}}(t+\mathrm{i}y)]_{kk}$$
$$= \sum_{k=k_1+\cdot+k_{l-1}+1}^{k_1+\cdot+k_l} \mathbb{E}[G_N^{\mathscr{G}}(t+\mathrm{i}y)]_{kk}.$$

Now, by considering

$$\tau' = (1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))z - \frac{\theta_l}{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)} - \sigma^2(1-c)$$

instead of dealing with $\tilde{\tau}_N$ defined in (4.72) at the end of the proof of Proposition 4.3, one can prove that there is a polynomial P with nonnegative coefficients, a sequence $(u_N)_N$ of nonnegative real numbers converging to zero when N goes to infinity and some nonnegative real number s, such that for any k in $\{k_1+\ldots+k_{l-1}+1,\ldots,k_1+\ldots+k_l\}$, for all $z\in\mathbb{C}\setminus\mathbb{R}$,

$$\mathbb{E}\left(\left(G_N^{\mathscr{G}}(z)\right)_{kk}\right) = \frac{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z)}{\omega_{\sigma,\nu,c}(z)-\theta_l} + \Delta_{k,N}(z), \tag{4.78}$$

with

$$\left|\Delta_{k,N}(z)\right| \le (1+|z|)^s P(|\Im z|^{-1})u_N.$$

Thus,

$$\mathbb{E}\mathrm{Tr}\left[G_N^{\mathscr{G}}(t+\mathrm{i}y)f_{\eta,l}(A_NA_N^*)\right] = k_l\frac{1-\sigma^2 c g_{\mu,\sigma,\nu}(t+\mathrm{i}y)}{\omega_{\sigma,\nu,c}(z)-\theta_l} + \Delta_N(t+\mathrm{i}y),$$

where for all $z\in\mathbb{C}\setminus\mathbb{R}$, $\Delta_N(z)=\sum_{k=k_1+\cdot+k_{l-1}+1}^{k_1+\cdot+k_l}\Delta_{k,N}(z)$, and $|\Delta_N(z)|\le k_l(1+|z|)^s P(|\Im z|^{-1})u_N$.

First let us compute

$$\lim_{y\downarrow 0}\frac{k_l}{\pi}\int_{\rho_{\theta_j}-\varepsilon}^{\rho_{\theta_j}+\varepsilon}\Im\frac{h_{\varepsilon,j}(t)(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t+\mathrm{i}y))}{\theta_l-\omega_{\sigma,\nu,c}(t+\mathrm{i}y)}\,dt.$$

The function $\omega_{\sigma,\nu,c}$ satisfies $\omega_{\sigma,\nu,c}(\overline{z})=\overline{\omega_{\sigma,\nu,c}(z)}$ and $g_{\mu_{\sigma,\nu,c}}(\overline{z})=\overline{g_{\mu_{\sigma,\nu,c}}(z)}$, so that $\Im\frac{(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t+\mathrm{i}y))}{\theta_l-\omega_{\sigma,\nu,c}(t+\mathrm{i}y)}=\frac{1}{2i}[\frac{(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t+\mathrm{i}y))}{\theta_l-\omega_{\sigma,\nu,c}(t+\mathrm{i}y)}-\frac{(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t-iy))}{\theta_l-\omega_{\sigma,\nu,c}(t-iy)}]$. As in [10], the above integral is split into three pieces, namely $\int_{\rho_{\theta_j}-\varepsilon}^{\rho_{\theta_j}-\varepsilon/2}+\int_{\rho_{\theta_j}-\varepsilon/2}^{\rho_{\theta_j}+\varepsilon/2}+\int_{\rho_{\theta_j}+\varepsilon/2}^{\rho_{\theta_j}+\varepsilon}$.

Each of the first and third integrals are easily seen to go to zero when $y \downarrow 0$ by a direct application of the definition of the functions involved and of the (Riemann) integral. As $h_{\varepsilon,j}$ is constantly equal to one on $[\rho_{\theta_j} - \epsilon/2; \rho_{\theta_j} + \epsilon/2]$, the second (middle) term is simply the integral

$$\frac{k_l}{2\pi i}\int_{\rho_{\theta_j}-\varepsilon/2}^{\rho_{\theta_j}+\varepsilon/2} \frac{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t+\mathrm{i}y)}{\theta_l - \omega_{\sigma,\nu,c}(t+\mathrm{i}y)} - \frac{1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(t-\mathrm{i}y)}{\theta_l - \omega_{\sigma,\nu,c}(t-\mathrm{i}y)}\, dt.$$

Completing this to a contour integral on the rectangular with corners $\rho_{\theta_j} \pm \varepsilon/2 \pm iy$ and noting that the integrals along the vertical lines tend to zero as $y \downarrow 0$ allows a direct application of the residue theorem for the final result, if $l = j$,

$$\frac{k_j(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(\rho_{\theta_j}))}{\omega'_{\sigma,\nu,c}(\rho_{\theta_j})}.$$

If we consider θ_l for some $l \neq j$, then $z \mapsto (1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(z))(\theta_l - \omega_{\sigma,\nu,c}(z))^{-1}$ is analytic around ρ_{θ_j}, so its residue at ρ_{θ_j} is zero, and the above argument provides zero as answer.

Now, according to Lemma 4.19, we have

$$\limsup_{y\to 0^+} (u_N)^{-1}\left|\int h_{\varepsilon,j}(t)\Delta_N(t+\mathrm{i}y)dt\right| < +\infty$$

so that

$$\lim_{N\to+\infty}\limsup_{y\to 0^+}\left|\int h_{\varepsilon,j}(t)\Delta_N(t+\mathrm{i}y)dt\right| = 0. \tag{4.79}$$

This concludes the proof of (4.11).

Step B In the second, and final, step, we shall use a perturbation argument identical to the one used in [10] to reduce the problem to the case of a spike with multiplicity one, case that follows trivially from Step A. A further property of eigenvectors of Hermitian matrices which are close to each other in the norm will be important in the analysis of the behaviour of the eigenvectors of our matrix models. Given a Hermitian matrix $M \in M_N(\mathbb{C})$ and a Borel set $S \subseteq \mathbb{R}$, we denote by $E_M(S)$ the spectral projection of M associated to S. In other words, the range of $E_M(S)$ is the vector space generated by the eigenvectors of M corresponding to eigenvalues in S. The following lemma can be found in [6].

Lemma 4.15 *Let M and M_0 be $N \times N$ Hermitian matrices. Assume that $\alpha, \beta, \delta \in \mathbb{R}$ are such that $\alpha < \beta$, $\delta > 0$, M and M_0 has no eigenvalues in $[\alpha-\delta, \alpha] \cup [\beta, \beta+\delta]$. Then,*

$$\|E_M((\alpha,\beta)) - E_{M_0}((\alpha,\beta))\| < \frac{4(\beta-\alpha+2\delta)}{\pi\delta^2}\|M - M_0\|.$$

In particular, for any unit vector $\xi \in E_{M_0}((\alpha,\beta))(\mathbb{C}^N)$,

$$\|(I_N - E_M((\alpha,\beta)))\xi\|_2 < \frac{4(\beta-\alpha+2\delta)}{\pi\delta^2}\|M-M_0\|.$$

Assume that θ_i is in $\Theta_{\sigma,\nu,c}$ defined in (4.7) and $k_i \neq 1$. Let us denote by $V_1(i), \ldots, V_{k_i}(i)$, an orthonormal system of eigenvectors of $A_N A_N^*$ associated with $\alpha_i(N)$. Consider a Singular Value Decomposition $A_N = U_N D_N V_N$ where V_N is a $N \times N$ unitary matrix, U_N is a $n \times n$ unitary matrix whose k_i first columns are $V_1(i), \ldots, V_{k_i}(i)$ and D_N is as (4.14) with the first k_i diagonal elements equal to $\sqrt{\alpha_i(N)}$.

Let δ_0 be as in Theorem 4.2. Almost surely, for all N large enough, there are k_i eigenvalues of M_N in $(\rho_{\theta_i} - \frac{\delta_0}{4}, \rho_{\theta_i} + \frac{\delta_0}{4})$, namely $\lambda_{n_{i-1}+q}(M_N)$, $q = 1, \ldots, k_i$ (where $n_{i-1}+1, \ldots, n_{i-1}+k_i$ are the descending ranks of $\alpha_i(N)$ among the eigenvalues of $A_N A_N^*$), which are moreover the only eigenvalues of M_N in $(\rho_{\theta_i} - \delta_0, \rho_{\theta_i} + \delta_0)$. Thus, the spectrum of M_N is split into three pieces:

$$\{\lambda_1(M_N), \ldots, \lambda_{n_{i-1}}(M_N)\} \subset (\rho_{\theta_i} + \delta_0, +\infty[,$$

$$\{\lambda_{n_{i-1}+1}(M_N), \ldots, \lambda_{n_{i-1}+k_i}(M_N)\} \subset (\rho_{\theta_i} - \frac{\delta_0}{4}, \rho_{\theta_i} + \frac{\delta_0}{4}),$$

$$\{\lambda_{n_{i-1}+k_i+1}(M_N), \ldots, \lambda_N(M_N)\} \subset [0, \rho_{\theta_i} - \delta_0).$$

The distance between any of these components is equal to $3\delta_0/4$. Let us fix ϵ_0 such that $0 \le \theta_i(2\epsilon_0 k_i + \epsilon_0^2 k_i^2) < dist(\theta_i, \mathrm{supp}\,\nu \cup_{i\neq s} \theta_s)$ and such that $[\theta_i; \theta_i + \theta_i(2\epsilon_0 k_i + \epsilon_0^2 k_i^2)] \subset \mathscr{E}_{\sigma,\nu,c}$ defined by (4.6). For any $0 < \epsilon < \epsilon_0$, define the matrix $A_N(\epsilon)$ as $A_N(\epsilon) = U_N D_N(\epsilon) V_N$ where

$$(D_N(\epsilon))_{m,m} = \sqrt{\alpha_i(N)}[1 + \epsilon(k_i - m + 1)], \quad \text{for } m \in \{1, \ldots, k_i\},$$

and $(D_N(\epsilon))_{pq} = (D_N)_{pq}$ for any $(p,q) \notin \{(m,m), m \in \{1, \ldots, k_i\}\}$.

Set

$$M_N(\epsilon) = \left(\sigma\frac{X_N}{\sqrt{N}} + A_N(\epsilon)\right)\left(\sigma\frac{X_N}{\sqrt{N}} + A_N(\epsilon)\right)^*.$$

For N large enough, for each $m \in \{1, \ldots, k_i\}$, $\alpha_i(N)[1 + \epsilon(k_i - m + 1)]^2$ is an eigenvalue of $A_N A_N^*(\epsilon)$ with multiplicity one. Note that, since $\sup_N \|A_N\| < +\infty$, it is easy to see that there exist some constant C such that for any N and for any $0 < \epsilon < \epsilon_0$,

$$\|M_N(\epsilon) - M_N\| \le C\epsilon\left(\left\|\frac{X_N}{\sqrt{N}}\right\| + 1\right).$$

Applying Remark 4.3 to the $(n+N)\times(n+N)$ matrix $\tilde{X}_N = \begin{pmatrix} 0_{n\times n} & X_N \\ X_N^* & 0_{N\times N} \end{pmatrix}$ (see also Appendix B of [14]), it readily follows that there exists some constant C' such that a.s for all large N, for any $0 < \epsilon < \epsilon_0$,

$$\|M_N(\epsilon) - M_N\| \leq C'\epsilon. \tag{4.80}$$

Therefore, for ϵ sufficiently small such that $C'\epsilon < \delta_0/4$, by Theorem A.46 [2], there are precisely n_{i-1} eigenvalues of $M_N(\epsilon)$ in $[0, \rho_{\theta_i} - 3\delta_0/4)$, precisely k_i in $(\rho_{\theta_i} - \delta_0/2, \rho_{\theta_i} + \delta_0/2)$ and precisely $N-(n_{i-1}+k_i)$ in $(\rho_{\theta_i} + 3\delta_0/4, +\infty[$. All these intervals are again at strictly positive distance from each other, in this case $\delta_0/4$.

Let ξ be a normalized eigenvector of M_N relative to $\lambda_{n_{i-1}+q}(M_N)$ for some $q \in \{1,\ldots,k_i\}$. As proved in Lemma 4.15, if $E(\epsilon)$ denotes the subspace spanned by the eigenvectors associated to $\{\lambda_{n_{i-1}+1}(M_N(\epsilon)),\ldots,\lambda_{n_{i-1}+k_i}(M_N(\epsilon))\}$ in $\mathbb{C}^N$, then there exists some constant C (which depends on δ_0) such that for ϵ small enough, almost surely for large N,

$$\left\| P_{E(\epsilon)^\perp}\xi \right\|_2 \leq C\epsilon. \tag{4.81}$$

According to Theorem 4.2, for any j in $\{1,\ldots,k_i\}$, for large enough N, $\lambda_{n_{i-1}+j}(M_N(\epsilon))$ separates from the rest of the spectrum and belongs to a neighborhood of $\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))$ where

$$\theta_i^{(j)}(\epsilon) = \theta_i\left(1+\epsilon(k_i-j+1)\right)^2.$$

If $\xi_j(\epsilon,i)$ denotes a normalized eigenvector associated to $\lambda_{n_{i-1}+j}(M_N(\epsilon))$, Step A above implies that almost surely for any $p \in \{1,\ldots,k_i\}$, for any $\gamma > 0$, for all large N,

$$\left| \left|\langle V_p(i), \xi_j(\epsilon,i)\rangle\right|^2 - \frac{\delta_{jp}\left(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon)))\right)}{\omega'_{\sigma,\nu,c}\left(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))\right)} \right| < \gamma. \tag{4.82}$$

The eigenvector ξ decomposes uniquely in the orthonormal basis of eigenvectors of $M_N(\epsilon)$ as $\xi = \sum_{j=1}^{k_i} c_j(\epsilon)\xi_j(\epsilon,i) + \xi(\epsilon)^\perp$, where $c_j(\epsilon) = \langle \xi | \xi_j(\epsilon,i)\rangle$ and $\xi(\epsilon)^\perp = P_{E(\epsilon)^\perp}\xi$; necessarily $\sum_{j=1}^{k_i}|c_j(\epsilon)|^2 + \|\xi(\epsilon)^\perp\|_2^2 = 1$. Moreover, as indicated in relation (4.81), $\|\xi(\epsilon)^\perp\|_2 \leq C\epsilon$. We have

$$\begin{aligned} P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi = & \sum_{j=1}^{k_i} c_j(\epsilon) P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi_j(\epsilon,i) \\ & + P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi(\epsilon)^\perp \end{aligned}$$

$$= \sum_{j=1}^{k_i} c_j(\epsilon) \sum_{l=1}^{k_i} \langle \xi_j(\epsilon, i) | V_l(i) \rangle V_l(i)$$
$$+ P_{\ker(\alpha_i(N) I_N - A_N A_N^*)} \xi(\epsilon)^\perp.$$

Take in the above the scalar product with $\xi = \sum_{j=1}^{k_i} c_j(\epsilon)\xi_j(\epsilon, i) + \xi(\epsilon)^\perp$ to get

$$\langle P_{\ker(\alpha_i(N) I_N - A_N A_N^*)} \xi | \xi \rangle$$
$$= \sum_{j,l,s=1}^{k_i} c_j(\epsilon) \langle \xi_j(\epsilon, i) | V_l(i) \rangle \overline{c_s(\epsilon)} \langle V_l(i) | \xi_s(\epsilon, i) \rangle$$
$$+ \sum_{j=1}^{k_i} c_j(\epsilon) \sum_{l=1}^{k_i} \langle \xi_j(\epsilon, i) | V_l(i) \rangle \langle V_l(i) | \xi(\epsilon)^\perp \rangle$$
$$+ \langle P_{\ker(\alpha_i(N) I_N - A_N A_N^*)} \xi(\epsilon)^\perp | \xi \rangle.$$

Relation (4.82) indicates that

$$\sum_{j,l,s=1}^{k_i} c_j(\epsilon) \langle \xi_j(\epsilon, i) | V_l(i) \rangle \overline{c_s(\epsilon)} \langle V_l(i) | \xi_s(\epsilon, i) \rangle$$
$$= \sum_{j=1}^{k_i} |c_j(\epsilon)|^2 |\langle V_j(i) | \xi_j(\epsilon, i) \rangle|^2 + \Delta_1$$
$$= \sum_{j=1}^{k_i} |c_j(\epsilon)|^2 \frac{\left(1 - \sigma^2 c g_{\mu_{\sigma,\nu,c}}(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon)))\right)}{\omega'_{\sigma,\nu,c}\left(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))\right)} + \Delta_1 + \Delta_2,$$

where for all large N, $|\Delta_1| \leq \sqrt{\gamma} k_i^3$ and $|\Delta_2| \leq \gamma$. Since $\|\xi(\epsilon)^\perp\|_2 \leq C\epsilon$,

$$\Bigg| \sum_{j=1}^{k_i} c_j(\epsilon) \sum_{l=1}^{k_i} \langle \xi_j(\epsilon, i) | V_l(i) \rangle \langle V_l(i) | \xi(\epsilon)^\perp \rangle$$
$$+ \langle P_{\ker(\alpha_i(N) I_N - A_N A_N^*)} \xi(\epsilon)^\perp | \xi \rangle \Bigg| \leq \left(k_i^2 + 1\right) C\epsilon.$$

Thus, we conclude that almost surely for any $\gamma > 0$, for all large N,

$$\left| \langle P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi|\xi\rangle - \sum_{j=1}^{k_i} \frac{|c_j(\epsilon)|^2 \left(1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon)))\right)}{\omega'_{\sigma,\nu,c}\left(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))\right)} \right|$$

$$\leq (k_i^2+1)C\epsilon + \sqrt{\gamma}k_i^3 + \gamma. \tag{4.83}$$

Since we have the identity

$$\langle P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi|\xi\rangle = \| P_{\ker(\alpha_i(N)I_N - A_N A_N^*)}\xi\|_2^2$$

and the three obvious convergences $\lim_{\epsilon\to 0}\omega'_{\sigma,\nu,c}\left(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))\right) = \omega'_{\sigma,\nu,c}(\rho_{\theta_i})$, $\lim_{\epsilon\to 0} g_{\mu_{\sigma,\nu,c}}\left(\Phi_{\sigma,\nu,c}(\theta_i^{(j)}(\epsilon))\right) = g_{\mu_{\sigma,\nu,c}}(\rho_{\theta_i})$ and $\lim_{\epsilon\to 0}\sum_{j=1}^{k_i}|c_j(\epsilon)|^2 = 1$, relation (4.83) concludes Step B and the proof of Theorem 4.3. (Note that we use (2.9) of [11] which is true for any $x\in\mathbb{C}\setminus\mathbb{R}$ to deduce that $1-\sigma^2 c g_{\mu_{\sigma,\nu,c}}(\Phi_{\sigma,\nu,c}(\theta_i)) = \frac{1}{1+\sigma^2 c g_\nu(\theta_i)}$ by letting x goes to $\Phi_{\sigma,\nu,c}(\theta_i)$).

Acknowledgements The author is very grateful to Charles Bordenave and Serban Belinschi for several fruitful discussions and thanks Serban Belinschi for pointing out Lemma 4.14. The author also wants to thank an anonymous referee who provided a much simpler proof of Lemma 4.13 and encouraged the author to establish the results for non diagonal perturbations, which led to an overall improvement of the paper.

Appendix 1

We present alternative versions on the one hand of the result in [3] about the lack of eigenvalues outside the support of the deterministic equivalent measure, and on the other hand of the result in [11] about the exact separation phenomenon. These new versions (Theorems 4.5 and 4.6 below) deal with random variables whose imaginary and real parts are independent, but remove the technical assumptions ((1.10) and "$b_1 > 0$" in Theorem 1.1 in [3] and "$\omega_{\sigma,\nu,c}(b) > 0$" in Theorem 1.2 in [11]). The proof of Theorem 4.5 is based on the results of [5]. The arguments of the proof of Theorem 1.2 in [11] and Theorem 4.5 lead to the proof of Theorem 4.6.

Theorem 4.4 *Consider*

$$M_N = (\sigma\frac{X_N}{\sqrt{N}} + A_N)(\sigma\frac{X_N}{\sqrt{N}} + A_N)^*, \tag{4.84}$$

and assume that

1. $X_N = [X_{ij}]_{1\le i\le n, 1\le j\le N}$ *is a* $n \times N$ *random matrix such that* $[X_{ij}]_{i\ge 1, j\ge 1}$ *is an infinite array of random variables which satisfy* (4.1) *and* (4.2) *and such that* $\Re(X_{ij})$, $\Im(X_{ij})$, $(i, j) \in \mathbb{N}^2$, *are independent, centered with variance* $1/2$.
2. A_N *is an* $n \times N$ *nonrandom matrix such that* $\|A_N\|$ *is uniformly bounded.*
3. $n \le N$ *and, as* N *tends to infinity,* $c_N = n/N \to c \in]0, 1]$.
4. $[x, y]$, $x < y$, *is such that there exists* $\delta > 0$ *such that for all large* N, $]x-\delta; y+\delta[\subset \mathbb{R} \setminus \mathrm{supp}(\mu_{\sigma,\mu_{A_N A_N^*}, c_N})$ *where* $\mu_{\sigma,\mu_{A_N A_N^*}, c_N}$ *is the nonrandom distribution which is characterized in terms of its Stieltjes transform which satisfies Eq.* (4.4) *where we replace* c *by* c_N *and* ν *by* $\mu_{A_N A_N^*}$.

Then, we have

$$\mathbb{P}[\textit{for all large } N, \mathrm{spect}(M_N) \subset \mathbb{R} \setminus [x, y]] = 1.$$

Since, in the proof of Theorem 4.4, we will use tools from free probability theory, for the reader's convenience, we recall the following basic definitions from free probability theory. For a thorough introduction to free probability theory, we refer to [29].

- A $\mathscr{C}^*$-probability space is a pair $(\mathscr{A}, \tau)$ consisting of a unital $\mathscr{C}^*$-algebra $\mathscr{A}$ and a state τ on $\mathscr{A}$ i.e. a linear map $\tau : \mathscr{A} \to \mathbb{C}$ such that $\tau(1_{\mathscr{A}}) = 1$ and $\tau(aa^*) \ge 0$ for all $a \in \mathscr{A}$. τ is a trace if it satisfies $\tau(ab) = \tau(ba)$ for every $(a, b) \in \mathscr{A}^2$. A trace is said to be faithful if $\tau(aa^*) > 0$ whenever $a \neq 0$. An element of $\mathscr{A}$ is called a noncommutative random variable.
- The noncommutative $\star$-distribution of a family $a = (a_1, \dots, a_k)$ of noncommutative random variables in a $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ is defined as the linear functional $\mu_a : P \mapsto \tau(P(a, a^*))$ defined on the set of polynomials in $2k$ noncommutative indeterminates, where (a, a^*) denotes the $2k$-uple $(a_1, \dots, a_k, a_1^*, \dots, a_k^*)$. For any selfadjoint element a_1 in $\mathscr{A}$, there exists a probability measure ν_{a_1} on $\mathbb{R}$ such that, for every polynomial P, we have

$$\mu_{a_1}(P) = \int P(t) \mathrm{d}\nu_{a_1}(t).$$

  Then we identify μ_{a_1} and ν_{a_1}. If τ is faithful then the support of ν_{a_1} is the spectrum of a_1 and thus $\|a_1\| = \sup\{|z|, z \in \mathrm{support}(\nu_{a_1})\}$.
- A family of elements $(a_i)_{i\in I}$ in a $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ is free if for all $k \in \mathbb{N}$ and all polynomials $p_1, \dots, p_k$ in two noncommutative indeterminates, one has

$$\tau(p_1(a_{i_1}, a_{i_1}^*) \cdots p_k(a_{i_k}, a_{i_k}^*)) = 0 \tag{4.85}$$

  whenever $i_1 \neq i_2, i_2 \neq i_3, \dots, i_{k-1} \neq i_k$, $(i_1, \dots i_k) \in I^k$, and $\tau(p_l(a_{i_l}, a_{i_l}^*)) = 0$ for $l = 1, \dots, k$.

- A noncommutative random variable x in a $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ is a standard semicircular random variable if $x = x^*$ and for any $k \in \mathbb{N}$,

$$\tau(x^k) = \int t^k d\mu_{sc}(t)$$

where $d\mu_{sc}(t) = \frac{1}{2\pi}\sqrt{4-t^2}\mathbb{1}_{[-2;2]}(t)dt$ is the semicircular standard distribution.
- Let k be a nonnull integer number. Denote by $\mathscr{P}$ the set of polynomials in $2k$ noncommutative indeterminates. A sequence of families of variables $(a_n)_{n\geq 1} = (a_1(n), \ldots, a_k(n))_{n\geq 1}$ in C^*-probability spaces $(\mathscr{A}_n, \tau_n)$ converges in $\star$-distribution, when n goes to infinity, to some k-tuple of noncommutative random variables $a = (a_1, \ldots, a_k)$ in a $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ if the map $P \in \mathscr{P} \mapsto \tau_n(P(a_n, a_n^*))$ converges pointwise towards $P \in \mathscr{P} \mapsto \tau(P(a, a^*))$.
- k noncommutative random variables $a_1(n), \ldots, a_k(n)$, in C^*-probability spaces $(\mathscr{A}_n, \tau_n)$, $n \geq 1$, are said asymptotically free if $(a_1(n), \ldots, a_k(n))$ converges in $\star$-distribution, as n goes to infinity, to some noncommutative random variables $(a_1, \ldots, a_k)$ in a $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ where $a_1, \ldots, a_k$ are free.

We will also use the following well known result on asymptotic freeness of random matrices. Let $\mathscr{A}_n$ be the algebra of $n \times n$ matrices with complex entries and endow this algebra with the normalized trace defined for any $M \in \mathscr{A}_n$ by $\tau_n(M) = \frac{1}{n}\mathrm{Tr}(M)$. Let us consider a $n \times n$ so-called standard G.U.E matrix, i.e. a random Hermitian matrix $\mathscr{G}_n = [\mathscr{G}_{jk}]_{j,k=1}^n$, where $\mathscr{G}_{ii}$, $\sqrt{2}\Re(\mathscr{G}_{ij})$, $\sqrt{2}\Im(\mathscr{G}_{ij})$, $i < j$ are independent centered Gaussian random variables with variance 1. For a fixed real number t independent from n, let $H_n^{(1)}, \ldots, H_n^{(t)}$ be deterministic $n \times n$ Hermitian matrices such that $\max_{i=1}^t \sup_n \|H_n^{(i)}\| < +\infty$ and $(H_n^{(1)}, \ldots, H_n^{(t)})$, as a t-tuple of noncommutative random variables in $(\mathscr{A}_n, \tau_n)$, converges in distribution when n goes to infinity. Then, according to Theorem 5.4.5 in [1], $\frac{\mathscr{G}_n}{\sqrt{n}}$ and $(H_n^{(1)}, \ldots, H_n^{(t)})$ are almost surely asymptotically free i.e. almost surely, for any polynomial P in t+1 noncommutative indeterminates,

$$\tau_n\left\{P\left(H_n^{(1)}, \ldots, H_n^{(t)}, \frac{\mathscr{G}_n}{\sqrt{n}}\right)\right\} \to_{n\to+\infty} \tau\left(P(h_1, \ldots, h_t, s)\right) \tag{4.86}$$

where $h_1, \ldots, h_t$ and s are noncommutative random variables in some $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ such that $(h_1, \ldots, h_t)$ and s are free, s is a standard semi-circular noncommutative random variable and the distribution of $(h_1, \ldots, h_t)$ is the limiting distribution of $(H_n^{(1)}, \ldots, H_n^{(t)})$.

Finally, the proof of Theorem 4.4 is based on the following result which can be established by following the proof of Theorem 1.1 in [5]. First, note that the algebra of polynomials in non-commuting indeterminates $X_1, \ldots, X_k$, becomes a $\star$-algebra by anti-linear extension of $(X_{i_1}X_{i_2}\ldots X_{i_m})^* = X_{i_m}\ldots X_{i_2}X_{i_1}$.

Theorem 4.5 *Let us consider three independent infinite arrays of random variables,* $[W_{ij}^{(1)}]_{i\geq 1, j\geq 1}$, $[W_{ij}^{(2)}]_{i\geq 1, j\geq 1}$ *and* $[X_{ij}]_{i\geq 1, j\geq 1}$ *where*

- *for* $l = 1, 2$, $W_{ii}^{(l)}$, $\sqrt{2}\Re(W_{ij}^{(l)})$, $\sqrt{2}\Im(W_{ij}^{(l)})$, $i < j$, *are i.i.d centered and bounded random variables with variance 1 and* $W_{ji}^{(l)} = \overline{W_{ij}^{(l)}}$,
- $\{\Re(X_{ij}), \Im(X_{ij}), i \in \mathbb{N}, j \in \mathbb{N}\}$ *are independent centered random variables with variance* $1/2$ *and satisfy* (4.1) *and* (4.2).

For any $(N, n) \in \mathbb{N}^2$, *define the* $(n+N) \times (n+N)$ *matrix:*

$$W_{n+N} = \begin{pmatrix} W_n^{(1)} & X_N \\ X_N^* & W_N^{(2)} \end{pmatrix} \tag{4.87}$$

where $X_N = [X_{ij}]_{1 \leq i \leq n, 1 \leq j \leq N}$, $W_n^{(1)} = [W_{ij}^{(1)}]_{1\leq i,j\leq n}$, $W_N^{(2)} = [W_{ij}^{(2)}]_{1\leq i,j\leq N}$.

Assume that $n = n(N)$ *and* $\lim_{N\to+\infty} \frac{n}{N} = c \in]0, 1]$.

Let t *be a fixed integer number and* P *be a selfadjoint polynomial in* $t + 1$ *noncommutative indeterminates.*

For any $N \in \mathbb{N}^2$, *let* $(B_{n+N}^{(1)}, \ldots, B_{n+N}^{(t)})$ *be a* $t-$*tuple of* $(n+N) \times (n+N)$ *deterministic Hermitian matrices such that for any* $u = 1, \ldots, t$, $\sup_N \|B_{n+N}^{(u)}\| < \infty$. *Let* $(\mathscr{A}, \tau)$ *be a* C^**-probability space equipped with a faithful tracial state and* s *be a standard semi-circular noncommutative random variable in* $(\mathscr{A}, \tau)$. *Let* $b_{n+N} = (b_{n+N}^{(1)}, \ldots, b_{n+N}^{(t)})$ *be a t-tuple of noncommutative selfadjoint random variables which is free from* s *in* $(\mathscr{A}, \tau)$ *and such that the distribution of* b_{n+N} *in* $(\mathscr{A}, \tau)$ *coincides with the distribution of* $(B_{n+N}^{(1)}, \ldots, B_{n+N}^{(t)})$ *in* $(M_{n+N}(\mathbb{C}), \frac{1}{n+N}\mathrm{Tr})$.

Let $[x, y]$ *be a real interval such that there exists* $\delta > 0$ *such that, for any large* N, $[x - \delta, y + \delta]$ *lies outside the support of the distribution of the noncommutative random variable* $P\left(s, b_{n+N}^{(1)}, \ldots, b_{n+N}^{(t)}\right)$ *in* $(\mathscr{A}, \tau)$. *Then, almost surely, for all large* N,

$$\mathrm{spect} P\left(\frac{W_{n+N}}{\sqrt{n+N}}, B_{n+N}^{(1)}, \ldots, B_{n+N}^{(t)})\right) \subset \mathbb{R} \setminus [x, y].$$

Proof We start by checking that a truncation and Gaussian convolution procedure as in Section 2 of [5] can be handled for such a matrix as defined by (4.87), to reduce the problem to a fit framework where,

(H) for any N, $(W_{n+N})_{ii}$, $\sqrt{2}\Re((W_{n+N})_{ij})$, $\sqrt{2}\Im((W_{n+N})_{ij})$, $i < j, i \leq n + N$, $j \leq n + N$, are independent, centered random variables with variance 1, which satisfy a Poincaré inequality with common fixed constant C_{PI}.

Note that, according to Corollary 3.2 in [24], (H) implies that for any $p \in \mathbb{N}$,

$$\sup_{N\geq 1} \sup_{1\leq i,j\leq n+N} \mathbb{E}\left(|(W_{n+N})_{ij}|^p\right) < +\infty. \tag{4.88}$$

Remark 4.3 Following the proof of Lemma 2.1 in [5], one can establish that, if $(V_{ij})_{i\geq 1, j\geq 1}$ is an infinite array of random variables such that $\{\Re(V_{ij}), \Im(V_{ij}), i \in \mathbb{N}, j \in \mathbb{N}\}$ are independent centered random variables which satisfy (4.1) and (4.2), then almost surely we have

$$\limsup_{N\to+\infty} \left\| \frac{Z_{n+N}}{\sqrt{N+n}} \right\| \leq 2\sigma^*$$

where

$$Z_{n+N} = \begin{pmatrix} (0) & V_N \\ V_N^* & (0) \end{pmatrix} \text{ with } V_N = [V_{ij}]_{1 \leq i \leq n, 1 \leq j \leq N} \text{ and } \sigma^* = \left\{ \sup_{(i,j)\in\mathbb{N}^2} \mathbb{E}(|V_{ij}|^2) \right\}^{1/2}.$$

Then, following the rest of the proof of Section 2 in [5], one can prove that for any polynomial P in $1+t$ noncommutative variables, there exists some constant $L > 0$ such that the following holds. Set $\theta^* = \sup_{i,j} \mathbb{E}\left(\left|X_{ij}\right|^3\right)$. For any $0 < \epsilon < 1$, there exist $C_\epsilon > 8\theta^*$ (such that $C_\epsilon > \max_{l=1,2} |W_{11}^{(l)}|$ a.s.) and $\delta_\epsilon > 0$ such that almost surely for all large N,

$$\left\| P\left(\frac{W_{n+N}}{\sqrt{n+N}}, B_{n+N}^{(1)}, \ldots, B_{n+N}^{(t)}\right) - P\left(\frac{\tilde{W}_{n+N}^{C_\epsilon,\delta_\epsilon}}{\sqrt{n+N}}, B_{n+N}^{(1)}, \ldots, B_{n+N}^{(t)}\right) \right\| \leq L\epsilon, \tag{4.89}$$

where, for any $C > 8\theta^*$ such that $C > \max_{l=1,2} |W_{11}^{(l)}|$ a.s., and for any $\delta > 0$, $\tilde{W}_{N+n}^{C,\delta}$ is a $(n+N) \times (n+N)$ matrix which is defined as follows. Let $(\mathscr{G}_{ij})_{i\geq 1, j\geq 1}$ be an infinite array which is independent of $\{X_{ij}, W_{ij}^{(1)}, W_{ij}^{(2)}, (i,j) \in \mathbb{N}^2\}$ and such that $\sqrt{2}\Re\mathscr{G}_{ij}$, $\sqrt{2}\Im\mathscr{G}_{ij}$, $i < j$, $\mathscr{G}_{ii}$, are independent centred standard real gaussian variables and $\mathscr{G}_{ij} = \overline{\mathscr{G}}_{ji}$. Set $\mathscr{G}_{n+N} = [\mathscr{G}_{ij}]_{1\leq i,j\leq n+N}$ and define $X_N^C = [X_{ij}^C]_{1 \leq i \leq n, 1 \leq j \leq N}$ as in (4.18). Set

$$\tilde{W}_{n+N}^C = \begin{pmatrix} W_n^{(1)} & X_N^C \\ (X_N^C)^* & W_N^{(2)} \end{pmatrix} \text{ and } \tilde{W}_{N+n}^{C,\delta} = \frac{\tilde{W}_{n+N}^C + \delta\mathscr{G}_{n+N}}{\sqrt{1+\delta^2}}.$$

$\tilde{W}_{N+n}^{C,\delta}$ satisfies (H) (see the end of Section 2 in [5]). (4.89) readily yields that it is sufficient to prove Theorem 4.5 for $\tilde{W}_{N+n}^{C,\delta}$.

Therefore, assume now that W_{N+n} satisfies (H). As explained in Section 6.2 in [5], to establish Theorem 4.5, it is sufficient to prove that for all $m \in \mathbb{N}$, all self-adjoint matrices $\gamma, \alpha, \beta_1, \ldots, \beta_t$ of size $m \times m$ and all $\epsilon > 0$, almost surely, for all large N, we have

$$spect(\gamma \otimes I_{n+N} + \alpha \otimes \frac{W_{n+N}}{\sqrt{n+N}} + \sum_{u=1}^{t} \beta_u \otimes B_{n+N}^{(u)})$$

$$\subset spect(\gamma \otimes 1_{\mathscr{A}} + \alpha \otimes s + \sum_{u=1}^{t} \beta_u \otimes b_{n+N}^{(u)}) +]-\epsilon, \epsilon[. \tag{4.90}$$

((4.90) is the analog of Lemma 1.3 for $r = 1$ in [5]). Finally, one can prove (4.90) by following Section 5 in [5].

We will need the following lemma in the proof of Theorem 4.4.

Lemma 4.16 *Let A_N and c_N be defined as in Theorem 4.4. Define the following $(n+N) \times (n+N)$ matrices: $P = \begin{pmatrix} I_n & (0) \\ (0) & (0) \end{pmatrix}$ $Q = \begin{pmatrix} (0) & (0) \\ (0) & I_N \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} (0) & A_N \\ (0) & (0) \end{pmatrix}$. Let $s, p_N, q_N, \boldsymbol{a}_N$ be noncommutative random variables in some $\mathscr{C}^*$-probability space $(\mathscr{A}, \tau)$ such that s is a standard semi-circular variable which is free with $(p_N, q_N, \boldsymbol{a}_N)$ and the $\star$-distribution of $(\mathbf{A}, P, Q)$ in $\left(M_{N+n}(\mathbb{C}), \frac{1}{N+n}\mathrm{Tr}\right)$ coincides with the $\star$-distribution of $(\boldsymbol{a}_N, p_N, q_N)$ in $(\mathscr{A}, \tau)$. Then, for any $\epsilon \geq 0$, the distribution of $(\sqrt{1+c_N}\sigma p_N s q_N + \sqrt{1+c_N}\sigma q_N s p_N + \boldsymbol{a}_N + \boldsymbol{a}_N^*)^2 + \epsilon p_N$ is $\frac{n}{N+n} T_\epsilon \star \mu_{\sigma, \mu_{A_N A_N^*}, c_N} + \frac{n}{N+n} \mu_{\sigma, \mu_{A_N A_N^*}, c_N} + \frac{N-n}{N+n} \delta_0$ where $T_\epsilon \star \mu_{\sigma, \mu_{A_N A_N^*}, c_N}$ is the pushforward of $\mu_{\sigma, \mu_{A_N A_N^*}, c_N}$ by the map $z \mapsto z + \epsilon$.*

Proof Here N and n are fixed. Let $k \geq 1$ and C_k be the $k \times k$ matrix defined by

$$C_k = \begin{pmatrix} (0) & & 1 \\ & \iddots & \\ 1 & & (0) \end{pmatrix}.$$

Define the $k(n+N) \times k(n+N)$ matrices

$$\hat{A}_k = C_k \otimes \mathbf{A}, \;\; \hat{P}_k = I_k \otimes P, \;\; \hat{Q}_k = I_k \otimes Q.$$

For any $k \geq 1$, the $\star$-distributions of $(\hat{A}_k, \hat{P}_k, \hat{Q}_k)$ in $(M_{k(N+n)}(\mathbb{C}), \frac{1}{k(N+n)}\mathrm{Tr})$ and $(\mathbf{A}, P, Q)$ in $(M_{(N+n)}(\mathbb{C}), \frac{1}{(N+n)}\mathrm{Tr})$ respectively, coincide. Indeed, let $\mathscr{K}$ be a noncommutative monomial in $\mathbb{C}\langle X_1, X_2, X_3, X_4\rangle$ and denote by q the total number of occurrences of X_3 and X_4 in $\mathscr{K}$. We have

$$\mathscr{K}(\hat{P}_k, \hat{Q}_k, \hat{A}_k, \hat{A}_k^*) = C_k^q \otimes \mathscr{K}(P, Q, \mathbf{A}, \mathbf{A}^*),$$

so that

$$\frac{1}{k(n+N)}\mathrm{Tr}\left[\mathscr{K}(\hat{P}_k,\hat{Q}_k,\hat{A}_k,\hat{A}_k^*)\right]=\frac{1}{k}\mathrm{Tr}(C_k^q)\frac{1}{(n+N)}\mathrm{Tr}\left[\mathscr{K}(P,Q,\mathbf{A},\mathbf{A}^*)\right].$$

Note that if q is even then $C_k^q = I_k$ so that

$$\frac{1}{k(n+N)}\mathrm{Tr}\left[\mathscr{K}(\hat{P}_k,\hat{Q}_k,\hat{A}_k,\hat{A}_k^*)\right]=\frac{1}{(n+N)}\mathrm{Tr}\left[\mathscr{K}(P,Q,\mathbf{A},\mathbf{A}^*)\right]. \tag{4.91}$$

Now, assume that q is odd. Note that $PQ = QP = 0$, $\mathbf{A}Q = \mathbf{A}$, $Q\mathbf{A} = 0$, $\mathbf{A}P = 0$ and $P\mathbf{A} = \mathbf{A}$ (and then $Q\mathbf{A}^* = \mathbf{A}^*$, $\mathbf{A}^*Q = 0$, $P\mathbf{A}^* = 0$ and $\mathbf{A}^*P = \mathbf{A}^*$). Therefore, if at least one of the terms X_1X_2, X_2X_1, X_2X_3, X_3X_1, X_4X_2 or X_1X_4 appears in the noncommutative product in $\mathscr{K}$, then $\mathscr{K}(P,Q,\mathbf{A},\mathbf{A}^*) = 0$, so that (4.91) still holds. Now, if none of the terms X_1X_2, X_2X_1, X_2X_3, X_3X_1, X_4X_2 or X_1X_4 appears in the noncommutative product in $\mathscr{K}$, then we have $\mathscr{K}(P,Q,\mathbf{A},\mathbf{A}^*) = \tilde{\mathscr{K}}(\mathbf{A},\mathbf{A}^*)$ for some noncommutative monomial $\tilde{\mathscr{K}} \in \mathbb{C}\langle X, Y\rangle$ with degree q. Either the noncommutative product in $\tilde{\mathscr{K}}$ contains a term such as X^p or Y^p for some $p \geq 2$ and then, since $\mathbf{A}^2 = (\mathbf{A}^*)^2 = 0$, we have $\tilde{\mathscr{K}}(\mathbf{A},\mathbf{A}^*) = 0$, or $\tilde{\mathscr{K}}(X,Y)$ is one of the monomials $(XY)^{\frac{q-1}{2}}X$ or $Y(XY)^{\frac{q-1}{2}}$. In both cases, we have $\mathrm{Tr}\tilde{\mathscr{K}}(\mathbf{A},\mathbf{A}^*) = 0$ and (4.91) still holds.

Now, define the $k(N+n)\times k(N+n)$ matrices

$$\tilde{P}_k=\begin{pmatrix} I_{kn} & (0)\\ (0) & (0)\end{pmatrix},\quad \tilde{Q}_k=\begin{pmatrix} (0) & (0)\\ (0) & I_{kN}\end{pmatrix},\quad \tilde{A}_k=\begin{pmatrix} (0) & \check{A}\\ (0) & (0)\end{pmatrix}$$

where $\check{A}$ is the $kn \times kN$ matrix defined by

$$\check{A}=\begin{pmatrix} (0) & & A_N\\ & \iddots & \\ A_N & & (0)\end{pmatrix}.$$

It is clear that there exists a real orthogonal $k(N+n)\times k(N+n)$ matrix O such that $\tilde{P}_k = O\hat{P}_kO^*$, $\tilde{Q}_k = O\hat{Q}_kO^*$ and $\tilde{A}_k = O\hat{A}_kO^*$. This readily yields that the noncommutative $\star$-distributions of $(\hat{A}_k,\hat{P}_k,\hat{Q}_k)$ and $(\tilde{A}_k,\tilde{P}_k,\tilde{Q}_k)$ in $(M_{k(N+n)}(\mathbb{C}), \frac{1}{k(N+n)}\mathrm{Tr})$ coincide. Hence, for any $k \geq 1$, the distribution of $(\tilde{A}_k,\tilde{P}_k,\tilde{Q}_k)$ in $(M_{k(N+n)}(\mathbb{C}), \frac{1}{k(N+n)}\mathrm{Tr})$ coincides with the distribution of $(\mathbf{a}_N, p_N, q_N)$ in $(\mathscr{A},\tau)$. By Theorem 5.4.5 in [1], it readily follows that the distribution of $(\sqrt{1+c_N}\sigma p_N s q_N + \sqrt{1+c_N}\sigma q_N s p_N + \mathbf{a}_N + \mathbf{a}_N^*)^2 + \epsilon p_N$ is the almost sure limiting distribution, when k goes to infinity, of $(\sqrt{1+c_N}\sigma \tilde{P}_k \frac{\mathscr{G}}{\sqrt{k(N+n)}}\tilde{Q}_k + \sqrt{1+c_N}\sigma \tilde{Q}_k \frac{\mathscr{G}}{\sqrt{k(N+n)}}\tilde{P}_k + \tilde{A}_k + \tilde{A}_k^*)^2 + \epsilon\tilde{P}_k$ in $(M_{k(N+n)}(\mathbb{C}), \frac{1}{k(N+n)}\mathrm{Tr})$, where

$\mathscr{G}$ is a $k(N+n) \times k(N+n)$ GUE matrix with entries with variance 1. Now, note that

$$\left[\sqrt{1+c_N}\sigma\left\{\tilde{P}_k \frac{\mathscr{G}}{\sqrt{k(N+n)}}\tilde{Q}_k + \tilde{Q}_k \frac{\mathscr{G}}{\sqrt{k(N+n)}}\tilde{P}_k\right\} + \tilde{A}_k + \tilde{A}_k^*\right]^2 + \epsilon \tilde{P}_k$$

$$= \begin{pmatrix} (\sigma \frac{\mathscr{G}_{kn\times kN}}{\sqrt{kN}} + \check{A})(\sigma \frac{\mathscr{G}_{kn\times kN}}{\sqrt{kN}} + \check{A})^* + \epsilon I_{kn} & (0) \\ (0) & (\sigma \frac{\mathscr{G}_{kn\times kN}}{\sqrt{kN}} + \check{A})^*(\sigma \frac{\mathscr{G}_{kn\times kN}}{\sqrt{kN}} + \check{A}) \end{pmatrix}$$

where $\mathscr{G}_{kn\times kN}$ is the upper right $kn \times kN$ corner of $\mathscr{G}$. Thus, noticing that $\mu_{\check{A}\check{A}^*} = \mu_{A_N A_N^*}$, the lemma follows from [15].

Proof of Theorem 4.4 Let W be a $(n+N) \times (n+N)$ matrix as defined by (4.87) in Theorem 4.5. Note that, with the notations of Lemma 4.16, for any $\epsilon \geq 0$,

$$\begin{pmatrix} (\sigma \frac{X_N}{\sqrt{N}} + A_N)(\sigma \frac{X_N}{\sqrt{N}} + A_N)^* + \epsilon I_n & (0) \\ (0) & (\sigma \frac{X_N}{\sqrt{N}} + A_N)^*(\sigma \frac{X_N}{\sqrt{N}} + A_N) \end{pmatrix}$$

$$= \begin{pmatrix} (0) & (\sigma \frac{X_N}{\sqrt{N}} + A_N) \\ (\sigma \frac{X_N}{\sqrt{N}} + A_N)^* & (0) \end{pmatrix}^2 + \epsilon P$$

$$= \left(\sqrt{1+c_N}P\frac{\sigma W}{\sqrt{N+n}}Q + \sqrt{1+c_N}Q\frac{\sigma W}{\sqrt{N+n}}P + \mathbf{A} + \mathbf{A}^*\right)^2 + \epsilon P.$$

Thus, for any $\epsilon \geq 0$,

$$\text{spect}\left\{(\sigma \frac{X_N}{\sqrt{N}} + A)(\sigma \frac{X_N}{\sqrt{N}} + A)^* + \epsilon I_n\right\}$$
$$\subset \text{spect}\left\{\left(\sqrt{1+c_N}P\frac{\sigma W}{\sqrt{N+n}}Q + \sqrt{1+c_N}Q\frac{\sigma W}{\sqrt{N+n}}P + \mathbf{A} + \mathbf{A}^*\right)^2 + \epsilon P\right\}. \tag{4.92}$$

Let $[x, y]$ be such that there exists $\delta > 0$ such that for all large N, $]x-\delta; y+\delta[\subset \mathbb{R} \setminus \text{supp}(\mu_{\sigma,\mu_{A_N A_N^*},c_N})$.

(i) Assume $x > 0$. Then, according to Lemma 4.16 with $\epsilon = 0$, there exists $\delta' > 0$ such that for all large n, $]x-\delta'; y+\delta'[$ is outside the support of the distribution of $(\sqrt{1+c_N}\sigma p_N s q_N + \sqrt{1+c_N}\sigma q_N s p_N + \mathbf{a}_N + \mathbf{a}_N^*)^2$. We readily deduce that almost surely for all large N, according to Theorem 4.5, there is no eigenvalue of $(\sqrt{1+c_N}P\frac{\sigma W}{\sqrt{N+n}}Q + \sqrt{1+c_N}Q\frac{\sigma W}{\sqrt{N+n}}P + \mathbf{A} + \mathbf{A}^*)^2$ in $[x, y]$. Hence, by (4.92) with $\epsilon = 0$, almost surely for all large N, there is no eigenvalue of M_N in $[x, y]$.

(ii) Assume $x = 0$ and $y > 0$. There exists $0 < \delta' < y$ such that $[0, 3\delta']$ is for all large N outside the support of $\mu_{\sigma,\mu_{A_N A_N^*},c_N}$. Hence, according to Lemma 4.16, $[\delta'/2, 3\delta']$ is outside the support of the distribution of $(\sqrt{1+c_N}\sigma p_N s q_N + \sqrt{1+c_N}\sigma q_N s p_N + \mathbf{a}_N + \mathbf{a}_N^*)^2 + \delta' p_N$. Then, almost surely for all large N, according to Theorem 4.5, there is no eigenvalue of $(\sqrt{1+c_N} P \frac{\sigma W}{\sqrt{N+n}} Q + \sqrt{1+c_N} Q \frac{\sigma W}{\sqrt{N+n}} P + \mathbf{A} + \mathbf{A}^*)^2 + \delta' P$ in $[\delta', 2\delta']$ and thus, by (4.92), no eigenvalue of $(\sigma \frac{X}{\sqrt{N}} + A_N)(\sigma \frac{X_N}{\sqrt{N}} + A_N)^* + \delta' I_n$ in $[\delta', 2\delta']$. It readily follows that, almost surely for all large N, there is no eigenvalue of $(\sigma \frac{X_N}{\sqrt{N}} + A_N)(\sigma \frac{X_N}{\sqrt{N}} + A_N)^*$ in $[0, \delta']$. Since moreover, according to (i), almost surely for all large N, there is no eigenvalue of $(\sigma \frac{X_N}{\sqrt{N}} + A_N)(\sigma \frac{X_N}{\sqrt{N}} + A_N)^*$ in $[\delta', y]$, we can conclude that there is no eigenvalue of M_N in $[x, y]$.

The proof of Theorem 4.4 is now complete. □

We are now in a position to establish the following exact separation phenomenon.

Theorem 4.6 *Let M_n as in (4.84) with assumptions [1–4] of Theorem 4.4. Assume moreover that the empirical spectral measure $\mu_{A_N A_N^*}$ of $A_N A_N^*$ converges weakly to some probability measure ν. Then for N large enough,*

$$\omega_{\sigma,\nu,c}([x, y]) = [\omega_{\sigma,\nu,c}(x); \omega_{\sigma,\nu,c}(y)] \subset \mathbb{R} \setminus supp(\mu_{A_N A_N^*}), \tag{4.93}$$

where $\omega_{\sigma,\nu,c}$ is defined in (4.5). With the convention that $\lambda_0(M_N) = \lambda_0(A_N A_N^*) = +\infty$ and $\lambda_{n+1}(M_N) = \lambda_{n+1}(A_N A_N^*) = -\infty$, for N large enough, let $i_N \in \{0, \ldots, n\}$ be such that

$$\lambda_{i_N+1}(A_N A_N^*) < \omega_{\sigma,\nu,c}(x) \quad \textit{and} \quad \lambda_{i_N}(A_N A_N^*) > \omega_{\sigma,\nu,c}(y). \tag{4.94}$$

Then

$$P[\textit{for all large } N, \lambda_{i_N+1}(M_N) < x \textit{ and } \lambda_{i_N}(M_N) > y] = 1. \tag{4.95}$$

Remark 4.4 Since $\mu_{\sigma,\mu_{A_N A_N^*},c_N}$ converges weakly towards $\mu_{\sigma,\nu,c}$ assumption 4. implies that $\forall 0 < \tau < \delta$, $[x - \tau; y + \tau] \subset \mathbb{R} \setminus \operatorname{supp} \mu_{\sigma,\nu,c}$.

Proof (4.93) is proved in Lemma 3.1 in [11].

- If $\omega_{\sigma,\nu,c}(x) < 0$, then $i_N = n$ in (4.94) and moreover we have, for all large N, $\omega_{\sigma,\mu_{A_N A_N^*},c_N}(x) < 0$. According to Lemma 2.7 in [11], we can deduce that, for all large N, $[x, y]$ is on the left hand side of the support of $\mu_{\sigma,\mu_{A_N A_N^*},c_N}$ so that $]-\infty; y + \delta]$ is on the left hand side of the support of $\mu_{\sigma,\mu_{A_N A_N^*},c_N}$. Since $[-|y| - 1, y]$ satisfies the assumptions of Theorem 4.4, we readily deduce that almost surely, for all large N, $\lambda_n(M_N) > y$. Hence (4.95) holds true.

- If $\omega_{\sigma,\nu,c}(x) \geq 0$, we first explain why it is sufficient to prove (4.95) for x such that $\omega_{\sigma,\nu,c}(x) > 0$. Indeed, assume for a while that (4.95) is true whenever $\omega_{\sigma,\nu,c}(x) > 0$. Let us consider any interval $[x, y]$ satisfying condition 4. of Theorem 4.4 and such that $\omega_{\sigma,\nu,c}(x) = 0$; then $i_N = n$ in (4.94). According to Proposition 4.1, $\omega_{\sigma,\nu,c}(\frac{x+y}{2}) > 0$ and then almost surely for all large N, $\lambda_n(M_N) > y$. Finally, sticking to the proof of Theorem 1.2 in [11] leads to (4.95) for x such that $\omega_{\sigma,\nu,c}(x) > 0$.

Appendix 2

We first recall some basic properties of the resolvent (see [12, 22]).

Lemma 4.17 *For a $N \times N$ Hermitian matrix M, for any $z \in \mathbb{C} \setminus \mathrm{spect}(M)$, we denote by $G(z) := (zI_N - M)^{-1}$ the resolvent of M.*

Let $z \in \mathbb{C} \setminus \mathbb{R}$,

(i) $\|G(z)\| \leq |\Im z|^{-1}$.
(ii) $|G(z)_{ij}| \leq |\Im z|^{-1}$ *for all* $i, j = 1, \ldots, N$.
(iii) $G(z)M = MG(z) = -I_N + zG(z)$.

Moreover, for any $N \times N$ Hermitian matrices M_1 and M_2,

$$(zI_N - M_1)^{-1} - (zI_N - M_2)^{-1} = (zI_N - M_1)^{-1}(M_1 - M_2)(zI_N - M_2)^{-1}.$$

The following technical lemmas are fundamental in the approach of the present paper.

Lemma 4.18 (Lemma 4.4 in [6]) *Let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function with compact support. Let B_N be a $N \times N$ Hermitian matrix and C_N be a $N \times N$ matrix. Then*

$$\mathrm{Tr}\left[h(B_N)C_N\right] = - \lim_{y \to 0^+} \frac{1}{\pi} \int \Im \mathrm{Tr}\left[(t + \mathrm{i}y - B_N)^{-1} C_N\right] h(t)dt. \tag{4.96}$$

Moreover, if B_N is random, we also have

$$\mathbb{E}\mathrm{Tr}\left[h(B_N)C_N\right] = - \lim_{y \to 0^+} \frac{1}{\pi} \int \Im \mathbb{E}\mathrm{Tr}\left[(t + \mathrm{i}y - B_N)^{-1} C_N\right] h(t)dt. \tag{4.97}$$

Lemma 4.19 *Let f be an analytic function on $\mathbb{C} \setminus \mathbb{R}$ such that there exist some polynomial P with nonnegative coefficients, and some positive real number α such that*

$$\forall z \in \mathbb{C} \setminus \mathbb{R}, \quad |f(z)| \leq (|z| + 1)^{\alpha} P(|\Im z|^{-1}).$$

Then, for any h in $\mathscr{C}^\infty(\mathbb{R},\mathbb{R})$ with compact support, there exists some constant τ depending only on h, α and P such that

$$\limsup_{y\to 0^+} \left| \int_{\mathbb{R}} h(x) f(x+\mathrm{i}y)dx \right| < \tau.$$

We refer the reader to the Appendix of [12] where it is proved using the ideas of [20].

Finally, we recall some facts on Poincaré inequality. A probability measure μ on $\mathbb{R}$ is said to satisfy the Poincaré inequality with constant C_{PI} if for any $\mathscr{C}^1$ function $f:\mathbb{R}\to\mathbb{C}$ such that f and f' are in $L^2(\mu)$,

$$\mathbf{V}(f) \leq C_{PI} \int |f'|^2 d\mu,$$

with $\mathbf{V}(f) = \int |f - \int f d\mu|^2 d\mu$.

We refer the reader to [9] for a characterization of the measures on $\mathbb{R}$ which satisfy a Poincaré inequality.

If the law of a random variable X satisfies the Poincaré inequality with constant C_{PI} then, for any fixed $\alpha \neq 0$, the law of αX satisfies the Poincaré inequality with constant $\alpha^2 C_{PI}$.

Assume that probability measures $\mu_1,\ldots,\mu_M$ on $\mathbb{R}$ satisfy the Poincaré inequality with constant $C_{PI}(1),\ldots,C_{PI}(M)$ respectively. Then the product measure $\mu_1\otimes\cdots\otimes\mu_M$ on $\mathbb{R}^M$ satisfies the Poincaré inequality with constant $C^*_{PI} = \max_{i\in\{1,\ldots,M\}} C_{PI}(i)$ in the sense that for any differentiable function f such that f and its gradient grad$f$ are in $L^2(\mu_1\otimes\cdots\otimes\mu_M)$,

$$\mathbf{V}(f) \leq C^*_{PI} \int \|\mathrm{grad} f\|_2^2 d\mu_1\otimes\cdots\otimes\mu_M$$

with $\mathbf{V}(f) = \int |f - \int f d\mu_1\otimes\cdots\otimes\mu_M|^2 d\mu_1\otimes\cdots\otimes\mu_M$ (see Theorem 2.5 in [18]) .

Lemma 4.20 (Theorem 1.2 in [4]) *Assume that the distribution of a random variable X is supported in $[-C; C]$ for some constant $C > 0$. Let g be an independent standard real Gaussian random variable. Then $X+\delta g$ satisfies a Poincaré inequality with constant $C_{PI} \leq \delta^2 \exp\left(4C^2/\delta^2\right)$.*

References

1. G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press, Cambridge, 2009)
2. Z.D. Bai, J.W. Silverstein, *Spectral Analysis of Large-Dimensional Random Matrices*. Mathematics Monograph Series, vol. 2 (Science Press, Beijing, 2006)
3. Z. Bai, J.W. Silverstein, No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. Random Matrices Theory Appl. **1**(1), 1150004, 44 (2012)
4. J.B. Bardet, N. Gozlan, F. Malrieu, P.-A. Zitt, Functional inequalities for Gaussian convolutions of compactly supported measures: explicit bounds and dimension dependence (2018). Bernoulli **24**(1), 333–353 (2018)
5. S.T. Belinschi, M. Capitaine, Spectral properties of polynomials in independent Wigner and deterministic matrices. J. Funct. Anal. **273**, 3901–3963 (2017). https://doi.org/10.1016/j.jfa.2017.07.010
6. S.T. Belinschi, H. Bercovici, M. Capitaine, M. Février, Outliers in the spectrum of large deformed unitarily invariant models. Ann. Probab. **45**(6A), 3571–3625 (2017)
7. F. Benaych-Georges, R.N. Rao, The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. Adv. Math. **227**(1), 494–521 (2011)
8. F. Benaych-Georges, R.N. Rao, The singular values and vectors of low rank perturbations of large rectangular random matrices (2011). J. Multivar. Anal. (111), 120–135 (2012)
9. S.G. Bobkov, F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal. **163**(1), 1–28 (1999)
10. M. Capitaine, Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices. J. Theor. Probab. **26**(3), 595–648 (2013)
11. M. Capitaine, Exact separation phenomenon for the eigenvalues of large Information-Plus-Noise type matrices. Application to spiked models. Indiana Univ. Math. J. **63**(6), 1875–1910 (2014)
12. M. Capitaine, C. Donati-Martin, Strong asymptotic freeness for Wigner and Wishart matrices. Indiana Univ. Math. J. **56**(2), 767–803 (2007)
13. F. Benaych-Georges, C. Bordenave, M. Capitaine, C. Donati-Martin, A. Knowles, Spectrum of deformed random matrices and free probability, in *Advanced Topics in Random Matrices*, ed. by F. Benaych-Georges, D. Chafaï, S. Péché, B. de Tiliére. Panoramas et syntheses, vol. 53 (2018)
14. R. Couillet, J.W. Silverstein, Z. Bai, M. Debbah, Eigen-inference for energy estimation of multiple sources. IEEE Trans. Inf. Theory **57**(4), 2420–2439 (2011)
15. R.B. Dozier, J.W. Silverstein, On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. J. Multivar. Anal. **98**(4), 678–694 (2007)
16. R.B. Dozier, J.W. Silverstein, Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. J. Multivar. Anal. **98**(6), 1099–1122 (2007)
17. J. Dumont, W. Hachem, S. Lasaulce, Ph. Loubaton, J. Najim, On the capacity achieving covariance matrix for Rician MIMO channels: an asymptotic approach. IEEE Trans. Inf. Theory **56**(3), 1048–1069 (2010)
18. A. Guionnet, B. Zegarlinski, Lectures on logarithmic Sobolev inequalities, in *Séminaire de Probabilités, XXXVI*. Lecture Notes in Mathematics, vol. 1801 (Springer, Berlin, 2003)
19. U. Haagerup, S. Thorbjørnsen, Random matrices with complex Gaussian entries. Expo. Math. **21**, 293–337 (2003)
20. U. Haagerup, S. Thorbjørnsen, A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. Ann. Math. (2) **162**(2), 711–775 (2005)
21. W. Hachem, P. Loubaton, J. Najim, Deterministic Equivalents for certain functionals of large random matrices. Ann. Appl. Probab. **17**(3), 875–930 (2007)

22. A.M. Khorunzhy, B.A. Khoruzhenko, L.A. Pastur, Asymptotic properties of large random matrices with independent entries. J. Math. Phys. **37**(10), 5033–5060 (1996)
23. O. Ledoit, S. Péché, Eigenvectors of some large sample covariance matrix ensembles. Probab. Theory Relat. Fields **151**, 233–264 (2011)
24. M. Ledoux, *The Concentration of Measure Phenomenon* (American Mathematical Society, Providence, 2001)
25. P. Loubaton, P. Vallet, Almost sure localization of the eigenvalues in a Gaussian information-plus-noise model. Application to the spiked models. Electron. J. Probab. **16**, 1934–1959 (2011)
26. L.A. Pastur, M. Shcherbina, *Eigenvalue Distribution of Large Random Matrices*. Mathematical Surveys and Monographs (American Mathematical Society, Providence, 2011)
27. D. Paul, Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. Stat. Sin. **17**(4), 1617–1642 (2007)
28. P. Vallet, P. Loubaton, X. Mestre, Improved subspace estimation for multivariate observations of high dimension: the deterministic signal case. IEEE Trans. Inf. Theory **58**(2), 1043–1068 (2012)
29. D.V. Voiculescu, K. Dykema, A. Nica, *Free Random Variables: A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras and Harmonic Analysis on Free Groups*. CRM Monograph Series, vol. 1 (American Mathematical Society, Providence, 1992). ISBN 0-8218-6999-X
30. J.-s. Xie, The convergence on spectrum of sample covariance matrices for information-plus-noise type data. Appl. Math. J. Chinese Univ. Ser. B **27**(2), 181191 (2012)

Chapter 5
Criteria for Exponential Convergence to Quasi-Stationary Distributions and Applications to Multi-Dimensional Diffusions

Nicolas Champagnat, Koléhè Abdoulaye Coulibaly-Pasquier, and Denis Villemonais

Abstract We consider general Markov processes with absorption and provide criteria ensuring the exponential convergence in total variation of the distribution of the process conditioned not to be absorbed. The first one is based on two-sided estimates on the transition kernel of the process and the second one on gradient estimates on its semigroup. We apply these criteria to multi-dimensional diffusion processes in bounded domains of $\mathbb{R}^d$ or in compact Riemannian manifolds with boundary, with absorption at the boundary.

Keywords Markov processes · Diffusions in Riemannian manifolds · Diffusions in bounded domains · Absorption at the boundary · Quasi-stationary distributions · Q-process · Uniform exponential mixing · Two-sided estimates · Gradient estimates

2010 Mathematics Subject Classification Primary: 60J60, 37A25, 60B10, 60F99; Secondary: 60J75, 60J70

5.1 Introduction

Let X be a Markov process evolving in a measurable state space $E \cup \{\partial\}$ absorbed at $\partial \notin E$ at time $\tau_\partial = \inf\{t \geq 0,\ X_t = \partial\}$. We assume that $\mathbb{P}_x(t < \tau_\partial) > 0$, for all $x \in E$ and all $t \geq 0$, where $\mathbb{P}_x$ is the law of X with initial position x. We consider

N. Champagnat (✉) · K. A. Coulibaly-Pasquier · D. Villemonais
Université de Lorraine, IECL, UMR 7502, Vandœuvre-lès-Nancy, France

CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, France

Inria, Villers-lès-Nancy, France
e-mail: Nicolas.Champagnat@inria.fr; Kolehe.Coulibaly@univ-lorraine.fr; Denis.Villemonais@univ-lorraine.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_5

the problem of existence of a probability measure α on E and of positive constants $B, \gamma > 0$ such that, for all initial distribution μ on E,

$$\left\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \alpha(\cdot)\right\|_{TV} \leq Be^{-\gamma t}, \quad \forall t \geq 0, \tag{5.1}$$

where $\mathbb{P}_\mu$ is the law of X with initial distribution μ and $\|\cdot\|_{TV}$ is the total variation norm on finite signed measures. It is well known that (5.1) entails that α is the unique quasi-stationary distribution for X, that is the unique probability measure satisfying

$$\alpha(\cdot) = \mathbb{P}_\alpha\left(X_t \in \cdot \mid t < \tau_\partial\right), \quad \forall t \geq 0.$$

Our goal is to provide sufficient conditions for (5.1) with applications when X is a diffusion process, absorbed at the boundary of a domain of $\mathbb{R}^d$ or of a Riemannian manifold. Our first result (Theorem 5.1) shows that a two-sided estimate for the transition kernel of a general absorbed Markov process is sufficient to ensure (5.1). This criterion applies in particular to diffusions with smooth coefficients in bounded domains of $\mathbb{R}^d$ with irregular boundary. Our second result (Theorem 5.2) concerns Markov processes satisfying gradient estimates (as in Wang [29] and Priola and Wang [26]), irreducibility conditions and controlled probability of absorption near the boundary. It applies to diffusions with less regular coefficients in smooth domains of $\mathbb{R}^d$ and to drifted Brownian motions in compact Riemannian manifolds with C^2 boundary.

Convergence of conditioned diffusion processes have been already obtained for diffusions in domains of $\mathbb{R}^d$, mainly using spectral theoretic arguments (see for instance [4, 5, 14, 19, 23, 24] for $d = 1$ and [3, 12, 18] for $d \geq 2$). Among these references, [12, 18] give the most general criteria for diffusions in dimension 2 or more. Using two-sided estimates and spectral properties of the infinitesimal generator of X, Knobloch and Partzsch [18] proved that (5.1) holds for a class of diffusion processes evolving in $\mathbb{R}^d$ ($d \geq 3$) with C^1 diffusion coefficient, drift in a Kato class and $C^{1,1}$ domain. In [12], the authors obtain (5.1) for diffusions with global Lipschitz coefficients (and additional local regularity near the boundary) in a domain with C^2 boundary. These results can be recovered with our method (see Sects. 5.2 and 5.3.2 respectively). When the diffusion is a drifted Brownian motion with drift deriving from a potential, the authors of [3] obtain existence and uniqueness results for the quasi-stationary distribution in cases with singular drifts and unbounded domains with non-regular boundary that do not enter the settings of this paper.

Usual tools to prove convergence in total variation for processes without absorption involve coupling arguments: for example, contraction in total variation norm for the non-conditioned semi-group can be obtained using mirror and parallel coupling, see [22, 26, 29], or lower bounds on the density of the process that could be obtained for example using Aronson-type estimates or Malliavin calculus [1, 25, 28, 30]. However, on the one hand, lower bounds on transition densities are not sufficient to control conditional distributions, and on the other hand, the process conditioned not to be killed up to a given time $t > 0$ is a time-inhomogeneous diffusion process with a singular drift for which these methods fail. For instance, a standard d-dimensional Brownian motion $(B_t)_{t\geq 0}$ conditioned not to exit a smooth domain $D\subset\mathbb{R}^d$ up to a time $t>0$ has the law of the solution $(X^{(t)}_s)_{s\in[0,t]}$ to the stochastic differential equation

$$dX^{(t)}_s = dB_s + [\nabla \ln \mathbb{P}_\cdot(t-s<\tau_\partial)]\,(X^{(t)}_s)ds.$$

where the drift term is singular near the boundary. Our approach is thus to use the following condition, which is actually equivalent to the exponential convergence (5.1) (see [6, Theorem 2.1]).

Condition (A)

There exist $t_0, c_1, c_2>0$ and a probability measure ν on E such that

(A1) for all $x\in E$,

$$\mathbb{P}_x(X_{t_0}\in\cdot\mid t_0<\tau_\partial)\geq c_1\nu(\cdot)$$

(A2) for all $z\in E$ and all $t\geq 0$,

$$\mathbb{P}_\nu(t<\tau_\partial)\geq c_2\mathbb{P}_z(t<\tau_\partial).$$

More precisely, if Condition (A) is satisfied, then, for all probability measure π on E,

$$\|\mathbb{P}_\pi(X_t\in\cdot\mid t<\tau_\partial)-\alpha(\cdot)\|_{TV}\leq 2(1-c_1c_2)^{\lfloor t/t_0\rfloor}$$

and it implies that, for all probability measures π_1 and π_2 on E,

$$\begin{aligned}&\left\|\mathbb{P}_{\pi_1}(X_t\in\cdot\mid t<\tau_\partial)-\mathbb{P}_{\pi_2}(X_t\in\cdot\mid t<\tau_\partial)\right\|_{TV}\\&\quad\leq\frac{(1-c_1c_2)^{\lfloor t/t_0\rfloor}}{c(\pi_1)\vee c(\pi_2)}\|\pi_1-\pi_2\|_{TV},\end{aligned}\tag{5.2}$$

where $c(\pi_i)=\inf_{t\geq 0}\mathbb{P}_{\pi_i}(t<\tau_\partial)/\sup_{z\in E}\mathbb{P}_z(t<\tau_\partial)$ (see Appendix for a proof of this improvement of [6, Corollary 2.2], where the same inequality is obtained with $c(\pi_1)\wedge c(\pi_2)$ instead of $c(\pi_1)\vee c(\pi_2)$).

Several other properties can also be deduced from Condition (A). For instance, $e^{\lambda_0 t}\mathbb{P}_x(t<\tau_\partial)$ converges when $t\to+\infty$, uniformly in x, to a positive eigenfunction η of the infinitesimal generator of $(X_t, t\geq 0)$ for the eigenvalue $-\lambda_0$ characterized by the relation $\mathbb{P}_\alpha(t<\tau_\partial)=e^{-\lambda_0 t}$, $\forall t\geq 0$ [6, Proposition 2.3]. Moreover, it implies a spectral gap property [6, Corollary 2.4], the existence and exponential ergodicity of the so-called Q-process, defined as the process X conditioned to never hit the boundary [6, Theorem 3.1] and a conditional ergodic property [7]. Note that we do not assume that $\mathbb{P}_x(\tau_\partial<\infty)=1$, which is only required in the proofs of [6] in order to obtain $\lambda_0>0$. Indeed, the above inequalities remain true under Condition (A), even if $\mathbb{P}_x(\tau_\partial < +\infty) < 1$ for some $x \in E$. The only difference is that, in this case, $E' := \{x \in E,\ \mathbb{P}_x(\tau_\partial < +\infty) = 0\}$ is non-empty, α is a classical stationary distribution such that $\alpha(E') = 1$ and $\lambda_0 = 0$.

The paper is organized as follows. In Sect. 5.2, we state and prove a sufficient criterion for (5.1) based on two-sided estimates. In Sect. 5.3.1, we prove (5.1) for Markov processes satisfying gradient estimates, irreducibility conditions and controlled probability of absorption near the boundary. In Sect. 5.3.2, we apply this result to diffusions in smooth domains of $\mathbb{R}^d$ and to drifted Brownian motions in compact Riemannian manifolds with smooth boundary. Section 5.3.3 is devoted to the proof of the criterion of Sect. 5.3.1. Finally, Appendix gives the proof of (5.2).

5.2 Quasi-Stationary Behavior Under Two-Sided Estimates

In this section, we consider as in the introduction a general absorbed Markov process X in $E \cup \{\partial\}$ satisfying two-sided estimates: there exist a time $t_0 > 0$, a constant $c > 0$, a positive measure μ on E and a measurable function $f : E \to (0, +\infty)$ such that

$$c^{-1} f(x)\mu(\cdot) \le \mathbb{P}_x(X_{t_0} \in \cdot) \le c f(x)\mu(\cdot),\ \forall x \in E. \tag{5.3}$$

Note that this implies that $f(x)\mu(E) \le c$ for all $x \in E$, hence μ is finite and f is bounded. As a consequence, one can assume without loss of generality that μ is a probability measure and then $\|f\|_\infty \le c$. Note also that $f(x) > 0$ for all $x \in E$ entails that $\mathbb{P}_x(t_0 < \tau_\partial) > 0$ for all $x \in E$ and hence, by Markov property, that $\mathbb{P}_x(t < \tau_\partial) > 0$ for all $x \in E$ and all $t > 0$, as needed to deduce (5.1) from Condition (A) (see [6]).

Estimates of the form (5.3) are well known for diffusion processes in a bounded domain of $\mathbb{R}^d$ since the seminal paper of Davies and Simon [11]. The case of standard Brownian motion in a bounded $C^{1,1}$ domain of $\mathbb{R}^d$, $d \ge 3$ was studied in [31]. This result has then been extended in [17] to diffusions in a bounded $C^{1,1}$ domain in $\mathbb{R}^d$, $d \ge 3$, with infinitesimal generator

$$L = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_i\partial_j + \sum_{i=1}^{d} b_i\partial_i,$$

with symmetric, uniformly elliptic and C^1 diffusion matrix $(a_{ij})_{1\le i,j\le d}$, and with drift $(b_i)_{1\le i\le d}$ in the Kato class $K_{d,1}$, which contains $L^p(dx)$ functions for $p > d$. Diffusions on bounded, closed Riemannian manifolds with irregular boundary and with generator

$$L = \Delta + X,$$

where Δ is the Laplace-Beltrami operator and X is a smooth vector field, were also studied in [21]. Two-sided estimates are also known for processes with jumps [2, 8–10, 16].

Theorem 5.1 *Assume that there exist a time $t_0 > 0$, a constant $c > 0$, a probability measure μ on E and a measurable function $f : E \to (0, +\infty)$ such that (5.3) holds. Then Condition (A) is satisfied with $\nu = \mu$, $c_1 = c^{-2}$ and $c_2 = c^{-3}\mu(f)$. In addition, for all probability measures π_1 and π_2 on E, we have*

$$\begin{aligned}\left\|\mathbb{P}_{\pi_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\pi_2}(X_t \in \cdot \mid t < \tau_\partial)\right\|_{TV} \\ \le c^3 \frac{(1 - c^{-5}\mu(f))^{\lfloor t/t_0 \rfloor}}{\pi_1(f) \vee \pi_2(f)} \|\pi_1 - \pi_2\|_{TV}, \end{aligned} \tag{5.4}$$

Moreover, the unique quasi-stationary distribution α for X satisfies

$$c^{-2}\mu \le \alpha \le c^2 \mu. \tag{5.5}$$

Remark 5.1 Recall that to any quasi-stationary distribution α is associated an eigenvalue $-\lambda_0 \le 0$. We deduce from the two-sided estimate (5.3) and [6, Corollary 2.4] an explicit estimate on the second spectral gap of the infinitesimal generator L of X (defined as acting on bounded measurable functions on $E \cup \{\partial\}$): for all λ in the spectrum of L such that $\lambda \notin \{0, \lambda_0\}$, the real part of λ is smaller than $-\lambda_0 + t_0^{-1} \log(1 - c^{-5}\mu(f))$.

Remark 5.2 In particular, we recover the results of Knobloch and Partzsch [18]. They proved that (5.1) holds for a class of diffusion processes evolving in $\mathbb{R}^d$ ($d \ge 3$), assuming continuity of the transition density, existence of ground states and the existence of a two-sided estimate involving the ground states of the generator. Similar results were obtained in the one-dimensional case in [24].

Proof (Proof of Theorem 5.1) We deduce from (5.3) that, for all $x \in E$,

$$c^{-2}\,\mu(\cdot) \le \mathbb{P}_x(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) = \frac{\mathbb{P}_x(X_{t_0} \in \cdot)}{\mathbb{P}_x(X_{t_0} \in E)} \le c^2\mu(\cdot). \tag{5.6}$$

We thus obtain (A1) with $c_1 = c^{-2}$ and $\nu = \mu$.

Moreover, for any probability measure π on E and any $z \in E$,

$$\begin{aligned}\mathbb{P}_\pi(X_{t_0} \in \cdot) &\ge c^{-1}\pi(f)\mu(\cdot) \\ &\ge \frac{f(z)}{\|f\|_\infty} c^{-1}\pi(f)\mu(\cdot) \\ &\ge c^{-3}\pi(f)\mathbb{P}_z(X_{t_0} \in \cdot).\end{aligned}$$

Hence, for all $t \geq t_0$, we have by Markov's property

$$\begin{aligned}
\mathbb{P}_\pi(t < \tau_\partial) &= \mathbb{E}_\pi\left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial)\right) \\
&\geq c^{-3}\pi(f)\,\mathbb{E}_z\left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial)\right) \\
&= c^{-3}\pi(f)\,\mathbb{P}_z(t < \tau_\partial).
\end{aligned}$$

When $t \leq t_0$, we have $\mathbb{P}_\pi(t < \tau_\partial) \geq \mathbb{P}_\pi(t_0 < \tau_\partial) \geq c^{-1}\pi(f) \geq c^{-3}\pi(f)$ and hence $\mathbb{P}_\pi(t < \tau_\partial) \geq c^{-3}\pi(f)\mathbb{P}_z(t < \tau_\partial)$, so that

$$c(\pi) := \inf_{t \geq 0} \frac{\mathbb{P}_\pi(t < \tau_\partial)}{\sup_{z \in E} \mathbb{P}_z(t < \tau_\partial)} \geq c^{-3}\pi(f).$$

Taking $\pi = \nu = \mu$, this entails (A2) for $c_2 = c^{-3}\mu(f)$ and (5.2) implies (5.4). The inequality (5.5) then follows from (5.6).

5.3 Quasi-Stationary Behavior Under Gradient Estimates

In this section, we explain how gradient estimates on the semi-group of the Markov process $(X_t, t \geq 0)$ imply the exponential convergence (5.1).

5.3.1 A General Result

We assume that the process X is a strong Markov, continuous[1] process and we assume that its state space $E \cup \{\partial\}$ is a compact metric space with metric ρ equipped with its Borel σ-field. Recall that ∂ is absorbing and that we assume that $\mathbb{P}_x(t < \tau_\partial) > 0$ for all $x \in E$ and $t \geq 0$. Our result holds true under three conditions: first, we assume that there exists $t_1 > 0$ such that the process satisfies a gradient estimate of the form: for all bounded measurable function $f : E \cup \{\partial\} \to \mathbb{R}$

$$\|\nabla P_{t_1} f\|_\infty \leq C\|f\|_\infty, \tag{5.7}$$

[1]The assumption of continuity is only used to ensure that the entrance times in compact sets are stopping times for the natural filtration (cf. e.g. [20, p. 48]), and hence that the strong Markov property applies at this time. Our result would also hold true for càdlàg (weak) Markov processes provided that the strong Markov property applies at the hitting times of compact sets.

where $P_t f(x) = \mathbb{E}_x(f(X_t)\mathbb{1}_{t<\tau_\partial})$ denotes the Dirichlet semi-group of X and the (a bit informal in such a general setting) notation $\|\nabla P_{t_1} f\|_\infty$ has to be understood as

$$\|\nabla P_{t_1} f\|_\infty := \sup_{x,y\in E\cup\{\partial\}} \frac{|P_{t_1} f(x) - P_{t_1} f(y)|}{\rho(x,y)}.$$

Second, we assume that there exist a compact subset K of E and a constant $C' > 0$ such that, for all $x \in E$,

$$\mathbb{P}_x(T_K \le t_1 < \tau_\partial) \ge C' \rho_\partial(x), \tag{5.8}$$

where $\rho_\partial(x) := \rho(x, \partial)$ and $T_K = \inf\{t \ge 0,\ X_t \in K\}$. Finally, we need the following irreducibility condition: for all $x, y \in E$ and all $r > 0$,

$$\mathbb{P}_x(X_s \in B(y,r),\ \forall s \in [t_1, 2t_1]) > 0, \tag{5.9}$$

where $B(y, r)$ denotes the ball of radius r centered at y.

Theorem 5.2 *Assume that the process* $(X_t, t \ge 0)$ *satisfies* (5.7)*,* (5.8) *and* (5.9) *for some constant* $t_1 > 0$*. Then Condition (A) and hence* (5.1) *are satisfied. Moreover, there exist two constants* $B, \gamma > 0$ *such that, for any initial distributions* μ_1 *and* μ_2 *on E,*

$$\begin{aligned} &\left\|\mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial)\right\|_{TV} \\ &\quad \le \frac{Be^{-\gamma t}}{\mu_1(\rho_\partial) \vee \mu_2(\rho_\partial)} \|\mu_1 - \mu_2\|_{TV}. \end{aligned} \tag{5.10}$$

The proof of this result is given in Sect. 5.3.3.

5.3.2 *The Case of Diffusions in Compact Riemannian Manifolds*

In this section, we provide two sets of assumptions for diffusions in compact manifolds with boundary M absorbed at the boundary ∂M (i.e. $E = M \setminus \partial M$ and $\partial = \{\partial M\}$) to which the last theorem applies:

S1. M is a bounded, connected and closed C^2 Riemannian manifold with C^2 boundary ∂M and the infinitesimal generator of the diffusion process X is given by $L = \frac{1}{2}\Delta + Z$, where Δ is the Laplace-Beltrami operator and Z is a C^1 vector field.

S2. M is a compact subset of $\mathbb{R}^d$ with non-empty, connected interior and C^2 boundary ∂M and X is solution to the SDE $dX_t = s(X_t)dB_t + b(X_t)dt$, where $(B_t, t \ge 0)$ is a r-dimensional Brownian motion, $b : M \to \mathbb{R}^d$ is bounded and continuous and $s : M \to \mathbb{R}^{d\times r}$ is continuous, ss^* is uniformly elliptic and for all $r > 0$,

$$\sup_{x,y\in M,\ |x-y|=r} \frac{|s(x)-s(y)|^2}{r} \le g(r) \tag{5.11}$$

for some function g such that $\int_0^1 g(r)dr < \infty$.

Note that (5.11) is satisfied as soon as s is uniformly α-Hölder on M for some $\alpha > 0$.

Let us now check that Theorem 5.2 applies in both situations.

First, the gradient estimate (5.7) is satisfied (see Wang in [29] and Priola and Wang in [26], respectively). These two references actually give a stronger version of (5.7):

$$\|\nabla P_t f\|_\infty \le \frac{c}{1 \wedge \sqrt{t}} \|f\|_\infty, \quad \forall t > 0. \tag{5.12}$$

The set of assumptions S2 is not exactly the same as in [26], but they clearly imply (i), (ii), (iv) of [26, Hyp. 4.1] (see [26, Lemma 3.3] for the assumption on s) and, since we assume that M is bounded and C^2, assumptions (iii') and (v) are also satisfied (see [26, Rk. 4.2]). Moreover, the gradient estimate of [26] is stated for $x \in M \setminus \partial M \mapsto P_t f(x)$, but can be easily extended to $x \in M$ since $P_t f(x) \to 0$ when $x \to \partial M$. Note also that in both references, the gradient estimates are obtained for not necessarily compact manifolds.

The irreducibility assumption (5.9) is an immediate consequence of classical support theorems [27, Exercise 6.7.5] for any value of $t_1 > 0$.

It only remains to prove the next lemma.

Lemma 5.1 *There exist $t_1, \varepsilon, C' > 0$ such that, for all $x \in M$,*

$$\mathbb{P}_x(T_\varepsilon \le t_1 < \tau_\partial) \ge C' \rho_{\partial M}(x), \tag{5.13}$$

where $\rho_{\partial M}(x)$ is the distance between x and ∂M, $T_\varepsilon = \inf\{t \ge 0,\ X_t \in M_\varepsilon\}$ and the compact set M_ε is defined as $\{x \in M : \rho_{\partial M}(x) \ge \varepsilon\}$.

Proof (Proof of Lemma 5.1) Let $\varepsilon_0 > 0$ be small enough for $\rho_{\partial M}$ to be C^2 on $M \setminus M_{\varepsilon_0}$. For all $t < T_{\varepsilon_0}$, we define $Y_t = \rho_{\partial M}(X_t)$. In both situations S1 and S2, we have

$$dY_t = \sigma_t dW_t + b_t dt,$$

where W is a standard Brownian motion, where $\sigma_t \in [\underline{\sigma}, \bar{\sigma}]$ and $|b_t| \le \bar{b}$ are adapted continuous processes, with $0 < \underline{\sigma}, \bar{\sigma}, \bar{b} < \infty$. There exists a differentiable time-change $\tau(s)$ such that $\tau(0) = 0$ and

$$\widetilde{W}_s := \int_0^{\tau(s)} \sigma_t dW_t$$

is a Brownian motion and $\tau'(s) \in [\bar{\sigma}^{-2}, \underline{\sigma}^{-2}]$. In addition,

$$\int_0^{\tau(s)} b_t\, dt \geq -\bar{b}\tau(s) \geq -\bar{b}\underline{\sigma}^{-2}s.$$

As a consequence, setting $Z_s = Y_0 + \widetilde{W}_s - \bar{b}\underline{\sigma}^{-2}s$, we have almost surely $Z_s \leq Y_{\tau(s)}$ for all s such that $\tau(s) \leq T_{\varepsilon_0}$.

Setting $a = \bar{b}\underline{\sigma}^{-2}$, the function

$$f(x) = \frac{e^{2ax} - 1}{2a}$$

is a scale function for the drifted Brownian motion Z. The diffusion process defined by $N_t = f(Z_t)$ is a martingale and its speed measure is given by $s(dv) = \frac{dv}{(1+2av)^2}$. The Green formula for one-dimensional diffusion processes [15, Lemma 23.10] entails, for $\varepsilon_1 = f(\varepsilon_0)$ and all $u \in (0, \varepsilon_1/2)$ (in the following lines, $\mathbb{P}_u^N$ denotes the probability with respect to N with initial position $N_0 = u$),

$$\begin{aligned}\mathbb{P}_u^N(t \leq T_0^N \wedge T_{\varepsilon_1/2}^N) \leq \frac{\mathbb{E}_u^N(T_0^N \wedge T_{\varepsilon_1/2}^N)}{t} &= \frac{2}{t}\int_0^{\varepsilon_1/2}\left(1 - \frac{u \vee v}{\varepsilon_1/2}\right)(u \wedge v)s(dv) \\ &\leq u\,\frac{C_{\varepsilon_1}}{t}, \quad \text{where } C_{\varepsilon_1} = 2\int_0^{\varepsilon_1/2} \frac{dv}{(1+2av)^2}, \end{aligned} \tag{5.14}$$

where we set $T_\varepsilon^N = \inf\{t \geq 0,\ N_t = \varepsilon\}$. Let us fix $s_1 = \varepsilon_1 C_{\varepsilon_1}$. Since N is a martingale, we have, for all $u \in (0, \varepsilon_1/2)$,

$$\begin{aligned} u = \mathbb{E}_u^N(N_{s_1 \wedge T_{\varepsilon_1/2}^N \wedge T_0^N}) &\leq \frac{\varepsilon_1}{2}\mathbb{P}_u^N(T_{\varepsilon_1/2}^N \leq s_1 \wedge T_0^N) + \frac{\varepsilon_1}{2}\mathbb{P}_u^N(s_1 < T_{\varepsilon_1/2}^N \wedge T_0^N) \\ &\leq \frac{\varepsilon_1}{2}\mathbb{P}_u^N(T_{\varepsilon_1/2}^N \leq s_1 \wedge T_0^N) + \frac{u}{2}. \end{aligned}$$

Hence there exists a constant $A > 0$ such that $\mathbb{P}_u^N(T_{\varepsilon_1/2}^N \leq s_1 \wedge T_0^N) \geq A\,u$, or, in other words,

$$\mathbb{P}_x(T_\varepsilon^Z \leq \underline{\sigma}^2 t_1 \wedge T_0^Z) \geq A\, f(\rho_{\partial M}(x)) \geq A\, \rho_{\partial M}(x)$$

for all $x \in M \setminus M_\varepsilon$, where $t_1 = s_1\underline{\sigma}^{-2}$ and $\varepsilon = f^{-1}(\varepsilon_1/2)$.

Now, using the fact that the derivative of the time change $\tau(s)$ belongs to $[\bar{\sigma}^{-2}, \underline{\sigma}^{-2}]$ and that $Z_s \leq Y_{\tau(s)}$, it follows that for all $x \in M \setminus M_\varepsilon$,

$$\mathbb{P}_x(T_\varepsilon^Y \leq t_1 \wedge T_0^Y) \geq \mathbb{P}_x(T_\varepsilon^Z \leq \underline{\sigma}^2 t_1 \wedge T_0^Z) \geq A\rho_{\partial M}(x).$$

Therefore,

$$\begin{aligned}\mathbb{P}_x(T_\varepsilon^Y \leq t_1 < T_0^Y) &\geq \mathbb{E}_x\left[\mathbb{1}_{T_\varepsilon^Y \leq t_1 \wedge T_0^Y} \mathbb{P}_{X_{T_\varepsilon^Y}}(t_1 < \tau_\partial)\right] \\ &\geq \mathbb{P}_x(T_\varepsilon^Y \leq t_1 \wedge T_0^Y) \inf_{y \in M_\varepsilon} \mathbb{P}_y(t_1 < \tau_\partial) \geq C'\rho_{\partial M}(x),\end{aligned}$$

where we used that $\inf_{y\in M_\varepsilon} \mathbb{P}_y(t_1 < \tau_\partial) > 0$. This last fact follows from the inequality $\mathbb{P}_y(t_1 < \tau_\partial) > 0$ for all $y \in M \setminus \partial M$, consequence of (5.9) and from the Lipschitz-continuity of $y \mapsto \mathbb{P}_y(t_1 < \tau_\partial) = P_{t_1}\mathbb{1}_E(y)$, consequence of (5.12).

Finally, since $T_\varepsilon = 0$ under $\mathbb{P}_x$ for all $x \in M_\varepsilon$, replacing C' by $C' \wedge [\inf_{y\in M_\varepsilon} \mathbb{P}_y(t_1 < \tau_\partial)/\mathrm{diam}(M)]$ entails (5.13) for all $x \in M$.

Remark 5.3 The gradient estimates of [26] are proved for diffusion processes with space-dependent killing rate $V : M \to [0, \infty)$. More precisely, they consider infinitesimal generators of the form

$$L = \frac{1}{2}\sum_{i,j=1}^{d} [ss^*]_{ij}\partial_i\partial_j + \sum_{i=1}^{d} b_i\partial_i - V$$

with V bounded measurable. Our proof also applies to this setting.

Remark 5.4 We have proved in particular that Condition (A1) is satisfied in situations S1 and S2. This is a minoration of conditional distributions of the diffusion. For initial positions in compact subsets of $M \setminus \partial M$, this reduces to a lower bound for the (unconditioned) distribution of the process. Such a result could be obtained from density lower bounds using number of techniques, for example Aronson-type estimates [1, 28, 30] or continuity properties [13]. Note that our result does not rely on such techniques, since it will appear in the proof that Conditions (5.7) and (5.9) are sufficient to obtain $\mathbb{P}_x(X_{t_0} \in \cdot) \geq \widetilde{\nu}$ for all $x \in M_\varepsilon$ for some positive measure $\widetilde{\nu}$.

5.3.3 Proof of Theorem 5.2

The proof is based on the following equivalent form of Condition (A) (see [6, Thm. 2.1])

Condition (A')

There exist $t_0, c_1, c_2 > 0$ such that

(A1') for all $x, y \in E$, there exists a probability measure $\nu_{x,y}$ on E such that

$$\mathbb{P}_x(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu_{x,y}(\cdot) \quad \text{and} \quad \mathbb{P}_y(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu_{x,y}(\cdot)$$

(A2') for all $x, y, z \in E$ and all $t \geq 0$,

$$\mathbb{P}_{\nu_{x,y}}(t < \tau_\partial) \geq c_2 \mathbb{P}_z(t < \tau_\partial).$$

Note that (A1') is a kind of coupling for conditional laws of the Markov process starting from different initial conditions. It is thus natural to use gradient estimates to prove such conditions since they are usually obtained by coupling of the paths of the process (see [26, 29]).

We divide the proof into four steps. In the first one, we obtain a lower bound for $\mathbb{P}_x(X_{2t_1} \in K \mid 2t_1 < \tau_\partial)$. The second and third ones are devoted to the proof of (A1') and (A2'), respectively. The last one gives the proof of (5.10).

5.3.3.1 Return to a Compact Conditionally on Non-absorption

The gradient estimate (5.7) applied to $f = \mathbb{1}_E$ implies that $P_{t_1}\mathbb{1}_E$ is Lipschitz. Since $\mathbb{P}_\partial(t_1 < \tau_\partial) = 0$, we obtain, for all $x \in E$,

$$\mathbb{P}_x(t_1 < \tau_\partial) \leq C\, \rho_\partial(x). \tag{5.15}$$

Combining this with Assumption (5.8), we deduce that, for all $x \in E$,

$$\mathbb{P}_x(T_K \leq t_1 \mid t_1 < \tau_\partial) = \frac{\mathbb{P}_x(T_K \leq t_1 < \tau_\partial)}{\mathbb{P}_x(t_1 < \tau_\partial)} \geq \frac{C'}{C}.$$

Fix $x_0 \in K$ and let $r_0 = d(x_0, \partial)/2$. We can assume without loss of generality that $B(x_0, r_0) \subset K$, since Assumption (5.8) remains true if one replaces K by the set $K \cup \overline{B(x_0, r_0)} \subset E$ (which is also a compact set, as a closed subset of the compact set $E \cup \{\partial\}$). Then, it follows from (5.9) that, for all $x \in E$,

$$\mathbb{E}_x\left[\mathbb{P}_{X_{t_1}}(X_s \in B(x_0, r_0),\ \forall s \in [0, t_1])\right] = \mathbb{P}_x(X_s \in B(x_0, r_0),\ \forall s \in [t_1, 2t_1]) > 0.$$

Because of (5.7), the left-hand side is continuous w.r.t. $x \in M$, and hence

$$\inf_{x\in K} \mathbb{P}_x(X_s \in K,\ \forall s \in [t_1, 2t_1]) \geq \inf_{x\in K} \mathbb{P}_x(X_s \in B(x_0, r_0),\ \forall s \in [t_1, 2t_1]) > 0.$$

Therefore, it follows from the strong Markov property at time T_K that

$$\begin{aligned}\mathbb{P}_x(X_{2t_1} \in K \mid 2t_1 < \tau_\partial) &\geq \frac{\mathbb{P}_x(X_{2t_1} \in K)}{\mathbb{P}_x(t_1 < \tau_\partial)} \\ &\geq \frac{\mathbb{P}_x(T_K \leq t_1 \text{ and } X_{T_K+s} \in K,\ \forall s \in [t_1, 2t_1])}{\mathbb{P}_x(t_1 < \tau_\partial)} \\ &\geq \inf_{x \in K} \mathbb{P}_x(X_s \in K,\ \forall s \in [t_1, 2t_1]) \frac{\mathbb{P}_x(T_K \leq t_1)}{\mathbb{P}_x(t_1 < \tau_\partial)}.\end{aligned}$$

Therefore, we have proved that, for all $x \in E$,

$$\mathbb{P}_x(X_{2t_1} \in K \mid 2t_1 < \tau_\partial) \geq A, \tag{5.16}$$

for the positive constant $A := \inf_{x \in K} \mathbb{P}_x(X_s \in K,\ \forall s \in [t_1, 2t_1])C'/C$.

5.3.3.2 Proof of (A1')

For all $x, y \in E$, let $\mu_{x,y}$ be the infimum measure of $\delta_x P_{2t_1}$ and $\delta_y P_{2t_1}$, i.e. for all measurable $A \subset E$,

$$\mu_{x,y}(A) := \inf_{A_1 \cup A_2 = A,\ A_1, A_2 \text{ measurable}} (\delta_x P_{2t_1} \mathbb{1}_{A_1} + \delta_y P_{2t_1} \mathbb{1}_{A_2}).$$

The proof of (A1') is based on the following lemma.

Lemma 5.2 *For all bounded continuous function $f : E \to \mathbb{R}_+$ not identically 0, the function $(x, y) \in E^2 \mapsto \mu_{x,y}(f)$ is Lipschitz and positive.*

Proof By (5.7), for all bounded measurable $g : E \to \mathbb{R}$,

$$\|\nabla P_{2t_1} g\|_\infty = \|\nabla P_{t_1}(P_{t_1} g)\|_\infty \leq C \|P_{t_1} g\|_\infty \leq C \|g\|_\infty. \tag{5.17}$$

Hence, for all $x, y \in E$,

$$|P_{2t_1} g(x) - P_{2t_1} g(y)| \leq C \|g\|_\infty \rho(x, y). \tag{5.18}$$

This implies the uniform Lipschitz-continuity of $P_{2t_1} g$. In particular, we deduce that

$$\mu_{x,y}(f) = \inf_{A_1 \cup A_2 = E} \left\{ P_{2t_1}(f \mathbb{1}_{A_1})(x) + P_{2t_1}(f \mathbb{1}_{A_2})(y) \right\}$$

is continuous w.r.t. $(x, y) \in E^2$ (and even Lipschitz).

Let us now prove that $\mu_{x,y}(f) > 0$. Let us define $\bar{\mu}_{x,y}$ as the infimum measure of $\delta_x P_{t_1}$ and $\delta_y P_{t_1}$: for all measurable $A \subset E$,

$$\bar{\mu}_{x,y}(A) := \inf_{A_1 \cup A_2 = A} (\delta_x P_{t_1} \mathbb{1}_{A_1} + \delta_y P_{t_1} \mathbb{1}_{A_2}).$$

The continuity of $(x, y) \mapsto \bar{\mu}_{x,y}(f)$ on E^2 holds as above.

Fix $x_1 \in E$ and $d_1 > 0$ such that $\inf_{x \in B(x_1,d_1)} f(x) > 0$. Then (5.9) entails

$$\bar{\mu}_{x_1,x_1}(f) = \delta_{x_1} P_{t_1} f \geq \mathbb{P}_{x_1}(X_{t_1} \in B(x_1, d_1)) \inf_{x \in B(x_1,d_1)} f(x) > 0.$$

Therefore, there exist $r_1, a_1 > 0$ such that $\bar{\mu}_{x,y}(f) \geq a_1$ for all $x, y \in B(x_1, r_1)$.

Hence, for all nonnegative measurable $g : E \to \mathbb{R}_+$ and for all $x, y \in E$ and all $u' \in E$,

$$\begin{aligned} \delta_x P_{2t_1} g &\geq \int_E \mathbb{1}_{u \in B(x_1,r_1)} P_{t_1} g(u)\, \delta_x P_{t_1}(du) \\ &\geq \int_E \mathbb{1}_{u,u' \in B(x_1,r_1)} \bar{\mu}_{u,u'}(g)\, \delta_x P_{t_1}(du). \end{aligned} \tag{5.19}$$

Integrating both sides of the inequality w.r.t. $\delta_y P_{t_1}(du')$, we obtain

$$\delta_x P_{2t_1} g \geq \delta_x P_{2t_1} g \delta_y P_{t_1}(E) \geq \iint_{E \times E} \mathbb{1}_{u,u' \in B(x_1,r_1)} \bar{\mu}_{u,u'}(g)\, \delta_x P_{t_1}(du)\, \delta_y P_{t_1}(du').$$

Since this holds for all nonnegative measurable g and since $\mu_{x,y}$ is the infimum measure between $\delta_x P_{2t_1}$ and $\delta_y P_{2t_1}$, by symmetry, we have proved that

$$\mu_{x,y}(\cdot) \geq \iint_{E \times E} \mathbb{1}_{u,u' \in B(x_1,r_1)} \bar{\mu}_{u,u'}(\cdot)\, \delta_x P_{t_1}(du)\, \delta_y P_{t_1}(du').$$

Therefore, (5.9) entails

$$\mu_{x,y}(f) \geq a_1 \mathbb{P}_x(X_{t_1} \in B(x_1, r_1)) \mathbb{P}_y(X_{t_1} \in B(x_1, r_1)) > 0.$$

We now construct the measure $\nu_{x,y}$ of Condition (A1'). Using a similar computation as in (5.19) and integrating with respect to $\delta_y P_{2t_1}(du')/\delta_y P_{2t_1} \mathbb{1}_E$, we obtain for all $x, y \in E$ and all nonnegative measurable $f : E \to \mathbb{R}_+$

$$\delta_x P_{4t_1} f \geq \iint_{K \times K} \mu_{u,u'}(f)\, \delta_x P_{2t_1}(du)\, \frac{\delta_y P_{2t_1}(du')}{\delta_y P_{2t_1} \mathbb{1}_E}.$$

Since $\delta_x P_{4t_1}\mathbb{1}_E \leq \delta_x P_{2t_1}\mathbb{1}_E$,

$$\begin{aligned}\frac{\delta_x P_{4t_1} f}{\delta_x P_{4t_1}\mathbb{1}_E} &\geq \iint_{K\times K} \mu_{u,u'}(f)\,\frac{\delta_x P_{2t_1}(du)}{\delta_x P_{2t_1}\mathbb{1}_E}\,\frac{\delta_y P_{2t_1}(du')}{\delta_y P_{2t_1}\mathbb{1}_E}\\ &= m_{x,y}\nu_{x,y}(f),\end{aligned}$$

where

$$m_{x,y} := \iint_{K\times K} \mu_{u,u'}(E)\,\frac{\delta_x P_{2t_1}(du)}{\delta_x P_{2t_1}\mathbb{1}_E}\,\frac{\delta_y P_{2t_1}(du')}{\delta_y P_{2t_1}\mathbb{1}_E}$$

and

$$\nu_{x,y} := \frac{1}{m_{x,y}} \iint_{K\times K} \mu_{u,u'}(\cdot)\,\frac{\delta_x P_{2t_1}(du)}{\delta_x P_{2t_1}\mathbb{1}_E}\,\frac{\delta_y P_{2t_1}(du')}{\delta_y P_{2t_1}\mathbb{1}_E}. \tag{5.20}$$

Note that

$$\begin{aligned}m_{x,y} &\geq \inf_{u,u'\in K^2} \mu_{u,u'}(E) \iint_{K\times K} \frac{\delta_x P_{2t_1}(du)}{\delta_x P_{2t_1}\mathbb{1}_E}\,\frac{\delta_y P_{2t_1}(du')}{\delta_y P_{2t_1}\mathbb{1}_E}\\ &\geq A^2 \inf_{u,u'\in K^2} \mu_{u,u'}(E) > 0,\end{aligned} \tag{5.21}$$

because of (5.16) and Lemma 5.2. Hence the probability measure $\nu_{x,y}$ is well-defined and we have proved (A1') for $t_0 = 4t_1$ and $c_1 = A^2 \inf_{u,u'\in K^2} \mu_{u,u'}(E)$.

5.3.3.3 Proof of (A2')

Our goal is now to prove Condition (A2'). We first prove the following gradient estimate for $f = \mathbb{1}_E$.

Lemma 5.3 *There exists a constant $C'' > 0$ such that, for all $t \geq 4t_1$,*

$$\|\nabla P_t \mathbb{1}_E\|_\infty \leq C'' \|P_t \mathbb{1}_E\|_\infty. \tag{5.22}$$

Note that, compared to (5.7), the difficulty is that we replace $\|\mathbb{1}_E\|_\infty$ by the smaller $\|P_{t_1}\mathbb{1}_E\|_\infty$ and that we extend this inequality to any time t large enough.

Proof We first use (5.16) to compute

$$P_{4t_1}\mathbb{1}_E(x) \geq \mathbb{P}_x(X_{2t_1} \in K) \inf_{y\in K} \mathbb{P}_y(2t_1 < \tau_\partial) \geq mA\mathbb{P}_x(2t_1 < \tau_\partial),$$

where $m := \inf_{y\in K} \mathbb{P}_y(2t_1 < \tau_\partial)$ is positive because of Lemma 5.2. Integrating the last inequality with respect to $(\delta_y P_{t-4t_1})(dx)$ for any fixed $y \in E$ and $t \geq 4t_1$, we deduce that

$$\|P_t\mathbb{1}_E\|_\infty \geq mA\|P_{t-2t_1}\mathbb{1}_E\|_\infty.$$

Hence it follows from (5.12) that, for all $t \geq 4t_1$,

$$\begin{aligned}\|\nabla P_t\mathbb{1}_E\|_\infty = \|\nabla P_{t_1}(P_{t-t_1}\mathbb{1}_E)\|_\infty &\leq C\|P_{t-t_1}\mathbb{1}_E\|_\infty \\ &\leq C\|P_{t-2t_1}\mathbb{1}_E\|_\infty \leq \frac{C}{mA}\|P_t\mathbb{1}_E\|_\infty.\end{aligned}$$

This concludes the proof of Lemma 5.3.

This lemma implies that the function

$$h_t : x \in E \cup \{\partial\} \mapsto \frac{P_t\mathbb{1}_E(x)}{\|P_t\mathbb{1}_E\|_\infty} \tag{5.23}$$

is C''-Lipschitz for all $t \geq 4t_1$. Since this function vanishes on ∂ and its maximum over E is 1, we deduce that, for any $t \geq 4t_1$, there exists at least one point $z_t \in E$ such that $h_t(z_t) = 1$. Since h_t is C''-Lipschitz, we also deduce that $\rho(z_t, \partial) \geq 1/C''$. Moreover, for all $x \in E$,

$$\frac{P_t\mathbb{1}_E(x)}{\|P_t\mathbb{1}_E\|_\infty} \geq f_{z_t}(x), \tag{5.24}$$

where, for all $z \in E$ and $x \in E$, $f_z(x) = \big(1 - C''\rho(x,z)\big) \vee 0$. We define the compact set $K' = \{x \in E : \rho(x,\partial) \geq 1/C''\}$ so that $z_t \in K'$ for all $t \geq 4t_1$. Then, for all $x, y \in E$ and for all $t \geq 4t_1$, using the definition (5.20) of $\nu_{x,y}$,

$$\begin{aligned}\mathbb{P}_{\nu_{x,y}}(t < \tau_\partial) &\geq \|P_t\mathbb{1}_E\|_\infty\, \nu_{x,y}(f_{z_t}) \\ &= \frac{\|P_t\mathbb{1}_E\|_\infty}{m_{x,y}} \iint_{K\times K} \mu_{z,z'}(f_{z_t})\, \frac{\delta_x P_{2t_1}(dz)}{\delta_x P_{2t_1}\mathbb{1}_E}\, \frac{\delta_y P_{2t_1}(dz')}{\delta_y P_{2t_1}\mathbb{1}_E}.\end{aligned}$$

Since $z \mapsto f_z$ is Lipschitz for the $\|\cdot\|_\infty$ norm (indeed, $|f_z(x) - f_{z'}(x)| \leq C''|\rho(x,z) - \rho(x,z')| \leq C''\rho(z,z')$ for all $x, z, z' \in E^3$), it follows from Lemma 5.2 that $(x,y,z) \mapsto \mu_{x,y}(f_z)$ is positive and continuous on E^3. Hence $c := \inf_{x\in K,\, y\in K,\, z\in K'} \mu_{x,y}(f_z) > 0$ and, using that $m_{x,y} \leq 1$,

$$\mathbb{P}_{\nu_{x,y}}(t < \tau_\partial) \geq c\|P_t\mathbb{1}_E\|_\infty \iint_{K\times K} \frac{\delta_x P_{2t_1}(dz)}{\delta_x P_{2t_1}\mathbb{1}_E}\, \frac{\delta_y P_{2t_1}(dz')}{\delta_y P_{2t_1}\mathbb{1}_E} \geq cA^2\|P_t\mathbb{1}_E\|_\infty,$$

where the last inequality follows from (5.16).

This entails Condition (A2') for all $t \geq 4t_1$. For $t \leq 4t_1$,

$$\mathbb{P}_{\nu_{x,y}}(t < \tau_\partial) \geq \mathbb{P}_{\nu_{x,y}}(4t_1 < \tau_\partial) \geq cA^2 \|P_{4t_1}\mathbb{1}_E\|_\infty$$
$$\geq cA^2 \|P_{4t_1}\mathbb{1}_E\|_\infty \sup_{z\in E} \mathbb{P}_z(t < \tau_\partial) > 0.$$

This ends the proof of (A2') and hence of (5.1).

5.3.3.4 Contraction in Total Variation Norm

It only remains to prove (5.10). By (5.2), we need to prove that there exists a constant $a > 0$ such that, for all probability measure μ on E,

$$c(\mu) := \inf_{t\geq 0} \frac{\mathbb{P}_\mu(t < \tau_\partial)}{\|P_t \mathbb{1}_E\|_\infty} \geq a\mu(\rho_\partial). \tag{5.25}$$

Because of the equivalence between (A) and (A') [6, Theorem 2.1], enlarging t_0 and reducing c_1 and c_2, one can assume without loss of generality that $\nu = \nu_{x,y}$ does not depend on $x, y \in E$. Then, using (A1) and (A2), we deduce that, for all $t \geq t_0 \geq 4t_1$,

$$\mathbb{P}_\mu(t < \tau_\partial) = \mu(P_{t_0} P_{t-t_0} \mathbb{1}_E) \geq c_1 \mathbb{P}_\mu(t_0 < \tau_\partial)\mathbb{P}_\nu(t - t_0 < \tau_\partial)$$
$$\geq c_1 c_2 \|P_{t-t_0}\mathbb{1}_E\|_\infty \mathbb{P}_\mu(t_0 < \tau_\partial) \geq c_1 c_2 \|P_t \mathbb{1}_E\|_\infty \mathbb{P}_\mu(t_0 < \tau_\partial).$$

Now, using Assumption (5.8), we deduce that

$$\mathbb{P}_\mu(t_0 < \tau_\partial) \geq \mathbb{E}_\mu\left(\mathbb{1}_{T_K < t_1} \inf_{y\in K} \mathbb{P}_y(t_0 < \tau_\partial)\right) \geq C'\mu(\rho(\partial, \cdot)) \inf_{y\in K} \mathbb{P}_y(t_0 < \tau_\partial),$$

where the constant $C'' := C' \inf_{y\in K} \mathbb{P}_y(t_0 < \tau_\partial)$ is positive. For $t \leq t_0$, the last inequality entails

$$\mathbb{P}_\mu(t < \tau_\partial) \geq \mathbb{P}_\mu(t_0 < \tau_\partial) \geq C''\mu(\rho(\partial, \cdot)) \geq C''\mu(\rho(\partial, \cdot))\|P_t \mathbb{1}_E\|_\infty.$$

Hence (5.25) holds true with $a = c_1 c_2 C''$. This ends the proof of Theorem 5.2.

Appendix: Proof of (5.2)

Let us assume that Condition (A) is satisfied. For all $t \geq 0$ and all probability measure π on E, let $c_t(\pi) := \frac{\pi(P_t \mathbb{1}_E)}{\|P_t \mathbb{1}_E\|_\infty}$. In the proof of [6, Corollary 2.2], it is proved that, for all probability measures π_1, π_2 on E

$$\left\|\mathbb{P}_{\pi_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\pi_2}(X_t \in \cdot \mid t < \tau_\partial)\right\|_{TV} \leq \frac{(1 - c_1 c_2)^{\lfloor t/t_0 \rfloor}}{c_t(\pi_1) \vee c_t(\pi_2)} \|\pi_1 - \pi_2\|_{TV}.$$

But

$$\inf_{t \geq 0} c_t(\pi_1) \vee c_t(\pi_2) \geq (\inf_{t \geq 0} c_t(\pi_1)) \vee (\inf_{t \geq 0} c_t(\pi_2)) = c(\pi_1) \vee c(\pi_2).$$

This ends the proof of (5.2).

References

1. D.G. Aronson, Non-negative solutions of linear parabolic equations. Ann. Scuola Norm. Sup. Pisa (3) **22**, 607–694 (1968)
2. K. Bogdan, T. Grzywny, M. Ryznar, Heat kernel estimates for the fractional Laplacian with Dirichlet conditions. Ann. Probab. **38**(5), 1901–1923 (2010)
3. P. Cattiaux, S. Méléard, Competitive or weak cooperative stochastic lotka-volterra systems conditioned to non-extinction. J. Math. Biol. **60**(6), 797–829 (2010)
4. P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, J.S. Martín, Quasi-stationary distributions and diffusion models in population dynamics. Ann. Probab. **37**(5), 1926–1969 (2009)
5. N. Champagnat, D. Villemonais, Uniform convergence of conditional distributions for absorbed one-dimensional diffusions (2015). ArXiv e-prints
6. N. Champagnat, D. Villemonais, Exponential convergence to quasi-stationary distribution and Q-process. Probab. Theory Relat. Fields **164**(1–2), 243–283 (2016)
7. N. Champagnat, D. Villemonais, Uniform convergence to the Q-process. Electron. Commun. Probab. **22**, 1–7 (2017)
8. Z.-Q. Chen, P. Kim, R. Song, Heat kernel estimates for $\Delta + \Delta^{\alpha/2}$ in $C^{1,1}$ open sets. J. Lond. Math. Soc. (2) **84**(1), 58–80 (2011)
9. Z.-Q. Chen, P. Kim, R. Song, Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation. Ann. Probab. **40**(6), 2483–2538 (2012)
10. Z.-Q. Chen, P. Kim, R. Song, Stability of Dirichlet heat kernel estimates for non-local operators under Feynman-Kac perturbation. Trans. Am. Math. Soc. **367**(7), 5237–5270 (2015)
11. E.B. Davies, B. Simon, Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. J. Funct. Anal. **59**(2), 335–395 (1984)
12. P. Del Moral, D. Villemonais, Exponential mixing properties for time inhomogeneous diffusion processes with killing. Bernoulli **24**(2), 1010–1032 (2018)
13. E.B. Dynkin, *Diffusions, Superdiffusions and Partial Differential Equations*. American Mathematical Society Colloquium Publications, vol. 50 (American Mathematical Society, Providence, 2002)
14. A. Hening, M. Kolb, Quasistationary distributions for one-dimensional diffusions with singular boundary points (2014). ArXiv e-prints
15. O. Kallenberg, *Foundations of Modern Probability*. Probability and Its Applications (New York), 2nd edn. (Springer, New York, 2002)
16. K.-Y. Kim, P. Kim, Two-sided estimates for the transition densities of symmetric Markov processes dominated by stable-like processes in $C^{1,\eta}$ open sets. Stoch. Process. Appl. **124**(9), 3055–3083 (2014)

17. P. Kim, R. Song, Estimates on Green functions and Schrödinger-type equations for non-symmetric diffusions with measure-valued drifts. J. Math. Anal. Appl. **332**(1), 57–80 (2007)
18. R. Knobloch, L. Partzsch, Uniform conditional ergodicity and intrinsic ultracontractivity. Potential Anal. **33**, 107–136 (2010)
19. M. Kolb, A. Wübker, Spectral analysis of diffusions with jump boundary. J. Funct. Anal. **261**(7), 1992–2012 (2011)
20. J.-F. Le Gall, *Brownian Motion, Martingales, and Stochastic Calculus*. Graduate Texts in Mathematics, vol. 274, French edn. (Springer, Berlin, 2016)
21. J. Lierl, L. Saloff-Coste, The Dirichlet heat kernel in inner uniform domains: local results, compact domains and non-symmetric forms. J. Funct. Anal. **266**(7), 4189–4235 (2014)
22. T. Lindvall, L.C.G. Rogers, Coupling of multidimensional diffusions by reflection. Ann. Probab. **14**(3), 860–872 (1986)
23. J.C. Littin, Uniqueness of quasistationary distributions and discrete spectra when ∞ is an entrance boundary and 0 is singular. J. Appl. Probab. **49**(3), 719–730 (2012)
24. Y. Miura, Ultracontractivity for Markov semigroups and quasi-stationary distributions. Stoch. Anal. Appl. **32**(4), 591–601 (2014)
25. D. Nualart, *The Malliavin Calculus and Related Topics*. Probability and Its Applications (New York), 2nd edn. (Springer, Berlin, 2006)
26. E. Priola, F.-Y. Wang, Gradient estimates for diffusion semigroups with singular coefficients. J. Funct. Anal. **236**(1), 244–264 (2006)
27. D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233 (Springer, Berlin, 1979)
28. F.-Y. Wang, Sharp explicit lower bounds of heat kernels. Ann. Probab. **25**(4), 1995–2006 (1997)
29. F.-Y. Wang, Gradient estimates of Dirichlet heat semigroups and application to isoperimetric inequalities. Ann. Probab. **32**(1A), 424–440 (2004)
30. Q.S. Zhang, Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u) + B\nabla u - u_t = 0$. Manuscr. Math. **93**(3), 381–390 (1997)
31. Q.S. Zhang, The boundary behavior of heat kernels of Dirichlet Laplacians. J. Differ. Equ. **182**(2), 416–430 (2002)

Chapter 6
Bismut-Elworthy-Li Formulae for Bessel Processes

Henri Elad Altman

Abstract In this article we are interested in the differentiability property of the Markovian semi-group corresponding to the Bessel processes of nonnegative dimension. More precisely, for all $\delta \geq 0$ and $T > 0$, we compute the derivative of the function $x \mapsto P_T^\delta F(x)$, where $(P_t^\delta)_{t\geq 0}$ is the transition semi-group associated to the δ-dimensional Bessel process, and F is any bounded Borel function on $\mathbb{R}_+$. The obtained expression shows a nice interplay between the transition semi-groups of the δ—and the $(\delta + 2)$-dimensional Bessel processes. As a consequence, we deduce that the Bessel processes satisfy the strong Feller property, with a continuity modulus which is independent of the dimension. Moreover, we provide a probabilistic interpretation of this expression as a Bismut-Elworthy-Li formula.

Keywords Bismut-Elworthy-Li formula · Strong Feller property · Bessel processes

6.1 Introduction

Bessel processes are a one-parameter family of nonnegative diffusion processes with a singular drift, which present a reflecting behavior when they hit the origin. The smaller the parameter (called dimension), the more intense the reflection. Hence, studying the dynamics of these processes is a non-trivial problem, especially when the dimension is small. Despite these apparent difficulties, Bessel processes have remarkably nice properties. Therefore they provide an instructive insight in the study of stochastic differential equations (SDEs) with a singular drift, as well as the study of reflected SDEs.

H. E. Altman (✉)
Université Pierre et Marie Curie, LPMA, Paris, France
e-mail: henri.elad_altman@upmc.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_6

For all $x \geq 0$ and $\delta \geq 0$, the squared Bessel process of dimension δ started at x^2 is the unique strong solution of the equation:

$$X_t = x^2 + 2 \int_0^t \sqrt{X_s} dB_s + \delta t. \tag{6.1}$$

Such a process X is nonnegative, and the law of its square-root $\rho = \sqrt{X}$ is, by definition, the δ-dimensional Bessel process started at x (see [10, section XI], or Chapter 3 of [14] for an introduction to Bessel processes). The process ρ satisfies the following SDE before its first hitting time T_0 of 0:

$$\forall t \in [0, T_0), \qquad \rho_t = x + \frac{\delta - 1}{2} \int_0^t \frac{ds}{\rho_s} + B_t.$$

This is an SDE with non-Lipschitz continuous drift term given by the function $x \mapsto \frac{\delta-1}{2} \frac{1}{x}$ on $(0, +\infty)$. Note that when $\delta < 1$, this function is nondecreasing on $\mathbb{R}_+$, and as δ decreases this "wrong" monotonicity becomes more and more acute. As a consequence, for δ small, the process ρ is not mean-square differentiable, so that classical criteria for the Bismut-Elworthy-Li formula to hold (see [4, Section 1.5, and Section 2 below]) do not apply here. Hence, one would not even expect such a formula to hold for $\delta < 1$. For instance, even continuity of the flow is not known in this regime, see Remark 6.10 below.

The aim of the present paper is to study the derivative in space of the family of transition kernels $(P_T^\delta)_{T \geq 0}$ of the δ-dimensional Bessel process. In a first part, we show that this derivative can be expressed in terms of the transition kernels of the δ- and the $(\delta + 2)$-dimensional Bessel processes. More precisely, we prove that, for all function $F : \mathbb{R}_+ \to \mathbb{R}$ bounded and Borel, all $T > 0$ and all $x \geq 0$, we have:

$$\frac{d}{dx} P_T^\delta F(x) = \frac{x}{T} \left(P_T^{\delta+2} F(x) - P_T^\delta F(x) \right). \tag{6.2}$$

As a consequence, the Bessel processes satisfy the strong Feller property uniformly in δ. In a second part, we interpret the above result probabilistically as a Bismut-Elworthy-Li formula. More precisely, given a realization ρ of the Bessel process through the SDE (6.1), we introduce the derivative η_t of ρ_t with respect to the initial condition x, and show that when $\delta > 0$, the stochastic integral $\int_0^t \eta_s dB_s$ is well-defined as an L^p martingale, for some $p > 1$ depending on δ. Moreover, it turns out that $\int_0^T \eta_s dB_s$ is (up to a constant) the Radon-Nikodym derivative of the $(\delta + 2)$-dimensional Bessel process over the interval $[0, T]$ w.r.t. the δ-dimensional one. As a consequence, we deduce that the above equation can be rewritten:

$$\frac{d}{dx} P_T^\delta F(x) = \frac{1}{T} \mathbb{E} \left[F(\rho_t(x)) \left(\int_0^t \eta_s(x) dB_s \right) \right] \tag{6.3}$$

which is an apparition, in an unexpected context, of the well-known Bismut-Elworthy-Li formula (see [7] for a precise statement and proof of the Bismut-Elworthy-Li formula in the case of diffusions with smooth coefficients).

One surprising feature is that, while (6.2) is very easy to prove whatever the value of $\delta \geq 0$, on the other hand, the process $(\int_0^t \eta_s(x) dB_s)_{t \geq 0}$ has less and less finite moments as δ decreases, which makes the proof of (6.3) more involved for small δ. In particular, this process is not in L^2 for $\delta < 2(\sqrt{2}-1)$, and when $\delta = 0$, we do not even know whether the stochastic integral $\int_0^t \eta_s(x) dB_s$ is well-defined as a local martingale.

This article was originally motivated by the hope to prove the strong Feller property for some singular reflected SDEs or SPDEs. Recently, several works have brought about new techniques to prove the strong Feller property for singular SPDEs. Thus, in [11], the authors established this property for the $P(\Phi)_2$ equation, and in [8], the authors established it for a large class of singular semilinear SPDEs. The fact, mentioned above, that blowup of η does not affect the strong Feller property of Bessel processes is reminiscent of the latter article, where the setting used to prove the strong Feller property allows blowup in finite time of the solution. Also, we hope that the techniques used in the present article might give inspiration to treat more general cases. Note that, even in the present context, where many computations can be performed explicitly, we still have an open problem concerning the Strong Feller bounds for Bessel processes of dimension $\delta \leq 2(\sqrt{2}-1)$ (see Remark 6.14 below).

The plan of our paper is as follows. In Sect. 6.2 we recall the classical Bismut-Elworthy-Li formula for diffusions in $\mathbb{R}$ with a dissipative drift, and show how this implies the strong Feller property. In Sect. 6.3 we recall the definition of Bessel processes and their basic properties. In Sect. 6.4 we compute the derivative of the Bessel semi-group. In Sect. 6.5 we establish the differentiability of the Bessel flow at any given point in $\mathbb{R}_+^*$, and we give an expression for (some modification of) the derivative. In Sect. 6.6, we show that this derivative is not bounded in time when $\delta < 1$. We prove, however, that it is linked to an interesting martingale corresponding to the family of Radon-Nikodym derivatives of the $(\delta+2)$-dimensional Bessel process w.r.t. the δ-dimensional one. In Sect. 6.7 we prove the Bismut-Elworthy-Li formula for the Bessel processes of dimension $\delta > 0$.

6.2 Classical Bismut-Elworthy-Li Formula for One-Dimensional Diffusions

In this section we recall very briefly the Bismut-Elworthy-Li formula in the case of one-dimensional diffusions, and the way this formula implies the strong Feller property.

Consider an SDE on $\mathbb{R}$ of the form:

$$dX_t = b(X_t)dt + dB_t, \quad X_0 = x \tag{6.4}$$

where $b : \mathbb{R} \to \mathbb{R}$ is smooth and satisfies:

$$\begin{aligned} |b(x) - b(y)| &\leq C|x - y|, & x, y \in \mathbb{R} \\ b'(x) &\leq L, & x \in \mathbb{R} \end{aligned} \tag{6.5}$$

where $C > 0$, $L \in \mathbb{R}$ are some constants. By the classical theory of SDEs, for all $x \in \mathbb{R}$, there exists a unique continuous, square-integrable process $(X_t(x))_{t\geq 0}$ satisfying (6.4). Actually, by the Lipschitz assumption on b, there even exists a bi-continuous process $(X_t(x))_{t\geq 0, x\in\mathbb{R}}$ such that, for all $x \in \mathbb{R}$, $(X_t(x))_{t\geq 0}$ solves (6.4).

Let $x \in \mathbb{R}$. Consider the solution $(\eta_t(x))_{t\geq 0}$ to the variation equation obtained by formally differentiating (6.4) with respect to x:

$$d\eta_t(x) = b'(X_t)\eta_t(x)dt, \quad \eta_0(x) = 1$$

Note that this is a (random) linear ODE with explicit solution given by:

$$\eta_t(x) = \exp\left(\int_0^t b'(X_s)ds\right)$$

It is easy to prove that, for all $t \geq 0$ and $x \in \mathbb{R}$, the map $y \to X_t(y)$ is a.s. differentiable at x and:

$$\frac{dX_t}{dx} \overset{a.s.}{=} \eta_t(x) \tag{6.6}$$

Remark 6.1 Note that $\eta_t(x) > 0$ for all $t \geq 0$ and $x \in \mathbb{R}$. This reflects the fact that, for all $x \leq y$, by a comparison theorem for SDEs (see Theorem 3.7 in Chapter IX in [10]), one has $X_t(x) \leq X_t(y)$.

Recall that a Markovian semi-group $(P_t)_{t\geq 0}$ on a Polish space E is said to satisfy the strong Feller property if, for all $t > 0$ and $\varphi : E \to \mathbb{R}$ bounded and Borel, the function $P_t\varphi : E \to \mathbb{R}$ defined by:

$$P_t\varphi(x) = \int \varphi(y)P_t(x, dy), \quad x \in \mathbb{R}$$

is continuous.

The strong Feller property is very useful in the study of SDEs and SPDEs, namely for the proof of ergodicity (see, e.g., the monographs [4, 5] and [14], as well as the recent articles [8] and [11], for applications of the strong Feller property in the context of SPDEs).

Let $(P_t)_{t\geq 0}$ be the Markovian semi-group associated to the SDE (6.9). We are interested in proving the strong Feller property for $(P_t)_{t\geq 0}$. Note that, by assumption (6.5), $\eta_t(x) \leq e^{Lt}$ for all $t \geq 0$ and $x \in \mathbb{R}$. Therefore, by (6.6) and the dominated convergence theorem, for all $\varphi : \mathbb{R} \to \mathbb{R}$ differentiable with a bounded derivative, one has:

$$\frac{d}{dx}\left(P_t\varphi\right)(x) = \frac{d}{dx}\mathbb{E}\left[\varphi(X_t(x))\right] = \mathbb{E}\left[\varphi(X_t(x))\eta_t(x)\right]$$

As a consequence, for all $t \geq 0$, P_t preserves the space $C^1_b(\mathbb{R})$ of bounded, continuously differentiable functions on $\mathbb{R}$ with a bounded derivative. It turns out that, actually, for all $t > 0$, P_t maps the space $C_b(\mathbb{R})$ of bounded and continuous functions into $C^1_b(\mathbb{R})$. This is a consequence of the following, nowadays well-known, result:

Theorem 6.1 (Bismut-Elworthy-Li Formula) *For all $T > 0$ and $\varphi \in C_b(\mathbb{R})$, the function $P_T\varphi$ is differentiable and we have:*

$$\frac{d}{dx}P_T\varphi(x) = \frac{1}{T}\mathbb{E}\left[\varphi(X_T(x))\int_0^T \eta_s(x)dB_s\right] \tag{6.7}$$

Proof See [7, Theorem 2.1], or [14, Lemma 5.17] for a proof.

Corollary 6.1 *The semi-group $(P_t)_{t\geq 0}$ satisfies the strong Feller property and, for all $T > 0$ and $\varphi : \mathbb{R} \to \mathbb{R}$ bounded and Borel, one has:*

$$\forall x, y \in \mathbb{R}, \quad |P_T\varphi(x) - P_T\varphi(y)| \leq e^L\frac{||\varphi||_\infty}{\sqrt{T \wedge 1}}|x-y|, \tag{6.8}$$

where $||\cdot||_\infty$ denotes the supremum norm.

The following remark is crucial.

Remark 6.2 Inequality (6.8) involves only the dissipativity constant L, not the Lipschitz constant C. This makes the Bismut-Elworthy-Li formula very useful in the study of SPDEs with a dissipative drift.

Proof (Proof of Corollary 6.1) By approximation, it suffices to prove (6.8) for $\varphi \in C_b(\mathbb{R})$. For such a φ and for all $T > 0$, by the Bismut-Elworthy-Li formula, one has:

$$\left|\frac{d}{dx}P_T\varphi(x)\right| \leq \frac{||\varphi||_\infty}{T}\mathbb{E}\left[\left|\int_0^T \eta_s(x)dB_s\right|\right]$$

Remark that the process $(\eta_t(x))_{t\geq 0}$ is locally bounded since it is dominated by $(e^{Lt})_{t\geq 0}$, so that the stochastic integral $\left(\int_0^t \eta_s(x)dB_s\right)_{t\geq 0}$ is an L^2 martingale.

Hence using Jensen's inequality as well as Itô's isometry formula, we obtain:

$$\mathbb{E}\left[|\int_0^T \eta_s(x)dB_s|\right] \leq \sqrt{\mathbb{E}\left[\int_0^T \eta_s(x)^2 ds|\right]}$$
$$\leq \sqrt{\int_0^T e^{2Ls}\, ds}$$

and the last quantity is bounded by $\sqrt{e^{2L}T} = e^L\sqrt{T}$ for all $T \in (0, 1]$. Therefore, we deduce that:

$$\forall x \in \mathbb{R}, \quad \left|\frac{d}{dx} P_T\varphi(x)\right| \leq e^L \frac{||\varphi||_\infty}{\sqrt{T}}$$

so that:

$$\forall x, y \in \mathbb{R}, \quad |P_T\varphi(x) - P_T\varphi(y)| \leq e^L \frac{||\varphi||_\infty}{\sqrt{T}} |x - y|$$

for all $\varphi \in C_b(\mathbb{R})$ and $T \in (0, 1]$. The case $T > 1$ follows at once by using the semi-group property of $(P_t)_{t\geq 0}$:

$$|P_T\varphi(x) - P_T\varphi(y)| = |P_1\left(P_{T-1}\varphi\right)(x) - P_1\left(P_{T-1}\varphi\right)(y)|$$
$$\leq e^L \frac{||P_{T-1}\varphi||_\infty}{\sqrt{1}} |x - y|$$
$$\leq e^L ||\varphi||_\infty |x - y|$$

The claim follows.

Remark 6.3 (A Brief History of the Bismut-Elworthy-Li Formula) A particular form of this formula had originally been derived by Bismut in [2] using Malliavin calculus in the framework of the study of the logarithmic derivative of the fundamental solution of the heat equation on a compact manifold. In [7], Elworthy and Li used a martingale approach, instead of a Malliavin calculus method, to generalize this formula to a large class of diffusion processes on noncompact manifolds with smooth coefficients, and gave also variants of this formula to higher-order derivatives. The key to their proof is to select a stochastic process, which in this case is the stochastic flow, to give a probabilistic representation for the derivative of the semigroup.

The key property allowing the analysis performed in this section is the dissipativity property (6.5). Without this property being true, one would not even expect the Bismut-Elworthy-Li formula to hold. However, in the sequel, we shall prove that results such as Theorem 6.1 and Corollary 6.1 above can also be obtained for a family of diffusions with a non-dissipative drift (informally $L = +\infty$), namely for the Bessel processes of dimension smaller than 1.

6.3 Bessel Processes: Notations and Basic Facts

In the sequel, for any subinterval I of $\mathbb{R}_+$, $C(I)$ will denote the set of continuous functions $I \to \mathbb{R}$. We shall consider this set endowed with the topology of uniform convergence on compact sets, and will denote by $\mathscr{B}(C(I))$ the corresponding Borel σ-algebra.

Consider the canonical measurable space $(C(\mathbb{R}_+), \mathscr{B}(C(I)))$ endowed with the canonical filtration $(\mathscr{F}_t)_{t\geq 0}$. Let $(B_t)_{t\geq 0}$ be a standard linear $(\mathscr{F}_t)_{t\geq 0}$-Brownian motion. For all $x \geq 0$ and $\delta \geq 0$, there exists a unique continuous, predictable, nonnegative process $(X_t^\delta(x))_{t\geq 0}$ satisfying:

$$X_t = x^2 + 2\int_0^t \sqrt{X_s} dB_s + \delta t. \tag{6.9}$$

$(X_t^\delta(x))_{t\geq 0}$ is a squared Bessel process of dimension δ started at x^2, and the process $\rho_t^\delta(x) := \sqrt{X_t^\delta(x)}$ is a δ-dimensional Bessel process started at x. In the sequel, we will also write the latter process as $(\rho_t(x))_{t\geq 0}$, or ρ, when there is no risk of ambiguity.

We recall the following monotonicity property of the family of Bessel processes:

Lemma 6.1 *For all couples $(\delta, \delta'), (x, x') \in \mathbb{R}_+$ such that $\delta \leq \delta'$ and $x \leq x'$, we have, a.s.:*

$$\forall t \geq 0, \qquad \rho_t^\delta(x) \leq \rho_t^{\delta'}(x').$$

Proof By Theorem (3.7) in [10, Section IX], applied to Eq. (6.9), the following property holds a.s.:

$$\forall t \geq 0, \qquad X_t^\delta(x) \leq X_t^{\delta'}(x').$$

Taking the square root on both sides above, we deduce the result.

For all $a \geq 0$, let $T_a(x)$ denote the $(\mathscr{F}_t)_{t\geq 0}$ stopping time defined by:

$$T_a(x) := \inf\{t > 0, \rho_t(x) \leq a\}$$

(we shall also write T_a). We recall the following fact, (see e.g. Proposition 3.6 of [14]):

Proposition 6.1 *The following dichotomy holds:*

- $T_0(x) = +\infty$ *a.s., if* $\delta \geq 2$,
- $T_0(x) < +\infty$ *a.s., if* $0 \leq \delta < 2$.

Applying Itô's lemma to $\rho_t = \sqrt{X_t^\delta(x)}$, we see that ρ satisfies the following relation on the interval $[0, T_0)$:

$$\forall t \in [0, T_0), \qquad \rho_t = x + \frac{\delta - 1}{2} \int_0^t \frac{ds}{\rho_s} + B_t. \tag{6.10}$$

6.4 Derivative in Space of the Bessel Semi-group

Let $\delta \geq 0$. We denote by P_x^δ the law, on $(C(\mathbb{R}_+), \mathscr{B}(C(\mathbb{R}_+)))$, of the δ-dimensional Bessel process started at x, and we write E_x^δ for the corresponding expectation operator. We also denote by $(P_t^\delta)_{t\geq 0}$ the family of transition kernels associated with the δ-dimensional Bessel process, defined by

$$P_t^\delta F(x) := E_x^\delta(F(\rho_t))$$

for all $t \geq 0$ and all $F : \mathbb{R}_+ \to \mathbb{R}$ bounded and Borel. The aim of this section is to prove the following:

Theorem 6.2 *For all $T > 0$ and all $F : \mathbb{R}_+ \to \mathbb{R}$ bounded and Borel, the function $x \to P_t^\delta F(x)$ is differentiable on $\mathbb{R}_+$, and for all $x \geq 0$:*

$$\frac{d}{dx} P_T^\delta F(x) = \frac{x}{T} \left(P_T^{\delta+2} F(x) - P_T^\delta F(x) \right). \tag{6.11}$$

In particular, the function $x \to P_t^\delta F(x)$ satisfies the Neumann boundary condition at 0:

$$\frac{d}{dx} P_T^\delta F(x) \Big|_{x=0} = 0.$$

Remark 6.4 By Theorem 6.2, the derivative of the function $x \mapsto P_T^\delta F(x)$ is a smooth function of $P_T^{\delta+2} F(x)$ and $P_T^\delta F(x)$. Hence, reasoning by induction, we deduce that the function $x \mapsto P_T^\delta F(x)$ is actually smooth on $\mathbb{R}_+$.

Proof The proof we propose here relies on the explicit formula for the transition semi-group of the Bessel processes. We first treat the case $\delta > 0$.

Given $\delta > 0$, let $\nu := \frac{\delta}{2} - 1$, and denote by I_ν the modified Bessel function of index ν. We have (see, e.g., Chap. XI.1 in [10]) :

$$P_t^\delta F(x) = \int_0^\infty p_T^\delta(x, y) F(y) dy$$

where, for all $y \geq 0$:

$$p_t^\delta(x, y) = \frac{1}{T}\left(\frac{y}{x}\right)^\nu y \exp\left(-\frac{x^2+y^2}{2T}\right) I_\nu\left(\frac{xy}{T}\right), \quad \text{if} \quad x > 0,$$
$$p_t^\delta(0, y) = \frac{2^{-\nu}T^{-(\nu+1)}}{\Gamma(\nu+1)} y^{2\nu+1} \exp\left(-\frac{y^2}{2T}\right)$$

where Γ denotes the gamma function. By the power series expansion of the function I_ν we have, for all $x, y \geq 0$:

$$p_T^\delta(x, y) = \frac{1}{T} \exp\left(-\frac{x^2+y^2}{2T}\right) \tilde{p}_T^\delta(x, y) \tag{6.12}$$

with:

$$\tilde{p}_T^\delta(x, y) := \sum_{k=0}^{\infty} \frac{y^{2k+2\nu+1}\, x^{2k}\, (1/2T)^{2k+\nu}}{k!\, \Gamma(k+\nu+1)}.$$

Note that $\tilde{p}_T^\delta(x, y)$ is the sum of a series with infinite radius of convergence in x, hence we can compute its derivative by differentiating under the sum. We have:

$$\begin{aligned}
\frac{\partial}{\partial x}\tilde{p}_T^\delta(x, y) &= \frac{\partial}{\partial x}\left(\sum_{k=0}^{\infty} \frac{x^{2k}\, y^{2k+2\nu+1}\, (1/2T)^{2k+\nu}}{k!\, \Gamma(k+\nu+1)}\right) \\
&= \sum_{k=0}^{\infty} \frac{2k\, x^{2k-1}\, y^{2k+2\nu+1}\, (1/2T)^{2k+\nu}}{k!\, \Gamma(k+\nu+1)} \\
&= \frac{x}{T} \sum_{k=1}^{\infty} \frac{x^{2k-2}\, y^{2k+2\nu+1}\, (1/2T)^{2k+\nu-1}}{(k-1)!\, \Gamma(k+\nu+1)}.
\end{aligned}$$

Hence, performing the change of variable $j = k - 1$, and remarking that $\nu + 1 = \frac{\delta+2}{2} - 1$, we obtain:

$$\begin{aligned}
\frac{\partial}{\partial x}\tilde{p}_T^\delta(x, y) &= \frac{x}{T} \sum_{j=0}^{\infty} \frac{x^{2(j+1)-2}\, y^{2(j+1)+2\nu+1}\, (1/2T)^{2(j+1)+\nu-1}}{j!\, \Gamma((j+1)+\nu+1)} \\
&= \frac{x}{T} \sum_{j=0}^{\infty} \frac{x^{2j}\, y^{2j+2(\nu+1)+1}\, (1/2T)^{2j+(\nu+1)}}{j!\, \Gamma(j+(\nu+1)+1)} \\
&= \frac{x}{T}\, \tilde{p}_T^{\delta+2}(x, y).
\end{aligned}$$

As a consequence, differentiating equality (6.12) with respect to x, we obtain:

$$\frac{\partial}{\partial x} p_T^\delta(x,y) = \left(-\frac{x}{T}\ \tilde{p}_T^\delta(x,y) + \frac{\partial}{\partial x}\tilde{p}_T^\delta(x,y)\right)\frac{1}{T}\exp\left(-\frac{x^2+y^2}{2T}\right)$$
$$= \frac{x}{T}\ \left(-p_T^\delta(x,y) + p_T^{\delta+2}(x,y)\right).$$

Hence, we deduce that the function $x \mapsto P_T^\delta F(x)$ is differentiable, with a derivative given by (6.11).

Now suppose that $\delta = 0$. We have, for all $x \geq 0$:

$$P_T^0 F(x) = \exp\left(-\frac{x^2}{2T}\right) F(0) + \int_0^\infty p_T(x,y)F(y)dy \tag{6.13}$$

where, for all $y \geq 0$:

$$p_T(x,y) = \frac{1}{T}\exp\left(-\frac{x^2+y^2}{2T}\right)\tilde{p}_T(x,y)$$

with:

$$\tilde{p}_T(x,y) := x\ I_1\left(\frac{xy}{T}\right) = \sum_{k=0}^\infty \frac{x^{2k+2}\ (y/2T)^{2k+1}}{k!(k+1)!}.$$

Here again, we can differentiate the sum term by term, so that, for all $x, y \geq 0$:

$$\frac{\partial}{\partial x}\tilde{p}_T(x,y) = \frac{x}{T}\ \sum_{k=0}^\infty \frac{x^{2k}y^{2k+1}(1/2T)^{2k}}{k!^2}$$
$$= \frac{x}{T}\ \tilde{p}_T^2(x,y).$$

Therefore, for all $x, y \geq 0$, we have:

$$\frac{\partial}{\partial x} p_T(x,y) = \left(-\frac{x}{T}\tilde{p}_T(x,y) + \frac{\partial}{\partial x}\tilde{p}_T(x,y)\right)\frac{1}{T}\exp\left(-\frac{x^2+y^2}{2T}\right)$$
$$= \frac{x}{T}\left(-\tilde{p}_T(x,y) + \tilde{p}_T^2(x,y)\right)\frac{1}{T}\exp\left(-\frac{x^2+y^2}{2T}\right)$$
$$= \frac{x}{T}\left(-p_T(x,y) + p_T^2(x,y)\right)$$

Hence, differentiating (6.13) with respect to x, and using the dominated convergence theorem to differentiate inside the integral, we obtain:

$$\begin{aligned}\frac{\partial}{\partial x} P_T^0 F(x) &= -\frac{x}{T} \exp(-\frac{x^2}{2T}) F(0) + \frac{x}{T} \int_0^\infty \left(-p_T(x,y) + p_T^2(x,y)\right) F(y) dy \\ &= \frac{x}{T} \left(-P_T^0 F(x) + P_T^2 F(x)\right),\end{aligned}$$

which yields the claim.

Remark 6.5 Formula (6.11) can also be derived using the Laplace transform of the one-dimensional marginals of the squared Bessel processes. Indeed, denote by $(Q_t^\delta)_{t\geq 0}$ the family of transition kernels of the δ-dimensional squared Bessel process. Then for all $\delta \geq 0$, $x \geq 0$, $T > 0$, and all function f of the form $f(x) = \exp(-\lambda x)$ with $\lambda \geq 0$, one has:

$$Q_T^\delta f(x) = \exp\left(-\frac{\lambda x}{1 + 2\lambda T}\right) (1 + 2\lambda T)^{-\delta/2}$$

(see [10, Chapter XI, Cor. (1.3)]). For such test functions f, we check at once that the following equality holds:

$$\frac{d}{dx} Q_T^\delta f(x) = \frac{1}{2T} \left(Q_T^{\delta+2} f(x) - Q_T^\delta f(x)\right).$$

By linearity and by the Stone-Weierstrass theorem, we deduce that this equality holds for all bounded, continuous functions f. Then an approximation argument enables to deduce the equality for all functions $f : \mathbb{R}_+ \to \mathbb{R}$ Borel and bounded. Finally, remarking that for all bounded Borel function F on $\mathbb{R}_+$ we have

$$P_T^\delta F(x) = Q_T^\delta f(x^2)$$

with $f(x) := F(\sqrt{x})$, we deduce that:

$$\begin{aligned}\frac{d}{dx} P_T^\delta F(x) &= 2x \, \frac{d}{dx} (Q_T^\delta f)(x^2) \\ &= \frac{x}{T} \left(Q_T^{\delta+2} f(x^2) - Q_T^\delta f(x^2)\right) \\ &= \frac{x}{T} \left(P_T^{\delta+2} F(x) - P_T^\delta F(x)\right)\end{aligned}$$

which yields the equality (6.11).

Corollary 6.2 *The semi-group $(P_t^\delta)_{t\geq 0}$ has the strong Feller property. More precisely, for all $T > 0$, $R > 0$, $x, y \in [0, R]$ and $F : \mathbb{R}_+ \to \mathbb{R}$ bounded and Borel,* *we have:*

$$|P_T^\delta F(x) - P_T^\delta F(y)| \leq \frac{2R||F||_\infty}{T}|y-x|. \tag{6.14}$$

Proof By Theorem 6.2, for all $x, y \in [0, R]$ such that $x \leq y$, we have:

$$\begin{aligned} |P_t^\delta F(x) - P_t^\delta F(y)| &= \left| \int_x^y \frac{u}{T} \left(P_T^{\delta+2} F(u) - P_T^\delta F(u) \right) du \right| \\ &\leq \frac{2||F||_\infty}{T} \int_x^y u \, du \\ &\leq \frac{2R||F||_\infty}{T}|y-x|. \end{aligned}$$

Remark 6.6 The bound (6.14) is in $1/T$, which is not very satisfactory for T small. However, in the sequel, we will improve this bound by getting a better exponent on T, at least for $\delta \geq 2(\sqrt{2}-1)$ (see inequality (6.28) below).

6.5 Differentiability of the Flow

In the following, we are interested in finding a probabilistic interpretation of Theorem 6.2, in terms of the Bismut-Elworthy-Li formula. To do so we study, for all $\delta \geq 0$, and all couple $(t, x) \in \mathbb{R}_+ \times \mathbb{R}_+^*$, the differentiability at x of the function:

$$\begin{aligned} \rho_t : \mathbb{R}_+ &\to \mathbb{R}_+ \\ y &\mapsto \rho_t^\delta(y). \end{aligned}$$

In this endeavour, we first need to choose an appropriate modification of the process $(\rho_t(x))_{t\geq 0, x>0}$. We have the following result:

Proposition 6.2 *Let $\delta \geq 0$ be fixed. There exists a modification $(\tilde{\rho}_t^\delta(x))_{x,t\geq 0}$ of the process $(\rho_t^\delta(x))_{x,t\geq 0}$ such that, a.s., for all $x, x' \in \mathbb{R}_+$ with $x \leq x'$, we have:*

$$\forall t \geq 0, \qquad \tilde{\rho}_t^\delta(x) \leq \tilde{\rho}_t^\delta(x'). \tag{6.15}$$

Proof For all $q, q' \in \mathbb{Q}_+$, such that $q \leq q'$, by Lemma 6.1, the following property holds a.s.:

$$\forall t \geq 0, \qquad \rho_t^\delta(q) \leq \rho_t^\delta(q').$$

For all $x \in \mathbb{R}_+$, we define the process $\tilde{\rho}^\delta(x)$ by:

$$\forall t \geq 0, \qquad \tilde{\rho}_t^\delta(x) := \inf_{q \in \mathbb{Q}_+, q \geq x} \rho_t^\delta(q).$$

Then $(\tilde{\rho}_t^\delta(x))_{x,t\geq 0}$ yields a modification of the process $(\rho_t^\delta(x))_{x,t\geq 0}$ with the requested property.

In the sequel, when $\delta \geq 0$ is fixed and there is no ambiguity, we shall write $\tilde{\rho}$ instead of $\tilde{\rho}^\delta$.

Remark 6.7 Given $\delta \geq 0$, we may not have, almost-surely, joint continuity of all the functions $t \mapsto \tilde{\rho}_t(x)$, $x \geq 0$. Note however that, by definition, for all $x \geq 0$, $x \in \mathbb{Q}$, we have a.s.:

$$\forall t \geq 0, \qquad \tilde{\rho}_t(x) = \rho_t(x),$$

so that, a.s., $t \mapsto \tilde{\rho}_t(x)$ is continuous and satisfies:

$$\forall t \in [0, T_0(x)), \qquad \tilde{\rho}_t(x) = x + \frac{\delta - 1}{2} \int_0^t \frac{ds}{\tilde{\rho}_s(x)} + B_t.$$

As a consequence, by countability of $\mathbb{Q}$, there exists an almost sure event $\mathscr{A} \in \mathscr{F}$ on which, for all $x \in \mathbb{Q}_+$, the function $t \mapsto \tilde{\rho}_t(x)$ is continuous and satisfies:

$$\forall t \in [0, T_0(x)), \qquad \tilde{\rho}_t(x) = x + \frac{\delta - 1}{2} \int_0^t \frac{ds}{\tilde{\rho}_s(x)} + B_t.$$

Actually, in Corollary 6.5 of the Appendix, we will prove the stronger fact that, almost-surely, we have:

$$\forall x \geq 0, \quad \forall t \in [0, \tilde{T}_0(x)), \qquad \tilde{\rho}_t(x) = x + \frac{\delta - 1}{2} \int_0^t \frac{ds}{\tilde{\rho}_s(x)} + B_t,$$

where, for all $x \geq 0$:

$$\tilde{T}_0(x) := \inf\{t > 0, \tilde{\rho}_t(x) = 0\}$$

In this section, as well as the Appendix, we always work with the modification $\tilde{\rho}$. Similarly, we work with $\tilde{T}_0(x)$ instead of $T_0(x)$, for all $\delta, x \geq 0$. We will write again ρ and T_0 instead of $\tilde{\rho}$ and $\tilde{T}_0$. Note that, a.s., the function $x \mapsto T_0(x)$ is non-decreasing on $\mathbb{R}_+$.

Proposition 6.3 *Let $\delta \geq 0$, $t > 0$ and $x > 0$. Then, a.s., the function ρ_t is differentiable at x, and its derivative there is given by:*

$$\left.\frac{d\rho_t(y)}{dy}\right|_{y=x} \overset{a.s.}{=} \eta_t(x) := \mathbf{1}_{t<T_0(x)} \exp\left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2}\right). \tag{6.16}$$

The proof of this proposition is quite technical. Since, moreover, the result will not be necessary in the sequel, we prefer to postpone the proof to the Appendix of the article.

Remark 6.8 In particular, when $\delta = 1$, the above formula reduces to:

$$\frac{d\rho_t(y)}{dy}\Big|_{y=x} \overset{a.s.}{=} \mathbf{1}_{t<T_0(x)} \tag{6.17}$$

a formula which was already well-known (see e.g. [1, Lemma A.1]).

Remark 6.9 Note that the indicator function $\mathbf{1}_{t<T_0(x)}$ in the right-hand side of (6.16) is related to the behavior of the Bessel process at the boundary 0. It is reminiscent of Theorem 1 in [6], where a similar indicator function appears in the expression of the spatial derivative of the flow of vector-valued solutions to SDEs with reflection.

Remark 6.10 Proposition 6.3 shows that, for all $t, x > 0$, the function ρ_t is almost-surely differentiable at x. We may, however, ask if, a.s., the function ρ_t is differentiable on the whole of $\mathbb{R}_+^*$. The case where $\delta > 1$ was treated in detail in [13], where it was shown that, a.s., for all $t \geq 0$ the function $x \mapsto \rho_t(x)$ is differentiable on $\mathbb{R}_+^*$, and that the derivative $\frac{d\rho_t(x)}{dx}$ is continuous in $(t, x) \in \mathbb{R}_+ \times \mathbb{R}_+^*$. However, as δ gets smaller than 1, the regularity of the process $(\rho_t(x))_{t\geq 0, x>0}$ becomes much worse. Note that $\delta = 1$ corresponds to the case of the flow of reflected Brownian motion on the half-line; in that case the flow is no longer continuously differentiable as suggested by (6.17). Many works have been carried out on the study of the flow of reflected Brownian motion on domains in higher dimension (see e.g. [3] and [12]) or on manifolds with boundary (see e.g. [1]). By contrast, the regularity of Bessel flows of dimension $\delta < 1$ seems to be a very open problem.

In the remainder of the article, however, we shall not need any regularity results on the Bessel flow. Instead, for all fixed $x > 0$, we shall study the process $(\eta_t(x))_{t\geq 0}$ defined above as a process in itself.

6.6 Properties of η

In the sequel, for all $x \geq 0$, we shall consider the process $(\eta_t(x))_{t\geq 0}$ defined as above:

$$\eta_t(x) := \mathbf{1}_{t<T_0(x)} \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right). \tag{6.18}$$

When there is no ambiguity we shall drop the x from our notation and denote this process by η.

6.6.1 Regularity of the Sample Paths of η

We are interested in the continuity property of the process η. It turns out that, as δ decreases, η becomes more and more singular, as shown by the following result.

Proposition 6.4 *If $\delta > 1$, then a.s. η is bounded and continuous on $\mathbb{R}_+$.*

If $\delta = 1$, then a.s. η is constant on $[0, T_0)$ and $[T_0, +\infty)$, but has a discontinuity at T_0.

If $\delta \in [0, 1)$, then a.s. η is continuous away from T_0, but it diverges to $+\infty$ as $t \uparrow T_0$.

Proof When $\delta \geq 2$, $T_0 = \infty$ almost-surely, so that, by (6.18), the following equality of processes holds:

$$\eta_t = \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right).$$

Hence, a.s., η takes values in [0, 1] and is continuous on $\mathbb{R}_+$. To treat the case $\delta < 2$ we need a lemma:

Lemma 6.2 *Let $\delta < 2$ and $x > 0$. Then the integral:*

$$\int_0^{T_0} \frac{ds}{(\rho_s(x))^2}$$

is infinite a.s.

We admit this result for the moment. Then, when $\delta \in (1, 2)$, η takes values in [0, 1], is continuous away from T_0 and, almost-surely, as $t \uparrow T_0$:

$$\eta_t = \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right) \longrightarrow 0.$$

Since, $\eta_t = 0$ for all $t \geq T_0$, η is continuous and the claim follows. When $\delta = 1$,

$$\eta_t(x) := \mathbf{1}_{t<T_0(x)}$$

so the claim follows at once. Finally, if $\delta \in [0, 1)$, then η is continuous away from T_0, but by the above lemma, a.s., as $t \uparrow T_0$:

$$\eta_t = \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right) \longrightarrow +\infty$$

so the claim follows.

We now prove Lemma 6.2

Proof (Proof of Lemma 6.2) The proof is in two steps. In a first step we prove the lemma when ρ is replaced with a Brownian motion started at some positive point, and in a second step we invoke a representation theorem of Bessel processes as time-changes of some power of the Brownian motion to conclude.

First Step Let (β_t) be a Brownian motion started from some $y > 0$, and let T_0 denote its hitting time of the origin. Then the integral:

$$\int_0^{T_0} \frac{ds}{(\beta_s(y))^2}$$

is a.s. infinite. Indeed, denote by $h : [0, \infty) \to \mathbb{R}_+$ the function given by:

$$h(t) := \begin{cases} \sqrt{t|\log(1/t)|}, & \text{if } t > 0, \\ 0, & \text{if } t = 0. \end{cases}$$

Let $A > 0$. By Levy's modulus of continuity (see Theorem (2.7), Chapter I, in [10]), there exists a $\kappa > 0$, such that the event

$$\mathscr{M} := \{ \forall s, t \in [0, 1], \quad |\beta_t - \beta_s| \leq \kappa \; h(|s-t|) \}$$

has probability one. Therefore, by scale invariance of Brownian motion, setting $\kappa_A := \sqrt{A}\kappa$, one deduces that the event

$$\mathscr{M}_A := \{ \forall s, t \in [0, A], \quad |\beta_t - \beta_s| \leq \kappa_A \; h(|s-t|) \}$$

also has probability one. Moreover, under the event $\{T_0 < A\} \cap \mathscr{M}_A$, we have, for small $h > 0$.

$$\beta_{T_0-h}^2 = |\beta_{T_0-h} - \beta_{T_0}|^2 \leq {\kappa_A}^2 \; h \log(1/h).$$

Since $\frac{1}{h\log(1/h)}$ is not integrable as $h \to 0^+$, we deduce that, under the event $\{T_0 < A\} \cap \mathscr{M}_A$, we have $\int_0^{T_0} \frac{ds}{(\beta_s)^2} = +\infty$. Therefore:

$$\mathbb{P}[T_0 < A] = \mathbb{P}[\{T_0 < A\} \cap \mathscr{M}_A] \leq \mathbb{P}\left(\int_0^{T_0} \frac{ds}{(\beta_s)^2} = +\infty \right).$$

Since $T_0 < +\infty$ a.s., we have $\lim_{A\to\infty} \mathbb{P}[T_0 < A] = 1$. Hence, letting $A \to \infty$ in the above, we deduce that:

$$\mathbb{P}\left(\int_0^{T_0} \frac{ds}{(\beta_s)^2} = +\infty \right) = 1$$

as claimed.

Second Step Now consider the original Bessel process $(\rho_t(x))_{t\geq 0}$. Suppose that $\delta \in (0,2)$. Then, by Thm 3.5 in [14], the process $(\rho_t(x))_{t\geq 0}$ is equal in law to $(|\beta_{\gamma(t)}|^{\frac{1}{2-\delta}})_{t\geq 0}$, where β is a Brownian motion started from $y := x^{2-\delta}$, and $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is the inverse of the increasing function $A : \mathbb{R}_+ \to \mathbb{R}_+$ given by:

$$\forall u \geq 0, \qquad A(u) = \frac{1}{(2-\delta)^2} \int_0^u |\beta_s|^{\frac{2(\delta-1)}{2-\delta}} ds.$$

Therefore, denoting by T_0^β the hitting time of 0 by the Brownian motion β, we have:

$$\begin{aligned} \int_0^{T_0} \frac{ds}{(\rho_s(x))^2} &\overset{(d)}{=} \int_0^{A(T_0^\beta)} \frac{ds}{|\beta_{\gamma(s)}|^{\frac{2}{2-\delta}}} \\ &= \int_0^{T_0^\beta} \frac{1}{|\beta_u|^{\frac{2}{2-\delta}}} \frac{1}{(2-\delta)^2} |\beta_u|^{\frac{2(\delta-1)}{2-\delta}} du \\ &= \frac{1}{(2-\delta)^2} \int_0^{T_0^\beta} \frac{du}{{\beta_u}^2} \end{aligned}$$

where we have used the change of variable $u = \gamma(s)$ to get from the first line to the second one. By the first step, the last integral is infinite a.s., so the claim follows.

There still remains to treat the case $\delta = 0$. By Thm 3.5 in [14], in that case, the process $(\rho_t(x))_{t\geq 0}$ is equal in law to $\left(\left(\beta_{\gamma(t)\wedge T_0^\beta}\right)^{1/2}\right)_{t\geq 0}$, where β is a Brownian motion started from $y := x^2$, T_0^β is its hitting time of 0 and $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is the inverse of the increasing function $A : \mathbb{R}_+ \to \mathbb{R}_+$ given by:

$$\forall u \geq 0, \qquad A(u) = \frac{1}{4} \int_0^{u \wedge T_0^\beta} \beta_s^{-1} ds.$$

Then, the same computations as above yield the equality in law:

$$\int_0^{T_0} \frac{ds}{(\rho_s(x))^2} \overset{(d)}{=} \frac{1}{4} \int_0^{T_0^\beta} \frac{du}{{\beta_u}^2}$$

so the result follows as well.

6.6.2 Study of a Martingale Related to η

Let $\delta \in [0, 2)$ and $x > 0$ be fixed. In the previous section, we have shown that, a.s.:

$$\int_0^t \frac{ds}{\rho_s(x)^2} \underset{t \to T_0(x)}{\longrightarrow} +\infty$$

As a consequence, for $\delta \in [0, 1)$, a.s., the modification η_t of the derivative at x of the stochastic flow ρ_t diverges at $T_0(x)$:

$$\eta_t(x) = \mathbf{1}_{t<T_0(x)} \exp\left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2}\right) \underset{t \uparrow T_0(x)}{\longrightarrow} +\infty.$$

However, since $\rho_t(x) \longrightarrow 0$ as $t \to T_0(x)$, this does not exclude the possibility that the product $\rho_t(x)\eta_t(x)$ converges as $t \to T_0(x)$. This motivates to study the process:

$$D_t := \rho_t(x)\eta_t(x) = \mathbf{1}_{t<T_0(x)}\rho_t(x) \exp\left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2}\right). \tag{6.19}$$

As a matter of fact, we will show that $(D_t)_{t\geq 0}$ is an L^p continuous martingale for some $p \geq 1$.

Remark 6.11 The process $(D_t)_{t\geq 0}$ appears as (one half times) the derivative of the stochastic flow associated with the squared Bessel process $X_t(x) = (\rho_t(x))^2$. Indeed, by applying formally the chain rule, we have, for all $t \geq 0$ and $x > 0$:

$$\frac{dX_t(x)}{dx} = 2\rho_t(x)\eta_t(x).$$

6.6.3 Continuity of $(D_t)_{t\geq 0}$

In this subsection we show that the process $(D_t)_{t\geq 0}$ has a.s. continuous sample paths. By the expression (6.19), continuity holds as soon as $T_0(x) = \infty$ a.s., i.e. as soon as $\delta \geq 2$. On the other hand, if $\delta \in [0, 2)$ it suffices to prove that, a.s., $D_t \to 0$ as $t \uparrow T_0(x)$. This is the content of the following proposition.

Proposition 6.5 *For all $\delta \in [0, 2)$ and $x > 0$, with probability one:*

$$\rho_t(x) \exp\left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2}\right) \underset{t \to T_0(x)}{\longrightarrow} 0.$$

Proof If $\delta \in [1,2)$, then $\exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right) \le 1$ for all $t \ge 0$. Since $\rho_t \longrightarrow 0$ as $t \to T_0(x)$, the claim follows at once.

If $\delta \in [0,1)$, on the other hand, $\exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right) \underset{t\uparrow T_0(x)}{\longrightarrow} +\infty$ whereas $\rho_t \underset{t\to T_0(x)}{\longrightarrow} 0$ so a finer analysis is needed. We have:

$$\log\left[\frac{\rho_t}{x}\exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right)\right] = \log\frac{\rho_t}{x} + \frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s^2}$$

Now, recall that a.s., for all $t < T_0$, we have:

$$\rho_t = x + \frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s} + B_t$$

Hence, defining for all integer $n \ge 1$ the $(\mathscr{F}_t)_{t\ge 0}$-stopping time τ_n as:

$$\tau_n := \inf\{t > 0,\ \rho_t \le 1/n\} \wedge n,$$

we have:

$$\rho_{t\wedge\tau_n} = x + \frac{\delta-1}{2}\int_0^{t\wedge\tau_n} \frac{ds}{\rho_s} + B_{t\wedge\tau_n}.$$

Hence, by Itô's lemma, we deduce that:

$$\log\frac{\rho_{t\wedge\tau_n}}{x} = \frac{\delta-1}{2}\int_0^{t\wedge\tau_n}\frac{ds}{\rho_s^2} + \int_0^{t\wedge\tau_n}\frac{dB_s}{\rho_s} - \frac{1}{2}\int_0^{t\wedge\tau_n}\frac{ds}{\rho_s^2}$$

so that:

$$\log\frac{\rho_{t\wedge\tau_n}}{x} + \frac{1-\delta}{2}\int_0^{t\wedge\tau_n}\frac{ds}{\rho_s^2} = \int_0^{t\wedge\tau_n}\frac{dB_s}{\rho_s} - \frac{1}{2}\int_0^{t\wedge\tau_n}\frac{ds}{\rho_s^2}. \tag{6.20}$$

Consider now the random time change:

$$A\colon [0,T_0) \to \mathbb{R}_+$$
$$t \mapsto A_t := \int_0^t \frac{ds}{\rho_s^2}.$$

Note that A is differentiable with strictly positive derivative. Moreover, since $A_t \underset{t\to T_0}{\longrightarrow} +\infty$ a.s. by Lemma 6.2, we deduce that A is a.s. onto. Hence, a.s., A

is a diffeomorphism $[0, T_0) \to \mathbb{R}_+$, the inverse of which we denote by

$$C : \mathbb{R}_+ \to [0, T_0)$$
$$u \mapsto C_u.$$

Let $\beta_u := \int_0^{C_u} \frac{dB_r}{\rho_r}, u \geq 0$. Then β is a local martingale started at 0 with quadratic variation $\langle \beta, \beta \rangle_u = u$, so by Lévy's theorem it is a Brownian motion. The equality (6.20) can now be rewritten:

$$\log \frac{\rho_{t \wedge \tau_n}}{x} + \frac{1-\delta}{2} \int_0^{t \wedge \tau_n} \frac{ds}{\rho_s^2} = \beta_{A_{t \wedge \tau_n}} - \frac{1}{2} A_{t \wedge \tau_n}.$$

Letting $n \to \infty$, we obtain, for all $t < T_0$:

$$\log \frac{\rho_t}{x} + \frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s^2} = \beta_{A_t} - \frac{1}{2} A_t.$$

By the asymptotic properties of Brownian motion (see Corollary (1.12), Chapter II in [10]), we know that, a.s.:

$$\limsup_{s \to +\infty} \frac{\beta_s}{h(s)} = 1$$

where $h(s) := \sqrt{2s \log \log s}$. In particular, a.s., there exists $T > 0$ such that, for all $t \geq T$, we have $\beta_t \leq 2h(t)$. Since, a.s., $A_t \underset{t \to T_0}{\longrightarrow} +\infty$, we deduce that:

$$\limsup_{t \to +\infty} \left(\beta_{A_t} - \frac{1}{2} A_t \right) \leq \limsup_{t \to +\infty} \left(2h(A_t) - \frac{1}{2} A_t \right)$$
$$= -\infty.$$

Hence, a.s.:

$$\log \left[\frac{\rho_t}{x} \exp \left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2} \right) \right] \underset{t \uparrow T_0(x)}{\longrightarrow} -\infty$$

i.e.

$$\rho_t \exp \left(\frac{1-\delta}{2} \int_0^t \frac{ds}{\rho_s(x)^2} \right) \underset{t \uparrow T_0(x)}{\longrightarrow} 0$$

as claimed.

6.6.4 Martingale Property of $(D_t)_{t\geq 0}$

Let $\delta \geq 0$ and $x > 0$ be fixed. We show in this section that $(D_t)_{t\geq 0}$ is an $(\mathscr{F}_t)_{t\geq 0}$ martingale which, up to a positive constant, corresponds to a Girsanov-type change of probability measure.

Recall that, by definition:

$$D_t = \mathbf{1}_{t<T_0(x)}\rho_t(x)\exp\left(-\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right). \tag{6.21}$$

Notation For all $a \geq 0$ and $t \geq 0$, we denote by $P_x^a\big|_{\mathscr{F}_t}$ the image of the probability measure P_x^a under the restriction map:

$$\begin{aligned}(C(\mathbb{R}_+), \mathscr{B}(C(\mathbb{R}_+))) &\to (C([0,t]), \mathscr{F}_t)\\ w &\mapsto w|_{[0,t]}\end{aligned}$$

The following proposition is a generalization of the absolute continuity results obtained in [9].

Proposition 6.6 *Let $\delta \geq 0$ and $x > 0$. Then, for all $t \geq 0$, the law $P_x^{\delta+2}{}_{|\mathscr{F}_t}$ is absolutely continuous w.r.t. the law $P_x^{\delta}{}_{|\mathscr{F}_t}$, and the corresponding Radon-Nikodym derivative is given by:*

$$\left.\frac{dP_x^{\delta+2}}{dP_x^{\delta}}\right|_{\mathscr{F}_t}(\rho) \overset{a.s.}{=} \mathbf{1}_{t<T_0(x)}\frac{\rho_t(x)}{x}\exp\left(-\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right).$$

Proof Fix $\epsilon > 0$. Under $P_x^{\delta}{}_{|\mathscr{F}_t}$, the canonical process ρ stopped at T_ϵ satisfies the following SDE on $[0,t]$:

$$\rho_{s\wedge T_\epsilon} = x + \frac{\delta-1}{2}\int_0^{s\wedge T_\epsilon}\frac{ds}{\rho_s} + B_{s\wedge T_\epsilon}.$$

Consider the process M^ϵ defined on $[0,t]$ by:

$$M_s^\epsilon := \int_0^{s\wedge T_\epsilon}\frac{dB_u}{\rho_u}$$

M^ϵ is an L^2 martingale on $[0,t]$. The exponential local martingale thereto associated is:

$$\mathscr{E}(M^\epsilon)_s = \exp\left(\int_0^{s\wedge T_\epsilon}\frac{dB_u}{\rho_u} - \frac{1}{2}\int_0^{s\wedge T_\epsilon}\frac{du}{\rho_u^2}\right).$$

Since, by Itô's lemma:

$$\log\left(\frac{\rho_{s\wedge T_\epsilon}}{x}\right) = \int_0^{s\wedge T_\epsilon} \frac{dB_u}{\rho_u} + \left(\frac{\delta}{2}-1\right)\int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u^2},$$

we have:

$$\begin{aligned}\mathscr{E}(M^\epsilon)_s &= \exp\left[\log\left(\frac{\rho_{s\wedge T_\epsilon}}{x}\right) - \frac{\delta-1}{2}\int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u^2}\right] \\ &= \frac{\rho_{s\wedge T_\epsilon}}{x}\exp\left[-\frac{\delta-1}{2}\int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u^2}\right].\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M^\epsilon, M^\epsilon\rangle_t\right)\right] &\leq \exp\left(\frac{t}{2\epsilon}\right) \\ &< \infty\end{aligned}$$

so that, by Novikov's criterion, $\mathscr{E}(M^\epsilon)$ is a uniformly integrable martingale on $[0, t]$. So we may consider the probability measure $\mathscr{E}(M^\epsilon)_t \ P_x^\delta\big|_{\mathscr{F}_t}$.

Note also that:

$$\langle M^\epsilon, B\rangle_t = \int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u}.$$

Hence, by Girsanov's theorem, under the probability measure $\mathscr{E}(M^\epsilon)_t \ P_x^\delta\big|_{\mathscr{F}_t}$, the process:

$$\rho_{s\wedge T_\epsilon} - x - \frac{\delta+1}{2}\int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u}$$

is a local martingale, with quadratic variation given by $s \wedge T_\epsilon$. Therefore, by Theorem (1.7) in Chapter V of [10], there exists, on some enlarged probability space, a Brownian motion β such that, a.s.:

$$\forall s \in [0, t], \ \rho_{s\wedge T_\epsilon} = x + \frac{\delta+1}{2}\int_0^{s\wedge T_\epsilon} \frac{du}{\rho_u} + \beta_{s\wedge T_\epsilon}.$$

Denote by $\bar{\rho}$ the unique strong solution on $[0, t]$ of the SDE:

$$\bar{\rho}_s = x + \frac{\delta+1}{2}\int_0^s \frac{du}{\bar{\rho}_u} + \beta_s.$$

Then, by strong uniqueness of the solution to this SDE, we deduce that, under $\mathscr{E}(M^\epsilon)_t \; P_x^\delta\big|_{\mathscr{F}_t}$, a.s.:

$$\forall s \in [0,t], \; s < T_\epsilon \implies \rho_s = \bar{\rho}_s.$$

Since $\bar{\rho}$ has the law of a $\delta+2$-dimensional Bessel process started at x, we deduce that, for all $F : C([0,T],\mathbb{R}_+) \to \mathbb{R}_+$ Borel, we have:

$$E_x^\delta\left[\mathscr{E}(M^\epsilon)_t F(\rho)\mathbf{1}_{t<T_\epsilon}\right] = E_x^{\delta+2}[F(\rho)\mathbf{1}_{t<T_\epsilon}]$$

i.e.:

$$E_x^\delta\left[\frac{\rho_t}{x}\exp\left(-\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right)F(\rho)\mathbf{1}_{t<T_\epsilon}\right] = E_x^{\delta+2}[F(\rho)\mathbf{1}_{t<T_\epsilon}].$$

Letting $\epsilon \to 0$, by the monotone convergence theorem, we obtain:

$$E_x^\delta\left[\frac{\rho_t}{x}\exp\left(-\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right)F(\rho)\mathbf{1}_{t<T_0}\right] = E_x^{\delta+2}[F(\rho)\mathbf{1}_{t<T_0}].$$

But, since $P_x^{\delta+2}[T_0 < +\infty] = 0$, this yields:

$$E_x^\delta\left[\frac{\rho_t}{x}\exp\left(-\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right)F(\rho)\mathbf{1}_{t<T_0}\right] = E_x^{\delta+2}[F(\rho)]$$

as stated.

Remark 6.12 Proposition 6.6 is actually a particular case of a more general result. Indeed, for all $x > 0$, $t \geq 0$, and $\delta' \geq \delta \geq 0$, such that $\delta' \geq 2$, $P_x^{\delta'}{}_{|\mathscr{F}_t}$ is absolutely continuous w.r.t. the law $P_x^\delta{}_{|\mathscr{F}_t}$, and the corresponding Radon-Nikodym derivative is given by:

$$\left.\frac{dP_x^{\delta'}}{dP_x^\delta}\right|_{\mathscr{F}_t}(\rho) \overset{a.s.}{=} \mathbf{1}_{t<T_0(x)}\left(\frac{\rho_t(x)}{x}\right)^{\frac{\delta'-\delta}{2}}\exp\left[-\frac{\delta'-\delta}{2}\left(\frac{\delta'+\delta}{4}-1\right)\int_0^t \frac{ds}{\rho_s^2}\right]. \tag{6.22}$$

The proof of this fact is in all respect similar to that of Proposition 6.6 above.

Corollary 6.3 $(D_t)_{t\geq 0}$ *is an* $(\mathscr{F}_t)_{t\geq 0}$ *continuous martingale*

Proof The process $(D_t)_{t\geq 0}$ is continuous. Moreover, for all $t \geq 0$, $\frac{1}{x}D_t$ is the Radon-Nikodym derivative of $P_x^{\delta+2}{}_{|\mathscr{F}_t}$ w.r.t. $P_x^\delta{}_{|\mathscr{F}_t}$. Therefore $(\frac{1}{x}D_t)_{t\geq 0}$ is an $(\mathscr{F}_t)_{t\geq 0}$ martingale, so $(D_t)_{t\geq 0}$ is a martingale as well, and the claim follows.

6.6.5 Moment Estimates for the Martingale $(D_t)_{t\geq 0}$

In this section, we prove that the martingale $(D_t)_{t\geq 0}$ is actually in L^p for some $p \geq 1$. We first recall the following fact:

Lemma 6.3 *For all $a \geq 0$, $t \geq 0$, and $m \geq 0$, we have:*

$$E_x^a(\rho_t^m) < \infty.$$

Proof Denote by d any integer such that $d \geq a$. By Lemma 6.1, we have:

$$E_x^a(\rho_t^m) \leq E_x^d(\rho_t^m)$$

Since P_x^d is the law of $(||B_s||)_{s\geq 0}$, where $(B_s)_{s\geq 0}$ is a d-dimensional Brownian motion and $||\cdot||$ is the Euclidean norm in $\mathbb{R}^d$ (see [10, Chapter 11]), this inequality can be rewritten as:

$$E_x^a(\rho_t^m) \leq \mathbb{E}\left(||B_t||^m\right)$$

Since B_t is a Gaussian random variable, $\mathbb{E}\left(||B_t||^m\right)$ is finite, and the result follows.

Proposition 6.7 *$(D_t)_{t\geq 0}$ is an L^p martingale for all finite positive number p such that $p \leq p(\delta)$, where $p(\delta) \in [1, +\infty]$ is given by:*

$$p(\delta) := \begin{cases} \frac{(2-\delta)^2}{4(1-\delta)} & \text{if } \delta < 1, \\ +\infty & \text{if } \delta \geq 1. \end{cases} \tag{6.23}$$

Moreover the above statement is sharp: for $\delta < 1$ and $t > 0$, the random variable D_t is not in L^p for $p > p(\delta)$.

Remark 6.13 We emphasize that p is *finite* in the above result. Indeed D_t is never in L^∞ even if $\delta \geq 1$; for example, when $\delta = 1$, $D_t = \rho_t \mathbf{1}_{t<T_0(x)}$ which is clearly not bounded a.s.

Proof (Proof of Proposition 6.7) If $\delta \geq 1$, then, for all $t \geq 0$, $D_t \leq \rho_t$. Hence, for all $p \in (0, +\infty)$:

$$\mathbb{E}\left(D_t^p\right) \leq E_x^\delta({\rho_t}^p)$$

which is finite by Lemma 6.3.

On the other hand, if $\delta \in [0, 1)$, then, for all $t > 0$ and $p > 0$, we have:

$$\mathbb{E}\left(D_t^p\right) = E_x^\delta\left[\mathbf{1}_{t<T_0}\, {\rho_t}^p\, \exp\left(-p\frac{\delta-1}{2}\int_0^t \frac{ds}{\rho_s^2}\right)\right].$$

By the absolute continuity relation (6.22) applied with $\delta' := 2$, the latter equals:

$$E_x^2 \left[x^{\frac{2-\delta}{2}} \, {\rho_t}^{p+\frac{\delta-2}{2}} \, \exp\left(\underbrace{\left(-p\frac{\delta-1}{2} - \frac{(\delta-2)^2}{8} \right)}_{:=A(p)} \int_0^t \frac{ds}{\rho_s^2} \right) \right].$$

For $p = p(\delta)$, $A(p) = 0$, so that:

$$\mathbb{E}\left[D_t^{p(\delta)} \right] = E_x^2 \left[x^{\frac{2-\delta}{2}} {\rho_t}^{p(\delta)+\frac{\delta-2}{2}} \right] \\ = x^{1-\frac{\delta}{2}} \, E_x^2 \left[{\rho_t}^{p(\delta)+\frac{\delta}{2}-1} \right].$$

Since $\frac{\delta}{2} + p(\delta) - 1 \geq 0$, by Lemma 6.3, the last quantity is finite. Hence D_t is indeed in $L^{p(\delta)}$.

Suppose now that $p = p(\delta) + r$ for some $r > 0$. We show that $D_t \notin L^p$. We have:

$$\mathbb{E}\left[D_t^p \right] = E_x^2 \left[x^{\frac{2-\delta}{2}} \, {\rho_t}^{p+\frac{\delta-2}{2}} \, \exp\left(\left(-p\frac{\delta-1}{2} - \frac{(\delta-2)^2}{8} \right) \int_0^t \frac{ds}{{\rho_s}^2} \right) \right] \\ = x^{1-\frac{\delta}{2}} E_x^2 \left[{\rho_t}^{p+\frac{\delta}{2}-1} \, \exp\left(\frac{1-\delta}{2} r \int_0^t \frac{ds}{{\rho_s}^2} \right) \right]$$

We claim that the last quantity is infinite. Indeed, first note that by Jensen's inequality and Fubini, for any $C > 0$ we have:

$$E_x^2 \left[\exp\left(C \int_0^t \frac{ds}{{\rho_s}^2} \right) \right] \geq \exp\left(C \int_0^t E_x^2 \left(\rho_s^{-2} \right) ds \right)$$

and the right-hand side is infinite since, for all $s > 0$, $E_x^2 \left(\rho_s^{-2} \right) = +\infty$ (indeed, by formula (6.12), the transition density $p_s^2(x, y)$ does not integrate y^{-2} as $y \to 0$). Therefore:

$$E_x^2 \left[\exp\left(C \int_0^t \frac{ds}{{\rho_s}^2} \right) \right] = +\infty \tag{6.24}$$

Consider now any $c > 0$ and $a, b > 0$ such that $\frac{1}{a} + \frac{1}{b} = 1$. By (6.24) and Hölder's inequality, we have:

$$+\infty = E_x^2 \left[\exp\left(\frac{1-\delta}{2a} r \int_0^t \frac{ds}{{\rho_s}^2} \right) \right] \\ \leq E_x^2 \left[{\rho_t}^{ac} \exp\left(\frac{1-\delta}{2} r \int_0^t \frac{ds}{{\rho_s}^2} \right) \right]^{1/a} E_x^2 \left[\rho_t^{-bc} \right]^{1/b}$$

Set $c = \frac{\frac{\delta}{2}+p-1}{\frac{\delta}{2}+p}$, $a = \frac{\delta}{2} + p$, and $b = \frac{\frac{\delta}{2}+p}{\frac{\delta}{2}+p-1}$. Remark that $\frac{\delta}{2} + p - 1 > 0$ since $p > p(\delta) \geq 1$, so that this choice for c, a, and b makes sense. We obtain:

$$E_x^2\left[\rho_t{}^{\frac{\delta}{2}+p-1} \exp\left(\frac{1-\delta}{2} r \int_0^t \frac{ds}{\rho_s{}^2}\right)\right]^{\frac{1}{\frac{\delta}{2}+p}} E_x^2\left[\rho_t^{-1}\right]^{\frac{\frac{\delta}{2}+p-1}{\frac{\delta}{2}+p}} = +\infty$$

By the comparison Lemma 6.1 and the expression (6.12) for the transition density of the Bessel process, we have

$$E_x^2\left[\rho_t^{-1}\right] \leq E_0^2\left[\rho_t^{-1}\right] = \int_0^\infty \frac{1}{t} \exp\left(-\frac{y^2}{2t}\right) dy$$

so that $E_x^2\left[\rho_t^{-1}\right] < +\infty$. Therefore, we deduce that:

$$E_x^2\left[\rho_t{}^{\frac{\delta}{2}+p-1} \exp\left(\frac{1-\delta}{2} r \int_0^t \frac{ds}{\rho_s{}^2}\right)\right] = +\infty$$

as claimed. Hence $D_t \notin L^p$ for $p > p(\delta)$.

6.7 A Bismut-Elworthy-Li Formula for the Bessel Processes

We are now in position to provide a probabilistic interpretation of the right-hand-side of Eq. (6.11) in Theorem 6.2.

Let $\delta > 0$, and $x > 0$. As we saw in the previous section, the process $(\eta_t(x))_{t\geq 0}$ may blow up at time T_0, so that the stochastic integral $\int_0^t \eta_s(x) dB_s$ is a priori ill-defined, at least for $\delta \in (0, 1)$. However, it turns out that we can define the latter process rigorously as a local martingale.

Proposition 6.8 *Suppose that $\delta > 0$. Then the stochastic integral process $\int_0^t \eta_s dB_s$ is well-defined as a local martingale and is indistinguishable from the continuous martingale $D_t - x$.*

Proof We first treat the case $\delta \geq 2$, which is much easier to handle. In that case, $\eta_t \in [0, 1]$ for all $t \geq 0$, so that the stochastic integral $\int_0^t \eta_s dB_s$ is clearly well-defined as an L^2 martingale. Moreover, since $T_0 = +\infty$ a.s., by Itô's lemma we have:

$$\begin{aligned} D_t = \rho_t \eta_t &= x + \int_0^t \eta_s \, d\rho_s + \int_0^t \rho_s \, d\eta_s \\ &= x + \int_0^t \eta_s \left(\frac{\delta-1}{2} \frac{ds}{\rho_s} + dB_s\right) - \int_0^t \rho_s \frac{\delta-1}{2} \frac{\eta_s}{\rho_s{}^2} ds \end{aligned}$$

$$= x + \int_0^t \eta_s dB_s$$

so the claim follows.

Now suppose that $\delta \in (0, 2)$ and fix an $\epsilon > 0$. Recall that $T_\epsilon(x) := \inf\{t \geq 0, \rho_t(x) \leq \epsilon\}$ and note that, since $T_\epsilon < T_0$, the stopped process η^{T_ϵ} is continuous on $\mathbb{R}_+$, so that the stochastic integral $\int_0^{t\wedge T_\epsilon(x)} \eta_s(x) dB_s$ is well-defined as a local martingale. Using as above Itô's lemma, but this time with the stopped processes ρ^{T_ϵ} and η^{T_ϵ}, we have:

$$\int_0^{t\wedge T_\epsilon} \eta_s dB_s = D_{t\wedge T_\epsilon} - x. \tag{6.25}$$

Our aim would be to pass to the limit $\epsilon \to 0$ in this equality. By continuity of D, as $\epsilon \to 0$, $D_{t\wedge T_\epsilon}$ converges to $D_{t\wedge T_0} = D_t$ almost-surely. So the right-hand side of (6.25) converges to $D_t - x$ almost-surely.

The convergence of the left-hand side to a stochastic integral is more involved, since we first have to prove that the stochastic integral $\int_0^t \eta_s dB_s$ is indeed well-defined as a local martingale. For this, it suffices to prove that, almost-surely:

$$\forall t \geq 0, \qquad \int_0^t \eta_s^2 \, ds < \infty.$$

We actually prove the following stronger fact. For all $t \geq 0$

$$\mathbb{E}\left[\left(\int_0^t \eta_s^2 ds\right)^{p/2}\right] < \infty \tag{6.26}$$

for all finite positive number p such that $p \in (1, p(\delta)]$. Indeed, applying successively the Burkholder-Davis-Gundy (BDG) inequality and Doob's inequality to the martingale $\int_0^{T_\epsilon\wedge\cdot} \eta_s dB_s$, we have:

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^{t\wedge T_\epsilon} \eta_s^2 ds\right)^{p/2}\right] &\leq C_p \, \mathbb{E}\left[\sup_{s\leq t\wedge T_\epsilon} \left|\int_0^s \eta_u dB_u\right|^p\right] \\
&= C_p \, \mathbb{E}\left[\sup_{s\leq t\wedge T_\epsilon} |D_s - x|^p\right] \\
&\leq C_p \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|D_{t\wedge T_\epsilon} - x|^p\right]
\end{aligned}$$

where C_p is a constant depending only on p. Now, since $(D_t - x)_{t\geq 0}$ is a continuous martingale, by the optional stopping theorem and Jensen's inequality, we have:

$$\mathbb{E}\left[|D_{t\wedge T_\epsilon} - x|^p\right] \leq \mathbb{E}(|D_t - x|^p)$$

and the right-hand side is finite because D_t is in L^p. Hence, letting $\epsilon \to 0$ in the above, by the monotone convergence theorem we deduce that:

$$\mathbb{E}\left[\left(\int_0^{t\wedge T_0} \eta_s^2 ds\right)^{p/2}\right] < \infty$$

But since $\eta_t = 0$ for all $t \geq T_0$, this implies the bound (6.26), and hence the stochastic integral $\int_0^t \eta_s dB_s$ is well-defined as a local martingale. Moreover, for all $t \geq 0$, by the BDG inequality, we have:

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t \eta_s dB_s - \int_0^{t\wedge T_\epsilon} \eta_s dB_s\right)^p\right] &= \mathbb{E}\left[\left(\int_{t\wedge T_\epsilon}^{t\wedge T_0} \eta_s dB_s\right)^p\right] \\ &\leq c_p\, \mathbb{E}\left[\left(\int_{t\wedge T_\epsilon}^{t\wedge T_0} \eta_s^2 ds\right)^{p/2}\right]\end{aligned}$$

where c_p is some constant depending only on p. Now, by the dominated convergence theorem, the last quantity above goes to 0 as $\epsilon \to 0$, and hence:

$$\int_0^{t\wedge T_\epsilon} \eta_s dB_s \underset{\epsilon\to 0}{\longrightarrow} \int_0^t \eta_s dB_s$$

in L^p. Hence, the left-hand side of equality (6.25) converges in L^p to the stochastic integral $\int_0^t \eta_s dB_s$. Letting $\epsilon \to 0$ in that equality, we thus obtain:

$$\int_0^t \eta_s dB_s = D_t - x$$

as claimed.

Using the above proposition, Theorem 6.2 can now be interpreted probabilistically as a Bismut-Elworthy-Li formula.

Theorem 6.3 (Bismut-Elworthy-Li Formula) *Let $\delta > 0$. Then, for all $T > 0$, and all $F : \mathbb{R}_+ \to \mathbb{R}$ bounded and Borel, the function $x \to P_t^\delta F(x)$ is differentiable on $\mathbb{R}_+$, and for all $x > 0$:*

$$\frac{d}{dx} P_T^\delta F(x) = \frac{1}{T}\, \mathbb{E}\left[F(\rho_t(x))\left(\int_0^T \eta_s(x) dB_s\right)\right]. \tag{6.27}$$

Proof By Theorem 6.2, the differentiability property holds, and we have:

$$\frac{d}{dx} P_T^\delta F(x) = \frac{x}{T} \left[P_T^{\delta+2} F(x) - P_T^\delta F(x) \right].$$

Moreover, by Proposition 6.6, for all $x > 0$:

$$P_T^{\delta+2} F(x) - P_T^\delta F(x) = E_x^\delta \left[F(\rho_T) \left(\frac{D_T}{x} - 1 \right) \right]$$

and, by Proposition 6.8, we have:

$$E_x^\delta \left[F(\rho_T) \left(\frac{D_T}{x} - 1 \right) \right] = \frac{1}{x} \mathbb{E} \left[F(\rho_T(x)) \left(\int_0^T \eta_s(x) dB_s \right) \right]$$

so equality (6.27) follows.

Using the Bismut-Elworthy-Li formula, we are now able to sharpen the Strong Feller estimate obtained in Eq. (6.14) above.

Corollary 6.4 *Let* $T > 0$ *and* $\delta \geq 2(\sqrt{2} - 1)$*. Then, for all* $R > 0$*, there exists a constant* $C > 0$ *such that, for all* $x, y \in [0, R]$ *and* $F : \mathbb{R}_+ \to \mathbb{R}$ *bounded and Borel, we have:*

$$|P_T^\delta F(x) - P_T^\delta F(y)| \leq \frac{C ||F||_\infty}{T^{\alpha(\delta)}} |y - x| \tag{6.28}$$

where the exponent $\alpha(\delta) \in [\frac{1}{2}, 1)$ *is given by:*

$$\alpha(\delta) := \begin{cases} \frac{1}{2} + \frac{1-\delta}{2-\delta} & \text{if } \delta \in [2(\sqrt{2} - 1), 1], \\ 1/2 & \text{if } \delta \geq 1. \end{cases}$$

Proof Let $x > 0$. By Theorem 6.3, we have:

$$\frac{d}{dx} P_T^\delta F(x) = \frac{1}{T} \mathbb{E} \left[F(\rho_t(x)) \left(\int_0^T \eta_s(x) dB_s \right) \right].$$

so that:

$$\left| \frac{d}{dx} P_T^\delta F(x) \right| \leq \frac{||F||_\infty}{T} \mathbb{E} \left[\left| \int_0^T \eta_s(x) dB_s \right| \right].$$

We now bound the quantity $\mathbb{E} \left[\left| \int_0^T \eta_s(x) dB_s \right| \right]$. If $\delta \geq 1$, then the process $(\eta_s(x))_{s \geq 0}$ takes values in $[0, 1]$, so that, using the Cauchy-Schwarz inequality and

Itô's isometry formula, we have:

$$\mathbb{E}\left[\left|\int_0^T \eta_s(x)dB_s\right|\right] \leq \sqrt{\mathbb{E}\left(\int_0^T \eta_s(x)^2 ds\right)} \leq \sqrt{T}.$$

Therefore:

$$\left|\frac{d}{dx}P_T^\delta F(x)\right| \leq \frac{||F||_\infty}{\sqrt{T}}$$

and the claim follows with $C = 1$.

Suppose now that $\delta \in [2(\sqrt{2}-1), 1)$. Letting $p := p(\delta)$ as in (6.23), we have, by Jensen's inequality:

$$\mathbb{E}\left[\left|\int_0^T \eta_s(x)dB_s\right|\right] \leq \left(\mathbb{E}\left|\int_0^T \eta_s(x)dB_s\right|^p\right)^{1/p}$$

Now, applying successively the BDG inequality, Jensen's inequality and the absolute continuity relation (6.22) between P_x^2 and P_x^δ, we have, for some constant c_p depending only on p:

$$\begin{aligned}
\mathbb{E}\left[\left|\int_0^T \eta_s(x)dB_s\right|^p\right] &\leq c_p\, \mathbb{E}\left[\left(\int_0^T \eta_s(x)^2 ds\right)^{p/2}\right] \\
&\leq c_p\, T^{p/2-1}\, \mathbb{E}\left(\int_0^T \eta_s(x)^p\, ds\right) \\
&\leq c_p\, T^{p/2-1} \int_0^T E_x^\delta(\eta_s^p)\, ds \\
&= c_p\, T^{p/2-1} \int_0^T E_x^2\left[\left(\frac{\rho_s}{x}\right)^{\frac{\delta-2}{2}}\right. \\
&\qquad \left.\times \exp\left(\left(\frac{1-\delta}{2}p - \frac{(2-\delta)^2}{8}\right)\int_0^s \frac{du}{\rho_u^2}\right)\right] ds \\
&= c_p\, T^{p/2-1} \int_0^T E_x^2\left[\left(\frac{\rho_s}{x}\right)^{\frac{\delta-2}{2}}\right] ds
\end{aligned}$$

where the last equality follows from the fact that $\frac{1-\delta}{2}p - \frac{(2-\delta)^2}{8} = 0$ for $p = p(\delta)$. Now, since $\frac{\delta-2}{2} \leq 0$, by the comparison Lemma 6.1, as well as the scaling property of the Bessel processes (see, e.g., Remark 3.7 in [14]), for all $s \in [0, T]$, we have:

$$E_x^2\left[\rho_s^{\frac{\delta-2}{2}}\right] \leq E_0^2\left[\rho_s^{\frac{\delta-2}{2}}\right] = s^{\frac{\delta-2}{4}} E_0^2\left[\rho_1^{\frac{\delta-2}{2}}\right].$$

Let $c := E_0^2\left[\rho_1^{\frac{\delta-2}{2}}\right]$. Using formula (6.12), we have:

$$c = \int_0^\infty y^{\delta/2} \exp\left(-\frac{y^2}{2}\right) dy < \infty.$$

Hence:

$$\begin{aligned}\int_0^T E_x^2\left[\left(\frac{\rho_s}{x}\right)^{\frac{\delta}{2}-1}\right] ds &\leq c\, x^{1-\frac{\delta}{2}} \int_0^T s^{\frac{\delta-2}{4}} ds \\ &\leq \frac{4\,c}{\delta+2}\, x^{1-\frac{\delta}{2}}\, T^{\frac{\delta+2}{4}}.\end{aligned}$$

Therefore, we obtain:

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^T \eta_s(x) dB_s\right|^p\right] &\leq K\, x^{1-\frac{\delta}{2}}\, T^{\frac{p}{2}-1}\, T^{\frac{\delta+2}{4}} \\ &\leq K\, x^{1-\frac{\delta}{2}}\, T^{\frac{p}{2}+\frac{\delta-2}{4}}\end{aligned}$$

where K is a constant depending only on δ. Hence

$$\mathbb{E}\left[\left|\int_0^T \eta_s(x) dB_s\right|\right] \leq K^{1/p}\, x^{\frac{1}{p}(1-\frac{\delta}{2})}\, T^{\frac{1}{2}+\frac{\delta-2}{4p}}.$$

Note that, since $p = p(\delta)$, we have $\frac{1}{p}(1-\frac{\delta}{2}) = \frac{2(1-\delta)}{2-\delta}$, and $\frac{\delta-2}{4p} = -\frac{1-\delta}{2-\delta}$. Therefore, we obtain:

$$\left|\frac{d}{dx} P_T^\delta F(x)\right| \leq K^{1/p}\, x^{2\frac{1-\delta}{2-\delta}}\, ||F||_\infty\, T^{-\frac{1}{2}-\frac{1-\delta}{2-\delta}}.$$

Therefore, given $R > 0$, one has for all $x \in [0, R]$:

$$\left|\frac{d}{dx} P_T^\delta F(x)\right| \leq C \frac{||F||_\infty}{T^{\alpha(\delta)}}$$

with $C := K^{1/p}\, R^{2\frac{1-\delta}{2-\delta}}$. This yields the claim.

Remark 6.14 In the above proposition, the value $2(\sqrt{2}-1)$ that appears is the smallest value of δ for which η is in L^2. For $\delta < 2(\sqrt{2}-1)$, η is no longer in L^2 but only in L^p for $p = p(\delta) < 2$, so that we cannot apply Jensen's inequality to bound the quantity $\mathbb{E}\left(\int_0^T \eta_s(x)^2 ds\right)^{p/2}$ anymore. It seems reasonable to expect that the bound (6.28) holds also for $\delta < 2(\sqrt{2}-1)$, although we do not have a proof of this fact.

Acknowledgements I would like to thank Lorenzo Zambotti, my Ph.D. advisor, for all the time he patiently devotes in helping me with my research. I would also like to thank Thomas Duquesne and Nicolas Fournier, who helped me solve a technical problem, as well as Yves Le Jan for a helpful discussion on the Bessel flows of low dimension, and Lioudmila Vostrikova for answering a question on this topic.

Appendix

In this Appendix, we prove Proposition 6.3. Recall that we still denote by $(\rho_t(x))_{t,x\geq 0}$ the process $(\tilde{\rho}_t^\delta(x))_{t,x\geq 0}$ constructed in Proposition 6.2.

Lemma 6.4 *For all rational numbers $\epsilon, \gamma > 0$, let:*

$$\mathscr{U}_\gamma^\epsilon := [0, T_\epsilon(\gamma)) \times (\gamma, +\infty)$$

and set:

$$\mathscr{U} := \bigcup_{\epsilon,\gamma\in\mathbb{Q}_+^*} \mathscr{U}_\gamma^\epsilon .$$

Then, a.s., the function $(t, x) \mapsto \rho_t(x)$ is continuous on the open set $\mathscr{U}$.

Proof By patching, it suffices to prove that, a.s., the function $(t, x) \mapsto \rho_t(x)$ is continuous on each $\mathscr{U}_\gamma^\epsilon$, where $\epsilon, \gamma \in \mathbb{Q}_+^*$.

Fix $\epsilon, \gamma \in \mathbb{Q}_+^*$, and let $x, y \in (\gamma, +\infty) \cap \mathbb{Q}$. We proceed to show that, a.s., for all $t \leq s < T_\epsilon(\gamma)$ the following inequality holds:

$$|\rho_t(x) - \rho_s(y)| \leq |x-y| \exp\left(\frac{|\delta-1|}{2\epsilon^2} t\right) + \frac{|\delta-1|}{2\epsilon}|s-t| + |B_s - B_t|. \tag{6.29}$$

Since $T_\epsilon(\gamma) < T_0(\gamma)$, a.s., for all $t \leq s \leq T_\epsilon(\gamma)$, we have:

$$\forall \tau \in [0, t], \qquad \rho_\tau(x) = x + \frac{\delta-1}{2}\int_0^\tau \frac{du}{\rho_u(x)} + B_\tau$$

as well as

$$\forall \tau \in [0, s], \qquad \rho_\tau(y) = y + \frac{\delta - 1}{2} \int_0^\tau \frac{du}{\rho_u(y)} + B_\tau$$

and hence:

$$\forall \tau \in [0, t], \qquad |\rho_\tau(x) - \rho_\tau(y)| \leq |x - y| + \frac{|\delta - 1|}{2} \int_0^\tau \frac{|\rho_u(x) - \rho_u(y)|}{\rho_u(x)\rho_u(y)} du.$$

By the monotonicity property of ρ, we have, a.s., for all t, s as above and $u \in [0, s]$:

$$\rho_u(x) \wedge \rho_u(y) \geq \rho_u(\gamma) \geq \epsilon \tag{6.30}$$

so that:

$$\forall \tau \in [0, t], \quad |\rho_\tau(x) - \rho_\tau(y)| \leq |x - y| + \frac{|\delta - 1|}{2} \int_0^\tau \frac{|\rho_u(x) - \rho_u(y)|}{\epsilon^2} du,$$

which, by Grönwall's inequality, implies that:

$$|\rho_t(x) - \rho_t(y)| \leq |x - y| \exp\left(\frac{|\delta - 1|}{2\epsilon^2} t\right). \tag{6.31}$$

Moreover, we have:

$$\rho_s(y) - \rho_t(y) = \frac{\delta - 1}{2} \int_t^s \frac{du}{\rho_u(y)} + B_s - B_t$$

which, by (6.30), entails the inequality:

$$|\rho_s(y) - \rho_t(y)| \leq \frac{|\delta - 1|}{2\epsilon} |s - t| + |B_s - B_t|. \tag{6.32}$$

Putting inequalities (6.31) and (6.32) together yields the claimed inequality (6.29). Hence, we have, a.s., for all rationals $x, y > \gamma$ and all $t \leq s < T_\epsilon(\gamma)$:

$$|\rho_t(x) - \rho_s(y)| \leq |x - y| \exp\left(\frac{|\delta - 1|}{2\epsilon^2} t\right) + \frac{\delta - 1}{2} |s - t| + |B_s - B_t|$$

and, by density of $\mathbb{Q} \cap (\gamma, +\infty)$ in $(\gamma, +\infty)$, this inequality remains true for all $x, y > \gamma$. Since, a.s., $t \mapsto B_t$ is continuous on $\mathbb{R}_+$, the continuity of ρ on $\mathscr{U}_\gamma^\epsilon$ is proved.

Corollary 6.5 *Almost-surely, we have:*

$$\forall x \geq 0, \quad \forall t \in [0, T_0(x)), \qquad \rho_t(x) = x + \frac{\delta - 1}{2} \int_0^t \frac{du}{\rho_u(x)} + B_t. \tag{6.33}$$

Remark 6.15 We have already remarked in Sect. 6.3 that, for all fixed $x \geq 0$, the process $(\rho_t(x))_{t \geq 0}$ satisfies the SDE (6.10). By contrast, the above Corollary shows the stronger fact that, considering the modification $\tilde{\rho}$ of the Bessel flow constructed in Proposition 6.2 above, a.s., for *each* $x \geq 0$, the path $(\tilde{\rho}_t(x))_{t \geq 0}$ still satisfies relation (6.10).

Proof Consider an almost-sure event $\mathscr{A} \in \mathscr{F}$ as in Remark 6.7. On the event $\mathscr{A}$, for all $r \in \mathbb{Q}_+$, we have:

$$\forall t \in [0, T_0(r)), \qquad \rho_t(r) = r + \frac{\delta - 1}{2} \int_0^t \frac{du}{\rho_u(r)} + B_t.$$

Denote by $\mathscr{B} \in \mathscr{F}$ any almost-sure event on which ρ satisfies the monotonicity property (6.15). We show that, on the event $\mathscr{A} \cap \mathscr{B}$, the property (6.33) is satisfied.

Suppose $\mathscr{A} \cap \mathscr{B}$ is fulfilled, and let $x \geq 0$. Then for all $r \in \mathbb{Q}$ such that $r \geq x$, we have:

$$\forall t \geq 0, \qquad \rho_t(x) \leq \rho_t(r)$$

so that $T_0(r) \geq T_0(x)$. Hence, for all $t \in [0, T_0(x))$, we have in particular $t \in [0, T_0(r))$, so that:

$$\rho_t(r) = r + \frac{\delta - 1}{2} \int_0^t \frac{du}{\rho_u(r)} + B_t.$$

Since, for all $u \in [0, t]$, $\rho_u(r) \downarrow \rho_u(x)$ as $r \downarrow x$ with $r \in \mathbb{Q}$, by the monotone convergence theorem, we deduce that:

$$\int_0^t \frac{du}{\rho_u(r)} \longrightarrow \int_0^t \frac{du}{\rho_u(x)}$$

as $r \downarrow x$ with $r \in \mathbb{Q}$. Hence, letting $r \downarrow x$ with $r \in \mathbb{Q}$ in the above equation, we obtain:

$$\rho_t(x) = x + \frac{\delta - 1}{2} \int_0^t \frac{du}{\rho_u(x)} + B_t.$$

This yields the claim.

One of the main difficulties for proving Proposition 6.3 arises from the behavior of $\rho_t(x)$ at $t = T_0(x)$. However we will circumvent this problem by working away from the event $t = T_0(x)$. To do so, we will make use of the following property.

Lemma 6.5 *Let $\delta < 2$ and $x \geq 0$. Then the function $y \mapsto T_0(y)$ is a.s. continuous at x.*

Proof The function $y \mapsto T_0(y)$ is nondecreasing over $\mathbb{R}_+$. Hence, if $x > 0$, it has left- and right-sided limits at x, $T_0(x^-)$ and $T_0(x^+)$, satisfying:

$$T_0(x^-) \leq T_0(x) \leq T_0(x^+). \tag{6.34}$$

Similarly, if $x = 0$, there exists a right-sided limit $T_0(0^+)$ satisfying $T_0(0) \leq T_0(0^+)$. Suppose, e.g., that $x > 0$. Then we have:

$$\mathbb{E}\left(e^{-T_0(x^+)}\right) \leq \mathbb{E}\left(e^{-T_0(x)}\right) \leq \mathbb{E}\left(e^{-T_0(x^-)}\right). \tag{6.35}$$

Now, by the scaling property of the Bessel processes (see, e.g., Remark 3.7 in [14]), for all $y \geq 0$, the following holds:

$$(y\rho_t(1))_{t\geq 0} \overset{(d)}{=} (\rho_{y^2 t}(y))_{t\geq 0},$$

so that $T_0(y) \overset{(d)}{=} y^2 T_0(1)$. Therefore, using the dominated convergence theorem, we have:

$$\begin{aligned}
\mathbb{E}\left(e^{-T_0(x^+)}\right) &= \lim_{y\downarrow x} \mathbb{E}\left(e^{-T_0(y)}\right) \\
&= \lim_{y\downarrow x} \mathbb{E}\left(e^{-y^2 T_0(1)}\right) \\
&= \mathbb{E}\left(e^{-x^2 T_0(1)}\right) \\
&= \mathbb{E}\left(e^{-T_0(x)}\right).
\end{aligned}$$

Similarly, we have $\mathbb{E}\left(e^{-T_0(x^-)}\right) = \mathbb{E}\left(e^{-T_0(x)}\right)$. Hence the inequalities (6.35) are actually equalities; recalling the original inequality (6.34), we deduce that $T_0(x^-) = T_0(x) = T_0(x^+)$ a.s.. Similarly, if $x = 0$, we have $T_0(0) = T_0(0^+)$ a.s.

Before proving Proposition 6.3, we need a coalescence lemma, which will help us prove that the derivative of ρ_t at x is 0 if $t > T_0(x)$:

Lemma 6.6 *Let $x, y \geq 0$, and let τ be a nonnegative $(\mathscr{F}_t)_{t\geq 0}$-stopping time. Then, almost-surely:*

$$\rho_\tau(x) = \rho_\tau(y) \quad \Rightarrow \quad \forall s \geq \tau,\ \rho_s(x) = \rho_s(y).$$

Proof On the event $\{\rho_\tau(x) = \rho_\tau(y)\}$, the processes $(X_t^\delta(x))_{t\geq 0} := (\rho_t(x)^2)_{t\geq 0}$ and $(X_t^\delta(y))_{t\geq 0} := (\rho_t(y)^2)_{t\geq 0}$ both satisfy, on $[\tau, +\infty)$, the SDE:

$$X_t = \rho_\tau(x)^2 + 2\int_\tau^t \sqrt{X_s} dB_s + \delta(t-\tau).$$

By pathwise uniqueness of this SDE (see [10, Theorem (3.5), Chapter IX]), we deduce that, a.s. on the event $\{\rho_\tau(x) = \rho_\tau(y)\}$, $X_t(x) = X_t(y)$, hence $\rho_t(x) = \rho_t(y)$ for all $t \geq \tau$.

Now we are able to prove Proposition 6.3.

Proof (Proof of Proposition 6.3) Let $t > 0$ and $x > 0$ be fixed. First remark that:

$$\mathbb{P}(T_0(x) = t) = 0.$$

Indeed, if $\delta > 0$, then:

$$\mathbb{P}(T_0(x) = t) \leq \mathbb{P}(\rho_t(x) = 0)$$

and the RHS is zero since the law of $\rho_t(x)$ has no atom on $\mathbb{R}_+$ (it has density $p_t^\delta(x,\cdot)$ w.r.t. Lebesgue measure on $\mathbb{R}_+$, where p_t^δ was defined in Eq. (6.12) above). On the other hand, if $\delta = 0$, then 0 is an absorbing state for the process ρ, so that, for all $s \geq 0$:

$$\mathbb{P}(T_0(x) \leq s) = \mathbb{P}(\rho_s(x) = 0)$$

and the RHS is continuous in s on $\mathbb{R}_+$, since it is given by $\exp(-\frac{x^2}{2s})$ (see [10, Chapter XI, Corollary 1.4]). Hence, also in the case $\delta = 0$ the law of $T_0(x)$ has no atom on $\mathbb{R}_+$. Hence, a.s., either $t < T_0(x)$ or $t > T_0(x)$.

First suppose that $t < T_0(x)$. A.s., the function $y \mapsto T_0(y)$ is continuous at x, so there exists a rational number $y \in [0, x)$ such that $t < T_0(y)$; since, by Remark (6.7), $t \mapsto \rho_t(y)$ is continuous, there exists $\epsilon \in \mathbb{Q}_+^*$ such that $t < T_\epsilon(y)$. By monotonicity of $z \mapsto \rho(z)$, for all $s \in [0, t]$ and $z \geq y$, we have:

$$\rho_s(z) \geq \rho_s(y) \geq \epsilon.$$

Hence, recalling Corollary 6.5, for all $s \in [0, t]$ and $h \in \mathbb{R}$ such that $|h| < |x - y|$:

$$\rho_s(x+h) = x + h + \int_0^s \frac{\delta - 1}{2} \frac{du}{\rho_u(x+h)} + B_s.$$

Hence, setting $\eta_s^h(x) := \frac{\rho_s(x+h)-\rho_s(x)}{h}$,we have:

$$\forall s \in [0,t], \qquad \eta_s^h(x) = 1 - \frac{\delta-1}{2}\int_0^t \frac{\eta_u^h(x)}{\rho_u(x)\rho_u(x+h)}du$$

so that:

$$\eta_t^h(x) = \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)\rho_s(x+h)}\right).$$

Note that, for all $s \in [0,t]$ and $h \in \mathbb{R}$ such that $|h| < |x-y|$, we have $(s, x+h) \in [0, T_\epsilon(y)) \times (y, +\infty) \subset \mathscr{U}$. Hence, by Lemma 6.4, we have, for all $s \in [0,t]$

$$\rho_s(x+h) \underset{h\to 0}{\longrightarrow} \rho_s(x)$$

with the domination property:

$$\frac{1}{\rho_s(x)\rho_s(x+h)} \leq \epsilon^{-2}$$

valid for all $|h| < |x-y|$. Hence, by the dominated convergence theorem, we deduce that:

$$\eta_t^h(x) \underset{h\to 0}{\longrightarrow} \exp\left(\frac{1-\delta}{2}\int_0^t \frac{ds}{\rho_s(x)^2}\right)$$

which yields the claimed differentiability of ρ_t at x.

We now suppose that $t > T_0(x)$. Since the function $y \mapsto T_0(y)$ is a.s. continuous at x, a.s. there exists $y > x$, $y \in \mathbb{Q}$, such that $t > T_0(y)$. By Remark (6.7), the function $t \mapsto \rho_t(y)$ is continuous, so that $\rho_{T_0(y)}(y) = 0$. By monotonicity of $z \mapsto \rho(z)$, we deduce that, for all $z \in [0,y]$, we have:

$$\rho_{T_0(y)}(z) = 0.$$

By Lemma 6.6, we deduce that, leaving aside some event of probability zero, all the trajectories $(\rho_t(z))_{t\geq 0}$ for $z \in [0,y] \cap \mathbb{Q}$ coincide from time $T_0(y)$ onwards. In particular, we have:

$$\forall z \in [0,y] \cap \mathbb{Q}, \qquad \rho_t(z) = \rho_t(x).$$

Since, moreover, the function $z \mapsto \rho_t(z)$ is nondecreasing, we deduce that it is constant on the whole interval $[0,y]$:

$$\forall z \in [0,y], \qquad \rho_t(z) = \rho_t(x).$$

In particular, the function $z \mapsto \rho_t(z)$ has derivative 0 at x. This concludes the proof.

References

1. M. Arnaudon, X.-M. Li et al., Reflected Brownian motion: selection, approximation and linearization. Electron. J. Probab. **22**, 1–55 (2017)
2. J.-M. Bismut, *Large Deviations and the Malliavin Calculus*, vol. 45 (Birkhauser, Basel, 1984)
3. K. Burdzy et al., Differentiability of stochastic flow of reflected Brownian motions. Electron. J. Probab. **14**(75), 2182–2240 (2009)
4. S. Cerrai, *Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach*, vol. 1762 (Springer, Berlin, 2001)
5. G. Da Prato, J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, vol. 229 (Cambridge University Press, Cambridge, 1996)
6. J.-D. Deuschel, L. Zambotti, Bismut–Elworthy's formula and random walk representation for SDES with reflection. Stochastic Process. Appl. **115**(6), 907–925 (2005)
7. K. Elworthy, X. Li, Formulae for the derivatives of heat semigroups. J. Funct. Anal. **125**(1), 252–286 (1994). ISSN 0022-1236. https://doi.org/10.1006/jfan.1994.1124. http://www.sciencedirect.com/science/article/pii/S0022123684711244
8. M. Hairer, J. Mattingly, The strong feller property for singular stochastic PDES (2016). arXiv preprint arXiv:1610.03415
9. J. Pitman, M. Yor, Bessel processes and infinitely divisible laws, in *Stochastic Integrals* (Springer, Berlin, 1981), pp. 285–370
10. D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, vol. 293 (Springer, New York, 2013)
11. P. Tsatsoulis, H. Weber, Spectral gap for the stochastic quantization equation on the 2-dimensional torus (2016). arXiv preprint arXiv:1609.08447
12. S.R. Varadhan, R.J. Williams, Brownian motion in a wedge with oblique reflection. Commun. Pure Appl. Math. **38**(4), 405–443 (1985)
13. L. Vostrikova, On regularity properties of Bessel flow. Stochastics: Int. J. Probab. Stochastics Process. **81**(5), 431–453 (2009)
14. L. Zambotti, *Random Obstacle Problems: École d'Été de Probabilités de Saint-Flour XLV-2015*, vol. 2181 (Springer, New York, 2017)

Chapter 7
Large Deviations for Infectious Diseases Models

Peter Kratz and Etienne Pardoux

Abstract We study large deviations of a Poisson driven system of stochastic differential equations modeling the propagation of an infectious disease in a large population, considered as a small random perturbation of its law of large numbers ODE limit. Since some of the rates vanish on the boundary of the domain where the solution takes its values, thus making the action functional possibly explode, our system does not obey assumptions which are usually made in the literature. We present the whole theory of large deviations for systems which include the infectious disease models, and apply our results to the study of the time taken for an endemic equilibrium to cease, due to random effects.

Keywords Poisson process driven SDE · Lareg deviations · Freidlin-Wentzell theory · Epidemic models

7.1 Introduction

Consider a model of infectious disease dynamics where the total number of individuals is constant over time, equal to N, and we denote by $Z^N(t)$ the vector of proportions of this population in each compartment (susceptible, infectious, removed, etc.). Our probabilistic model takes into account each event of infection, removal, etc. It takes the form

$$Z^{N,x}(t) := Z^N(t) := x + \frac{1}{N}\sum_{j=1}^{k} h_j P_j\Big(\int_0^t N\beta_j(Z^N(s))ds\Big). \tag{7.1}$$

P. Kratz · E. Pardoux (✉)
Aix-Marseille Univ, CNRS, Centrale Marseille, Marseille, France
e-mail: etienne.pardoux@univ-amu.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_7

Here, $(P_j)_{1\le j\le k}$ are i.i.d. standard Poisson processes. The $h_j \in \mathbb{Z}^d$ denote the k respective jump directions with jump rates $\beta_j(x)$ and $x \in A$ (where A is the "domain" of the process). The d components of the process denote the "proportions" of individuals in the various compartments. Usually A is a compact or at least a bounded set. For example, in the models we have in mind the compartment sizes are non-negative, hence $A \subset \mathbb{R}^d_+$.

As we shall prove below, $Z_t^{N,x} \to Y_t^x$ as $N \to \infty$, where Y_t^x is the solution of the ODE

$$Y^x(t) := Y(t) := x + \int_0^t b(Y^x(s))ds, \tag{7.2}$$

with

$$b(z) := \sum_{j=1}^{k} \beta_j(z) h_j.$$

This Law of Large Numbers result goes back to [9] (see the version in Theorem 7.3 below, where a rate of convergence is given).

Most of the literature on mathematical models of disease dynamics treats deterministic models of the type of (7.2). When an epidemics is established, and each compartment of the model contains a significant proportion of the total population, if N is large enough, the ODE (7.2) is a good model for the epidemics. The original stochastic model (7.1), which we believe to be more realistic than the (7.2), can be considered as a stochastic perturbation of (7.2). However, we know from the work of [7], that small Brownian perturbations of an ODE will eventually produce a large deviation from its law of large numbers limit. For instance, if the ODE starts in the basin of attraction of an locally stable equilibrium, the solution of the ODE converges to that equilibrium, and stays for ever close to that equilibrium. The theory of Freidlin and Wentzell, based upon the theory of Large Deviations, predicts that soon or later the solution of a random perturbation of that ODE will exit the basin of attraction of the equilibrium. The aim of this paper is to show that the Poissonian perturbation (7.1) of (7.2) behaves similarly. This should allow us to predicts the time taken by an endemic equilibrium to cease, and a disease-free equilibrium to replace it.

We shall apply at the end of this paper our results to the following example.

Example 7.1 We consider a so-called $SIRS$ model without demography ($S(t)$ being the number of susceptible individuals, $I(t)$ the number of infectious individuals and $R(t)$ the number of removed/immune individuals at time t). We let $\beta > 0$ and assume that the average number of new infections per unit time is $\beta S(t)I(t)/N$.[1]

---

[1]The reasoning behind this is the following. Assume that an infectious individual meets on average $\alpha > 0$ other individuals in unit time. If each contact of a susceptible and an infectious individual

For $\gamma, \nu > 0$, we assume that the average number of recoveries per unit time is $\gamma I(t)$ and the average number of individuals who lose immunity is $\nu R(t)$. As population size is constant, we can reduce the dimension of the model by solely considering the proportion of infectious and removed at time t. Using the notation of Eqs. (7.1) and (7.2), we have

$$A = \{x \in \mathbb{R}_+^2 | 0 \le x_1 + x_2 \le 1\}, \quad h_1 = (1,0)^\top, \quad h_2 = (-1,1)^\top, \quad h_3 = (0,-1)^\top,$$
$$\beta_1(z) = \beta z_1(1 - z_1 - z_2), \quad \beta_2(z) = \gamma z_1, \quad \beta_3(z) = \nu z_2.$$

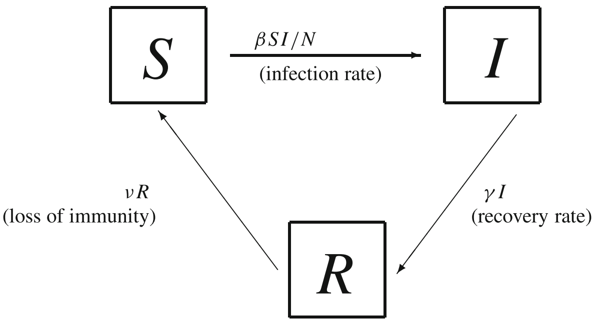
It is easy to see that in this example the ODE (7.2) has a disease free equilibrium $\bar{x} = (0,0)^\top$. This equilibrium is asymptotically stable if $R_0 = \beta/\gamma < 1$. R_0 is the so-called *basic reproduction number*. It denotes the average number of secondary cases infected by one primary case during its infectious period at the start of the epidemic (while essentially everybody is susceptible). If $R_0 > 1$, $\bar{x}$ is unstable and there exists a second, endemic equilibrium

$$x^* = \left(\frac{\nu(\beta - \gamma)}{\beta(\gamma + \nu)}, \frac{\gamma(\beta - \gamma)}{\beta(\gamma + \nu)} \right)$$

which is asymptotically stable. While in the deterministic model the proportion of infectious and removed individuals converges to the endemic equilibrium x^*, the disease will go extinct soon or later in the stochastic model (Fig. 7.1).

Our results also apply to other models like the SIV model (V like *vaccinated*) and the S_0IS_1 (with two levels of susceptibility), see [8] and the references therein. These two models have the property that for certain values of their parameters, both the disease-free equilibrium and one of the endemic equilibria are locally stable.

Fig. 7.1 Transmission diagram of the $SIRS$ model without demography



yields a new infection with probability p, the average number of new infections per unit time is $\beta S(t)I(t)/N$, where $\beta = p\alpha$ since all individuals are contacted with the same probability (hence $S(t)/N$ is the probability that a contacted individual is susceptible).

Our results predict the time taken by the solution of the stochastic equation to leave the basin of attraction of the endemic equilibrium. We shall discuss those and other applications elsewhere in the future.

There is already a vast literature on the theory of large deviations for systems with Poissonian inputs, see [4–6, 13], among others.

However, the assumptions in those works are not satisfied in our case. The difficulty is the following. For obvious reasons, the solution of our SDE (7.1) must remain in $\mathbb{R}_+^d$. This implies that some rates vanish when one of the components of $Z^{N,x}(t)$ vanishes. However, the expression of the large deviation functional (as well as the ratio of the probabilities in the Girsanov theorem) involves the logarithm of those rates, which hence explodes as the rate vanishes. The same happens with the computer network models which was the motivation of the work of [13], and this led them to consider situations with vanishing rates in [14]. However, even the assumptions in that paper are not fully satisfied in our models (see our discussion below in Sect. 7.2.3). For that reason, in order to avoid the awkward situation where we would have to cite both [13] and [14], and add some arguments to cope with our specific situation, we preferred to rewrite the whole theory, so as to cover the situation of the epidemiological models in a self-consistent way. We must however recognize that the work of Shwartz and Weiss has been an importance source of inspiration for this work.

Let us now discuss one subtlety of our models. In the models without demography, i.e. the models where the total population remains constant, then we choose N as this total population, so that the various components of the vector $Z_t^{N,x}$ are the proportions of the total population in the various compartments of the model, that is each component of $Z_t^{N,x}$ at any time is of the form k/N, where $k \in \mathbb{Z}_+$, and also $\sum_{i=1}^d Z_i^{N,x}(t) = 1$, if $Z_i^{N,x}(t)$ denotes the i-th component of the vector $Z^{N,x}(t)$, $1 \le i \le d$. In this case, provided we start our SDE from a point of the type $(k_1/N, \ldots, k_d/N)$, where $k_1, \ldots, k_d \in \mathbb{Z}_+$, then the solution visits only such points, and cannot escape the set $\mathbb{R}_+^d$ without hitting first its boundary, where the rates for exiting vanish. Consequently $Z_t^{N,x}$ remains in $\mathbb{R}_+^d$ for all time a.s. However, if we start our process outside the above grid, or if the total population size does not remain constant, the components of the vector $Z_t^{N,x}$ multiplied by N need not be integers. Then some of the components of $Z_t^{N,x}$ might become negative, and one can still continue to define $Z_t^{N,x}$ provided for any $1 \le i \le d$ and $1 \le j \le k$, the rule $x_i = 0 \Rightarrow \beta_j(x) = 0$ is extended to $x_i \le 0 \Rightarrow \beta_j(x) = 0$. However, in order to make things simpler, we restrict ourselves in this paper to the situation where all coordinates of the vector $NZ_0^{N,x}$ are integers, and the same is true with $NZ_t^{N,x}$ for all $t > 0$. In particular, we shall consider Eq. (7.1) only with a starting point x such that all coordinates of Nx are integers. This will be explicitly recalled in the main statements, and implicitly assumed everywhere. We shall consider more general situations in a further publication.

The paper is organized as follows. Our set-up is made precise and the general assumptions are formulated in Sect. 7.2. Section 7.3 is devoted to the law of large numbers. In Sect. 7.4 we study the rate function. The Large Deviations lower bound is established in Sect. 7.5 and the Large Deviations upper bound in Sect. 7.6. Section 7.7 treats the exit time from a domain, including the case of a characteristic boundary. Finally in Sect. 7.8 we show how our results apply to the SIRS model (which requires an additional argument), and a Girsanov theorem is formulated in the Appendix.

7.2 Set-Up

We consider a set $A \subset \mathbb{R}^d$ (whose properties we specify below) and define the grids

$$\mathbb{Z}^{d,N} := \{x \in \mathbb{R}^d \,\big|\, x_i = j/N \text{ for some } j \in \mathbb{Z}\}, \quad A^N := A \cap \mathbb{Z}^{d,N}.$$

We rewrite the process defined by Eq. (7.1) as

$$\begin{aligned} Z^{N,x}(t) := Z^N(t) &:= x + \frac{1}{N}\sum_{j=1}^{k} h_j P_j\Big(\int_0^t N\beta_j(Z^N(s))ds\Big) \qquad (7.3)\\ &= x + \int_0^t b(Z^N(s))ds + \frac{1}{N}\sum_j h_j M_j\Big(\int_0^t N\beta_j(Z^N(s))ds\Big), \end{aligned}$$

where the $M_j(t) = P_j(t) - t$, $j = 1, \ldots, k$. $M_j(t)$ is the so-called compensated Poisson processes corresponding to $P_j(t)$.

Let us shortly comment on this definition. In the models we have in mind, the components of Z^N usually denote the proportions of individuals in the respective compartments. It is hence plausible to demand that the starting point is in $\mathbb{Z}^{d,N}$. However, it is not sufficient to simply restrict our analysis to those starting points with $x \in A^N$ as this does obviously not imply $x \in A^M$ for all $M > N$. Note that $U^N(t) = NZ^N(t)$ would solve the SDE

$$U^N(t) = Nx + \sum_{j=1}^{k} h_j P_j\left(\int_0^t \beta_{j,N}(U^N(s))\right),$$

where $N\beta_j(x) = \beta_{j,N}(Nx)$. Here the coefficients of the vector $U^N(t)$ are the numbers of individuals from the population in each compartment. The equation for $U^N(t)$ is really the original model, where all events of infection, recovery, loss of immunity, etc. are modeled. Dividing by N leads to a process which has a law of large number limits as $N \to \infty$. The crucial assumption for this procedure to make sense is that $N^{-1}\beta_{j,N}(Nx)$ does not depend upon N, which is typically the case in the epidemics models, see in particular Example 7.1.

We first introduce the following notations. For $x \in A$ and $y \in \mathbb{R}^d$, let

$$V_x := \Big\{\mu \in \mathbb{R}_+^k \,\big|\, \mu_j > 0 \text{ only if } \beta_j(x) > 0\Big\},$$
$$V_{x,y} := \Big\{\mu \in V_x \,\big|\, y = \sum_j \mu_j h_j\Big\}.$$

As $V_{x,y}$ is sometimes independent of x or $V_{x,y} = \emptyset$, we also define for $y \in \mathbb{R}^d$,

$$\tilde{V}_y := \Big\{\mu \in \mathbb{R}_+^d \,\big|\, y = \sum_j \mu_j h_j\Big\}$$

We define the cone spanned by a (finite) set of vectors $(v_j)_j$ ($v_j \in \mathbb{R}^d$ by

$$\mathcal{C}((v_j)_j) := \Big\{v = \sum_j \alpha_j v_j \,\big|\, \alpha_j \geq 0\Big\}.$$

Similarly, we define the cone generated by the jump directions $(h_j)_j$ at $x \in A$ by

$$\mathcal{C}_x := \mathcal{C}((h_j)_{j:\beta_j(x)>0}) = \Big\{v = \sum_{j:\beta_j(x)>0} \mu_j h_j \,\big|\, \mu_j \geq 0\Big\}.$$

Note that

$$\mathcal{C}_x = \mathcal{C} = \{v = \sum_{j=1}^k \mu_j h_j \,\big|\, \mu_j \geq 0\} \tag{7.4}$$

whenever $x \in \mathring{A}$, since $\beta_j(x) > 0$ for all $1 \leq j \leq k$ if $x \in \mathring{A}$. Also, in part of this paper, we shall assume that the $\log \beta_j$'s are bounded, which then means that (7.4) is true for all $x \in A$.

We define the following upper and lower bounds of the rates. Let $\rho > 0$.

$$\bar{\beta} := \sup_{x \in A,\, j=1,\ldots,k} \beta_j(x) \in \bar{\mathbb{R}}_+,$$
$$\underline{\beta} := \inf_{x \in A,\, j=1,\ldots,k} \beta_j(x) \in \mathbb{R}_+,$$
$$\underline{\beta}(\rho) := \inf\big\{\beta_j(x) | j = 1, \ldots, k,\ x \in A$$
$$\text{and } |x - z| \geq \rho \forall z \in A \text{ with } \beta_j(z) = 0\big\} \in \mathbb{R}_+,$$
$$\bar{h} := \sup_{j=1,\ldots,k} |h_j| \in \mathbb{R}_+.$$

7.2.1 Cramer's Theorem for the Poisson Distribution

Let $X_1, X_2, \ldots, X_n, \ldots$ be mutually independent Poi(μ) r.v.'s. The Law of Large Numbers tells us that

$$\frac{1}{N}\sum_{i=1}^{N} X_i \to \mu \quad \text{a.s. as } N \to \infty.$$

Let us first define, for $X \sim \text{Poi}(\mu)$ the logarithm of its Laplace transform

$$\Lambda_\mu(\lambda) = \log \mathbb{E}[\exp(\lambda X)] = \mu(e^\lambda - 1),$$

and the Fenchel–Legendre transform of the latter

$$\Lambda_\mu^*(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \Lambda(\lambda)\} = x \log\left(\frac{x}{\mu}\right) - x + \mu.$$

Note that the minimum of Λ^* is achieved at $x = \mu$, and Λ^* is zero at that point.

Let ν_N denote the law of the r.v. $\frac{1}{N}\sum_{i=1}^{N} X_i$. We can now state Cramer's theorem, see e.g. Theorem 2.2.3 in [3].

Theorem 7.1 *For any closed set $F \subset \mathbb{R}$,*

$$\limsup_{N\to\infty} \frac{1}{N} \log \nu_N(F) \le -\inf_{x\in F} \Lambda^*(x).$$

For any open set $G \subset \mathbb{R}$,

$$\liminf_{N\to\infty} \frac{1}{N} \log \nu_N(G) \ge -\inf_{x\in G} \Lambda^*(x).$$

We deduce in particular (with $\mathring{A}$ the interior of A, $\bar{A}$ the closure of A)

Corollary 7.1 *For any Borel set $A \subset \mathbb{R}$,*

$$-\inf_{x\in\mathring{A}} \Lambda^*(x) \le \liminf_N \frac{1}{N}\log \nu_N(\mathring{A}) \le \limsup_N \frac{1}{N}\log\nu_N(\bar{A}) \le -\inf_{x\in\bar{A}} \Lambda^*(x).$$

In particular, if $\inf_{x\in\mathring{A}} \Lambda^*(x) = \inf_{x\in\bar{A}} \Lambda^*(x) = \inf_{x\in A} \Lambda^*(x)$,

$$\frac{1}{N}\log\nu_N(A) \to \inf_{x\in A} \Lambda^*(x) \text{ as } N \to \infty.$$

It is not surprising that the function $\Lambda_\mu^*(x)$ will appear again below, see the very beginning of the next section.

7.2.2 The Legendre-Fenchel Transform and the Rate Function

We define the following transforms. For $x \in A$, $y \in \mathbb{R}^d$, let

$$\ell(x,\mu) := \sum_j \left\{\beta_j(x) - \mu_j + \mu_j \log\left(\tfrac{\mu_j}{\beta_j(x)}\right)\right\},$$

with the convention $0 \log(0/\alpha) = 0$ for all $\alpha \in \mathbb{R}$, and

$$\overline{L}(x,y) := \begin{cases} \inf_{\mu \in V_{x,y}} \ell(x,\mu) & \text{if } V_{x,y} \neq \emptyset \\ +\infty & \text{otherwise.} \end{cases}$$

Now let, for x, y as above and $\theta \in \mathbb{R}^d$,

$$\tilde{\ell}(x,y,\theta) = \langle \theta, y \rangle - \sum_j \beta_j(x)\big(\mathrm{e}^{\langle \theta, h_j \rangle} - 1\big),$$

and define

$$\underline{L}(x,y) := \sup_{\theta \in \mathbb{R}^d} \tilde{\ell}(x,y,\theta),$$

Remark 7.1 For $\mu \in \tilde{V}_y \setminus V_{x,y}$, we have $\ell(x,\mu) = \infty$ and hence

$$\overline{L}(x,y) = \inf_{\mu \in \tilde{V}_y} \ell(x,\mu).$$

We first show

Lemma 7.1 *Let $x \in A$, $y \in \mathcal{C}_x$, $\theta \in \mathbb{R}^d$ and $\mu \in V_{x,y}$. Then*

$$\tilde{\ell}(x,y,\theta) \le \ell(x,\mu),$$

in particular

$$\underline{L}(x,y) \le \overline{L}(x,y).$$

Proof The result is obvious if $V_{x,y} = \emptyset$. If not, for $\mu \in V_{x,y}$, with

$$f_j(z) = \mu_j z - \beta_j(x)(e^z - 1),$$

$$\tilde{\ell}(x,y,\theta) = \sum_j \mu_j \langle \theta, h_j \rangle - \beta_j(x)(\exp(\langle \theta, h_j \rangle) - 1)$$

$$
\begin{aligned}
&= \sum_j f_j(\langle \theta, h_j \rangle) \\
&\le \sum_j f_j(\log \mu_j / \beta_j(x)) \\
&= \ell(x, \mu),
\end{aligned}
$$

since f_j achieves its maximum at $z = \log[\mu_j / \beta_j(x)]$.

We will show in Sect. 7.4.2 that under appropriate assumptions, $\underline{L}(x, y) = \overline{L}(x, y)$, and we shall write $L(x, y)$ for the common value of those two quantities.

For any $T > 0$, we define

$$
\begin{aligned}
C([0, T]; A) &:= \{\phi : [0, T] \to A \mid \phi \text{ continuous}\}, \\
D([0, T]; A) &:= \{\phi : [0, T] \to A \mid \phi \text{ càdlàg}\}.
\end{aligned}
$$

On $C([0, T]; A)$ (or $D([0, T]; A)$), d_C denotes the metric corresponding to the supremum-norm, denoted by $\|\cdot\|$. Whenever the context is clear, we write $\mathrm{d} := \mathrm{d}_C$. On $D([0, T]; A)$ we denote by d_D the metric given, e.g., in [2], Sections 12.1 and 12.2 which defines the Skorohod topology in such a way that the resulting space is Polish. The resulting metric spaces are denoted by $C([0, T]; A; \mathrm{d}_C)$, $D([0, T]; A; \mathrm{d}_C)$ and $D([0, T]; A; \mathrm{d}_D)$, respectively (where the metrics are omitted, whenever they are clear from the context).

We now introduce a candidate I for the rate function. For $\phi : [0, T] \to A$, let

$$
I_T(\phi) := \begin{cases} \int_0^T L(\phi(t), \phi'(t)) dt & \text{if } \phi \text{ is absolutely continuous} \\ \infty & \text{otherwise.} \end{cases}
$$

For $x \in A$ and $\phi : [0, T] \to A$, let

$$
I_{T,x}(\phi) := \begin{cases} I_T(\phi) & \text{if } \phi(0) = x \\ \infty & \text{otherwise.} \end{cases}
$$

7.2.3 *Assumptions on the Process* Z^N

We do not assume that the rates are bounded away form zero (as in [13]) and allow for them to vanish near the boundary (cf. the discussion in the introduction). Shwartz and Weiss [14] allow for vanishing rates. We generalize these assumptions as we outline below. The difference is essentially Assumption 7.2 (C) below.

Assumption 7.2

(A) Assumptions on the set A.

(A1) The set A is compact and $A = \overline{\mathring{A}}$*. Furthermore, there exists a constant* $\lambda_0 > 0$ *such that for all* $N \in \mathbb{N}$*,* $z \in A^N$ *and* $j = 1, \dots, k$ *with* $\beta_j(z) > 0$*,*[2]

$$z + \frac{h_j}{N} \in A^N \quad \textit{and} \quad |\tilde{z} - z| \geq \frac{\lambda_0}{N} \textit{ for all } \tilde{z} \textit{ with } \beta_j(\tilde{z}) = 0.$$

(A2) There exist open balls[3] $B_i = B(x_i, r_i)$*,* $i = 1, \dots, I_1, \dots, I$ *(*$0 < I_1 < I$*) such that*

$$x_i \in \partial A \textit{ for } i \leq I_1 \quad \textit{and} \quad x_i \in \mathring{A} \textit{ for } i > I_1$$

and

$$A \subset \bigcup_{i \leq I} B_i, \quad \partial A \subset \bigcup_{i \leq I_1} B_i \quad \textit{and} \quad B_i \cap \partial A = \emptyset \textit{ for } i > I_1.$$

(A3) There exist (universal) constants $\lambda_1, \lambda_2 > 0$ *and vectors* v_i *(*$i \leq I_1$*, w.l.o.g., we assume* $0 < |v_i| \leq 1$*; for notational reasons, we set* $v_i = 0$ *for* $i > I_1$*) such that for all* $x \in B_i \cap A$*,*

$$B(x + tv_i, \lambda_1 t) \subset A \quad \textit{for all } t \in (0, \lambda_2).$$

and $\operatorname{dist}(x + tv_i, \partial A)$ *is increasing for* $t \in (0, \lambda_2)$*.*

(A4) There exists a Lipschitz continuous mapping $\psi_A : \mathbb{R}^d \to A$ *such that* $\psi_A(x) = x$ *whenever* $x \in A$*.*

(B) Assumptions on the rates β_j*.*

(B1) The rates $\beta_j : A \to \mathbb{R}_+$ *are Lipschitz continuous.*

(B2) For $x \in \mathring{A}$*,* $j = 1, \dots, k$*,* $\beta_j(x) > 0$ *and* $\mathcal{C}((h_j)_j) = \mathbb{R}^d$*.*

(B3) For all $x \in \partial A$ *there exists a constant* $\lambda_3 = \lambda_3(x) > 0$ *such that*

$$y \in \mathcal{C}_x, |y| \leq \lambda_3 \Rightarrow x + y \in A.$$

[2] This implies that for $z \in A^N$ with $\beta_j(z) > 0$, we have $\beta_j(z) \geq \underline{\beta}(\lambda_0/N)$.

[3] Here (and later) $B(x, r)$ denotes the open ball around x with radius r.

(B4) There exists a (universal) constant $\lambda_4 > 0$ *such that for all* $i \leq I_1$, $x \in B_i \cap A$ *and*

$$v \in \mathcal{C}_{1,i} := \left\{ \frac{\tilde{v}}{|\tilde{v}|} \middle| \tilde{v} = v_i + w \text{ for } w \in \mathbb{R}^d, |w| \leq \frac{\lambda_1}{3 - \lambda_1} \right\},$$

$$\beta_j(x) < \lambda_4 \Rightarrow \beta_j(x + \cdot v) \text{ is increasing in } (0, \lambda_2).$$

(C) There exists an $\eta_0 > 0$ *such that for all* $N \in \mathbb{N}$, $\epsilon > 0$ *there exists a constant* $\delta(N, \epsilon) > 0$ *(decreasing in* N *and in* ϵ*) such that for all* $i \leq I_1$, $x \in B_i$ *there exists a* $\mu^i = \mu^i(x) \in \tilde{V}_{v_i}$ *and*

$$\mathbb{P}\Big[\sup_{t \in [0, \eta_0]} |\tilde{Z}^{N,x}(t) - \phi^x(t)| \geq \epsilon \Big] \leq \delta(N, \epsilon),$$

where $\tilde{Z}^{N,x}$ *denotes the solution of* (7.3) *if the rates* β_j *are replaced by the rates* $\tilde{\mu}^i_j$ *for*

$$\tilde{\mu}^i_j(z) := \begin{cases} \mu^i_j & \text{if } z + \epsilon h_j \in A \text{ for all } \epsilon \text{ small enough} \\ 0 & \text{else} \end{cases}$$

and $\phi^x = x + t v_i$ *as before.*[4]

Furthermore, there exists a constant $\alpha \in (0, 1/2)$ *and a sequence* ϵ_N *such that*

$$\epsilon_N < \frac{1}{N^\alpha} \quad \text{and} \quad \frac{\delta(N, \epsilon_N)}{\epsilon_N} \to 0 \text{ as } N \to \infty. \tag{7.5}$$

and

$$\rho^\alpha \log \underline{\beta}(\rho) \to 0 \quad \text{as } \rho \to 0. \tag{7.6}$$

Let us comment on Assumption 7.2. Assumption (A) is essentially Assumption 2.1 of [14]. We want to remark that Assumption 2.1 (iv) of [14] is not included here as it is redundant (see Lemma 3.5 of [14]; cf. also the discussion preceding Lemma 7.23). In the epidemiological models we want to consider, A is a compact, convex d-polytope, i.e., ∂A is composed by $d - 1$-dimensional hyperplanes. For example for the SIRS model in Example 7.1,

$$A = \{x \in \mathbb{R}^2 | 0 \leq x_1 + x_2 \leq 1\}.$$

[4]We do not necessarily have $\mu^i \in V_{x,v_i}$ for all $x \in B_i$; this might not be the case if $x \in \partial A$. In such a case $V_{x,v_i} = \emptyset$ is possible, cf. the discussion about $x = (1, 0)^\top$ for the SIRS model below.

In line with the Assumption (A1), let us note that we always want to choose the starting point x of Eq. (7.1) to belong to A^N. If that would not be the case, then in our simplest models the solution $Z^{N,x}$ might exit the domain A. Choosing the starting point arbitrarily in A would force us to let the rates β_j depend upon N (and vanish) near the boundary. Note that the coordinates of the vector $Z_t^{N,x}$ are proportions of the population in various compartments. The coordinates of the vector $NZ_t^{N,x}$ are integers, while those of the vectors h_j belong to the set $\{-1, 0, 1\}$.

Note that in all situations we have in mind, both the set A itself and its boundary can be covered by a finite number of balls. These balls can furthermore be chosen in such a way that those centered in the interior do not intersect with the boundary. For the SIRS model, we can for instance define the balls covering the boundary by $B(x, 3/(4m))$ for large $m \in \mathbb{N}$ and $x = (i/m, j/m)^\top$ for $i = 0, \ldots, m$, $j = 0$ or $i = 0$, $j = 0, \ldots, m$ or $i + j = m$. The vectors v_i can be defined to be the inside normal vectors for those balls with $x \notin \{(0,0)^\top, (1,0)^\top, (0,1)^\top\}$. For the remaining three balls, we define v_i by the normalizations of

$$(1/2, 1/2)^\top, \quad (-1/2, 1/4)^\top \quad \text{respectively} \quad (1/2, -1/2)^\top.$$

In general, the constant λ_1 can be interpreted to be given via the "angle" of the vector v_i to the boundary. We have $\lambda_1 \leq 1$. It is straightforward that Assumption (A) is satisfied for the SIRS model. We also note that Assumption (A) is not very restrictive, see [14] Lemma 2.1. In particular, every convex, compact set with non-empty interior satisfies the assumption.

Most of Assumption (B) is taken from Assumption 2.2 of [14]. We outline the difference below. Assumption (B1) is quite standard and ensures in particular that the ODE (7.2) admits a unique solution. For the compartmental epidemiological models we consider, the rates are usually polynomials and hence this assumption is satisfied. Assumption (B2) implies that within $\mathring{A}$ it is possible to move into all directions. Only by approaching the boundary the rates are allowed to vanish. (B3) implies that at least locally the convex cone $x + \mathcal{C}_x$ is included in A. In particular, it is not possible to exit the set A from its boundary. Assumption (B4) differs slightly from the corresponding assumption in [14]. While in [14], it is implied that close to the boundary, "small" rates are increasing while following the vector v_i, we assume this for a set of vector in a "cone" around v_i. We note that for $i \leq I_1$, $x \in B_i$, $v \in \mathcal{C}_{1,i}$, we have (cf. Assumption (A3))

$$\operatorname{dist}(x + tv, \partial A) \geq \operatorname{dist}(x + tv_i, \partial A) - t\frac{\lambda_1}{3 - \lambda_1} \geq t\lambda_1\Big(\frac{2 - \lambda_1}{3 - \lambda_1}\Big).$$

It is easily seen that this assumption is satisfied for the SIRS model. In addition to this, [14] also require that (cf. the meaning of λ_4 in Assumption (B4))

$$v_i \in \mathcal{C}\big(\{h_j | \inf_{x \in B_i} \beta_j(x) > \lambda_4\}\big). \tag{7.7}$$

In order to apply the theory to epidemiological models, we have to remove this assumption. To see this, consider the SIRS model and the point $x = (1, 0)^\top$ with corresponding ball B containing it. We readily observe that a vector v pointing "inside" A (as required by Assumption (A3)) which is generated by only those h_j whose corresponding rates are bounded away from zero in B does not exist. We hence replace this assumption by Assumption (C), which follows from (7.7). Indeed, if Assumption (7.7) holds, the μ^i representing v_i can be chosen in such a way that the directions corresponding to components $\mu^i_j > 0$ do not point outside A in B_i. Hence, $\tilde{\mu}^i \equiv \mu^i$ (as long as the process is in B_i) and the LLN Theorem 7.3 can be applied. In general, Theorem 7.3 cannot be applied as the rates $\tilde{\mu}^i$ can be discontinuous. Note that the assumption can only fail if $x \in \partial A$. Else, the process is equal to the process with constant rates μ^i on the set

$$\left\{ \sup_{t\in[0,\eta_0]} |\tilde{Z}^{N,x}(t) - \phi^x(t)| < \epsilon \right\}$$

for all small enough $\epsilon > 0$ and Theorem 7.3 is applicable. We note that Assumption (C) implies that

$$\delta(N, \epsilon) \to 0 \text{ as } N \to \infty \text{ for all } \epsilon > 0.$$

Moreover, as $\delta(N, \cdot)$ is decreasing, we can choose ϵ_N in such a way that

$$\epsilon_N = \frac{1}{N^\alpha} \quad \text{for some } \alpha \in (0, 1). \tag{7.8}$$

We remark here (and further discuss this important issue below) that Assumption 7.2 (C) may well fail to be satisfied. To this end, we consider the SIRS model an $x \in A$ with $x_1 = 0$. We hence have $\beta_2(x) = 0$ and hence the process $Z^{N,x}$ cannot enter the interior of A. Therefore, we have

$$\mathbb{P}\Big[\sup_{t\in[0,\eta_0]} |Z^{N,x}(t) - \phi^x(t)| \geq \epsilon \Big] = 1$$

for ϵ small enough. Assumption (C) can hence be considered as a means to ensure that the process can enter the interior of A from every point on the boundary (7.6) implies that

$$\int_0^\eta |\log \underline{\beta}(\rho)| d\rho \to 0 \quad \text{as } \eta \to 0, \tag{7.9}$$

since $\rho^{\alpha/2}|\log \underline{\beta}(\rho)| \leq C$ for appropriate C and hence $|\log \underline{\beta}(\rho)| \leq C/\rho^{\alpha/2}$ is integrable, and hence in particular that the rate $I(\phi)$ of linear functions ϕ is finite, as it is shown in Lemma 7.25 below.

It remains to show that Assumption 7.2 (C) is satisfied for the SIRS model. This is accomplished in Sect. 7.8.

Exploiting (7.6), it is easy to prove

Lemma 7.2 *Under the Assumption 7.2 (C), for all $i \le I$, $x \in A \cap B_i$, let $\phi^x(t) := x + tv_i$. For all $\epsilon > 0$, there exists an $\eta > 0$ (independent of i, x) such that for all $i \le I$ and all $x \in A \cap B_i$,*

$$I_{\eta,x}(\phi^x) < \epsilon.$$

Note that for $i > I_1$, we have $\phi^x(t) = x$ for all t.

7.3 Law of Large Numbers

We first prove the law of large numbers as stated in Theorem 5.3 from [13], which adds a rate of convergence to the classical result by Kurtz [9].

Theorem 7.3 *Let $Z^{N.x}$ and Y^x be given as in Eqs. (7.3) and (7.2) respectively, and assume that the rates β_j are bounded and Lipschitz continuous. Then there exist constants $\tilde{C}_1 = \tilde{C}_1(T) > 0$ (independent of ϵ) and $\tilde{C}_2(\epsilon) = \tilde{C}_2(T, \epsilon) > 0$ with $\tilde{C}_2(\epsilon) = O(\epsilon^2)$ as $\epsilon \to 0$ such that*

$$\mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x}(t) - Y^x(t)| \ge \epsilon\Big] \le \tilde{C}_1 \exp\big(-N\tilde{C}_2(\epsilon)\big).$$

C_1 and C_2 can be chosen independently of x.

Before we prove Theorem 7.3, we require some auxiliary results. We first have

Lemma 7.3 *Let $T > 0$. Suppose that $f : D([0,T]; A) \times \mathbb{R} \to \mathbb{R}$ and $G : D([0,T]; A) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are such that for all $\rho > 0$,*

$$M(t) := \exp\big(\rho f(Z^{N,x}, t) - G(Z^{N,x}, t, \rho)\big)$$

is a right-continuous martingale with mean one. Suppose furthermore that $R : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is increasing in the first argument and

$$G(\phi, t, \rho) \le R(t, \rho)$$

for all $\phi \in D([0,T]; A)$ and $\rho > 0$. Then for all $\epsilon > 0$

$$\mathbb{P}\Big[\sup_{t\in[0,T]} f(Z^{N,x}, t) \ge \epsilon\Big] \le \inf_{\rho>0} \exp\big(R(T,\rho) - \rho\epsilon\big).$$

Proof Fix $\rho > 0$. Then by the assumptions of the lemma,

$$\begin{aligned}\mathbb{P}\Big[\sup_{t\in[0,T]} f(Z^{N,x},t) \geq \epsilon\Big] &= \mathbb{P}\Big[\sup_{t\in[0,T]} \exp\big(\rho f(Z^{N,x},t)\big) \geq \exp\big(\rho\epsilon\big)\Big] \\ &\leq \mathbb{P}\Big[\sup_{t\in[0,T]} \exp\big(\rho f(Z^{N,x},t) - G(Z^{N,x},t,\rho)\big) \\ &\geq \exp\big(\rho\epsilon - R(T,\rho)\big)\Big] \\ &\leq \exp\big(R(T,\rho) - \rho\epsilon\big)\end{aligned}$$

where the last inequality is Doob's martingale inequality, see, e.g. Theorem II.1.7 in [11].

The next result is an easy exercise which we leave to the reader.

Lemma 7.4 *Let Y be a d-dimensional random vector. Suppose that there exist positive numbers a and δ such that for all $\theta \in \mathbb{R}^d$ with $|\theta| = 1$*

$$\mathbb{P}\big[\langle\theta, Y\rangle \geq a\big] \leq \delta.$$

Then

$$\mathbb{P}\big[|Y| \geq a\sqrt{d}\big] \leq 2d\delta.$$

The main step towards the proof of Theorem 7.3 is the following Lemma

Lemma 7.5 *Assume that β_j ($j = 1,\ldots,k$) is bounded and that Y^x is a solution of (7.2). Then for all $\theta \in \mathbb{R}^d$ with $|\theta| = 1$ and all $T > 0$, there is a function $\tilde{C} : \mathbb{R}_+ \to \mathbb{R}_+$ (independent of x) such that*

$$\mathbb{P}\Big[\sup_{t\in[0,T]} \Big\{\langle Z^{N,x}(t) - Y^x(t), \theta\rangle - \int_0^t \sum_{j=1}^{k} \big(\beta_j(Z^{N,x}(s)) - \beta_j(Y^x(s))\big)\langle h_j, \theta\rangle ds\Big\} \geq \epsilon\Big] \leq \exp\big(-N\tilde{C}(\epsilon)\big),$$

and moreover

$$0 < \lim_{\epsilon\to 0} \tilde{C}(\epsilon)/\epsilon^2 < \infty, \quad \textit{and} \quad \lim_{\epsilon\to\infty} \tilde{C}(\epsilon)/\epsilon = \infty.$$

Proof Let

$$\mathcal{N}_t^\theta = \langle Z^{N,x}(t) - Y^x(t), \theta\rangle - \int_0^t \sum_{j=1}^k \big(\beta_j(Z^{N,x}(s)) - \beta_j(Y^x(s))\big)\langle\theta, h_j\rangle ds$$

$$= \frac{1}{N}\sum_{j=1}^k \langle h_j, \theta\rangle M_j \left(N\int_0^t \beta_j(Z_s^{N,x})ds\right).$$

We want to use Lemma 7.3, with $f(Z^{N,x}, t) = \mathcal{N}_t^\theta$. It is not hard to check that if we define

$$G(Z^{N,x}, t, \rho) = N\sum_{j=1}^k \left(e^{\frac{\rho}{N}\langle h_j,\theta\rangle} - 1 - \frac{\rho}{N}\langle h_j, \theta\rangle\right)\int_0^t \beta_j(Z^{N,x}(s))ds,$$

we have that

$$M(t) = \exp\left(\rho f(Z^{N,x}, t) - G(Z^{N,x}, t, \rho)\right)$$

is a martingale. Hence from Lemma 7.3, with $a = \rho/N$,

$$\mathbb{P}\left(\sup_{0\le t\le T} \mathcal{N}_t^\theta > \epsilon\right) \le \min_{a>0} \exp\left(N\overline{\beta}T\left[\sum_{j=1}^k \left\{e^{a\langle h_j,\theta\rangle} - 1 - a\langle h_j, \theta\rangle\right\} - a\epsilon\right]\right).$$

The main inequality of the Lemma is established, with

$$\tilde{C}(\epsilon) = \overline{\beta}T \max_{a>0}\left[a\epsilon - \sum_{j=1}^k \left\{e^{a\langle h_j,\theta\rangle} - 1 - a\langle h_j, \theta\rangle\right\}\right].$$

It is not hard to show that as $\epsilon \to 0$,

$$\frac{\tilde{C}(\epsilon)}{\epsilon^2} \to \frac{\overline{\beta}T}{2\sum_{j=1}^k \langle h_j, \theta\rangle^2}.$$

Consider now the case where ϵ is large. If $\langle h_j, \theta\rangle \le 0$ for $1 \le j \le k$, then for $\epsilon > -\sum_j \langle h_j, \theta\rangle$, $\tilde{C}(\epsilon) = +\infty$, which means that a certain event has probability zero. Now consider the more interesting case where $\langle h_j, \theta\rangle > 0$ for at least one $1 \le j \le k$. If we choose a_ϵ such that

$$\sum_{j=1}^k \left\{e^{a_\epsilon\langle h_j,\theta\rangle} - 1 - a_\epsilon\langle h_j, \theta\rangle\right\} = \epsilon,$$

then $a_\epsilon \to \infty$ as $\epsilon \to \infty$, while $\tilde{C}(\epsilon) \geq \epsilon(a_\epsilon - 1)$, which completes the proof of the Lemma.

Proof (Proof of Theorem 7.3) We deduce from Lemma 7.5 and a variant of Lemma 7.7 that

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\frac{1}{N}\left|\sum_{j=1}^{k} h_j M_j\left(N\int_0^t \beta_j(Z_s^{N,x})ds\right)\right| > \epsilon\right) \leq 2de^{-N\tilde{C}'(\epsilon)}, \tag{7.10}$$

where $\tilde{C}'(\epsilon) = \tilde{C}(\epsilon/\sqrt{d})$. In view of the Lipschitz property of b, we have

$$\begin{aligned} Z_t^{N,x} - Y_t^x &= \int_0^t \left[b(Z_s^{N,x}) - b(Y_s^x)\right]ds \\ &\quad + \frac{1}{N}\sum_{j=1}^{k} h_j M_j\left(N\int_0^t \beta_j(Z_s^{N,x})ds\right), \\ \sup_{0\leq s\leq t}\left|Z_s^{N,x} - Y_s^x\right| &\leq K\int_0^t \sup_{0\leq r\leq s}\left|Z_r^{N,x} - Y_r^x\right|ds \\ &\quad + \sup_{0\leq s\leq t}\frac{1}{N}\left|\sum_{j=1}^{k} h_j M_j\left(N\int_0^s \beta_j(Z_r^{N,x})dr\right)\right|. \end{aligned}$$

The result now follows from (7.10) and Gronwall's Lemma.

We can deduce from Theorem 7.3.

Corollary 7.2 *Let M be a compensated standard Poisson process. Then there exist constants $C_1 = C_1(T) > 0$ (independent of ϵ) and $C_2(\epsilon) = C_2(T,\epsilon) > 0$ with $C_2(\epsilon) = O(\epsilon^2)$ as $\epsilon \to 0$ such that*

$$\mathbb{P}\Big[\sup_{t\in[0,T]}\frac{|M(tN)|}{N} \geq \epsilon\Big] \leq C_1\exp\big(-NC_2(\epsilon)\big).$$

C_1 and C_2 can be chosen independently of x.

Proof We apply Theorem 7.3 to $d = k = 1$, $\beta_1(x) \equiv 1$ and $h_1 = 1$. Hence,

$$|Z^N(t) - Y(t)| = \frac{|M(tN)|}{N}.$$

The result follows directly.

We shall need below the

Lemma 7.6 *Let β_j $(j = 1, \dots, k)$ be bounded. Then there exist positive constants $\tilde{C}_1$ and $\tilde{C}_2$ independent of x such that for all $0 \le s < t \le T$ and for all $\epsilon > 0$,*

$$\mathbb{P}\Big[\sup_{r\in[s,t]} |Z^{N,x}(r) - Z^{N,x}(s)| \ge \epsilon\Big] \le \exp\Big(- N\epsilon\tilde{C}_1 \log\Big(\frac{\epsilon\tilde{C}_2}{t-s}\Big)\Big).$$

Proof Let $\xi^N_{s,t}$ denote the number of jumps of the process $Z^{N,x}$ on the time interval $[s, t]$. It is plain that

$$\{\sup_{r\in[s,t]} |Z^{N,x}(r) - Z^{N,x}(s)| \ge \epsilon\} \subset \{\xi^N_{s,t} \ge CN\epsilon\},$$

for some universal constant $C > 0$. Now $\xi^N_{s,t}$ is stochastically dominated by a Poisson random variable with parameter $C'N(t-s)$, for some other constant $C' > 0$. Now let Θ be a Poisson r.v. with parameter λ. We claim that for any $b > 0$,

$$\mathbb{P}(\Theta > b) \le \exp\left(-b\log\left(\frac{b}{e\lambda}\right)\right). \tag{7.11}$$

Indeed, if $b \le e\lambda$ this is obvious, and if $b > e\lambda$ it is obtained by applying Chebychef's inequality to the r.v. Θ and the function $\exp[x\log(b/\lambda)]$. The result follows by applying (7.11) with $\lambda = C'N(t-s)$, and $b = CN\epsilon$.

7.4 Properties of the Rate Function

7.4.1 Properties of the Legendre Fenchel Transform

In this subsection we assume that the β_j's are bounded and continuous. We recall that ℓ, $\overline{L}$, $\tilde{\ell}$ and $\underline{L}$ have been defined in Sect. 7.2.2, and start with

Lemma 7.7

1. *For all $x \in A$, $\underline{L}(x,\cdot) : \mathcal{C}_x \to \mathbb{R}_+$ is convex and lower semicontinuous.*
2. *For all $y \in \mathbb{R}^d$,*

$$\underline{L}(x, y) \ge \underline{L}\Big(x, \sum_j \beta_j(x)h_j\Big) = 0$$

with strict inequality if $y \ne \sum_j \beta_j(x)h_j$.

Proof

1. We readily observe that $\tilde{\ell}(x, \cdot, \theta)$ is linear and hence convex. As the supremum of these functions, the function $\underline{L}(x, \cdot)$ is convex.

   Lower semicontinuity follows as $\underline{L}(x, \cdot)$ is the supremum of a family of continuous functions.
2. Let first $y = \sum_j \beta_j(x)h_j$. We have

$$\begin{aligned}\underline{L}(x, y) &= \sup_\theta \Big\{ \sum_j \beta_j(x)\langle h_j, \theta\rangle - \sum_j \beta_j(x)\big(\exp\langle h_j, \theta\rangle - 1\big)\Big\} \\ &= \sup_\theta \Big\{ \sum_j \beta_j(x)\big(\langle h_j, \theta\rangle - \exp\langle h_j, \theta\rangle + 1\big)\Big\} \\ &= 0\end{aligned}$$

as $\beta_j(x) \geq 0$ and $e^z \geq z + 1$ for all $z \in \mathbb{R}$ with equality for $z = 0$.

Let now y be such that $\underline{L}(x, y) = 0$. This implies

$$\langle y, \theta\rangle - \sum_j \beta_j(x)\big(\exp\langle h_j, \theta\rangle - 1\big) \leq 0 \quad \text{for all } \theta \in \mathbb{R}^d,$$

in particular for $\theta = \epsilon e_i$ (where e_i is the ith unit-vector and $\epsilon > 0$; in the following h^i_j is the ith component of h_j),

$$\epsilon y_i \leq \sum_j \beta_j(x)\big(\exp(\epsilon h^i_j) - 1\big).$$

Dividing by ϵ and letting $\epsilon \to 0$, we deduce that

$$y_i \leq \sum_j \beta_j(x)h^i_j.$$

For $\theta = -\epsilon e_i$ the converse inequality follows accordingly.

Remark 7.2 The function $\underline{L}(x, \cdot)$ is even strictly convex, see Corollary 7.3 below.

Lemma 7.8 *Assume that β_j ($j = 1, \ldots, k$) is bounded. Then, there exist constants C_1 and B_1 such that for all $|y| \geq B_1$, $x \in \mathbb{R}^d$,*

$$\underline{L}(x, y) \geq C_1|y|\log(|y|).$$

Proof Let

$$\theta := y \frac{\log|y|}{\bar{h}|y|},$$

hence provided $|y| \geq 1$,

$$\underline{L}(x, y) \geq \frac{|y| \log|y|}{\bar{h}} - k\bar{\beta}|y|$$

which grows like $|y| \log|y|$ as $|y| \to \infty$.

We now have

Lemma 7.9 *There exists a constant C_2 such that for all $x \in A$, $y \in \mathcal{C}_x$, there exists a $\mu \in V_{x,y}$ with*

$$|\mu| \leq C_2|y|.$$

Proof We first note that there are only finitely many convex cones $\mathcal{C}_x$ and we can hence restrict our attention to a fixed $x \in A$.

We proceed by contradiction. Assume that for all n there exists a $y^n \in \mathcal{C}_x$ such that for all $\mu \in V_{x,y^n}$,

$$|\mu| \geq n|y^n|.$$

We note that for any $y \in \mathcal{C}_x$, there exists a minimal representation $\mu \in V_{x,y}$ (in the sense that $\tilde{\mu} \in V_{x,y} \Rightarrow \max_j \tilde{\mu}_j \geq \max_j \mu_j$). Indeed, let $\{\mu^n, : n \geq 1\} \subset V_{x,y}$ be such that, as $n \to \infty$,

$$\max_j \mu_j^n \downarrow \inf_{\mu \in V_{x,y}} \left(\max_j \mu_j \right).$$

There exists a subsequence along which $\mu^n \to \mu \in \mathbb{R}_+^d$ as $n \to \infty$. If $\mu_j > 0$, we have $\mu_j^n > 0$ for n large enough and hence $\beta_j(x) > 0$. Hence $\mu \in V_{x,y}$ since moreover

$$\sum_j \mu_j h_j = \lim_n \sum_j \mu_j^n h_j = y.$$

Given y^n, we denote this minimal representation by $\bar{\mu}^n$. We now define

$$\tilde{y}^n := \frac{y^n}{|\bar{\mu}^n|}, \quad \text{hence} \quad |\tilde{y}^n| \leq \frac{1}{n}.$$

Furthermore, it is easy to see that minimal representations for the $\tilde{y}^n$ are given by

$$\tilde{\mu}^n := \frac{\bar{\mu}^n}{|\bar{\mu}^n|}, \quad \text{hence} \quad |\tilde{\mu}^n| = 1.$$

Boundedness implies (after possibly the extraction of a subsequence) $\tilde{\mu}^n \to \tilde{\mu}$ with $|\tilde{\mu}| = 1$. We let n large enough such that for all j

$$\tilde{\mu}^n_j > 0 \Rightarrow \tilde{\mu}^n_j > \frac{\tilde{\mu}_j}{2}$$

(note that for at least one j, $\tilde{\mu}_j > 0$). We have

$$0 = \lim_n \tilde{y}^n = \lim_n \sum_j \tilde{\mu}^n_j h_j = \sum_j \tilde{\mu}_j h_j$$

and therefore

$$\tilde{y}^n = \sum_j \tilde{\mu}^n_j h_j = \sum_{j;\tilde{\mu}_j>0} \underbrace{\left(\tilde{\mu}^n_j - \frac{\tilde{\mu}_j}{2}\right)}_{=:\hat{\mu}^n_j>0} h_j + \sum_{j;\tilde{\mu}_j=0} \underbrace{\tilde{\mu}^n_j}_{=:\hat{\mu}^n_j} h_j,$$

a contradiction to the minimality of the $\tilde{\mu}^n_j$.

We require the following result

Lemma 7.10 *Let* $x \in A$.

1. $\ell(x,\mu) \geq 0$ *for* $\mu \in V_x$ *and* $\ell(x,\cdot) : V_x \longrightarrow \mathbb{R}_+$ *is strictly convex and has compact level sets* $\{\mu \in V_x | \ell(x,\mu) \leq \alpha\}$.
2. *Let* $y \in \mathcal{C}_x$. *Then there exists a unique* $\mu^* = \mu^*(y)$ *such that*

$$\ell(x,\mu^*) = \inf_{\mu \in V_{x,y}} \ell(x,\mu).$$

3. *There exist constants* $C_3, C_4, C_5, B_2 > 0$ *(which depend only upon* $\sup_{x\in A} \max_{i\leq j\leq k} \beta_j(x)$*), such that*

$$|\mu^*(y)| \leq C_3|y| \quad \text{if } |y| > B_2, \tag{7.12}$$

$$|\mu^*(y)| \leq C_4 \quad \text{if } |y| \leq B_2, \tag{7.13}$$

$$|\mu^*(y)| \geq C_5|y| \quad \text{for all } y. \tag{7.14}$$

4. $\overline{L}(x,\cdot), \mu^* : \mathcal{C}_x \to \mathbb{R}_+$ *are continuous.*

Proof

1. We define the function $f(z) = 1 - z + z \log z$ for $z \geq 0$ and note that for $\mu \in V_x$,

$$\ell(x, \mu) = \sum_{j, \beta_j(x) > 0} \beta_j(x) f\left(\frac{\mu_j}{\beta_j(x)}\right).$$

We readily observe (by differentiation) that $f \geq 0$ and that f is strictly convex. Thus the first two assertions follow.

As V_x is closed and $\ell(x, \cdot)$ is continuous, the level sets are closed. Compactness follows form the fact that $\lim_{x \to \infty} f(x) = \infty$.

2. Existence of a minimizer follows from the fact that $V_{x,y}$ is closed. Uniqueness follows from the strict convexity of $\ell(x, \cdot)$.
3. By the definition of ℓ, there exists a $B_2 = B_2(\bar{\beta}(x)) > 0$ and $C = C(\bar{\beta}(x)) > 0$ such that for $y \in \mathcal{C}_x$ with $|y| \geq B_2$ (and appropriate $\mu \in V_{x,y}$ according to Lemma 7.9),

$$\ell(x, \mu^*(y)) \leq \ell(x, \mu) \leq C|y| \log |y|.$$

On the other hand, assume that for all n there exists an $y^n \in \mathcal{C}_x$ with $|y^n| \geq B_2$ such that

$$|\mu^*(y^n)| \geq n|y^n|.$$

This implies for an appropriate constant $\tilde{C}$ and n large enough

$$\ell(x, \mu^*(y^n)) \geq n\tilde{C}|y| \log |y|,$$

a contradiction. Hence Inequality (7.12) follows.

Assume now that for all n there exists an $y^n \in \mathcal{C}_x$ with $|y^n| \leq B_2$,

$$|\mu^*(y^n)| \geq n, \quad \text{hence } \lim_{n \to \infty} \ell(x, \mu^*(y^n)) \to \infty.$$

However, Lemma 7.9 implies that there exists an $\mu^n \in V_{x,y^n}$ and a constant $C = C(\bar{\beta}(x), B_2)$ independent of n with

$$\ell(x, \mu^n) \leq C,$$

a contradiction. Hence Inequality (7.13) follows.

Finally, Inequality (7.14) follows from the definition of $V_{x,y}$.

4. Let $y, y^n \in \mathcal{C}_x$ with $y^n \to y$. By 3., the sequence $(\mu^*(y^n))_n$ is bounded and hence there exists a convergent subsequence, say (by abuse of notation)

$$\mu^*(y^n) \to \mu^* \quad \text{with } \mu^*_j \geq 0 \text{ for all } j.$$

In particular, we have

$$\sum_j \mu_j^* h_j = y. \tag{7.15}$$

We have

$$\begin{aligned} y^n &= \sum_j \mu_j^*(y^n) h_j \\ &= (1-\epsilon^n) \sum_j \mu_j^* h_j + \sum_j \big(\mu_j^*(y^n) - \mu_j^* + \epsilon^n \mu_j^*\big) h_j \\ &= (1-\epsilon^n) \sum_j \mu_j^*(y) h_j + \sum_j \tilde{\mu}_j^n h_j, \end{aligned} \tag{7.16}$$

where we have used (7.15), $\mu^*(y) = \operatorname{argmax}_\mu \ell(x, \mu)$, $\tilde{\mu}_j^n = \mu_j^*(y^n) - \mu_j^* + \epsilon^n \mu_j^*$, and

$$\epsilon^n = \begin{cases} \frac{2 \max_j |\mu_j^*(y^n) - \mu_j^*|}{\min_{j;\mu_j^*>0} \mu_j^*}, & \text{if } \min_{j;\mu_j^*>0} \mu_j^* > 0; \\ 1/n, & \text{otherwise.} \end{cases}$$

In particular, we have $0 \le \tilde{\mu}_j^n \to 0$ as $n \to \infty$. By Eq. (7.16), 2. and the continuity of ℓ, we have

$$\begin{aligned} \ell(x, \mu^*(y^n))) &\le \ell(x, (1-\epsilon^n)\mu^*(y) + \tilde{\mu}^n) \\ &\le \ell(x, \mu^*(y)) + \delta(n) \end{aligned} \tag{7.17}$$

with $\delta(n) \to \infty$ as $n \to \infty$. This implies (again by the continuity of ℓ)

$$\ell(x, \mu^*) \le \ell(x, \mu^*(y))$$

and hence $\mu^* = \mu^*(y)$ by 2. As this holds true for all convergent subsequences of $(\mu^*(y^n))_n$, this establishes the continuity of $\mu^*(\cdot)$.

The continuity of $\overline{L}(x, \cdot)$ follows directly from this and the continuity of ℓ.

Remark 7.3 Assume that for $x \in A$, $\mathcal{C}_x = \mathcal{C}_{\tilde{x}}$ for all $\tilde{x}$ in some neighborhood U of x. Then the function $\ell : U \times V_{x,y} \to \mathbb{R}_+$ is continuous and hence we have that $\mu^*(y) = \mu^*(x, y)$ as given in Lemma 7.10 is also continuous in x (as the argmin of a continuous function).

We have moreover

Lemma 7.11

1. *Let* $x \in A$*. For all* $B > 0$*, there exists a constant* $C_6 = C_6(x, B) > 0$ *such that for all* $y \in \mathcal{C}_x$ *with* $|y| \leq B$ *and* $\theta \in \mathbb{R}^d$ *with* $\tilde{\ell}(x, y, \theta) \geq -1$*,*[5]

$$\langle \theta, h_j \rangle \leq C_6 \quad \textit{for all } j \textit{ with } \beta_j(x) > 0.$$

If $\log \beta_j(\cdot)$ $(j = 1 \ldots, k)$ *is bounded,* C_6 *can be chosen independently of* x*.*

2. *Let* $x \in A$ *and* $y \in \mathcal{C}_x$*. If* $(\theta_n)_n$ *is a maximizing sequence of* $\tilde{\ell}(x, y, \cdot)$ *and for some* $j = 1, \ldots, k$*,*

$$\liminf_{n\to\infty} \langle \theta_n, h_j \rangle = -\infty,$$

then

$$\mu_j = 0 \quad \textit{for all } \mu \in V_{x,y}.$$

Conversely, there exists a constant $\tilde{C}_6 = \tilde{C}_6(B) > 0$ *such that if* $|y| \leq B$ *and* $\mu_j > 0$ *for some* $\mu \in V_{x,y}$*, then*

$$\liminf_{n\to\infty} \langle \theta_n, h_j \rangle > -\tilde{C}_6.$$

Proof

1. Let $|y| \leq B$, C_2 and $\mu \in V_{x,y}$ be according to Lemma 7.9. Define the functions from $\mathbb{R}$ into itself

$$f_j(z) := \mu_j z - \beta_j(x)(\mathrm{e}^z - 1).$$

Note that $f_j(z) = 0$ if $\beta_j(x) = 0$, and $\mathrm{argmax}_z\, f_j(z) = \log \mu_j / \beta_j(x)$ if $\beta_j(x) > 0$. Let

$$\tilde{C}(x, B) = \sup_{j;\ \beta_j(x)>0} \sup_{|\mu|\leq C_2 B} f_j\left(\log \frac{\mu_j}{\beta_j(x)}\right).$$

[5] The constant -1 can be replaced by any other constant $-C$ $(C > 0)$. Note that C_6 then depends on C with C_6 increasing in C.

If x, y and θ are as in the statement, and $1 \le j \le k$ is such that $\beta_j(x) > 0$ and $\langle \theta, h_j \rangle > 0$, then

$$\sum_{j' \neq j} f_{j'}(\langle \theta, h_{j'} \rangle) = \tilde{\ell}(x, y, \theta) - f_j(\langle \theta, h_j \rangle),$$

hence in view of the assumption,

$$f_j(\langle \theta, h_j \rangle) \ge -1 - (k-1)\tilde{C}(x, B),$$

As $f_j(z) \to -\infty$ as $z \to \infty$, the assertion follows.

2. If $\liminf_{n\to\infty} \langle \theta_n, h_j \rangle = -\infty$ and $\mu \in V_{x,y}$ with $\mu_j > 0$, then 1. implies that $\tilde{\ell}(x, y, \theta_n) \to -\infty$, a contradiction.

   The second assertion follows accordingly.

We now prove

Lemma 7.12

1. *Let $x \in A$ and $y \in \mathcal{C}_x$. Then there exists a maximizing sequence $(\theta_n)_n$ of $\tilde{\ell}(x, y, \cdot)$ and constants $\tilde{s}_j$ (for all $j = 1, \dots, k$ for which there exists a $\mu \in V_{x,y}$ with $\mu_j > 0$) such that*

$$\lim_{n\to\infty} \langle \theta_n, h_j \rangle = \tilde{s}_j \in \mathbb{R}.$$

   The constants $\tilde{s}_j$ are bounded uniformly over bounded sets of $y \in \mathcal{C}_x$.

   In particular, there exists a maximizing sequence $(\theta_n)_n$ such that for all $j = 1, \dots, k$ with $\beta_j(x) > 0$,[6]

$$\lim_{n\to\infty} \exp(\langle \theta_n, h_j \rangle) = s_j \in \mathbb{R}.$$

2. *Let $x \in A$ and $y \notin \mathcal{C}_x$. Then $\underline{L}(x, y) = \infty$.*

Proof

1. By Lemma 7.11,

$$-\tilde{C}_6 = -\tilde{C}_6(|y|) < \langle \theta_n, h_j \rangle \le C_6 = C_6(|y|)$$

   for all n and for all j with $\mu_j > 0$ for some $\mu \in V_{x,y}$. The first assertion follows by taking appropriate subsequences.

[6] Note that here, we also include those j with $\beta_j(x) > 0$ and $\mu_j = 0$ for all $\mu \in V_{x,y}$.

For the second assertion, we have to consider those j with $\mu_j = 0$ for all $\mu \in V_{x,y}$ although $\beta_j(x) > 0$. If $\liminf_{n\to\infty}\langle\theta_n, h_j\rangle = -\infty$, we take further subsequences and obtain (with a slight abuse of notation)

$$\lim_{n\to\infty} \exp(\langle\theta_n, h_j\rangle) = 0.$$

2. Let $y \notin \mathcal{C}_x$ and v be the projection of y on $\mathcal{C}_x$. Hence, $0 = \langle y-v, v\rangle \geq \langle y-v, \tilde{v}\rangle$ for all $\tilde{v} \in \mathcal{C}_x$. For $z = y - v$ ($\neq 0$ as $y \notin \mathcal{C}_x$), we have $\langle z, y\rangle = \langle z, z\rangle + \langle z, v\rangle > 0$ and $\langle z, h_j\rangle \leq 0$ for all j with $\beta_j(x) > 0$. If we set $\theta_n = nz$, we obtain $\ell(x, y, \theta_n) \to \infty$.

7.4.2 Equality of $\underline{L}$ and $\overline{L}$

We can now finally establish

Theorem 7.4 *For all $x \in A$, $y \in \mathbb{R}^d$,*

$$\underline{L}(x, y) = \overline{L}(x, y).$$

Proof In view of Lemma 7.1, it suffices to prove that $\overline{L}(x, y \leq \underline{L}(x, y)$. We first note that we have $\underline{L}(x, y) < \infty$ if and only if $y \in \mathcal{C}_x$ by Lemmas 7.12 2. and 7.13. As the same is true for $\overline{L}(x, y)$ by definition, we can restrict our attention to the case $y \in \mathcal{C}_x$.

We choose a maximizing sequence $(\theta_n)_n$ according to Lemma 7.11 and obtain

$$\lim_n \langle\theta_n, y\rangle = \underline{L}(x, y) + \sum_j \beta_j(x)(s_j - 1); \tag{7.18}$$

here we set $s_j = 0$ if $\beta_j(x) = 0$. We now differentiate with respect to θ and obtain for all n

$$\nabla_\theta \tilde{\ell}(x, y, \theta_n) = y - \sum_{j;\beta_j(x)>0} \beta_j(x) h_j \exp(\langle\theta_n, h_j\rangle);$$

hence (by the fact that $(\theta_n)_n$ is a maximizing sequence and the limit of $\nabla_\theta \tilde{\ell}(x, y, \theta_n)$ exists),

$$\lim_n \nabla_\theta \tilde{\ell}(x, y, \theta_n) = y - \sum_{j;\beta_j(x)>0} \beta_j(x) s_j h_j = 0.$$

We set,

$$\mu_j^* := \beta_j(x) s_j, \tag{7.19}$$

in particular

$$y = \sum_j \mu_j^* h_j \quad \text{and} \quad \mu^* \in V_{x,y}.$$

Therefore,

$$\begin{aligned}
\overline{L}(x, y) &\leq \ell(x, \mu^*) \\
&= \sum_j \beta_j(x) - \mu_j^* + \mu_j^* \log\left(\frac{\mu_j^*}{\beta_j(x)}\right) \\
&= \sum_j \beta_j(x)(1 - s_j) + \mu_j^* \log s_j \\
&= \underline{L}(x, y),
\end{aligned}$$

where we have used (7.18) and (7.19) for the last identity. The assertion follows.

From now on, we shall write $L(x, y)$ for the quantity $\underline{L}(x, y) = \overline{L}(x, y)$. We now prove the strict convexity of $L(x, \cdot)$.

Corollary 7.3 *For all $x \in A$, $L(x, \cdot) : \mathcal{C}_x \to \mathbb{R}_+$ is strictly convex.*

Proof For strict convexity, we exclude the case that $\beta_j(x) = 0$ for all j (as then $L(x, y) = \infty$ for all $y \neq 0$ and the assertion is trivial).

Convexity was proven in Lemma 7.7. Assume now that for $y, \tilde{y} \in \mathcal{C}_x$ and $\lambda \in (0, 1)$,

$$L(x, \lambda y + (1 - \lambda)\tilde{y}) = \lambda L(x, y) + (1 - \lambda)L(x, \tilde{y}).$$

In other words,

$$\begin{aligned}
&\sup_\theta \left\{\langle\theta, \lambda y + (1 - \lambda)\tilde{y}\rangle - \sum_j \beta_j(x)(e^{\langle\theta, h_j\rangle} - 1)\right\} \\
&= \sup_\theta \Big\{\lambda\Big[\langle\theta, y\rangle - \sum_j \beta_j(x)(e^{\langle\theta, h_j\rangle} - 1)\Big] \\
&\quad + (1 - \lambda)\Big[\langle\theta, \tilde{y}\rangle - \sum_j \beta_j(x)(e^{\langle\theta, h_j\rangle} - 1)\Big]\Big\} \\
&= \lambda \sup_\theta \left\{\langle\theta, y\rangle - \sum_j \beta_j(x)(e^{\langle\theta, h_j\rangle} - 1)\right\} \\
&\quad + (1 - \lambda) \sup_\theta \left\{\langle\theta, \tilde{y}\rangle - \sum_j \beta_j(x)(e^{\langle\theta, h_j\rangle} - 1)\right\}.
\end{aligned}$$

Hence, if $(\theta_n)_n$ is a maximizing sequence for $\tilde{\ell}(x, \lambda y + (1-\lambda)\tilde{y}, \cdot)$, it is also a maximizing sequence for $\tilde{\ell}(x, y, \cdot)$ and $\tilde{\ell}(x, \tilde{y}, \cdot)$. As in the proof of Theorem 7.4, this implies

$$\lim_{n\to\infty} \nabla_\theta \tilde{\ell}(x, y, \theta_n) = y - \lim_{n\to\infty} \sum_j \beta_j(x)(\mathrm{e}^{\langle\theta_n, h_j\rangle} - 1) = 0,$$

$$\lim_{n\to\infty} \nabla_\theta \tilde{\ell}(x, \tilde{y}, \theta_n) = \tilde{y} - \lim_{n\to\infty} \sum_j \beta_j(x)(\mathrm{e}^{\langle\theta_n, h_j\rangle} - 1) = 0.$$

Hence $y = \tilde{y}$ as required.

7.4.3 Further Properties of the Legendre Fenchel Transform

In this subsection, we assume that the $\log \beta_j$'s are bounded. In this case $\mathcal{C}_x = \mathcal{C} = \mathbb{R}^d$ for all x.

We have

Lemma 7.13 *Assume that* $\log \beta_j$ *(* $j = 1, \ldots, k$ *) is bounded. Then*

1. *for all* $B > 0$ *exists a constant* $C_7 = C_7(B) > 0$ *such that for all* $x \in A$, $y \in \mathcal{C}$ *with* $|y| \le B$,

$$L(x, y) \le C_7;$$

2. *for all* $x \in A$, $L(x, \cdot) : \mathcal{C} \to \mathbb{R}_+$ *is continuous.*

Proof

1. Let $x \in A$, $y \in \mathcal{C}$. By Lemma 7.9 and Theorem 7.4 above,[7] we obtain

$$L(x, y) \le \sum_{j, \beta_j(x)>0} \beta_j(x) - \mu_j + \mu_j \log \mu_j - \mu_j \log \beta_j(x)$$
$$\le k(\overline{\beta} + C|y|\log C + C|y|\log|y| + C|y||\log \underline{\beta}|).$$

The assertion follows.
2. The assertion follows directly from 1. and Lemma 7.7 1.

[7]Note that this result is not used for the proof of Theorem 7.4.

We have moreover

Lemma 7.14 *Assume that* $\log \beta_j$ *(*$j = 1, \dots, k$*) is bounded. For all* $\rho > 0$*,* $\epsilon > 0$*,* $C_8 > 0$*, there exists a constant* $B_3 = B(C_8, \epsilon)$ *such that for all* $x \in A$*,* $y \in \mathcal{C}$ *with* $|y| \leq C_8$*,*

$$\sup_{|\theta| \leq B} \tilde{\ell}(x, y, \theta) \geq \sup_{\theta \in \mathbb{R}^d} \tilde{\ell}(x, y, \theta) - \epsilon = L(x, y) - \epsilon.$$

Proof We first fix $x \in A$ and define the compact set

$$\tilde{\mathcal{C}} := \{y \in \mathcal{C} \,|\, |y| \leq C_8\}.$$

We fix $\delta > 0$ and define for $y \in \tilde{\mathcal{C}}$,

$$z(y, \delta) := y + \sum_j \delta h_j,$$

$$N^{y,\delta} := \Big\{y + \sum_j \alpha_j h_j | \alpha_j \in (-\delta, \delta)\Big\}.$$

For all $y \in \tilde{\mathcal{C}}$, $N^{y,\delta}$ is relatively open (with respect to $\mathcal{C}$) and $y \in N^{y,\delta}$. Hence there exists a finite cover $N_1, \dots, N_n$ of $\tilde{\mathcal{C}}$, where $N_i := N^{y_i,\delta}$ for appropriate $y_i \in \tilde{\mathcal{C}}$; we define $z_i := z(y_i, \delta)$.

From the continuity of $L(x, \cdot) : \mathcal{C} \to \mathbb{R}_+$ (cf. Lemma 7.13 2.) and the compactness of $\tilde{\mathcal{C}}$, we deduce that for δ small enough and for all $y \in \tilde{\mathcal{C}}$, $v \in N^{y,\delta}$,

$$|L(x, v) - L(x, z(y, \delta))| < \frac{\epsilon}{4}. \tag{7.20}$$

We let θ_i be almost optimal for z_i in the sense that

$$\tilde{\ell}(x, z_i, \theta_i) \geq L(x, z_i) - \frac{\epsilon}{4}. \tag{7.21}$$

We now set $B^x := \max_i |\theta_i|$ and let $y \in \tilde{\mathcal{C}}$, say $y \in N_i$. Then, making use successively of (7.20) and (7.21), we obtain

$$\begin{aligned} L(x, y) &\leq L(x, z_i) + \frac{\epsilon}{4} \\ &\leq \tilde{\ell}(x, z_i, \theta_i) + \frac{\epsilon}{2} \\ &= \tilde{\ell}(x, y, \theta_i) + \frac{\epsilon}{2} + \langle \theta_i, z_i - y \rangle. \end{aligned} \tag{7.22}$$

We have $z_i - y = \sum_j \mu_j h_j$ for appropriate $\mu_j = \alpha_j + \delta \in (0, 2\delta)$ and by Lemma 7.11 (cf. also Inequality (7.20)), $\langle \theta_i, h_j \rangle \le C_6$. Hence

$$\langle z_i - y, \theta_i \rangle = \sum_j \mu_j \langle h_j, \theta_i \rangle \le 2kC_6\delta \le \frac{\epsilon}{4}, \tag{7.23}$$

provided we choose δ such that $8kC_6\delta \le \epsilon$. Therefore by Inequalities (7.22) and (7.23) for all $y \in \tilde{\mathcal{C}}$,

$$L(x, y) \le \tilde{\ell}(x, y, \theta_i) + \epsilon \quad \text{(recall that } |\theta_i| \le B^x\text{)}. \tag{7.24}$$

Let now for all $x \in A$, B^x be the bound obtained above belonging to $\frac{\epsilon}{4}$.[8] Let furthermore $x, \tilde{x} \in A$ with $|\beta(x) - \beta(\tilde{x})| < \delta$ for some $\delta > 0$, $y \in \tilde{\mathcal{C}}$, $|y| \le C_8$ and $\theta \in \mathbb{R}^d$ such that $\tilde{\ell}(x, y, \theta) \ge -1$ (which implies $\langle \theta, h_j \rangle \le C_6$ by Lemma 7.11). This implies

$$|\tilde{\ell}(x, y, \theta) - \tilde{\ell}(\tilde{x}, y, \theta)| \le \sum_j |\beta_j(x) - \beta_j(\tilde{x})| \, \mathrm{e}^{C_6} < \frac{\epsilon}{4} \tag{7.25}$$

for δ small enough (and independent of $x, \tilde{x}, y, \theta$). Let now be $\tilde{\theta}$ be almost optimal for $\tilde{x}, y$. Using twice (7.25) and once (7.24), we obtain

$$\begin{aligned} L(\tilde{x}, y) &\le \tilde{\ell}(\tilde{x}, y, \tilde{\theta}) + \frac{\epsilon}{4} \\ &\le \tilde{\ell}(x, y, \tilde{\theta}) + \frac{\epsilon}{2} \\ &\le \sup_{|\theta| \le B^x} \tilde{\ell}(x, y, \theta) + \frac{3\epsilon}{4} \\ &\le \sup_{|\theta| \le B^x} \tilde{\ell}(\tilde{x}, y, \theta) + \epsilon \end{aligned}$$

We can cover the compact interval $[\underline{\beta}, \bar{\beta}]$ by finitely many $\tilde{\delta}$-neighborhoods of β^i. The assertion follows by taking the maximum of the corresponding B^i (cf. Footnote 8).

Lemma 7.15 *Assume that* $\log \beta_j$ *($j = 1, \ldots, k$) is bounded. There exist constants* B_4 *and* C_9 *such that for all,* $x \in A$ *and* $y \in \mathcal{C}$*,*

$$L(x, y) \le \begin{cases} C_9 & \text{if } |y| \le B_4 \\ C_9 |y| \log |y| & \text{if } |y| > B_4. \end{cases}$$

[8] Note that B^x depends on x only through $\beta(x)$.

Proof From the formula for $\overline{L}(x, y)$ and Lemma 7.9, we have

$$L(x, y) \le \sum_j \bar{\beta} + C|y| \log |y| + C|y|| \log \underline{\beta}|$$
$$\le k \cdot \big(\bar{\beta} + C|y| \log |y| + C|y|| \log \underline{\beta}|\big).$$

We also obtain the continuity of L in x

Lemma 7.16 *Assume that* $\log \beta_j$ *(*$j = 1, \ldots, k$*) is bounded and continuous. For all* $y \in \mathcal{C}$*,*

$$L(\cdot, y) : A \to \mathbb{R}_+$$

is continuous. The continuity is uniform over bounded y*.*

Proof We let $y \in \mathcal{C}$ with $|y| \le B$, $0 < \epsilon < 1$. and $x, \tilde{x} \in A$. Let θ such that

$$L(x, y) \le \tilde{\ell}(x, y, \theta) + \frac{\epsilon}{2}.$$

We have by the continuity of the β_j and Lemma 7.11,

$$|\tilde{\ell}(x, y, \theta) - \tilde{\ell}(\tilde{x}, y, \theta)| \le \sum_j |\beta_j(x) - \beta_j(\tilde{x})| e^{C_6} < \frac{\epsilon}{2}$$

if $|x - \tilde{x}| < \delta$ for appropriate $\delta > 0$ (independent of $x, \tilde{x} \in A$ and y with $|y| \le B$). Thus,

$$L(x, y) \le \tilde{\ell}(\tilde{x}, y, \theta) + \epsilon \le L(\tilde{x}, y) + \epsilon.$$

Reversing the roles of x and $\tilde{x}$ proves the assertion.

Combining Lemmas 7.16 and 7.13, we deduce the

Corollary 7.4 *Assume that* $\log \beta_j$ *(*$j = 1, \ldots, k$*) is bounded and continuous. Then* $L : A \times \mathbb{R}^d \to \mathbb{R}_+$ *is continuous.*

7.4.4 The Rate Function

Recall that for $\phi : [0, T] \to A$, we let

$$I_T(\phi) := \begin{cases} \int_0^T L(\phi(t), \phi'(t))dt & \text{if } \phi \text{ is absolutely continuous} \\ \infty & \text{otherwise.} \end{cases}$$

For $x \in A$ and $\phi : [0, T] \to A$, let

$$I_{T,x}(\phi) := \begin{cases} I_T(\phi) & \text{if } \phi(0) = x \\ \infty & \text{otherwise.} \end{cases}$$

We first have the following statement, which follows readily from point 2 in Lemma 7.7.

Lemma 7.17 *Assume that $x \in A$. If ϕ solves the ODE (7.2), then $I_{T,x}(\phi) = 0$. Conversely, if the ODE (7.2) admits a unique solution Y^x and $I_{T,x}(\phi) = 0$, then $\phi(t) = Y^x(t)$ for all $t \in [0, T]$.*

In the next statement, B_1 refers to the constant appearing in Lemma 7.8.

Lemma 7.18 *Assume that β_j ($j = 1, \dots, k$) is bounded.*

1. *Let $K, \epsilon > 0$. There exits $\delta > 0$ such that for all ϕ with $I_{T,x}(\phi) \le K$ and for all finite collections of non-overlapping subintervals of $[0, T]$, $[s_1, t_1], \dots, [s_J, t_J]$, with $\sum_i (t_i - s_i) = \delta$,*

$$\sum_i |\phi(t_i) - \phi(s_i)| < \epsilon.$$

2. *Let $K > 0$. Then, for all constants $B \ge B_1$ and for all ϕ with $I_{T,x}(\phi) \le K$,*

$$\int_0^T \mathbb{1}_{\{|\phi'(t)| \ge B\}} dt < \frac{K}{C_1 B \log B}.$$

Proof

1. Note first that

$$f(\alpha) := \inf_{x,y} \left\{ \frac{L(x, y)}{|y|} \,\middle|\, |y| \ge \alpha \right\} \to \infty, \tag{7.26}$$

as $\alpha \to \infty$ by Lemma 7.8. For $g(t) := \mathbb{1}_{\cup_j [s_j, t_j]}$ and $\alpha = 1/\sqrt{\delta}$, we obtain by the fact that ϕ is absolutely continuous

$$\begin{aligned} \sum_j |\phi(t_j) - \phi(s_j)| &\le \int_0^T |\phi'(t)| g(t) dt \\ &\le \int_0^T \alpha \mathbb{1}_{\{|\phi'(t)| \le \alpha\}} g(t) dt \\ &\quad + \int_0^T \mathbb{1}_{\{|\phi'(t)| > \alpha\}} \frac{L(\phi(t), \phi'(t))}{L(\phi(t), \phi'(t))/|\phi'(t)|} g(t) dt \end{aligned}$$

$$\leq \alpha\delta + \frac{I_{T,x}(\phi)}{f(\alpha)}$$
$$\leq \sqrt{\delta} + \frac{K}{f(1/\sqrt{\delta})} \to 0,$$

as $\delta \to 0$.

2. By Lemma 7.8, $f(B) \geq C_1 \log B$ for $B \geq B_1$, where f is again defined as in (7.26).

$$\begin{aligned}\int_0^T \mathbb{1}_{\{|\phi'(t)|\geq B\}}dt &\leq \frac{1}{B}\int_0^T |\phi'(t)|\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt \\ &= \frac{1}{B}\int_0^T \frac{L(\phi(t),\phi'(t))|\phi'(t)|}{L(\phi(t),\phi'(t))}\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt \\ &\leq \frac{K}{Bf(B)} \leq \frac{K}{C_1 B\log B}.\end{aligned}$$

Note that Lemma 7.18 1. says that the collection of functions ϕ satisfying $I_T(\phi) \leq K$ are uniformly absolutely continuous.

Theorem 7.5 *Assume that* $\log \beta_j$ *(*$j = 1, \ldots, k$*) is bounded and continuous. Let* $\phi \in D([0,T]; A)$ *with* $I_{T,x}(\phi) < \infty$. *For all* $\epsilon > 0$, *there exists a* $\delta > 0$ *such that for*

$$\tilde{\phi} : [0,T] \to A \quad \textit{with} \quad \sup_{0\leq t\leq T} |\tilde{\phi}(t) - \phi(t)| < \delta,$$

$$\left| \int_0^T \big(L(\tilde{\phi}(t),\phi'(t)) - L(\phi(t),\phi'(t))\big)dt \right| < \epsilon.$$

Proof We choose $B \geq B_1 \vee B_4$ large enough such that for $x \in A$, $y \in \mathcal{C}_x = \mathcal{C}$ (independent of x) with $|y| \geq B$ (cf. Lemmas 7.8 and 7.15),

$$C_1|y|\log|y| \leq L(x,y) \leq C_9|y|\log|y|.$$

As $I_{T,x}(\phi) < \infty$, the set $\{t|\phi'(t) \notin \mathcal{C}\}$ is a Lebesgue null-set and we assume w.l.o.g. that for all t, $\phi'(t) \in \mathcal{C}$. We hence obtain that

$$\begin{aligned}\int_0^T L(\tilde{\phi}(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt &\leq \int_0^T C_9|\phi'(t)|\log|\phi'(t)|\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt \\ &\leq \frac{C_9}{C_1}\int_0^T L(\phi(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt.\end{aligned}$$

From this and Lemma 7.18, we can choose B large enough such that

$$\sup\left(\int_0^T L(\tilde{\phi}(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt, \int_0^T L(\phi(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt\right) < \frac{\epsilon}{4}.$$

By Lemma 7.16, there exists an $\delta > 0$ such that for all $x, \tilde{x} \in A$ with $|x-\tilde{x}| < \delta$ and $y \in \mathcal{C}$ with $|y| \leq B$,

$$|L(x,y) - L(\tilde{x},y)| < \frac{\epsilon}{2T}.$$

We obtain for $\sup_{0\leq t\leq T} |\tilde{\phi}(t) - \phi(t)| < \delta$,

$$\begin{aligned}
&\Big|\int_0^T \big(L(\tilde{\phi}(t),\phi'(t)) - L(\phi(t),\phi'(t))\big)dt\Big| \\
&\quad \leq \Big|\int_0^T L(\tilde{\phi}(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt\Big| + \Big|\int_0^T L(\phi(t),\phi'(t))\mathbb{1}_{\{|\phi'(t)|\geq B\}}dt\Big| \\
&\qquad + \int_0^T \big|L(\tilde{\phi}(t),\phi'(t)) - L(\phi(t),\phi'(t))\big|\mathbb{1}_{\{|\phi'(t)|<B\}}dt \\
&\quad < \epsilon.
\end{aligned}$$

7.4.5 I is a Good Rate Function

We first have

Lemma 7.19 *For $\delta > 0$, $x \in A$ and $y \in \mathbb{R}^d$, we define*

$$L_\delta(x,y) := \sup_{\theta\in\mathbb{R}^d} \tilde{\ell}_\delta(x,y,\theta),$$

where

$$\tilde{\ell}_\delta(x,y,\theta) := \langle\theta, y\rangle - \sup_{z\in A;|z-x|\leq\delta} \sum_j \beta_j(z)\big(e^{\langle\theta,h_j\rangle} - 1\big).$$

Since the β_j are bounded and continuous, then

$$L_\delta(x,y) \uparrow L_0(x,y) = L(x,y)$$

and $L_\delta(x,y)$ is lower semicontinuous in (δ,x,y).

Proof It is easy to see that $\tilde{\ell}_\delta(x, y, \theta)$ is continuous is (x, y, δ), hence the first assertion follows. The second assertion follows from the fact that the supremum of a family of lower semicontinuous functions is lower semicontinuous.

We next establish (recall the metric d_D introduced in Sect. 7.2.2)

Lemma 7.20 *Let the β_j be bounded and continuous. Then, I_T is lower semicontinuous with respect to the metric d_D on $D([0, T]; A)$.*

Proof As $I_T(\phi) = \infty$ if ϕ is not absolutely continuous, we can restrict our attention to sequences of absolutely continuous functions. As the Skorohod topology relativized to $C([0, T]; A)$ coincides with the uniform topology (see, e.g., [2, Section 12, p. 124]), we can consider a sequence of functions $\phi_n \in C([0, T]; A)$ converging to a function ϕ under the uniform topology. We can furthermore assume that $I_T(\phi_n) \leq K$ for some K and all $n \geq 1$. By Lemma 7.18, the functions ϕ_n are hence uniformly absolutely continuous and therefore the limit ϕ is absolutely continuous.

Therefore, for any given $\delta > 0$, there exists a $\Delta > 0$ such that

$$|s - t| \leq \Delta \Rightarrow |\phi_n(s) - \phi_n(t)| \leq \delta \text{ for all } n.$$

We choose Δ smaller if necessary such that $T/\Delta =: J \in \mathbb{N}$ and divide $[0, T]$ into subintervals $[t_j, t_{j+1}]$, $j = 1, \ldots, J$ of length $\leq \Delta$. We note that for $|x' - x| \leq \delta$, we have $L_\delta(x', y) \leq L(x, y)$. Furthermore, we observe that $L_\delta(x, \cdot)$ is convex as a supremum of linear functions and hence by Jensen's Inequality

$$\begin{aligned}\int_0^T L(\phi_n(t), \phi_n'(t))dt &\geq \sum_{j=1}^{J} \int_{t_j}^{t_{j+1}} L_\delta(\phi_n(t_j), \phi_n'(t))dt \\ &\geq \sum_{j=1}^{J} \Delta L_\delta\Big(\phi_n(t_j), \frac{\phi_n(t_{j+1}) - \phi_n(t_j)}{\Delta}\Big). \qquad (7.27)\end{aligned}$$

We now further divide the interval $[0, T]$ into subintervals of length $\Delta_k := \Delta 2^{-k}$, $k \in \mathbb{N}$, $[t_j^k, t_{j+1}^k]$, $j = 1, \ldots, J_k := 2^k J$ and define the functions

$$\underline{\phi}^k(t) := \phi(t_j^k) \quad \text{if } t \in [t_j^k, t_{j+1}^k], \quad \overline{\phi}^k(t) := \underline{\phi}^k(t + \Delta_k).$$

Note that there exits a sequence $\delta_k \downarrow 0$ such that

$$|s - t| < \Delta_k \Rightarrow |\phi_n(s) - \phi_n(t)| < \delta_k \text{ for all } n.$$

Hence by Inequality (7.27) and Lemma 7.19 for all $k \in \mathbb{N}$,

$$\liminf_{n\to\infty} \int_0^T L(\phi_n(t), \phi_n'(t))dt \geq \sum_{j=1}^{J_k} \Delta_k \liminf_{n\to\infty} L_{\delta_k}\Big(\phi_n(t_j^k), \frac{\phi_n(t_{j+1}^k) - \phi_n(t_j^k)}{\Delta_k}\Big)$$
$$\geq \int_0^{T-\Delta_k} L_{\delta_k}\Big(\underline{\phi}^k(t), \frac{\overline{\phi}^k(t) - \underline{\phi}^k(t)}{\Delta_k}\Big)dt. \qquad (7.28)$$

As ϕ is absolutely continuous, we have that for almost all $t \in [0, T]$,

$$\frac{\overline{\phi}^k(t) - \underline{\phi}^k(t)}{\Delta_k} \to \phi'(t) \quad \text{as } k \to \infty.$$

We conclude by using Inequality (7.28), Fatou's Lemma and Lemma 7.19 again:

$$\begin{aligned}
&\liminf_{n\to\infty} \int_0^T L(\phi_n(t), \phi_n'(t))dt \\
&\geq \liminf_{k\to\infty} \int_0^{T-\Delta_k} L_{\delta_k}\Big(\underline{\phi}^k(t), \frac{\overline{\phi}^k(t) - \underline{\phi}^k(t)}{\Delta_k}\Big)dt \\
&\geq \int_0^T \liminf_{k\to\infty} \Big(\mathbb{1}_{[0,T-\Delta_k]}(t) L_{\delta_k}\Big(\underline{\phi}^k(t), \frac{\overline{\phi}^k(t) - \underline{\phi}^k(t)}{\Delta_k}\Big)\Big)dt \\
&\geq \int_0^T L(\phi(t), \phi'(t))dt
\end{aligned}$$

as required.

We define for $K > 0$, $x \in A$,

$$\Phi(K) = \big\{\phi \in D([0, T]; A) | I_T(\phi) \leq K\big\},$$
$$\Phi_x(K) = \big\{\phi \in D([0, T]; A) | I_{T,x}(\phi) \leq K\big\}.$$

We have moreover

Proposition 7.1 *Assume that β_j $(j = 1, \ldots, k)$ are bounded and continuous. Let furthermore $K > 0$ and $\tilde{A} \subset A$ be compact. Then, the set*

$$\bigcup_{x\in\tilde{A}} \Phi_x(K)$$

is compact in $C([0, T]; A)$.

Proof By Lemma 7.18, the set $\bigcup_{x\in\tilde{A}} \Phi_x(K)$ is equicontinuous. As $\tilde{A}$ is compact, the theorem of Arzelà-Ascoli hence implies that $\bigcup_{x\in\tilde{A}} \Phi_x(K)$ has compact closure. Now, the semicontinuity of I (cf. Lemma 7.20) implies that $\bigcup_{x\in\tilde{A}} \Phi_x(K)$ is closed, which ends the proof.

We define for $S \subset D([0,T];A)$,

$$I_x(S) := \inf_{\phi\in S} I_{T,x}(\phi).$$

Lemma 7.21 *Assume that β_j $(j = 1,\dots,k)$ are bounded and continuous. Let $F \subset C([0,T];A)$ be closed. Then $I_x(F)$ is lower semicontinuous in x.*

Proof We let $x_n \to x$ with $\liminf_{n\to\infty} I_{x_n}(F) =: K < \infty$. For simplicity, we assume that $I_{x_n}(F) \le K + \epsilon$ for some fixed $\epsilon > 0$ and for all n. By Proposition 7.1, we have that for all $\epsilon, \delta > 0$,

$$F \cap \Phi_{x_n}(K+\epsilon) \quad \text{and} \quad F \cap \bigcup_{|x-y|\le\delta} \Phi_y(K+\epsilon)$$

are compact. By the semicontinuity of $I_T(\cdot)$ (cf. Lemma 7.20) and the fact that a l.s.c. function attains its minimum on a compact set, there exist $\phi_n \in F$ such that $I_{x_n}(F) = I_{T,x_n}(\phi_n)$ (for n large enough). As the ϕ_n are in a compact set, there exists a convergent subsequence with limit ϕ, in particular $\phi(0) = x$. As F is closed, we have $\phi \in F$. We use Lemma 7.20 again and obtain

$$I_x(F) \le I_T(\phi) \le \liminf_{n\to\infty} I_T(\phi_n) = \liminf_{n\to\infty} I_{x_n}(F) = K$$

as required.

The following result is a direct consequence of Lemma 7.21.

Lemma 7.22 *Assume that β_j $(j = 1,\dots,k)$ is bounded and continuous. For $F \subset D([0,T];A)$ closed and $x \in A$, we have*

$$\lim_{\epsilon\downarrow 0} \inf_{y\in A,\, |x-y|<\epsilon} I_y(F) = I_x(F)$$

We can now establish the main result of this subsection.

Proposition 7.2 *Let the β_j be bounded and continuous. For all x, I_x is a good rate function on $C([0,T];A) \cap \{\phi \mid \phi(0) = x\}$.*

Proof It is clear that I_T is non-negative as L is non-negative. Furthermore, it is lower semicontinuous by Lemma 7.20. By Proposition 7.1 its level sets are compact.

We have moreover

Corollary 7.5 *Let the β_j be bounded and continuous. For all $x \in A$, $I_{T,x}$ is a good rate function on $D([0,T]; A) \cap \{\phi \mid \phi(0) = x\}$ under both metrics d_C and d_D.*

Proof Since $I_{t,x}$ is finite only for absolutely continuous functions, it suffices to consider sequences in $C([0,T]; A) \cap \{\phi|\phi(0) = x\}$. Limits of such sequences (under either metric) are continuous and convergence is equivalent for both metrics (see, e.g., [2]). Lower semicontinuity follows. Compactness of the level sets follows by Proposition 7.2 and the fact that the identity maps from $(C([0,T]; A), d_C)$ into $(D([0,T]; A), d_C)$ and $(D([0,T]; A), d_D)$ are continuous.

7.4.6 A Property of Non-exponential Equivalence

It is worth wondering whether or not $\{Z_t^{N,x_N},\ 0 \le t \le T\}$ and $\{Z_t^{N,x},\ 0 \le t \le T\}$ are exponentially equivalent, whenever $x_N \to x$ as $N \to \infty$. Indeed, [3] prove that property for diffusions with small noise and Lipschitz coefficients, and use it to establish certain results, of which we shall prove analogs below, but without that exponential equivalence, which fails to hold in our Poissonian case.

Let $x, y \in A$, and consider the processes

$$Z_t^{N,x} = x + \sum_{j=1}^{k} \frac{h_j}{N} P_j\left(N \int_0^t \beta_j(Z_s^{N,x})ds\right),$$

$$Z_t^{N,y} = y + \sum_{j=1}^{k} \frac{h_j}{N} P_j\left(N \int_0^t \beta_j(Z_s^{N,y})ds\right).$$

For any $\delta > 0$, as $|x - y| \to 0$, we ask what is the limit, as $|x - y| \to 0$, of

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\left(\sup_{0\le t\le T} |Z_t^{N,x} - Z_t^{N,y}| > \delta\right) ?$$

If that limit is $-\infty$, then we would have the above exponential equivalence. We now show on a particularly simple example that this is not the case. It is easy to infer that it in fact fails in the above generality, assuming that the β_j's are Lipschitz continuous and bounded. We consider the case $d = 1$, $A = \mathbb{R}_+$, $k = 1$, $\beta(x) = x$, $h = 1$. We could truncate $\beta(x)$ to make it bounded, in order to comply with our standing assumptions. The modifications below would be minor, but we prefer to keep the simplest possible notations. Assume $0 < x < y$ and consider the two processes

$$Z_t^{N,x} = x + \frac{1}{N} P\left(N \int_0^t Z_s^{N,x} ds\right),$$
$$Z_t^{N,y} = y + \frac{1}{N} P\left(N \int_0^t Z_s^{N,y} ds\right).$$

It is plain that $0 < Z_t^{N,x} < Z_t^{N,y}$ for all $N \geq 1$ and $t > 0$. Let $\Delta_t^{N,x,y} = Z_t^{N,y} - Z_t^{N,x}$. The law of $\{\Delta_t^{N,x,y},\ 0 \leq t \leq T\}$ is the same as that of the solution of

$$\Delta_t^{N,x,y} = y - x + \frac{1}{N} P\left(N \int_0^t \Delta_s^{N,x,y} ds\right).$$

We deduce from Theorem 7.9 below (which is established in case of a bounded coefficient $\beta(x)$, but it makes no difference here) that

$$\begin{aligned}\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}[\Delta_1^{N,x,y} > 1] &\geq - \inf_{\phi(0)=y-x,\ \phi(1)>1} I_{1,y-x}(\phi) \\ &\geq -I_{1,y-x}(\psi),\end{aligned}$$

with $\psi(t) = y - x + t$, hence

$$\begin{aligned}I_{1,y-x}(\psi) &= \int_0^1 L(y - x + t, 1)dt \\ &= \int_0^1 [y - x + t - 1 - \log(y - x + t)]dt \\ &= y - x + 1/2 - (y - x + 1)\log(y - x + 1) + (y - x)\log(y - x) \\ &\to 1/2,\end{aligned}$$

as $y - x \to 0$. This clearly contradicts the exponential equivalence.

We note that the above process $Z_t^{N,x}$ can be shown to be "close" (in a sense which is made very precise in [9]) to its diffusion approximation

$$X_t^{N,x} = x + \int_0^t X_s^{N,x} ds + \frac{1}{\sqrt{N}} \int_0^t \sqrt{X_s^{N,x}} dB_s,$$

where $\{B_t,\ t \geq 0\}$ is standard Brownian motion. One can study large deviations of this diffusion process from its Law of Large Numbers limit (which is the same as that of $Z_t^{N,x}$). The rate function on the time interval $[0, 1]$ is now

$$I(\phi) = \int_0^1 \frac{(\phi'(t) - \phi(t))^2}{\phi(t)} dt.$$

Let again $\psi(t) = y - x + t$, now with $0 = x < y$. $I(\psi) = \log(1+y) - \log(y) - 3/2 + y \to +\infty$, as $y \to 0$. We see here that the large deviations behaviour of the solution of the Poissonian SDE and of its diffusion approximation differ dramatically, as was already noted by Pakdaman [10] (see also the references in this paper).

7.5 Lower Bound

We first establish the LDP lower bound under the assumption that the rates are bounded away from zero, or in other words the $\log \beta_j$'s are bounded. From this, we will derive later the general result.

7.5.1 LDP Lower Bound If the Rates Are Bounded Away from Zero

We first note that if the β_j are bounded away from zero, then the convex cone $\mathcal{C}_x$ is dependent of x, $\mathcal{C}_x = \mathcal{C}$ for all x. Note that this implies that the "domain" A of the process cannot be bounded.

We require a LDP for linear functions. This follows from the LLN (Theorem 7.3).

Proposition 7.3 *Assume that* $\log \beta_j$ *(*$j = 1, \dots, k$*) is bounded and continuous. For any* $\epsilon > 0$, $\delta > 0$ *there exists an* $\tilde{\epsilon} > 0$ *such that for* $x \in A$, $y \in \mathcal{C}$ *and* $\mu \in V_{x,y} = \tilde{V}_y$,

$$\liminf_{N \to \infty} \frac{1}{N} \log \Big(\inf_{z \in A; |z-x| < \tilde{\epsilon}} \mathbb{P}\Big[\sup_{t \in [0,T]} |Z^{N,z}(t) - \phi^x(t)| < \epsilon \Big] \Big)$$
$$\geq - \int_0^T \ell(\phi^x(t), \mu) dt - \delta,$$

where

$$\phi^x(t) := x + ty = x + t \sum_j \mu_j h_j.$$

Proof We define

$$F^{N,z} := \Big\{ \sup_{t \in [0,T]} |Z^{N,z}(t) - \phi^z(t)| < \frac{\epsilon}{2} \Big\}$$

and let $\tilde{\epsilon} < \epsilon_1 = \epsilon/2$. Let now $\xi_T = \xi_T^{N,z} = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}|_{\mathcal{F}_T}$ be given as in Theorem 7.15 for initial value z and constant rates $\tilde{\beta}_j = \mu_j$. Then, with the notation $\tilde{\mathbb{E}}_{F^{N,z}}[X] :=$

$\tilde{\mathbb{E}}[X|F^{N,z}]$ and (recall that $\xi_T \neq 0$ $\tilde{\mathbb{P}}$-almost surely)

$$X_T^{N,z} := X_T := \log \xi_T^{-1} = \sum_{\tau \le T} \Big[\log \beta_{j(\tau)}(Z^{N,z}(\tau-)) - \log \mu_{j(\tau)} \Big]$$

$$-N \sum_j \int_0^T \big(\beta_j(Z^{N,z}(t)) - \mu_j\big) dt,$$

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N} \log \Big(\inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \mathbb{P}\big[\sup_{t\in[0,T]} |Z^{N,z}(t) - \phi^x(t)| < \epsilon \big] \Big) \\
&\ge \liminf_{N\to\infty} \frac{1}{N} \log \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \mathbb{P}[F^{N,z}] \\
&= \liminf_{N\to\infty} \frac{1}{N} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \log \mathbb{P}[F^{N,z}] \\
&\ge \liminf_{N\to\infty} \frac{1}{N} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \log \tilde{\mathbb{E}}\big[\xi_T^{-1} \mathbb{1}_{F^{N,z}}\big] \\
&= \liminf_{N\to\infty} \frac{1}{N} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \log \big(\tilde{\mathbb{P}}[F^{N,z}] \tilde{\mathbb{E}}_{F^{N,z}}[\exp(X_T)] \big) \\
&\ge \liminf_{N\to\infty} \frac{1}{N} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \log \tilde{\mathbb{P}}[F^{N,z}] \\
&\quad + \liminf_{N\to\infty} \frac{1}{N} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \log \tilde{\mathbb{E}}_{F^{N,z}}[\exp(X_T)] \\
&\ge \liminf_{N\to\infty} \inf_{z\in A,\, |x-z|<\tilde{\epsilon}} \tilde{\mathbb{E}}_{F^{N,z}}\Big[\frac{X_T}{N} \Big],
\end{aligned} \tag{7.29}$$

where we have used Corollary 7.10 for the second inequality, Theorem 7.3 and Jensen's inequality on the last line. Note the independence of the constants $\tilde{C}_1, \tilde{C}_2$ of z in Theorem 7.3 and hence

$$\tilde{\mathbb{P}}[F^{N,z}] \to 1 \quad \text{as } N \to \infty \text{ independently of } z.$$

We have

$$\frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \tilde{\mathbb{E}}\big[\mathbb{1}_{F^{N,z}} T \sum_j \mu_j\big] = T \sum_j \mu_j.$$

By the fact that the β_j's are bounded and continuous and by Theorem 7.3, we have for $j = 1, \dots, k$,

$$\sup_{t\in[0,T]} |\beta_j(Z^{N,z}(t)) - \beta_j(\phi^z(t))| \to 0 \quad \tilde{\mathbb{P}} - \text{a.s.}$$

as $N \to \infty$ uniformly in z. This implies

$$\frac{1}{\tilde{\mathbb{P}}[F^{N,z}]}\tilde{\mathbb{E}}\Big[\mathbb{1}_{F^{N,z}} \int_0^T \sum_j \beta_j(Z^{N,z}(t))dt\Big] \longrightarrow \sum_j \int_0^T \beta_j(\phi^z(t))dt$$

as $N \to \infty$ uniformly in z.

Let us now define the following processes. For $z \in A$, $j = 1, \dots, k$ and $0 \le t_1 < t_2 \le T$ let

$$Y_j^{N,z,t_1,t_2} := \frac{1}{N} \cdot \#\text{jumps of } Z^{N,z} \text{ in direction } h_j \text{ in } [t_1, t_2].$$

Let furthermore $\tau_j \in [0, T]$ denote the jump times of Z^N in direction h_j; we obtain

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j;\mu_j>0} \tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^{N,z}} \sum_{\tau_j} \log \mu_j\Big] \\
&= \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j;\mu_j>0} \log \mu_j \Big\{\tilde{\mathbb{E}}\big[Y_j^{N,z,0,T}\big]\tilde{\mathbb{P}}\big[F^{N,z}\big] + \widetilde{\text{Cov}}(\mathbb{1}_{F^{N,z}}, Y_j^{N,z,0,T})\Big\} \\
&\to T \sum_j \mu_j \log \mu_j,
\end{aligned} \tag{7.30}$$

since, for a given set F,

$$\begin{aligned}
\tilde{\mathbb{E}}[Y_j^{N,z,t_1,t_2}] &= (t_2 - t_1)\mu_j \\
\widetilde{\text{Var}}[Y_j^{N,z,t_1,t_2}] &= (t_2 - t_1)\mu_j \\
|\widetilde{\text{Cov}}(\mathbb{1}_F, Y_j^{N,z,t_1,t_2})| &\le \sqrt{\widetilde{\text{Var}}[\mathbb{1}_F]}\sqrt{\widetilde{\text{Var}}[Y_j^{N,z,t_1,t_2}]} = \sqrt{\tilde{\mathbb{P}}[F] - \tilde{\mathbb{P}}[F]^2}\sqrt{(t_2 - t_1)\mu_j}.
\end{aligned}$$

We now define the set

$$\tilde{F}^{N,z} := \Big\{\sup_{t\in[0,T]} |Z^{N,z}(t) - \phi^z(t)| < \epsilon_N\Big\} \quad \text{for } \epsilon_N := \epsilon \wedge \frac{1}{N^{1/3}};$$

we have (for N large enough)

$$\tilde{\mathbb{P}}[\tilde{F}^{N,z}] \geq 1 - \tilde{C}_1 \exp\big(-N\tilde{C}_2(\epsilon_N)\big) \longrightarrow 1$$

as $N \to \infty$ uniformly in z by Theorem 7.3. We furthermore let $\bar{A} \subset A$ be compact such that for all z, $|z - x| < \tilde{\epsilon}$ and all $t \in [0, T]$, $\phi^z(t) \in \bar{A}$ and $Z^{N,z}(t) \in \bar{A}$ on $F^{N,z}$. As the $\log \beta_j$ are bounded and uniformly continuous, there exit constants $\tilde{\delta}_N > 0$ with $\tilde{\delta}_N \downarrow 0$ such that

$$\tilde{x}, \bar{x} \in \bar{A},\ |\tilde{x} - \bar{x}| < \frac{2}{N^{1/3}} \Rightarrow |\log \beta_j(\tilde{x}) - \log \beta_j(\bar{x})| < \tilde{\delta}_N.$$

We define $\bar{\mu} = \max_j \mu_j$,

$$M = M(N) := \lfloor T N^{1/3} k\bar{h}\bar{\mu} + 1 \rfloor$$

and divide the interval $[0, T]$ into M subintervals $[t_r, t_{r+1}]$ ($r = 0, \ldots, M - 1$, $t_r = t_r(N)$) of length $\Delta = \Delta(N)$, i.e. for $N \geq N_0$ independent of z large enough,

$$\Delta < \frac{1}{N^{1/3} k\bar{\mu}\bar{h}}.$$

For $j, r = 0, \ldots, M - 1$ and $\tau_j, t \in [t_r, t_{r+1}]$, since for $|t - s| < \frac{1}{N^{1/3}k\bar{\mu}\bar{h}}$, $|\phi^z(t) - \phi^z(s)| < \frac{1}{N^{1/3}}$, we have on $\tilde{F}^{N,z}$

$$|Z^{N,z}(\tau_j-) - \phi^z(t)| \leq |Z^{N,z}(\tau_j-) - \phi^z(\tau_j)| + |\phi^z(\tau_j) - \phi^z(t)| \leq \frac{2}{N^{1/3}},$$

and hence

$$\inf_{t\in[t_r,t_{r+1}]} \log \beta_j(\phi^z(t)) - \tilde{\delta}_N \leq \log \beta_j(Z^{N,z}(\tau_j-)) \leq \sup_{t\in[t_r,t_{r+1}]} \log \beta_j(\phi^z(t)) + \tilde{\delta}_N.$$

We compute

$$\frac{1}{\tilde{\mathbb{P}}[F^{N,z}]}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^{N,z}} \sum_{\tau} \log \beta_{j(\tau)}(Z^{N,z}(\tau-))\Big]$$
$$= \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\, \mu_j>0} \sum_{r=0}^{M-1} \tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^{N,z}} \sum_{\tau_j\in[t_r,t_{r+1})} \log \beta_j(Z^{N,z}(\tau_j-))\Big]$$
$$+ \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\, \mu_j>0} \sum_{r=0}^{M-1} \tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^{N,z}\setminus\tilde{F}^{N,z}} \sum_{\tau_j\in[t_r,t_{r+1})} \log \beta_j(Z^{N,z}(\tau_j-))\Big]$$

$$
\begin{aligned}
&\leq \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \Big(\sup_{t\in[t_r,t_{r+1})} \log \beta_j(\phi^z(t)) + \tilde{\delta}_N \Big) \tilde{\mathbb{E}}\Big[\mathbb{1}_{\tilde{F}^{N,z}} Y_j^{N,z,t_r,t_{r+1}} \Big] \\
&\quad + \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \log \bar{\beta} \tilde{\mathbb{E}}\Big[\mathbb{1}_{F^{N,z}\setminus \tilde{F}^{N,z}} Y_j^{N,z,t_r,t_{r+1}} \Big] \\
&\leq \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \sup_{t\in[t_r,t_{r+1})} \log \beta_j(\phi^z(t)) \Big\{ \tilde{\mathbb{E}}\big[\mathbb{1}_{\tilde{F}^{N,z}}\big] \tilde{\mathbb{E}}\big[Y_j^{N,z,t_r,t_{r+1}}\big] \\
&\quad + \widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}}) \Big\} \\
&\quad + \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \log \bar{\beta} \Big\{ \tilde{\mathbb{E}}\big[\mathbb{1}_{F^{N,z}\setminus \tilde{F}^{N,z}}\big] \tilde{\mathbb{E}}\big[Y_j^{N,z,t_r,t_{r+1}}\big] \\
&\quad + \widetilde{\mathrm{Cov}}(\mathbb{1}_{F^{N,z}\setminus \tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}}) \Big\} \\
&\quad + \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \tilde{\delta}_N k \bar{\mu} T \\
&\leq \frac{\tilde{\mathbb{P}}[\tilde{F}^{N,z}]}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \mu_j \sum_{r=0}^{M-1} \Delta \sup_{t\in[t_r,t_{r+1})} \log \beta_j(\phi^z(t)) \\
&\quad + \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \Big\{ \log \bar{\beta} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \Big(|\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}})| \\
&\quad + |\widetilde{\mathrm{Cov}}(\mathbb{1}_{F^{N,z}\setminus \tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}})| \Big) \\
&\quad + \tilde{\delta}_N k \bar{\mu} T + \tilde{\mathbb{P}}[F^{N,z} \setminus \tilde{F}^{N,z}] k \bar{\mu} T \log \bar{\beta} \Big\} \qquad (7.31) \\
&=: \overline{S}^{N,z} + \overline{U}^{N,z},
\end{aligned}
$$

where $\overline{S}^{N,z}$ and $\overline{U}^{N,z}$ are the first respectively the second term in Inequality (7.31). In a similar fashion we obtain

$$
\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^{N,z}]} \tilde{\mathbb{E}}\Big[\frac{1}{N} \mathbb{1}_{F^{N,z}} \sum_{\tau} \log \beta_{j(\tau)}(Z^{N,z}(\tau-)) \Big] \\
&\qquad \geq \frac{\tilde{\mathbb{P}}[\tilde{F}^{N,z}]}{\tilde{\mathbb{P}}[F^{N,z}]} \sum_{j,\,\mu_j>0} \mu_j \sum_{r=0}^{M-1} \Delta \inf_{t\in[t_r,t_{r+1})} \log \beta_j(\phi^z(t))
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{1}{\tilde{\mathbb{P}}[F^{N,z}]}\Big\{ \log \underline{\beta} \sum_{j,\,\mu_j>0} \sum_{r=0}^{M-1} \Big(|\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}})| \\
&+ |\widetilde{\mathrm{Cov}}(\mathbb{1}_{F^{N,z}\setminus\tilde{F}^{N,z}}, Y_j^{N,z,t_r,t_{r+1}})|\Big) \\
&- \tilde{\delta}_N k\bar{\mu}T + \tilde{\mathbb{P}}[F^{N,z} \setminus \tilde{F}^{N,z}]k\bar{\mu}T \log \underline{\beta}\Big\} \\
=:\ & \underline{S}^{N,z} + \underline{U}^{N,z};
\end{aligned}
$$

we first note that $\overline{U}^{N,z}, \underline{U}^{N,z} \to 0$ as $N \to \infty$ uniformly in z, since $\tilde{\delta}_N \to 0$ and $\tilde{\mathbb{P}}[F^{N,z} \setminus \tilde{F}^{N,z}] \to 0$ as $N \to 0$ uniformly in z. Furthermore, as (up to a factor which converges to 1 uniformly in z) $\overline{S}^{N,z}$ and $\underline{S}^{N,z}$ are upper respectively lower Riemann sums, we obtain

$$
\overline{S}^{N,z}, \underline{S}^{N,z} \to \sum_j \mu_j \int_0^T \log \beta_j(\phi^z(t))dt \tag{7.32}
$$

as $N \to \infty$; since

$$
|\overline{S}^{N,z} - \underline{S}^{N,z}| \le 2\frac{\tilde{\mathbb{P}}[\tilde{F}^{N,z}]}{\tilde{\mathbb{P}}[F^{N,z}]}k\bar{\mu}\tilde{\delta}_N \to 0
$$

uniformly in z, the convergence in (7.32) is likewise uniform in z.

The uniform convergence implies (cf. (7.29)–(7.30) and the preceding discussion) that

$$
\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N} \log\Big(\inf_{z\in A,\,|x-z|<\tilde{\epsilon}} \mathbb{P}\big[\sup_{t\in[0,T]} |Z^{N,z}(t) - \phi^x(t)| < \epsilon\big]\Big) \\
&\ge - \sup_{z\in A,\,|x-z|<\tilde{\epsilon}} \int_0^T \ell(\phi^z(t), \mu)dt
\end{aligned}
$$

In combination with the uniform continuity of $\ell(\cdot, \mu)$ (recall the boundedness of the $\log \beta_j$) this proves the assertion.

The main building block for the lower bound is the following result.

Theorem 7.6 *Assume that* $\log \beta_j$ *(*$j = 1, \ldots, k$*) is bounded and continuous. Let* $\phi \in D([0, T]; A)$ *with* $\phi(0) = x$ *and* $\epsilon > 0$. *Then,*

$$
\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x}(t) - \phi(t)| < \epsilon\Big] \ge -I_{T,x}(\phi).
$$

The convergence is uniform in $x \in A$.

Proof We can w.l.o.g. assume that $I_{T,x}(\phi) < \infty$ (and hence ϕ is absolutely continuous) as else the assertion is trivial. We approximate the function ϕ by a continuous piecewise linear function and then apply the LDP for linear functions to each of these linear functions. To this end, we let $\delta > 0$ and divide the interval $[0, T]$ into J subintervals of length $\Delta = T/J$, $[t_{r-1}, t_r]$ $(r = 1, \dots, J)$ such that the resulting piecewise linear approximation

$$\tilde{\phi}(t) = \phi(t_{r-1}) + \frac{t - t_{r-1}}{\Delta}(\phi(t_r) - \phi(t_{r-1}))$$

satisfies

$$\sup_{t\in[0,T]} |\phi(t) - \tilde{\phi}(t)| < \frac{\epsilon}{2}$$

(recall that ϕ is continuous).

We now apply Theorem 7.5 twice (in Inequalities (7.33) and (7.35)) and choose J large enough in order to assure

$$\begin{aligned}
\int_0^T L(\phi(t), \phi'(t))dt &= \sum_{r=1}^{J} \int_{t_{r-1}}^{t_r} L(\phi(t), \phi'(t))dt \\
&\geq \sum_{r=1}^{J} \int_{t_{r-1}}^{t_r} L(\phi(t_{r-1}), \phi'(t))dt - \frac{\delta}{4} &(7.33)\\
&\geq \Delta \sum_{r=1}^{J} L\Big(\phi(t_{r-1}), \frac{\Delta\phi(t_r)}{\Delta}\Big) - \frac{\delta}{4} &(7.34)\\
&\geq \sum_{r=1}^{J} \int_{t_{r-1}}^{t_r} L\big(\tilde{\phi}(t), \tilde{\phi}'(t)\big)dt - \frac{\delta}{2}, &(7.35)
\end{aligned}$$

where

$$\Delta\phi(t_r) := \phi(t_r) - \phi(t_{r-1}).$$

Note that for Inequality (7.34), we have applied Jensen's inequality and the fact that L is convex in its second argument (cf. Corollary 7.3). As $I_{T,x}(\tilde{\phi}) < \infty$, this implies

$$\Delta\phi(t_r) \in \mathcal{C} \quad \text{for all } r.$$

We note that by the continuity of $L(\cdot, y)$, $\mu^*(x, y)$ (the minimizing $\mu \in V_{x,y} = \tilde{V}_y$ for $\ell(x, \cdot)$) is "almost optimal" for all $\tilde{x}$ sufficiently close to x (in the sense that $\ell(\tilde{x}, \mu^*(x, y))$ is close to $L(\tilde{x}, y)$). By dividing each interval $[t_{r-1}, t_r]$ into further subintervals $[s_{j-1}, s_j]$ if necessary, we can hence represent the directions $\Delta\phi(t_k)/\Delta$

by

$$\mu^j \in V_{\tilde{\phi}(t),\Delta\phi(t_r)/\Delta} = \tilde{V}_{\Delta\phi(t_r)/\Delta}$$

in such a way that

$$L\Big(\tilde{\phi}(t), \frac{\Delta\phi(t_r)}{\Delta}\Big) \geq \ell(\tilde{\phi}(t), \mu^j) - \frac{\delta}{4T} \quad \text{for all } t \in [s_{j-1}, s_j].$$

For simplicity of exposition, we assume that this further subdivision of the intervals $[t_{r-1}, t_r]$ is not required and denote the "almost optimal" μ's by μ^r $(r = 1, \dots, J)$. Hence

$$\int_0^T L(\phi(t), \phi'(t))dt \geq \sum_{r=1}^{J} \int_{t_{r-1}}^{t_r} \ell(\tilde{\phi}(t), \mu^r) - \frac{3\delta}{4}. \tag{7.36}$$

Choose now $\tilde{\epsilon} = \tilde{\epsilon}_{J-1}$ according to Proposition 7.3 corresponding to $\epsilon/2$, $\delta/(4J)$, initial value $\tilde{\phi}(t_{J-1})$ and time-horizon Δ. Using the Markov property of Z^N, we compute

$$\mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x}(t) - \phi(t)| < \epsilon\Big] \geq \mathbb{P}\Big[\sup_{t\in[0,t_{J-1}]} |Z^{N,x}(t) - \tilde{\phi}(t)| < \tilde{\epsilon}\Big]$$
$$\cdot \inf_{z\in A; |z-\tilde{\phi}(t_{J-1})|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[t_{J-1},T]} |Z^{N,z}(t) - \tilde{\phi}(t)| < \frac{\epsilon}{2}\Big];$$

here, we denote (by a slight abuse of notation) the process starting at z at time t_{J-1} by $Z^{N,z}$. Proposition 7.3 implies

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x}(t) - \phi(t)| < \epsilon\Big] \\
&\geq \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,t_{J-1}]} |Z^{N,x}(t) - \tilde{\phi}(t)| < \tilde{\epsilon}_{J-1}\Big] \\
&\quad + \liminf_{N\to\infty} \frac{1}{N} \log \Big(\inf_{z\in A; |z-\tilde{\phi}(t_{J-1})|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[t_{J-1},T]} |Z^{N,z}(t) - \tilde{\phi}(t)| < \frac{\epsilon}{2}\Big]\Big) \\
&\geq \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,t_{J-1}]} |Z^{N,x}(t) - \tilde{\phi}(t)| < \tilde{\epsilon}_{J-1}\Big] \\
&\quad - \int_{t_{J-1}}^{T} \ell(\tilde{\phi}(t), \mu^J)dt - \frac{\delta}{4J}.
\end{aligned}$$

Iterating this procedure, we obtain

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big[\sup_{t\in[0,T]}|Z^{N,x}(t)-\phi(t)|<\epsilon\Big]\geq-\sum_{r=1}^{J}\int_{t_{r-1}}^{t_r}\ell(\tilde{\phi}(t),\mu^r)dt-\frac{\delta}{4}$$

and the assertion follows from Inequality (7.36) if we let $\delta\to 0$.

We note that the convergence is uniform in x by the uniformity in Proposition 7.3.

Theorem 7.7 *Assume that* $\log\beta_j$ *(*$j=1,\ldots,k$*) is bounded and continuous. Let* $G\subset D([0,T];A)$ *be open and* $x\in A$*. Then,*

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}[Z^{N,x}\in G]\geq-\inf_{\phi\in G,}I_{T,x}(\phi).$$

The convergence is uniform in $x\in A$*.*

Proof Let $\inf_{\phi\in G}I_{T,x}(\phi)=:I^*<\infty$; hence, for $\delta>0$, there exists a $\phi^\delta\in G$ ($\phi(0)=x$) with $I_{T,x}(\phi^\delta)\leq I^*+\delta$. For small enough $\epsilon=\epsilon(\phi^\delta)>0$, we have

$$\Big\{\phi\in D([0,T];A)|\sup_{t\in[0,T]};|\phi^\delta(t)-\phi(t)|<\epsilon\Big\}\subset G$$

and therefore

$$\mathbb{P}\Big[\sup_{t\in[0,T]}|Z^{N,x}(t)-\phi^\delta(t)|<\epsilon\Big]\leq\mathbb{P}[Z^{N,x}\in G].$$

This implies by Theorem 7.6 that for all $\delta>0$,

$$\begin{aligned}\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}[Z^{N,x}\in G]&\geq\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big[\sup_{t\in[0,T]}|Z^{N,x}(t)-\phi^\delta(t)|<\epsilon\Big]\\&\geq-I_{T,x}(\phi^\delta)\\&\geq-I^*-\delta.\end{aligned}$$

This implies

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}[Z^{N,x}\in G]\geq-I^*$$

as desired.

We obtain the following result.

Corollary 7.6 *Assume that* $\log \beta_j$ *(*$j = 1, \ldots, k$*) is bounded and continuous. Then for all* $\phi \in D([0,T]; A)$ *with* $\phi(0) = x$ *and* $\epsilon, \delta > 0$*, there exists an* $\tilde{\epsilon} > 0$ *such that*

$$\liminf_{N\to\infty} \frac{1}{N} \log \Big(\inf_{z\in A;\, |x-z|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,z} - \phi(t)| < \epsilon \Big] \Big) \geq -I_{T,x}(\phi) - \delta.$$

Proof We assume w.l.o.g. that $I_{T,x}(\phi) < \infty$. By Theorem 7.7, there exists an N_0 and $\tilde{\epsilon}$ such that for $N \geq N_0$ and z with $|z - x| < \tilde{\epsilon}$,

$$\frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,z} - \phi(t)| < \epsilon \Big] \geq - \inf_{\tilde{\phi}:\|\phi-\tilde{\phi}\|<\epsilon} I_{T,x}(\tilde{\phi}) - \delta \geq -I_{T,x}(\phi) - \delta.$$

The assertion follows.

7.5.2 LDP Lower Bound with Vanishing Rates

In the following, we drop the assumption that the log-rates are bounded. Instead, we rather consider situations, where Assumption 7.2 is satisfied.

We start by some preliminary considerations and assume that Assumption 7.2 (A1) and (A2) are satisfied. We note that there exists a constant $\alpha > 0$ such that for all $x \in A$ there exists a $i \leq I$ such that $B(x, \alpha) \subset B_i$. Indeed, assume that this is incorrect and consider a sequence of points $x_n \in A$ such that $B(x_n, 1/n)$ is not contained in any B_i. W.l.o.g., we can assume that $x_n \to x \in A$ (recall that A is compact). As $x \in B_{i_0}$ for some i_0, we have $B(x_n, 1/n) \subset B_{i_0}$ for n large enough, a contradiction.

Lemma 7.23 *Assume that* β_j *(*$j = 1, \ldots, k$*) is bounded and that Assumption 7.2 (A1) and (A2) are satisfied. Then, for* $T > 0$*,* $K > 0$*, there exists a* $J = J(T, K) \in \mathbb{N}$ *such that for all* $\phi \in D([0,T]; A)$ *with* $I_T(\phi) \leq K$*, there exist*

$$0 = t_0 < t_1 < \cdots < t_J = T \text{ and } i_1, \ldots, i_J \text{ such that } \phi(t) \in B_{i_r} \text{ for } t \in [t_{r-1}, t_r].$$

Furthermore, for $r = 1, \ldots, J$*,*

$$\operatorname{dist}(\phi(t_{r-1}), \partial B_{i_r}) \geq \alpha \quad \text{and} \quad \operatorname{dist}(\phi(t), \partial B_{i_r}) \geq \alpha/2 \quad \text{for } t \in [t_{r-1}, t_r]$$

for α *as before.*

Proof By the considerations above, we have $B(x,\alpha) \subset B_{i_1}$ for an appropriate i_1. We define

$$\tilde{t}_1 := \inf\{t \geq 0 | B(\phi(t), \alpha/2) \not\subset B_{i_1}\} \wedge T > 0.$$

Now, there exists an i_2 such that $B(\phi(t_1), \alpha) \subset B_{i_2}$. If $\tilde{t}_1 < T$, we define

$$\tilde{t}_2 := \inf\{t \geq t_1 | B(\phi(t), \alpha/2) \not\subset B_{i_2}\} \wedge T > \tilde{t}_1.$$

In the same way, we proceed. By the uniform absolute continuity (Lemma 7.18) of all ϕ with $I_T(\phi) \leq K$, we have

$$\tilde{t}_r - \tilde{t}_{r-1} \geq \delta \quad \text{for a constant } \delta > 0 \text{ independent of } \phi.$$

The assertion hence follows for $J := \lfloor \frac{T}{\delta} \rfloor + 1$ and $t_r := r\delta$ $(r = 1, \ldots, J-1)$, $t_J := T$.

We now define a function ϕ^η which is close to a given function ϕ with $I_T(\phi) < \infty$. We assume that Assumption 7.2 (A) holds. Hence, for $x \in B_i \cap A$ and $t \in (0, \lambda_2)$,

$$\mathrm{d}(x + tv_i, \partial A) > \lambda_1 t.$$

Note that $\lambda_1 \leq 1$. Let $\eta > 0$ be small. We define for $r = 1, \ldots, J$, with the notation $\sum_{j=1}^{0} \cdots = 0$,

$$\eta_r := \eta \sum_{j=1}^{r} (\frac{3}{\lambda_1})^{j-1}.$$

- For $r = 1, \ldots, J$, $t \in \left[t_{r-1} + \eta_{r-1}, t_{r-1} + \eta_r\right]$,

$$\phi^\eta(t) := \phi(t_{r-1}) + \eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j} + \left(t - t_{r-1} - \eta_{r-1}\right) v_{i_r}.$$

- For $r = 1, \ldots, J$, $t \in \left[t_{r-1} + \eta_r, t_r + \eta_r\right]$,

$$\phi^\eta(t) := \phi\left(t - \eta_r\right) + \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j}.$$

We make the following assumptions on η:

$$\eta_J = \eta \sum_{r=1}^{J} \left(\frac{3}{\lambda_1}\right)^{r-1} \leq \frac{\alpha}{4} \wedge \min_{r=1,\dots,J} |t_r - t_{r-1}| \wedge \lambda_2.$$

Therefore, we have the following properties for ϕ^η:

- For $r = 1, \dots, J$, $t \in \left[t_{r-1} + \eta_{r-1}, t_{r-1} + \eta_r\right]$,

  $$|\phi(t) - \phi^\eta(t)| \leq |\phi(t) - \phi(t_{r-1})| + \eta_{r-1} + (\eta_r - \eta_{r-1}) = V_{t_r - t_{r-1}}(\phi) + \eta_r \to 0$$

  as $\eta \to 0$, where $V_\cdot(\phi)$ is the modulus of continuity of ϕ. Similarly, for $r = 1, \dots, J$, $t \in \left[t_{r-1} + \eta_r, t_r + \eta_r\right]$,

  $$|\phi(t) - \phi^\eta(t)| \leq |\phi(t) - \phi(t - \eta_r)| + \eta_r = V_{\eta_r}(\phi) + \eta_r \to 0$$

  as $\eta \to 0$.
- For $r = 1, \dots, J$, $t \in \left[t_{r-1} + \eta_{r-1}, t_{r-1} + \eta_r\right]$,

  $$\text{dist}(\phi^\eta(t), \partial B_{i_r}) \geq \text{dist}(\phi(t_{r-1}), \partial B_{i_r}) - \eta_r \geq \alpha - \frac{\alpha}{4}.$$

  Similarly, for $r = 1, \dots, J$, $t \in \left[t_{r-1} + \eta_r, t_r + \eta_r\right]$, hence $t - \eta_r \in [t_{r-1}, t_r]$,

  $$\text{dist}(\phi^\eta(t), \partial B_{i_r}) \geq \text{dist}(\phi(t - \eta_r), \partial B_{i_r}) - \eta_r \geq \frac{\alpha}{2} - \frac{\alpha}{4}.$$

  Hence, for $r = 1, \dots, J$, $t \in \left[t_{r-1} + \eta_{r-1}, t_r + \eta_r\right]$,

  $$\text{dist}(\phi^\eta(t), \partial B_{i_r}) \geq \frac{\alpha}{4}.$$

- For $t \in [0, \eta]$,

  $$\text{dist}(\phi^\eta(t), \partial A) \geq t\lambda_1. \tag{7.37}$$

  For $t \in \left[\eta, T + \eta_J\right]$,

  $$\text{dist}(\phi^\eta(t), \partial A) \geq \lambda_1 \eta. \tag{7.38}$$

  This can be seen by induction on $r = 1, \dots, J$ (the induction hypothesis is clear, cf. Inequality (7.37)). For $r = 1, \dots, J$, we have (by induction hypothesis and the assumptions on η)

  $$\phi^\eta\Big(t_{r-1} + \eta_{r-1}\Big) \in B_{i_r}, \text{ and for } r \geq 2, \text{ dist}\Big(\phi^\eta\Big(t_{r-1} + \eta_{r-1}\Big), \partial A\Big) \geq \eta\lambda_1.$$

From Assumption 7.2 (A3), the distance of $\phi^\eta(t)$ to the boundary is increasing for $t \in \left[t_{r-1} + \eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1}, t_{r-1} + \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1}\right]$, and is at least

$$\left(t - \left(t_{r-1} + \eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1}\right)\right) \cdot \lambda_1 \vee \eta\lambda_1.$$

In particular,

$$\operatorname{dist}\left(\phi^\eta\left(t_{r-1} + \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1}\right), \partial A\right) \geq \eta\lambda_1\left(\frac{3}{\lambda_1}\right)^{r-1}.$$

For $t \in \left[t_{r-1} + \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1}, t_r + \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1}\right]$, we have

$$\begin{aligned}
\phi^\eta(t) &= \phi\left(t - \eta \sum_{j=1}^{r} \left(\frac{3}{\lambda_1}\right)^{j-1}\right) + \eta\left(\frac{3}{\lambda_1}\right)^{r-1} v_{i_r} + \eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j} \\
&= \bar{\phi}(t) + \eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j}
\end{aligned}$$

and therefore (by elementary calculus and the fact that $|v_i| \leq 1$)

$$\begin{aligned}
\operatorname{dist}(\phi^\eta(t), \partial A) &\geq \operatorname{dist}(\bar{\phi}(t), \partial A) - \left|\eta \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j}\right| \\
&\geq \eta \left(\frac{3}{\lambda_1}\right)^{r-1} \lambda_1 \left(1 - \frac{\mathbf{1}_{r\geq 2}}{2}\right) \\
&\geq \eta\lambda_1.
\end{aligned}$$

We now have

Lemma 7.24 *Assume that Assumption 7.2 holds. Let $K > 0$ and $\epsilon > 0$. Then there exists an $\eta_0 = \eta_0(T, K, \epsilon) > 0$ such that for all $\phi \in D([0, T]; A)$ with $I_T(\phi) \leq K$ and all $\eta < \eta_0$,*

$$I_T(\phi^\eta) \leq I_T(\phi) + \epsilon,$$

where $\phi^\eta(t)$ is defined as above.

Proof We first use Lemma 7.2.3 and chose $\eta < \eta_1$ small enough (independent of i, ϕ) such that

$$\sum_{r=1}^{J} \int_{t_{r-1}+\eta_{r-1}}^{t_{r-1}+\eta_r} L(\phi^\eta(t), (\phi^\eta)'(t))dt = \sum_{r=1}^{J} I_{\eta_r - \eta_{r-1}}(\phi) < \frac{\epsilon}{2}.$$

We now denote by $\mu^*(t)$ the optimal μ corresponding to $(\phi(t), \phi'(t))$ (cf. Lemma 7.10). We let $r = 1, \ldots, J$ and $t \in [t_{r-1} + \eta_r, t_r + \eta_r]$ and note that $(\phi^\eta)'(t) = \phi'(t - \eta_r)$. By Theorem 7.4, we have

$$L(\phi^\eta(t), \phi'(t - \eta_r)) \leq \ell(\phi^\eta(t), \mu^*(t - \eta_r)). \tag{7.39}$$

By the Lipschitz continuity of the β_j, we have

$$|\beta_j(\phi^\eta(t)) - \beta_j(\phi(t - \eta_r))| \leq \delta_K(\eta) \tag{7.40}$$

where $\delta_K(\eta)$ is independent of ϕ and $\delta_K(\eta) \to 0$ as $\eta \to 0$. We deduce from (7.39) and (7.40)

$$\begin{aligned} L(\phi^\eta(t), \phi'(t - \eta_r)) - L(\phi(t - \eta_r), \phi'(t - \eta_r)) &\leq k\delta_K(\eta) \\ + \sum_j \mu_j^*(t - \eta_r) \log \frac{\beta_j(\phi(t - \eta_r))}{\beta_j(\phi^\eta(t))}&. \end{aligned} \tag{7.41}$$

Let

$$\tilde{v}_{i_r} = \left(\frac{3}{\lambda_1}\right)^{r-1} v_{i_r} + \sum_{j=1}^{r-1} \left(\frac{3}{\lambda_1}\right)^{j-1} v_{i_j}, \quad \text{and } \hat{v}_{i_r} = \frac{\tilde{v}_{i_r}}{|\tilde{v}_{i_r}|} \in \mathcal{C}_{1,i_r}.$$

By Assumption 7.2 (B4), there exists a constant $\lambda_4 > 0$ such that for $z \in B_{i_r}$ and $\eta < \eta_2 \leq \eta_1$ small enough (note that η_2 depends on λ_1 and λ_2 but not directly on ϕ, except through K),

$$\beta_j(z) < \lambda_4 \Rightarrow \beta_j(z + \eta\tilde{v}_{i_r}) \geq \beta_j(z),$$

hence

$$\log \frac{\beta_j(\phi(t - \eta_r))}{\beta_j(\phi^\eta(t))} < 0 \quad \text{if } \beta_j(\phi(t - \eta_r)) < \lambda_4. \tag{7.42}$$

If $\beta_j(\phi(t-\eta_r)) \geq \lambda_4$ (recall the definition of $\delta_K(\eta)$ and choose $\eta < \eta_3 < \eta_2$ small enough such that $\delta_K(\eta) < \lambda_4/2$)

$$\begin{aligned}\log \frac{\beta_j(\phi(t-\eta_r))}{\beta_j(\phi^\eta(t))} &\leq \log \frac{\beta_j(\phi(t-\eta_r))}{\beta_j(\phi(t-\eta_r)) - \delta_K(\eta)} \\ &\leq \log \frac{\lambda_4}{\lambda_4 - \delta_K(\eta)} \\ &= \log \frac{1}{1-\delta_K(\eta)/\lambda_4} \\ &\leq \frac{2\delta_K(\eta)}{\lambda_4}, \end{aligned} \tag{7.43}$$

since $\log(1/(1-x)) < 2x$ for $0 < x \leq 1/2$.

From Lemmas 7.8 and 7.10, there exist (universal, i.e., independent of x) constants $B \geq B_1 \vee B_2$, $B > 1$, C_1, C_3 such that for $|y| \geq B$, and $x \in A$,

$$L(x,y) \geq C_1|y|\log|y|, \tag{7.44}$$

$$|\mu^*| = |\mu^*(x,y)| \leq C_3|y|. \tag{7.45}$$

Hence if $|\phi'(t-\eta_r)| \geq B$, using (7.41)–(7.43) and (7.45) for the first inequality and (7.44) for the second, we get

$$\begin{aligned} L(\phi^\eta(t), \phi'(t-\eta_r)) - L(\phi(t-\eta_r), \phi'(t-\eta_r)) &\leq k\delta_K(\eta) + kC_3|\phi'(t-\eta_r)|\frac{2\delta_K(\eta)}{\lambda_4} \\ &\leq k\delta_K(\eta) + kC_3\frac{2\delta_K(\eta)L(\phi(t-\eta_r), \phi'(t-\eta_r))}{C_1\lambda_4\log|\phi'(t-\eta_r)|}. \end{aligned} \tag{7.46}$$

If however $|\phi'(t-\eta_r)| < B$, Lemma 7.10 implies similarly as before that $|\mu^*(t-\eta_r)| \leq C_3B$. From (7.41), we deduce

$$L(\phi^\eta(t), \phi'(t-\eta_r)) - L(\phi(t-\eta_r), \phi'(t-\eta_r)) \leq k\delta_K(\eta) + kC_3B\frac{2\delta_K(\eta)}{\lambda_4}. \tag{7.47}$$

Inequalities (7.46) and (7.47) imply

$$\begin{aligned} L(\phi^\eta(t), \phi'(t-\eta_r)) - L(\phi(t-\eta_r), \phi'(t-\eta_r)) &\leq \delta_{1,K}(\eta) \\ &+ \delta_{2,K}(\eta)L(\phi(t-\eta_r), \phi'(t-\eta_r)) \end{aligned}$$

with constants $\delta_{i,K}(\eta) \to 0$ as $\eta \to 0$. We can hence choose $\eta < \eta_4 < \eta_3$ small enough such that

$$\sum_{r=1}^{J} \int_{t_{r-1}+\eta_r}^{t_r+\eta_r} L(\phi^\eta(t), \phi'(t-\eta_r))dt - \sum_{r=1}^{J} \int_{t_{r-1}+\eta_r}^{t_r+\eta_r} L(\phi(t-\eta_r), \phi'(t-\eta_r))dt < \frac{\epsilon}{2}.$$

This yields the result.

The following lemma is the main difference with the corresponding result of [14]. We transform the LLN from Assumption 7.2 (C) to a LDP lower bound for linear functions following the vector v_i near the boundary.

In the next statement, α is the exponent which appears in the Assumption 7.2 (C).

Lemma 7.25 *Assume that Assumption 7.2 holds. Let $i \leq I_1$, $x \in A \cap B_i$ and $x^N \in A^N \cap B_i$ such that*

$$\limsup_{N\to\infty} |x^N - x| N^\alpha < 1.$$

Let furthermore $\epsilon > 0$ and define μ^i, ϕ^x and η_0 as in Assumption 7.2 (C). Then for all η small enough, all ϵ,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,\eta]} |Z^{N,x^N}(t) - \phi^x(t)| < \epsilon\Big] \geq -\int_0^\eta \ell(\phi^x(t), \mu^i)dt,$$

and the above convergence is uniform in $x \in A$.

Proof The proof follows the same line of reasoning as the proof of Proposition 7.3 but is technically more involved.

For simplicity, let N be large enough and $\eta < \eta_0$ (for η_0 as in Assumption 7.2 (C)) be small enough such that $\phi^{x^N}(t) \in B_i$ for all $t \leq \eta$. We furthermore let

$$\tilde{\epsilon} < \epsilon_1 := \epsilon \wedge \lambda_1 \eta,$$

and define the set

$$F^N := \Big\{\sup_{t\in[0,\eta]} |Z^{N,x^N}(t) - \phi^x(t)| < \tilde{\epsilon}\Big\}.$$

Let

$$\xi_\eta = \xi_\eta^N = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}|_{\mathcal{F}_\eta}$$

be given as in Theorem 7.15 for the rates $\tilde{\beta}_j = \tilde{\mu}^i_j$. We note that due to Assumption 7.2 (C),

$$\tilde{\mathbb{P}}[F^N] \geq 1 - \delta(N, \tilde{\epsilon}) \to 1 \quad \text{as } N \to \infty. \tag{7.48}$$

From Corollary 7.10, (7.48) and Jensen's inequality, we deduce that

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,\eta]} |Z^{N,x^N}(t) - \phi^x(t)| < \epsilon\Big] \\
&\quad\geq \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\big[F^N\big] \\
&\quad\geq \liminf_{N\to\infty} \frac{1}{N} \log \tilde{\mathbb{E}}\big[\xi_\eta^{-1} \mathbb{1}_{F^N}\big] \\
&\quad= \liminf_{N\to\infty} \frac{1}{N} \log \Big\{\tilde{\mathbb{P}}\big[F^N\big]\tilde{\mathbb{E}}_{F^N}\big[\exp(X_\eta)\big]\Big\} \\
&\quad\geq \liminf_{N\to\infty} \frac{1}{N} \log \tilde{\mathbb{P}}\big[F^N\big] + \liminf_{N\to\infty} \frac{1}{N} \log \tilde{\mathbb{E}}_{F^N}\big[\exp(X_\eta)\big] \\
&\quad\geq \liminf_{N\to\infty} \tilde{\mathbb{E}}_{F^N}\Big[\frac{X_\eta}{N}\Big],
\end{aligned} \tag{7.49}$$

where $\tilde{\mathbb{E}}_{F^N}[X] := \tilde{\mathbb{E}}[X|F^N]$ and

$$\begin{aligned}
X_\eta^N := X_\eta := \log \xi_\eta^{-1} &= \sum_{\tau\leq\eta} \Big[\log \beta_{j(\tau)}(Z^{N,x^N}(\tau-) - \log \tilde{\mu}^i_{j(\tau)}(Z^{N,x^N}(\tau-)\Big] \\
&\quad + N \sum_j \int_0^\eta \big(\tilde{\mu}^i_j(Z^{N,x^N}(t) - \beta_j(Z^{N,x^N}(t))\big)dt.
\end{aligned}$$

We have $\operatorname{dist}(\phi^{x^N}(t), \partial A) \geq \lambda_1 t$ (cf. Assumption 7.2 (A3)) and therefore on F^N,

$$\operatorname{dist}(Z^{N,x^N}(t), \partial A) > \lambda_1 t - \tilde{\epsilon} \quad \text{for } t \in \Big[\frac{\tilde{\epsilon}}{\lambda_1}, \eta\Big].$$

Consequently

$$\tilde{\mu}^i_j(Z^{N,x^N}(t)) = \mu^i_j \quad \text{for all } j \text{ and for all } t \in \Big[\frac{\tilde{\epsilon}}{\lambda_1}, \eta\Big].$$

We obtain

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\mathbb{E}}\Big[\mathbb{1}_{F^N}\int_0^\eta \sum_j \tilde{\mu}^i_j(Z^{N,x^N}(t))dt\Big] \\
&= \frac{1}{\tilde{\mathbb{P}}[F^N]}\Big(\tilde{\mathbb{E}}\Big[\mathbb{1}_{F^N}\int_{\tilde{\epsilon}/\lambda_1}^\eta \sum_{j=1}^k \tilde{\mu}^i_j(Z^{N,x^N}(t))dt\Big] \\
&\quad + \tilde{\mathbb{E}}\Big[\mathbb{1}_{F^N}\int_0^{\tilde{\epsilon}/\lambda_1} \sum_{j=1}^k \tilde{\mu}^i_j(Z^{N,x^N}(t))dt\Big]\Big) \\
&=: \sum_{j=1}^k \int_{\tilde{\epsilon}/\lambda_1}^\eta \mu^i_j dt + X_1^N(\tilde{\epsilon}), \qquad (7.50)
\end{aligned}$$

since $\mu^i_j(Z^{N,x^N}(t)) = \mu^i_j$ on F^N. We note that for all N,

$$|X_1^N(\tilde{\epsilon})| \le \frac{\tilde{\epsilon}}{\lambda_1}k\bar{\mu} \quad \text{where } \bar{\mu} := \max_{j=1,\dots,k} \mu^i_j. \qquad (7.51)$$

Since the β_j's are bounded and continuous and by Theorem 7.3, we have for $j = 1, \dots, k$,

$$\sup_{t\in[0,\eta]} |\beta_j(Z^{N,x^N}(t)) - \beta_j(\phi^x(t))| \to 0 \quad \text{a.s. as } N \to \infty.$$

Combined with (7.48), this implies

$$\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\mathbb{E}}\Big[\mathbb{1}_{F^N}\int_0^\eta \sum_{j=1}^k \beta_j(Z^{N,x^N}(t))dt\Big] \longrightarrow \sum_{j=1}^k \int_0^\eta \beta_j(\phi^x(t))dt \qquad (7.52)$$

as $N \to \infty$.

Let us now define the following processes. For $z \in A$, $j = 1, \dots, k$ and $0 \le s < t \le \eta$ let $\bar{Z}^{N,z}$ solves Eq. (7.3) with constant rates μ^i_j under $\tilde{\mathbb{P}}$, and

$$\begin{aligned}
Y_j^{N,z,s,t} &:= \frac{1}{N}\cdot \#\text{jumps of } Z^{N,z} \text{ in direction } h_j \text{ in } [s,t], \\
\bar{Y}_j^{N,z,s,t} &:= \frac{1}{N}\cdot \#\text{jumps of } \bar{Z}^{N,z} \text{ in direction } h_j \text{ in } [s,t],
\end{aligned}$$

We have for any event F, noting that $\tilde{\mathbb{P}}[F] - \tilde{\mathbb{P}}[F]^2 = \tilde{\mathbb{P}}[F^c] - \tilde{\mathbb{P}}[F^c]^2$,

$$\tilde{\mathbb{E}}[Y_j^{N,z,s,t}] \le (t-s)\mu_j^i = \tilde{\mathbb{E}}[\bar{Y}_j^{N,z,s,t}], \tag{7.53}$$

$$\widetilde{\mathrm{Var}}[Y_j^{N,z,s,t}] \le (t-s)\mu_j^i = \widetilde{\mathrm{Var}}[\bar{Y}_j^{N,z,s,t}],$$

$$\begin{aligned} |\widetilde{\mathrm{Cov}}(\mathbb{1}_F, Y_j^{N,z,s,t})|,\ |\widetilde{\mathrm{Cov}}(\mathbb{1}_F, \bar{Y}_j^{N,z,s,t})| &\le \sqrt{\widetilde{\mathrm{Var}}[\mathbb{1}_F]}\sqrt{\widetilde{\mathrm{Var}}[\bar{Y}_j^{N,z,s,t}]} \\ &\le \sqrt{\tilde{\mathbb{P}}[F] - \tilde{\mathbb{P}}[F]^2}\sqrt{(t-s)\mu_j^i}. \end{aligned} \tag{7.54}$$

We define the sets

$$F_1^N := \left\{ \left| Z^{N,x^N}\left(\frac{2\tilde{\epsilon}}{\lambda_1}\right) - \phi^x\left(\frac{2\tilde{\epsilon}}{\lambda_1}\right) \right| < \frac{\tilde{\epsilon}}{2} \right\} \in \mathcal{F}_{2\tilde{\epsilon}/\lambda_1}$$

and for $z \in A$ with $|z - \phi^x(2\tilde{\epsilon}/\lambda_1)| < \tilde{\epsilon}/2$,

$$F_2^{N,z} := \left\{ \sup_{t \in [0, \eta - 2\tilde{\epsilon}/\lambda_1]} |Z^{N,z}(t) - \phi^z(t)| < \frac{\tilde{\epsilon}}{2} \right\}.$$

Note that

$$\mathrm{dist}(\phi^x(t), \partial A) \ge 2\tilde{\epsilon} \quad \text{for } t \in \left[\frac{2\tilde{\epsilon}}{\lambda_1}, \eta\right]$$

and whenever $|z - \phi^x(2\tilde{\epsilon}/\lambda_1)| < \tilde{\epsilon}/2$,

$$\left| \phi^z(t) - \phi^x\left(t + \frac{2\tilde{\epsilon}}{\lambda_1}\right) \right| < \frac{\tilde{\epsilon}}{2} \quad \text{for } t \in \left[0, \eta - \frac{2\tilde{\epsilon}}{\lambda_1}\right].$$

Hence

$$\mathrm{dist}(Z^{N,z}(t), \partial A) \ge \tilde{\epsilon} \quad \text{for } t \in \left[0, \eta - \frac{2\tilde{\epsilon}}{\lambda_1}\right]$$

and therefore $Z^{N,z}(t) = \bar{Z}^{N,z}(t)$ on $F_2^{N,z}$. This implies

$$F_2^{N,z} = \left\{ \sup_{t \in [0, \eta - 2\tilde{\epsilon}/\lambda_1]} |\bar{Z}^{N,z}(t) - \phi^z(t)| < \frac{\tilde{\epsilon}}{2} \right\}.$$

We now let

$$\frac{2\tilde{\epsilon}}{\lambda_1} \le s < t \le \eta$$

and compute (by using the Markov property of Z^N and the fact that $Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1} = \bar{Y}_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}$ on the event $F_2^{N,z}$)

$$
\begin{aligned}
\tilde{\mathbb{E}}[Y_j^{N,x^N,s,t}] &= \tilde{\mathbb{E}}[\mathbb{1}_{F_1^N} Y_j^{N,x^N,s,t}] + \tilde{\mathbb{E}}[\mathbb{1}_{(F_1^N)^c} Y_j^{N,x^N,s,t}] \\
&\geq \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{E}}[Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}] \\
&\quad + \tilde{\mathbb{E}}[\mathbb{1}_{(F_1^N)^c} Y_j^{N,x^N,s,t}] \\
&\geq \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{E}}[\mathbb{1}_{F_2^{N,z}} Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}] \\
&\quad + \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{E}}[\mathbb{1}_{(F_2^{N,z})^c} Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}] \\
&\quad + \tilde{\mathbb{E}}[\mathbb{1}_{(F_1^N)^c} Y_j^{N,x^N,s,t}] \\
&= \mu_j^i(t-s)\tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{P}}[F_2^{N,z}] \\
&\quad + \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \widetilde{\mathrm{Cov}}(\mathbb{1}_{F_2^{N,z}}, Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}) \\
&\quad + \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{E}}[\mathbb{1}_{(F_2^{N,z})^c} Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}] \\
&\quad + \tilde{\mathbb{E}}[\mathbb{1}_{(F_1^N)^c} Y_j^{N,x^N,s,t}] \qquad (7.55) \\
&\geq \mu_j^i(t-s)\tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{P}}[F_2^{N,z}] \\
&\quad - \tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \left|\widetilde{\mathrm{Cov}}(\mathbb{1}_{F_2^{N,z}}, Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1})\right|
\end{aligned} \tag{7.56}
$$

as the third and the fourth term in (7.55) are non-negative. As furthermore

$$\tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \tilde{\mathbb{P}}[F_2^{N,z}] \to 1$$

and

$$\tilde{\mathbb{P}}[F_1^N] \cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \widetilde{\mathrm{Cov}}(\mathbb{1}_{F_2^{N,z}}, Y_j^{N,z,s-2\tilde{\epsilon}/\lambda_1,t-2\tilde{\epsilon}/\lambda_1}) \to 0$$

as $N \to \infty$ by Assumption 7.2 (C), Theorem 7.3 and (7.54). Combining the resulting inequality with (7.53) for all $\tilde{\epsilon} < \epsilon_1$ and $2\tilde{\epsilon}/\lambda_1 \leq s < t$, we deduce that

$$\lim_{N\to\infty} \tilde{\mathbb{E}}[Y_j^{N,x^N,s,t}] = \mu_j^i(t-s).$$

Note that for $0 \le s < t \le \eta$, and all $\tilde{\epsilon} < \epsilon_1$

$$Y_j^{N,x^N,s,t} = Y_j^{N,x^N,s,(s\vee 2\tilde{\epsilon}/\lambda_1)\wedge t} + Y_j^{N,x^N,(s\vee 2\tilde{\epsilon}/\lambda_1)\wedge t,t}$$

and hence also for $0 \le s < t \le \eta$, since when $s < 2\tilde{\epsilon}/\lambda_1$, $Y_j^{N,x^N,s,2\tilde{\epsilon}/\lambda_1\wedge t}$ is of the order of $\tilde{\epsilon}$,

$$\lim_{N\to\infty} \tilde{\mathbb{E}}[Y_j^{N,x^N,s,t}] = \mu_j^i(t-s). \tag{7.57}$$

Let now $\tau_j \in [0,\eta]$ denote the jump times of Z^{N,x^N} in direction h_j. Since $\tilde{\mu}_j^i(Z^{N,x^N}(\tau_j-)) = \log \mu_j^i$ $\tilde{\mathbb{P}}$ a.s.,

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]} \sum_{j;\mu_j^i>0} \tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^N} \sum_{\tau_j\le\eta} \log \tilde{\mu}_j^i(Z^{N,x^N}(\tau_j-))\Big] \\
&\quad= \frac{1}{\tilde{\mathbb{P}}[F^N]} \sum_{j;\mu_j^i>0} \log \mu_j^i \;\tilde{\mathbb{E}}\big[\mathbb{1}_{F^N} Y_j^{N,x^N,0,\eta}\big] \\
&\quad= \frac{1}{\tilde{\mathbb{P}}[F^N]} \sum_{j;\mu_j^i>0} \log \mu_j^i \Big\{\tilde{\mathbb{P}}[F^N]\cdot \tilde{\mathbb{E}}\big[Y_j^{N,x^N,0,\eta}\big] + \widetilde{\mathrm{Cov}}(\mathbb{1}_{F^N}, Y_j^{N,x^N,0,\eta})\Big\} \\
&\quad\longrightarrow \sum_{j;\mu_j^i>0} \eta\mu_j^i \log \mu_j^i = \sum_{j=1}^k \int_0^\eta \mu_j^i \log \mu_j^i dt
\end{aligned} \tag{7.58}$$

as $N \to \infty$ by (7.48), (7.54) and (7.57).

For the last and most extensive step of the proof, we define for $\tilde{\epsilon} < \epsilon_1$ and (cf. Assumption 7.2 (C) and (7.8))

$$\epsilon_N = \frac{1}{N^\alpha}$$

and the set

$$\tilde{F}^N := \Big\{ \sup_{t\in[0,\eta]} |Z^{N,x^N}(t) - \phi^x(t)| < \epsilon_N \Big\}.$$

We assume w.l.o.g. that from now on N is large enough (cf. Assumption 7.2 (C)) such that

$$\tilde{\mathbb{P}}\big[\tilde{F}^N\big] \ge 1 - \delta(N, \epsilon_N),$$

where $\delta(N,\epsilon_N)\to 0$ as $N\to\infty$. We note that we have for all $\tilde\epsilon\le\epsilon_1$, $j=1,\dots,k$ with $\mu_j^i>0$, $N\in\mathbb{N}$ and $t\in[2\tilde\epsilon/\lambda_1,\eta]$,

$$\log\beta_j(\phi^x(t)),\log\beta_j(Z^{N,x^N}(t))\ge\log\underline{\beta}(\tilde\epsilon)>0\quad\text{on } F^N \text{ and } \tilde F^N.$$

We compute

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^N}\sum_{\tau\le\eta}\log\beta_{j(\tau)}(Z^{N,x^N}(\tau-))\Big]\\
&=\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu_j^i>0}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde F^N}\sum_{\tau_j\in[2\tilde\epsilon/\lambda_1,\eta]}\log\beta_j(Z^{N,x^N}(\tau_j-))\Big]\\
&\quad+\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu_j^i>0}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde F^N}\sum_{\tau_j\in[0,2\tilde\epsilon/\lambda_1]}\log\beta_j(Z^{N,x^N}(\tau_j-))\Big]\\
&\quad+\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu_j^i>0}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^N\setminus\tilde F^N}\sum_{\tau_j\in[0,\eta]}\log\beta_j(Z^{N,x^N}(\tau_j-))\Big].
\end{aligned}\tag{7.59}$$

Let us first consider the first term in Eq. (7.59). As for all j,

$$\log\beta_j(\cdot):\tilde A(\tilde\epsilon):=\{z\in A\,|\,\mathrm{dist}(z,\partial A)\ge\tilde\epsilon\}\to\mathbb{R}$$

is uniformly continuous, there exit constants $\tilde\delta_N>0$ with $\tilde\delta_N\downarrow 0$ such that

$$z,\tilde z\in\tilde A(\tilde\epsilon),\ |\tilde z-z|<3\epsilon_N\Rightarrow|\log\beta_j(\tilde z)-\log\beta_j(z)|<\tilde\delta_N.\tag{7.60}$$

We define

$$M=M(N):=\lfloor(\eta-2\tilde\epsilon/\lambda_1)\epsilon_N^{-1}+1\rfloor$$

and divide the interval $[2\tilde\epsilon/\lambda_1,\eta]$ into M equidistant subintervals $[t_r,t_{r+1}]$ ($r=0,\dots,M-1$, $t_r=t_r(N)$) of length $\Delta=\Delta(N)$, i.e. (for N large enough),

$$\frac{\epsilon_N}{2}\le\Delta<\epsilon_N.$$

For $j=1,\dots,k$, $r=0,\dots,M-1$ and $\tau_j,t\in[t_r,t_{r+1}]$ we have,

$$|Z^{N,x^N}(\tau_j-)-\phi^x(t)|\le 2\epsilon_N\quad\text{on }\tilde F^N,$$

since $|\phi^x(\tau_j) - \phi^x(t)| \le |\tau_j - t|$ as $|v_i| \le 1$, and hence (cf. (7.60))

$$\inf_{t\in[t_r,t_{r+1}]} \log \beta_j(\phi^x(t)) - \tilde{\delta}_N \le \log \beta_j(Z^{N,x^N}(\tau_j-)).$$

From this inequality, we deduce

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau\in[2\tilde{\epsilon}/\lambda_1,\eta]}\log\beta_{j(\tau)}(Z^{N,x^N}(\tau-))\Big] \\
&= \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\sum_{r=0}^{M-1}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j\in[t_r,t_{r+1})}\log\beta_j(Z^{N,x^N}(\tau_j-))\Big] \\
&\ge \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\sum_{r=0}^{M-1}\Big(\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t)) - \tilde{\delta}_N\Big)\tilde{\mathbb{E}}\Big[\mathbb{1}_{\tilde{F}^N}Y_j^{N,x^N,t_r,t_{r+1}}\Big] \\
&\ge \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\sum_{r=0}^{M-1}\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t))\tilde{\mathbb{P}}[\tilde{F}^N]\tilde{\mathbb{E}}\big[Y_j^{N,x^N,t_r,t_{r+1}}\big] \\
&\quad - \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\sum_{r=0}^{M-1}\big|\log\underline{\beta}(\tilde{\epsilon})\big|\big|\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^N}, Y_j^{N,x^N,t_r,t_{r+1}})\big| \\
&\quad - \frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\delta}_N\sum_{j,\,\mu^i_j>0}\tilde{\mathbb{E}}\big[\mathbb{1}_{\tilde{F}^N}Y^{N,x^N,2\tilde{\epsilon}/\lambda_1,\eta}\big].
\end{aligned} \tag{7.61}$$

The second term in Inequality (7.61) satisfies (cf. Inequality (7.54) and Assumption 7.2 (C); we assume that N is sufficiently large such that $M \le 2\epsilon_N^{-1}\eta$),

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\sum_{r=0}^{M-1}\big|\log\underline{\beta}(\tilde{\epsilon})\big|\big|\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^N}, Y_j^{N,x^N,t_r,t_{r+1}})\big| \\
&\le \frac{1}{\tilde{\mathbb{P}}[F^N]}2k\eta\big|\log\underline{\beta}(\tilde{\epsilon})\big|\epsilon_N^{-1}\sqrt{\bar{\mu}\,\epsilon_N}\sqrt{\delta(N,\epsilon_N)} \\
&\to 0
\end{aligned}$$

as $N \to \infty$. The third term in Eq. (7.61) satisfies

$$\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\delta}_N\sum_{j,\,\mu^i_j>0}\tilde{\mathbb{E}}\big[\mathbb{1}_{\tilde{F}^N}Y^{N,x^N,2\tilde{\epsilon}/\lambda_1,\eta}\big] \le \frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\delta}_N k\bar{\mu}\eta \to 0$$

as $N \to \infty$. Finally, let us consider the first term in Eq. (7.61). Recall that by (7.53) and (7.56), we have

$$\begin{aligned}\mu_j^i(t_{r+1}-t_r) &\geq \tilde{\mathbb{E}}[Y_j^{N,x^N,t_r,t_{r+1}}]\\ &\geq \mu_j^i(t_{r+1}-t_r)\tilde{\mathbb{P}}[F_1^N]\cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2}\tilde{\mathbb{P}}[F_2^{N,z}]\\ &- \tilde{\mathbb{P}}[F_1^N]\cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2}\left|\widetilde{\mathrm{Cov}}(\mathbb{1}_{F_2^{N,z}}, Y_j^{N,z,t_r-2\tilde{\epsilon}/\lambda_1,t_{r+1}-2\tilde{\epsilon}/\lambda_1})\right|.\end{aligned}$$

We define

$$\begin{aligned}\alpha_1^N &:= \frac{\tilde{\mathbb{P}}[\tilde{F}^N]}{\tilde{\mathbb{P}}[F^N]},\\ \alpha_2^N &:= \tilde{\mathbb{P}}[F_1^N]\cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2}\tilde{\mathbb{P}}[F_2^{N,z}] < 1,\\ \alpha_3^N &:= \tilde{\mathbb{P}}[F_1^N]\cdot \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2}\left|\widetilde{\mathrm{Cov}}(\mathbb{1}_{F_2^{N,z}}, Y_j^{N,z,t_r-2\tilde{\epsilon}/\lambda_1,t_{r+1}-2\tilde{\epsilon}/\lambda_1})\right|,\\ \phi_j^r &:= \inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t)),\\ S^N &:= \sum_{j,\,\mu_j^i>0}\mu_j^i\sum_{r=0}^{M-1}(t_{r+1}-t_r)\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t)).\end{aligned}$$

We compute (for N large enough as before)

$$\begin{aligned}&\frac{\tilde{\mathbb{P}}[\tilde{F}^N]}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu_j^i>0}\sum_{r=0}^{M-1}\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t))\tilde{\mathbb{E}}\big[Y_j^{N,x^N,t_r,t_{r+1}}\big]\\ &\geq \alpha_1^N\sum_{j,\,\mu_j^i>0}\sum_{r=0}^{M-1}\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t))\cdot\Big\{\mathbb{1}_{\{\phi_j^r<0\}}\mu_j(t_{r+1}-t_r)\\ &\quad+\mathbb{1}_{\{\phi_j^r>0\}}\big(\alpha_2^N\mu_j(t_{r+1}-t_r)-\alpha_3^N\big)\Big\}\\ &\geq \alpha_1^N\sum_{j,\,\mu_j^i>0}\mu_j^i\sum_{r=0}^{M-1}(t_{r+1}-t_r)\inf_{t\in[t_r,t_{r+1})}\log\beta_j(\phi^x(t))\Big\{\mathbb{1}_{\{\phi_j^r<0\}}+\alpha_2^N\mathbb{1}_{\{\phi_j^r>0\}}\Big\}\\ &\quad-2\eta\alpha_1^N k|\log\bar{\beta}|\epsilon_N^{-1}\alpha_3^N\\ &\geq \alpha_1^N S^N-2\eta\alpha_1^N k|\log\bar{\beta}|\epsilon_N^{-1}\alpha_3^N-\alpha_1^N k|\log\bar{\beta}|\bar{\mu}\eta(1-\alpha_2^N).\end{aligned}$$

We readily observe that

$$\alpha_1^N, \alpha_2^N \to 1 \quad \text{as } N \to \infty$$

by Theorem 7.3 and Assumption 7.2 (C). We furthermore note that by Assumption 7.2 (C) and Theorem 7.3 (cf. also the comment corresponding to (7.7) and again the fact that the rate of convergence in Theorem 7.3 is independent of initial values),

$$\frac{\inf_z \sqrt{\tilde{\mathbb{P}}[F_2^{N,z}] - \tilde{\mathbb{P}}[F_2^{N,z}]^2}}{\sqrt{\epsilon_N}} \le \frac{\inf_z \sqrt{\tilde{\mathbb{P}}[(F_2^{N,z})^c]}}{\sqrt{\epsilon_N}} \to 0 \quad \text{as } N \to \infty.$$

Therefore (for N sufficiently large as before),

$$\begin{aligned} 2\eta\alpha_1^N k|\log\bar{\beta}|\epsilon_N^{-1}\alpha_3^N &\le 2\eta\alpha_1^N k|\log\bar{\beta}|\tilde{\mathbb{P}}[F_1^N]\sqrt{\bar{\mu}}\frac{1}{\sqrt{\epsilon_N}} \\ &\quad \inf_{z;|z-\phi^x(\tilde{\epsilon}/\lambda_1)|<\tilde{\epsilon}/2} \sqrt{\tilde{\mathbb{P}}[F_2^{N,z}] - \tilde{\mathbb{P}}[F_2^{N,z}]^2} \\ &\to 0 \end{aligned}$$

as $N \to \infty$. Finally, S^N is a Riemann sum and we have

$$S^N \to \sum_j \mu_j^i \int_{2\tilde{\epsilon}/\lambda_1}^{\eta} \log\beta_j(\phi^x(t))dt \quad \text{as } N \to \infty. \tag{7.62}$$

We observe that (7.61)–(7.62) yield

$$\begin{aligned} &\liminf_{N\to\infty} \frac{1}{\tilde{\mathbb{P}}[F^N]} \sum_{j,\, \mu_j^i>0} \tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N} \sum_{\tau_j\in[2\tilde{\epsilon}/\lambda_1,\eta]} \log\beta_j(Z^{N,x^N}(\tau_j-))\Big] \\ &\ge \sum_{j=1}^k \mu_j^i \int_{2\tilde{\epsilon}/\lambda_1}^{\eta} \log\beta_j(\phi^x(t))dt. \end{aligned} \tag{7.63}$$

We now consider the second term in the right and side of (7.59). We define

$$\tilde{M} = \tilde{M}(N) := \lfloor 2\tilde{\epsilon}\epsilon_N^{-1} + 1 \rfloor$$

and divide the interval $[0, 2\tilde{\epsilon}/\lambda_1]$ into $\tilde{M}$ subintervals $[\tilde{t}_r, \tilde{t}_{r+1}]$ ($r = 0, \dots, \tilde{M}-1$, $\tilde{t}_r = \tilde{t}_r(N)$) of length $\tilde{\Delta} = \tilde{\Delta}(N)$, i.e., for N large enough,

$$\frac{\epsilon_N}{2\lambda_1} \le \tilde{\Delta} < \frac{\epsilon_N}{\lambda_1}.$$

For $r = 0, \ldots, \tilde{M} - 1$ and $\tau_j \in [\tilde{t}_r, \tilde{t}_{r+1}]$, we obtain on $\tilde{F}^N$,

$$\begin{aligned}\operatorname{dist}(Z^{N,x^N}(\tau_j-), \partial A) &> \operatorname{dist}(\phi^x(\tau_j-), \partial A) - 2\epsilon_N \\ &\geq \operatorname{dist}(\phi^x(\tilde{t}_r), \partial A) - 2\epsilon_N \\ &\geq \lambda_1 \tilde{t}_r - 2\epsilon_N \\ &\geq \frac{r-4}{2}\epsilon_N.\end{aligned}$$

Hence, $\operatorname{dist}(Z^{N,x^N}(\tau_j-), \partial A) > \epsilon_N$ for $r \geq 6$. We compute for j with $\mu_j > 0$,

$$\begin{aligned}&\frac{1}{\tilde{\mathbb{P}}[F^N]}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j \in [0, 2\tilde{\epsilon}/\lambda_1]} \log \beta_j(Z^{N,x^N}(\tau_j-))\Big] \\ &= \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=0}^{\tilde{M}-1}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j \in [\tilde{t}_r, \tilde{t}_{r+1}]} \log \beta_j(Z^{N,x^N}(\tau_j-))\Big] \\ &= \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=6}^{\tilde{M}-1}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j \in [\tilde{t}_r, \tilde{t}_{r+1}]} \log \beta_j(Z^{N,x^N}(\tau_j-))\Big] \\ &\quad + \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=0}^{5}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j \in [\tilde{t}_r, \tilde{t}_{r+1}]} \log \beta_j(Z^{N,x^N}(\tau_j-))\Big].\end{aligned} \tag{7.64}$$

We note that for all j, $\beta_j(Z^{N,x^N}(\tau_j-)) \geq \underline{\beta}(\lambda_0/N)$ $\tilde{\mathbb{P}}$-a.s. by Assumption 7.2 (A1). The second term in the right hand side of (7.64) can be bounded from below (w.l.o.g. $\underline{\beta}(\lambda_0/N) < 1$):

$$\begin{aligned}&\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=0}^{5}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j \in [\tilde{t}_r, \tilde{t}_{r+1}]} \log \beta_j(Z^{N,x^N}(\tau_j-))\Big] \\ &\geq \frac{4}{\tilde{\mathbb{P}}[F^N]}\log \underline{\beta}\Big(\frac{\lambda_0}{N}\Big)\sum_{r=0}^{5}\Big\{\tilde{\mathbb{P}}[\tilde{F}^N]\tilde{\mathbb{E}}[Y_j^{N,x^N,\tilde{t}_r,\tilde{t}_{r+1}}] + \widetilde{\operatorname{Cov}}(\mathbb{1}_{\tilde{F}^N}, Y_j^{N,z,\tilde{t}_r,\tilde{t}_{r+1}})\Big\} \\ &\geq \frac{6\tilde{\mathbb{P}}[\tilde{F}^N]}{\tilde{\mathbb{P}}[F^N]}\log \underline{\beta}\Big(\frac{\lambda_0}{N}\Big)\bar{\mu}\frac{\epsilon_N}{\lambda_1} + \frac{4}{\tilde{\mathbb{P}}[F^N]}\log \underline{\beta}\Big(\frac{\lambda_0}{N}\Big)\sqrt{\frac{\bar{\mu}\,\epsilon_N}{\lambda_1}}\sqrt{\delta(N, \epsilon_N)} \\ &\to 0\end{aligned}$$

as $N \to \infty$ by Assumption 7.2 (C) (cf. also (7.6)). For the first term in Eq. (7.64), we compute for j with $\mu_j > 0$ (similarly as before, we assume w.l.o.g. that $\underline{\beta}(\tilde{\epsilon}) < 1$

and note that $\beta_j(Z^{N,x^N}(\tau_j-)) \geq \underline{\beta}(\lambda_1\tilde{t}_r - \epsilon_N) \geq \underline{\beta}((r-4)\lambda_1\tilde{\Delta}/2))$

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=6}^{\tilde{M}-1}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{\tilde{F}^N}\sum_{\tau_j\in[\tilde{t}_r,\tilde{t}_{r+1}]}\log\beta_j(Z^{N,x^N}(\tau_j-))\Big]\\
&\geq \frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=2}^{\tilde{M}-5}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2)\tilde{\mathbb{E}}[\mathbb{1}_{\tilde{F}^N}Y^{N,x^N,\tilde{t}_r,\tilde{t}_{r+1}}]\\
&= \frac{\tilde{\mathbb{P}}[\tilde{F}^N]}{\tilde{\mathbb{P}}[F^N]}\mu^i_j\sum_{r=2}^{\tilde{M}-5}\tilde{\Delta}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2)\\
&\quad+\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=2}^{\tilde{M}-5}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2)\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^N}, Y^{N,x^N,\tilde{t}_r,\tilde{t}_{r+1}}).
\end{aligned}\tag{7.65}$$

For the first term in Eq. (7.65), we have by Assumption 7.2 (C) (in particular by the fact that the integral below converges, cf. (7.9))

$$\frac{\tilde{\mathbb{P}}[\tilde{F}^N]}{\tilde{\mathbb{P}}[F^N]}\mu^i_j\sum_{r=2}^{\tilde{M}-5}\tilde{\Delta}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2) \to \mu^i_j\int_0^{2\tilde{\epsilon}/\lambda_1}\log\underline{\beta}(\lambda_1\rho/2)d\rho$$

as $N \to \infty$. Similarly, we obtain for the second term in Eq. (7.65),

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{r=2}^{\tilde{M}-5}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2)\widetilde{\mathrm{Cov}}(\mathbb{1}_{\tilde{F}^N}, Y^{N,x^N,\tilde{t}_r,\tilde{t}_{r+1}})\\
&\geq \frac{1}{\tilde{\mathbb{P}}[F^N]}\sqrt{\mu^i_j\frac{\delta(N,\epsilon_N)}{\tilde{\Delta}}}\sum_{r=2}^{\tilde{M}-5}\tilde{\Delta}\log\underline{\beta}(r\lambda_1\tilde{\Delta}/2)\\
&\to 0
\end{aligned}\tag{7.66}$$

as $N \to 0$ by Assumption 7.2 (C) (cf. (7.5) and (7.9)).

Finally, we consider the third term in Eq. (7.59). We obtain by Assumption (7.2) (C),

$$\begin{aligned}
&\frac{1}{\tilde{\mathbb{P}}[F^N]}\sum_{j,\,\mu^i_j>0}\tilde{\mathbb{E}}\Big[\frac{1}{N}\mathbb{1}_{F^N\setminus\tilde{F}^N}\sum_{\tau_j\in[0,\eta]}\log\beta_j(Z^{N,z}(\tau_j-))\Big]\\
&\geq \frac{1}{\tilde{\mathbb{P}}[F^N]}\log\underline{\beta}\Big(\frac{\lambda_0}{N}\Big)\sum_{j;\,\mu^i_j>0}\big\{\tilde{\mathbb{P}}[(\tilde{F}^N)^c]\tilde{\mathbb{E}}[Y_j^{N,z,0,\eta}]+\widetilde{\mathrm{Cov}}(\mathbb{1}_{(\tilde{F}^N)^c}, Y_j^{N,z,0,\eta})\big\}
\end{aligned}$$

$$\geq \frac{1}{\tilde{\mathbb{P}}[F^N]} \log \underline{\beta}\Big(\frac{\lambda_0}{N}\Big) k\, \bar{\mu}\, \eta\, \delta(N, \epsilon_N) + \frac{1}{\tilde{\mathbb{P}}[F^N]} \log \underline{\beta}\Big(\frac{\lambda_0}{N}\Big) k \sqrt{\bar{\mu}\, \eta\, \delta(N, \epsilon_N)}$$
$$\to 0 \tag{7.67}$$

as $N \to \infty$ similarly as before (cf. (7.5) and (7.6)).

We obtain by Eqs. (7.59) and (7.63), (7.64)–(7.67),

$$\liminf_{N\to\infty} \frac{1}{\tilde{\mathbb{P}}[F^N]} \tilde{\mathbb{E}}\Big[\frac{1}{N} \mathbb{1}_{F^N} \sum_{\tau \leq \eta} \log \beta_{j(\tau)}(Z^{N,x^N}(\tau-))\Big]$$
$$\geq \sum_{j=1}^{k} \mu_j^i \int_{2\tilde{\epsilon}/\lambda_1}^{\eta} |\log \beta_j(\phi^x(t))| dt - k\bar{\mu} \int_0^{2\tilde{\epsilon}/\lambda_1} |\log \underline{\beta}(\lambda_1 \rho/2)| d\rho. \tag{7.68}$$

We conclude by letting $\delta > 0$ and choosing $\tilde{\epsilon} < \epsilon_1$ small enough such that (cf. Eqs. (7.50), (7.51) and Inequality (7.68); note that we require the convergence of the integral in (7.9) of Assumption 7.2 (C) here)

$$\frac{\tilde{\epsilon}}{\lambda_1} k\bar{\mu},\ k\bar{\mu} \int_0^{2\tilde{\epsilon}/\lambda_1} |\log \beta_j(\phi^x(t))| dt,\, k\bar{\mu} \int_0^{2\tilde{\epsilon}/\lambda_1} |\log \underline{\beta}(\lambda_1 \rho/2)| d\rho < \frac{\delta}{4}.$$

The assertion now follows from Inequality (7.49) and (7.50)–(7.52), (7.58) and (7.68):

$$\liminf_{N\to\infty} \tilde{\mathbb{E}}_{F^N}\Big[\frac{X_\eta}{N}\Big] \geq -\int_0^{\eta} \ell(\phi^x(t), \mu) dt - \delta.$$

The uniformity of the convergence follows from the fact that we have used only Assumption 7.2 (C) and Theorem 7.3, where the convergences are uniform in x.

Again in the following result, the exponent α is the one from Assumption 7.2 (C).

Theorem 7.8 *Assume that Assumption 7.2 holds. Let $x \in A$ and $x^N \in A^N$ such that*

$$\limsup_{N\to\infty} |x^N - x| N^{\alpha} < 1.$$

Then, for $\phi \in D([0, T]; A)$ and $\epsilon > 0$,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x^N}(t) - \phi(t)| < \epsilon\Big] \geq -I_{T,x}(\phi).$$

Moreover the above convergence is uniform in $x \in A$

Proof We can w.l.o.g. assume that $I_{T,x}(\phi) \le K < \infty$. Let $\delta > 0$ and divide the interval $[0, T]$ into J subintervals as before. We define the function ϕ^η as before and choose η_1 small enough such that for all $\eta < \eta_1$ (cf. Lemma 7.24),

$$\int_\eta^T L(\phi^\eta(t), (\phi^\eta)'(t))dt < \int_0^T L(\phi(t), \phi'(t))dt + \frac{\delta}{3}. \tag{7.69}$$

We furthermore assume that $\eta < \eta_1$ is such that

$$\sup_{t\in[0,T]} |\phi(t) - \phi^\eta(t)| < \frac{\epsilon}{4}.$$

Hence,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x^N}(t) - \phi(t)| < \epsilon \Big]$$
$$\ge \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x^N}(t) - \tilde{\phi}^\eta(t)| < \frac{\epsilon}{2} \Big].$$

From (7.38), for $t \ge \eta$, $\mathrm{dist}(\phi^\eta(t), \partial A) \ge \eta\lambda_1$. We define

$$\epsilon_1 = \epsilon_1(\eta) = \frac{\epsilon}{2} \wedge \frac{\lambda_1\eta}{4},$$

$$\underline{\beta}^\eta := \inf\Big\{\beta_j(z) \Big| \ 1 \le j \le k, \ z \in A, \mathrm{dist}(z, \partial A) \ge \frac{\eta\lambda_1}{2}\Big\} > 0$$

and

$$\tilde{\beta}_j^\eta(z) := \begin{cases} \beta_j(z) \vee \underline{\beta}^\eta & \text{if } z \in A \\ \tilde{\beta}_j^\eta(\psi_A(z)) & \text{else,} \end{cases}$$

where the function ψ_A has been specified in Assumption (A4). We denote by $\tilde{Z}^{N,z,\eta}$ the process starting at z at time η with rates $\tilde{\beta}_j^\eta$. As the $\log \tilde{\beta}_j^\eta$ are bounded, we have by Theorem 7.6 that there exists an

$$\epsilon_2 = \epsilon_2(\eta) < \epsilon_1(\eta)$$

such that for all $\tilde{\epsilon} < \epsilon_2$,

$$\liminf_{N\to\infty} \frac{1}{N} \log\Big(\inf_{|z-\phi^\eta(\eta)|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[\eta,T]} |\tilde{Z}^{N,z,\eta}(t) - \tilde{\phi}^\eta(t)| < \epsilon_1 \Big]\Big)$$
$$\geq -\int_\eta^T \tilde{L}^\eta(\phi^\eta(t), (\phi^\eta)'(t))dt - \frac{\delta}{3},$$

where $\tilde{L}^\eta$ denotes the Legendre transform corresponding to the rates $\tilde{\beta}_j^\eta$. We readily observe that for all $t \in [\eta, T]$,

$$\tilde{L}^\eta(\phi^\eta(t), \phi^\eta(t)) = L(\phi^\eta(t), (\phi^\eta)'(t))$$

and that for $|z - \phi^\eta(\eta)| < \tilde{\epsilon}$, denoting by an abuse of notation $Z^{N,z}$ the process starting from z at time η,

$$\sup_{t\in[\eta,T]} |Z^{N,z}(t) - \phi^\eta(t)| < \epsilon_1 \Leftrightarrow \sup_{t\in[\eta,T]} |\tilde{Z}^{N,z,\eta}(t) - \phi^\eta(t)| < \epsilon_1.$$

and hence

$$\mathbb{P}\Big[\sup_{t\in[\eta,T]} |Z^{N,z}(t) - \tilde{\phi}^\eta(t)| < \frac{\epsilon}{2} \Big] \geq \mathbb{P}\Big[\sup_{t\in[\eta,T]} |Z^{N,z}(t) - \tilde{\phi}^\eta(t)| < \epsilon_1 \Big]$$
$$= \mathbb{P}\Big[\sup_{t\in[\eta,T]} |\tilde{Z}^{N,z,\eta}(t) - \tilde{\phi}^\eta(t)| < \epsilon_1 \Big]$$

consequently for $\tilde{\epsilon} < \epsilon_2$

$$\liminf_{N\to\infty} \frac{1}{N} \log\Big(\inf_{|z-\phi^\eta(\eta)|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[\eta,T]} |Z^{N,z}(t) - \tilde{\phi}^\eta(t)| < \epsilon_1 \Big]\Big)$$
$$\geq -\int_\eta^T L(\phi^\eta(t), (\phi^\eta)'(t))dt - \frac{\delta}{3}$$
$$\geq -\int_0^T L(\phi(t), \phi'(t))dt - \frac{2\delta}{3},$$

where we have used (7.69) for the second inequality. We use the Markov property of Z^N and obtain for $\tilde{\epsilon} < \epsilon_2$

$$\mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x^N}(t) - \phi(t)| < \epsilon \Big] \geq \mathbb{P}\Big[\sup_{t\in[0,\eta]} |Z^{N,x^N}(t) - \tilde{\phi}^\eta(t)| < \tilde{\epsilon} \Big]$$
$$\cdot \inf_{|z-\phi^\eta(\eta)|<\tilde{\epsilon}} \mathbb{P}\Big[\sup_{t\in[\eta,T]} |Z^{N,z}(t) - \tilde{\phi}^\eta(t)| < \epsilon_1 \Big].$$

Combining the last two inequalities with Lemma 7.25, we deduce that (i being the index of the ball B_i to which the starting point x belongs)

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big[\sup_{t\in[0,T]}|Z^{N,x^N}(t)-\phi(t)|<\epsilon\Big]$$
$$\geq -\int_0^\eta \ell(\phi^x(t),\mu^i)dt-\int_0^T L(\phi(t),\phi'(t))dt-\frac{2\delta}{3}$$
$$\geq -\int_0^T L(\phi(t),\phi'(t))dt-\delta$$

thanks to Lemma 7.26 below, provided η is small enough. The result follows since $\delta>0$ is arbitrary.

Lemma 7.26 *Let $x\in B_i$, where $i\leq I_1$, and suppose $\phi^x(t)=x+tv_i$. Let moreover μ^i be such that $\sum_{j=1}^k \mu^i_j h_j=v_i$. Then, uniformly in x, as $t\to 0$,*

$$\int_0^t \ell(\phi^x(s),\mu^i)ds\to 0.$$

Proof Since according to Assumption (A3) $d(\phi^x(t),\partial A)\geq\lambda_1 t$, the result follows from (7.6) from Assumption (C).

Theorem 7.9 *Assume that Assumption 7.2 as well as the assumptions from Theorem 7.8 hold. Then for any open set $G\subset D([0,T];A)$,*

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big[Z^{N,x^N}\in G\big]\geq-\inf_{\phi\in G}I_{T,x}(\phi).$$

Moreover the convergence is uniform in x.

Proof The proof follows the same line of reasoning as the proof of Theorem 7.7.

We will need the following stronger version. Recall the definition of A^N at the start of Sect. 7.2.

Theorem 7.10 *Assume that Assumption 7.2 holds. Then for any open set $G\subset D([0,T];A)$ and any compact subset $K\subset A$,*

$$\liminf_{N\to\infty}\frac{1}{N}\log\inf_{x\in K\cap A^N}\mathbb{P}\big[Z^{N,x}\in G\big]\geq-\sup_{x\in K}\inf_{\phi\in G}I_{T,x}(\phi).$$

Proof This follows readily from the uniformity in x of the convergence in Theorem 7.9.

7.6 LDP Upper Bound

We now prove the LDP upper bound. For reasons of readability, we split up the proof into four parts. In the first three parts, we prove the main auxiliary results required (Sects. 7.6.1–7.6.3). Finally, we prove the main results of the section in Sect. 7.6.4.

In this section, whenever we consider the process $Z^{N,x}$, we will mean that the process Z^N is started from the nearest point to x on the grid A^N (see the beginning of Sect. 7.2 for the definition of A^N).

7.6.1 Piecewise Linear Approximation

The goal of this section is to prove that $Z^{N,x}$ is exponentially close to its piecewise linear approximation. For $Z^{N,x}$, we define the piecewise linear interpolation $Y^{N,x}$. To this end, we divide $[0, T]$ into N subintervals $[t_{j-1}, t_j]$ with $t_j = \frac{jT}{N}$, $j = 1, \ldots, N$. We define $t \in [t_{j-1}, t_j]$

$$Y_t^{N,x} = Z^{N,x}(t_{j-1}) + \frac{t - t_{j-1}}{t_j - t_{j-1}}(Z^{N,x}(t_j) - Z^{N,x}(t_{j-1})). \tag{7.70}$$

We prove that $Y^{N,x}$ is exponentially close to $Z^{N,x}$.

Lemma 7.27 *Assume that β_j $(j = 1, \ldots, k)$ is bounded. Let $\delta > 0$. Then*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[\mathrm{d}(Y^{N,x}, Z^{N,x}) > \delta] = -\infty$$

uniformly in $x \in A$.

Proof For any $1 \le j \le [N/T]$, we have the inclusion

$$\{\sup_{t\in[t_{j-1},t_j]} |Y_t^{N,x} - Z_t^{N,x}| \ge \delta\} \subset \{\sup_{t\in[t_{j-1},t_j]} |Z_t^{N,x} - Z_{t_{j-1}}^{N,x}| \ge \delta/2\}.$$

It then follows from Lemma 7.6 that for some positive constant C and for each j,

$$\mathbb{P}(\sup_{t\in[t_{j-1},t_j]} |Y_t^{N,x} - Z_t^{N,x}| \ge \delta) \le \exp\left(-CN\delta \log(CN\delta)\right).$$

Consequently

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,T]} |Y_t^{N,x} - Z_t^{N,x}| \ge \delta\right) &= \mathbb{P}\left(\bigcup_{j=1}^{[N/T]} \left\{\sup_{t\in[t_{j-1},t_j]} |Y_t^{N,x} - Z_t^{N,x}| \ge \delta\right\}\right) \\ &\le N \exp\left(-CN\delta \log(CN\delta)\right).\end{aligned}$$

The result clearly follows.

7.6.2 The Modified Rate Function I^δ

In this section, we define a modified rate function I^δ and analyse how it relates to I. The main result is Corollary 7.8 below.

We define the following functional (Lemma 7.19 above). For $\delta > 0$, $x \in A$, $y, \theta \in \mathbb{R}^d$, let

$$\tilde{\ell}_\delta(x, y, \theta) := \langle\theta, y\rangle - \sum_{j=1}^{k} \sup_{z=z^j \in A; |z-x|<\delta} \beta_j(z)\big(\exp(\langle\theta, h_j\rangle) - 1\big),$$

$$L_\delta(x, y) := \sup_{\theta\in\mathbb{R}^d} \tilde{\ell}_\delta(x, y, \theta).$$

Obviously, we have

$$L_\delta(x, y) \le L(x, y)$$

and for the respectively defined functional, I^δ,

$$I^\delta \le I.$$

We obtain

$$\begin{aligned}
L_\delta(x, y) &= \sup_{\theta\in\mathbb{R}^d} \Big\{\langle\theta, y\rangle - \sum_{j=1}^{k} \sup_{z^j \in A; |z^j-x|<\delta} \beta_j(z^j)\big(\exp(\langle\theta, h_j\rangle) - 1\big)\Big\} \\
&= \sup_{\theta\in\mathbb{R}^d} \inf_{z^1,\ldots,z^k \in A; |z^j-x|<\delta} \Big\{\langle\theta, y\rangle - \sum_{j=1}^{k} \beta_j(z^j)\big(\exp(\langle\theta, h_j\rangle) - 1\big)\Big\} \\
&= \inf_{z^1,\ldots,z^k \in A; |z^j-x|<\delta} \sup_{\theta\in\mathbb{R}^d} \Big\{\langle\theta, y\rangle - \sum_{j=1}^{k} \beta_j(z^j)\big(\exp(\langle\theta, h_j\rangle) - 1\big)\Big\} \qquad (7.71) \\
&= \inf_{z^1,\ldots,z^k, |z^j-x|<\delta} \inf_{\mu\in\tilde{V}_{z_j,y}} \sum_{j=1}^{k} \big(\beta_j(z^j) - \mu_j + \mu_j \log\mu_j - \mu_j \log\beta_j(z^j)\big) \qquad (7.72) \\
&= \ell(z^*, \mu^*), \qquad (7.73)
\end{aligned}$$

where we use the slight abuse of notation: for $z = (z^1, \ldots, z^k)$, $\mu = (\mu_1, \ldots, \mu_k)$,

$$\ell(z, \mu) = \sum_j \beta_j(z^j) - \mu_j + \mu_j \log\Big(\frac{\mu_j}{\beta_j(z^j)}\Big)$$

Note that $\ell(x, \mu)$ depends on x only through the rates $\beta(x)$.

Here, Eq. (7.71) follows from Sion's min-max theorem,[9] see [15]. Indeed, since the variables of the β_j's are decoupled, we can as well decide that we are taking the inf-sup of

$$F(\beta_1, \ldots, \beta_k; \theta) = \langle \theta, y \rangle - \sum_{j=1}^{k} \beta_j (e^{\langle \theta, h_j \rangle} - 1).$$

F is defined on $X \times Y$, where $X = \prod_{j=1}^{k} \{\beta_j(z),\ |z - x| \le \delta\}$ and $Y = \mathbb{R}^d$. Now X is convex and compact, Y is convex, F is continuous, is a linear (hence convex) function of $(\beta_1, \ldots, \beta_k)$ for each θ, and a concave function of θ for each $(\beta_1, \ldots, \beta_k)$. Hence Sion's theorem applies. Equation (7.72) follows by Theorem 7.4. Equation (7.73) follows from Lemma 7.10 and the continuity of ℓ and μ^* (as a function in the state). We remark that $|z_j^* - x| = \delta$ is possible.

In a similar way as before (cf. Proposition 7.1), we define the sets

$$\Phi^\delta(K) := \{\phi \in D([0, T]; A) | I^\delta(\phi) \le K\},$$

$$\Phi_x^\delta(K) := \{\phi \in D([0, T]; A) | I_x^\delta(\phi) \le K\}.$$

In particular, we have $\Phi(K) \subset \Phi^\delta(K)$ and $\Phi_x(K) \subset \Phi_x^\delta(K)$ and $\Phi_x^\delta(K)$, $\Phi^\delta(K)$ are increasing in δ.

For technical reasons, we define for $m > 0$, $z \in A$ the rates

$$\beta_j^m(z) := \max\{\beta_j(z), 1/m\}$$

and the corresponding functionals L^m and I^m by replacing the rates β_j by the rates β_j^m in the respective definitions.

We will need the following slightly stronger version of Lemma 7.24, where again ϕ^η is defined from ϕ as in the lines before Lemma 7.24.

Lemma 7.28 *Assume that Assumption 7.2 holds. Let $K > 0$ and $\epsilon > 0$. Then there exists an $\eta_0 = \eta_0(T, K, \epsilon) > 0$ such that for all $\eta < \eta_0$ there exists an $m_0 > 0$ such that for all $m > m_0$ and for all $\phi \in D([0, T]; A)$ with $I_T^m(\phi) \le K$,*

$$I_T(\phi^\eta) < K + \epsilon,$$

where ϕ^η is defined before Lemma 7.24 and satisfies $\|\phi^\eta - \phi\| \le \epsilon$.

[9] This theorem says that if $F : X \times Y \to \mathbb{R}$, where X and Y are convex, one the two being compact, F being quasi-concave and u.s.c. with respect to its first variable, quasi-convex and l.s.c. with respect to the second, then $\sup_x \inf_y F(x, y) = \inf_y \sup_x F(x, y)$.

Proof We follows the first steps of the proof of Lemma 7.24, where we replace $\mu^*(t)$ by $\mu^{m,*}(t)$ the optimal μ corresponding to $(\phi(t), \phi'(t))$ and jump rates β_j^m. Now (7.39) is replaced by

$$L(\phi^\eta(t), \phi'(t-\eta_r)) \le \ell(\phi^\eta(t), \mu^{m,*}(t-\eta_r)). \tag{7.74}$$

We now choose $m > 1/\eta$ and deduce

$$|\beta_j(\phi^\eta(t)) - \beta_j^m(\phi(t-\eta_r))| \le \frac{1}{m} + |\beta_j(\phi^\eta(t)) - \beta_j(\phi(t-\eta_r))| \le \delta_K'(\eta), \tag{7.75}$$

where $\delta_K'(\eta) = \eta + \delta_K(\eta) \to 0$ as $\eta \to 0$ by the (uniform) continuity of the β_j. We deduce from (7.74) the following modified version of (7.41)

$$\begin{aligned} L(\phi^\eta(t), \phi'(t-\eta_r)) &- L^m(\phi(t-\eta_r), \phi'(t-\eta_r)) \\ &\le k\delta_K'(\eta) + \sum_j \mu^{m,*}(t-\eta_r) \log \frac{\beta_j^m(\phi(t-\eta_r))}{\beta_j(\phi^\eta(t))}, \end{aligned} \tag{7.76}$$

since $\beta_j^m(\phi(t)) > 0$ and $\beta_j(\phi^\eta(t)) > 0$ for $t \neq 0$.

We recall that λ_1 and the v_i's have been defined in Assumption 7.2 (A3), and that the $\tilde{v}_i$'s, $\bar{v}_i$'s and $\hat{v}_i$'s have been defined in the proof of Lemma 7.24.

By Assumption 7.2 (B4), there exists a constant $\lambda_4 > 0$ such that for $z \in B_{i_r}$ (and $\eta < \eta_2 \le \eta_1$ small enough, depending upon λ_1 and λ_2 but not on ϕ, except through K),

$$\beta_j(z) < \lambda_4 \Rightarrow \beta_j(z + \eta\tilde{v}_{i_r}) \ge \beta_j(z). \tag{7.77}$$

We now want to bound from above the second term in the right hand side of (7.76). If $\beta_j(\phi(t-\eta_r)) \ge \lambda_4$, then $\beta_j(\phi(t-\eta_r)) \ge 1/m$ and therefore $\beta_j^m(\phi(t-\eta_r)) = \beta_j(\phi(t-\eta_r))$, so that the bound (7.43) holds.

Now consider the case $\beta_j(\phi(t-\eta_r)) < \lambda_4$. We define the function

$$s(\delta) := \inf\{\beta_j(x) \mid \mathrm{d}(x, \partial A) \ge \delta\}.$$

From the continuity of the β_j and the compactness of A, $s(\delta) > 0$ for $\delta > 0$, and for any $x \in A$, $\beta_j(x) \ge s(\mathrm{d}(x, \partial A))$.

We furthermore let

$$m_0 = m_0(\eta, \lambda_4) > \max\{1/\lambda_4, 1/s(\lambda_1\eta)\};$$

and recall that $\mathrm{d}(\phi^\eta(t), \partial A) \ge \lambda_1\eta$ for $t \ge \eta$ (cf. the discussion preceding Lemma 7.24).

We let $m > m_0$. Since $\beta_j(\phi(t-\eta_r)) < \lambda_4$, by (7.77),

$$\beta_j(\phi^\eta(t)) \geq \beta_j(\phi(t-\eta_r)).$$

By the definition of s, we have furthermore

$$\beta_j(\phi^\eta(t))) \geq s(\lambda_1\eta) \geq 1/m.$$

Combining these observations, we obtain

$$\beta_j(\phi^\eta(t)) \geq \max\{\beta_j(\phi(t-\eta_r)), 1/m\} = \beta_j^m(\phi(t-\eta_r))$$

and therefore

$$\log \frac{\beta_j^m(\phi(t-\eta_r))}{\beta_j(\phi^\eta(t))} \leq 0.$$

From Lemmas 7.8 and 7.10, there exist (universal, i.e., independent of x and m) constants $B \geq B_1 \vee B_2$, $B > 1$, C_1, C_3 such that whenever $|y| > B$, for all $x \in A$ and $m \geq 1$,

$$L(x, y) \geq C_1|y| \log |y|, \quad L^m(x, y) \geq C_1|y| \log |y|,$$

$$|\mu^{m,*}| = |\mu^*(x, y, m)| \leq C_3|y|.$$

Hence if $|\phi'(t-\eta_r)| \geq B$, we get, instead of (7.47),

$$\begin{aligned}
&L(\phi^\eta(t), \phi'(t-\eta_r)) - L^m(\phi(t-\eta_r), \phi'(t-\eta_r)) \\
&\leq k\delta'_K(\eta) + kC_3|\phi'(t-\eta_r)|\frac{2\delta_K(\eta)}{\lambda_4} \\
&\leq k\delta'_K(\eta) + kC_3\frac{2\delta_K(\eta)L^m(\phi(t-\eta_r), \phi'(t-\eta_r))}{C_1\lambda_4 \log|\phi'(t-\eta_r)|}. \qquad (7.78)
\end{aligned}$$

If $|\phi'(t-\eta_r)| < B$, Lemma 7.10 implies that $|\mu^{m,*}(t-\eta_r)| \leq \tilde{C}B$ for a universal constant $\tilde{C} > 0$. Using Eqs. (7.75) and (7.76), we obtain

$$L(\phi^\eta(t), \phi'(t-\eta_r)) - L^m(\phi(t-\eta_r), \phi'(t-\eta_r)) \leq k\delta'_K(\eta) + k\tilde{C}B\frac{2\delta_K(\eta)}{\lambda_4}. \qquad (7.79)$$

Inequalities (7.78) and (7.79) imply

$$\begin{aligned}
&L(\phi^\eta(t), \phi'(t-\eta_r)) - L^m(\phi(t-\eta_r), \phi'(t-\eta_r)) \\
&\quad \le k(\eta + \delta_K(\eta)) + k\tilde{C}B\frac{2\delta_K(\eta)}{\lambda_4} + kC_3\frac{2\delta_K(\eta)L^m(\phi(t-\eta_r), \phi'(t-\eta_r))}{C_1\lambda_4 \log|\phi'(t-\eta_r)|} \\
&\quad =: \delta_1(\eta) + \delta_2(\eta)L^m(\phi(t-\eta_r), \phi'(t-\eta_r))
\end{aligned}$$

where $\delta_i(\eta) \to 0$ as $\eta \to 0$, $i = 1, 2$. We now choose $\eta > 0$ such that

$$\delta_2(\eta)K < \frac{\epsilon}{4} \quad \text{and} \quad T\delta_1(\eta) < \frac{\epsilon}{4},$$

and choose $m > m_0(\eta)$; this yields

$$I_T(\phi^\eta) < K + \epsilon.$$

In the following, we show a relation between L^m and L_δ.

Remark 7.4 It can easily be seen that Lemma 7.8 holds for L_δ (with exactly the same proof). The same holds true for Lemmas 7.14 and 7.18.

Lemma 7.29 *Let β_j ($j = 1, \ldots, k$) be bounded and $\epsilon > 0$. Then there exists an $m_0 > 0$ such that for all $m > m_0$, there exists an $\delta_0 > 0$ such that for all $\delta < \delta_0$ and all $x \in A$ $y \in \mathbb{R}^d$,*

$$L^m(x, y) \le \epsilon + (1+\epsilon)L_\delta(x, y)$$

Proof Let $m_0 > 0$, $m > m_0$ and $\delta > 0$. We let $\mu^* = \mu^*(z^*, y)$ be the optimal μ associated to the optimal z^* according to Eq. (7.73). Then

$$L^m(x, y) - L_\delta(x, y) \le \ell^m(x, \mu^*) - \ell(z^*, \mu^*). \tag{7.80}$$

Furthermore, we have by the uniform continuity of the β_j (cf. the proof of Lemma 7.28),

$$|\beta_j^m(x) - \beta_j(z^*)| \le \frac{1}{m} + K(\delta) =: K_1(m, \delta). \tag{7.81}$$

Moreover, we note that if $\beta_j(x) < \frac{1}{m} - K(\delta)$, then

$$\log\frac{\beta_j(z^*)}{\beta_j^m(x)} < 0. \tag{7.82}$$

On the other hand, if $\beta_j(x) \geq \frac{1}{m} - K(\delta)$, then

$$\log \frac{\beta_j(z^*)}{\beta_j^m(x)} \leq \log \frac{\beta_j^m(z^*)}{\beta_j^m(x)} \leq \log \frac{\frac{1}{m} + K(\delta)}{\frac{1}{m}} \leq mK(\delta) =: K_2(m, \delta). \tag{7.83}$$

By Lemmas 7.8 and 7.9, there exist constants B, C_1, C_3 and C_4 such that for all $x \in A$ and $y \in \mathcal{C}_x$,

$$L(x, y), L_\delta(x, y) > C_1|y| \log B \quad \text{if } |y| > B, \tag{7.84}$$

$$|\mu^*(y)| \leq C_3|y| \quad \text{if } |y| > B, \tag{7.85}$$

$$|\mu^*(y)| \leq C_4 \quad \text{if } |y| \leq B \tag{7.86}$$

(note that the constants in Inequality (7.84) are independent of δ).

For $|y| > B$, we have by Inequalities (7.80)–(7.85),

$$\begin{aligned} L^m(x, y) - L_\delta(x, y) &\leq kK_1(m, \delta) + kC_3|y|K_2(m, \delta) \\ &\leq kK_1(m, \delta) + \frac{kC_3K_2(m, \delta)}{C_1 \log B} L_\delta(x, y). \end{aligned} \tag{7.87}$$

For $|y| \leq B$, we have by Inequalities (7.80)–(7.83) and (7.86),

$$L^m(x, y) - L_\delta(x, y) \leq kK_1(m, \delta) + kC_4K_2(m, \delta) \tag{7.88}$$

The assertion now follows from Inequalities (7.87) and (7.88) by choosing m_0 large enough, $m > m_0$ and $\delta_0 = \delta_0(m)$ small enough such that

$$kK_1(m, \delta_0), kC_4K_2(m, \delta_0), \frac{kC_3K_2(m, \delta_0)}{C_1 \log B} < \frac{\epsilon}{2}.$$

We directly deduce the following result

Corollary 7.7 *Let β_j ($j = 1, \ldots, k$) be bounded and continuous. For all $\epsilon, K, T > 0$, there exists an $m_0 > 0$ such that for all $m > m_0$, there exists a $\delta_0 > 0$ such that for all $\delta < \delta_0$ and all functions ϕ with $I_T^\delta(\phi) \leq K - \epsilon$,*

$$I_T^m(\phi) < K.$$

We now deduce from Lemma 7.28 and Corollary 7.7 .

Corollary 7.8 *Assume that Assumption 7.2 holds. Then for all $\epsilon, K > 0$, there exists a $\delta_0 > 0$ such that for all $\delta < \delta_0$,*

$$\Phi_x^\delta(K - \epsilon) \subset \left\{\phi \in D([0, T]; A) | \, \mathrm{d}(\phi, \Phi_x(K)) \leq \epsilon\right\}.$$

Proof Let $\epsilon > 0$ and choose m_0, m, δ_0, δ according to Corollary 7.7 for $\epsilon/2$. Let $\phi \in \Phi_x^\delta(K-\epsilon)$. Then by Corollary 7.7, $\phi \in \Phi_x^m(K-\epsilon/2)$. By Lemma 7.28, there exists a $\tilde{\phi}$ such that

$$\|\tilde{\phi} - \phi\| < \epsilon \quad \text{and} \quad I_{T,x}(\tilde{\phi}) \le K.$$

7.6.3 Distance of Y^N to Φ^δ

In this section, we derive a result about the distance of Y^N, defined by (7.70), to Φ^δ (Lemma 7.37 below).

We state the following elementary result (see, e.g., [12, Chapter 3, Proposition 22]).

Lemma 7.30 *Let $f : [a,b] \to \mathbb{R}^d$ be measurable. For all $\epsilon > 0$, there exists a step function g such that $|g - f| < \epsilon$ except on a set with measure less than ϵ. Moreover the range of g is a subset of the convex hull of the range of f.*

We define for $\delta > 0$, $\phi : [0,T] \to A$ and Borel-measurable $\theta : [0,T] \to \mathbb{R}^d$,

$$I_T^\delta(\phi, \theta) := \int_0^T \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt.$$

Lemma 7.31 *Let $\log \beta_j$ ($j = 1, \dots, k$) be bounded. For all absolutely continuous $\phi : [0,T] \to A$ with $I_T^\delta(\phi) < \infty$ and all $\epsilon > 0$ there exists a step function $\theta : [0,T] \to \mathbb{R}^d$ such that*

$$I_T^\delta(\phi, \theta) \ge I_T^\delta(\phi) - \epsilon.$$

Proof As $I^\delta(\phi) < \infty$, there exists a large enough positive number B such that

$$\int_0^T \mathbb{1}_{\{|\phi'(t)|>B\}} L_\delta(\phi'(t), \phi(t))dt \le \frac{\epsilon}{3}$$

(cf. Lemma 7.18 and Remark 7.4). We set

$$\theta_1(t) := 0 \quad \text{if } |\phi'(t)| > B.$$

By Lemma 7.14 (which holds true with L replaced by L_δ, see Remark 7.4—this is where we need the assumption that the $\log \beta_j$ are bounded), there exists a constant $\tilde{B}$ such that for all $x \in A$ and $y \in \mathcal{C}$ with $|y| \le B$,

$$\sup_{|\theta| \le \tilde{B}} \tilde{\ell}_\delta(x, y, \theta) > L_\delta(x, y) - \frac{\epsilon}{6T}.$$

We set

$$D := \{(x, y, \theta) | \; x \in A, y \in \mathcal{C}, |y| \le B, |\theta| \le \tilde{B}\}.$$

The function $\tilde{\ell}_\delta$ is uniformly continuous on D. Hence there exists an $\eta > 0$ such that for $|x - \tilde{x}|, |y - \tilde{y}|, |\theta - \tilde{\theta}| < \eta$,

$$|\tilde{\ell}_\delta(x, y, \theta) - \tilde{\ell}_\delta(\tilde{x}, \tilde{y}, \tilde{\theta})| < \frac{\epsilon}{6T}.$$

By a compactness argument, we obtain a finite cover $\{\theta_{i,j}, x_i, y_j\}$ of D such that

$$\tilde{\ell}_\delta(x_i, y_j, \theta_{i,j}) \ge L_\delta(x, y) - \frac{\epsilon}{3T} \quad \text{for } |x - x_i|, |y - y_j| < \eta.$$

We set

$$\theta_1(t) := \theta_{i,j} \quad \text{if } |\phi(t) - x_i|, |\phi'(t) - y_j| < \eta$$

(with some kind of tie-breaking rule). Hence θ_1 only takes finitely many values. However, it is not clear whether θ_1 is piecewise constant.

We now choose $\tilde{\eta}$ small enough such that $\text{Leb}[E] < \frac{\tilde{\eta}}{2}$ implies that

$$\int_{[0,T]\cap E} L_\delta(\phi(t), \phi'(t))dt \vee \int_{[0,T]\cap E} \sup_{|\theta|\le\tilde{B}} \Big(-\tilde{\ell}_\delta(\phi(t), \phi'(t), \theta)\Big)\, dt \;<\; \frac{\epsilon}{3}. \tag{7.89}$$

and

$$\min |\theta_{i,j} - \theta_{l,m}| > \tilde{\eta}, \quad \min |\theta_{i,j}| > \tilde{\eta}.$$

By Lemma 7.30, there exists a step function θ_2 with $|\theta_1 - \theta_2| < \frac{\tilde{\eta}}{2}$ except on a set $\tilde{E}$ with Lebesgue measure $< \frac{\tilde{\eta}}{2}$.

Hence there exists a step function θ which agrees with θ_1 except on $\tilde{E}$ (modify θ_2 if necessary such that $|\theta_1 - \theta_2| < \frac{\tilde{\eta}}{2} \Rightarrow \theta_2 = \theta_1$ on $\tilde{E}^c$). Note that $|\theta(t)| \le \tilde{B}$, for all $t \in [0, T]$.

We conclude by collecting the approximations above:

$$\begin{aligned} I_T^\delta(\phi) &= \int_{[0,T]} L_\delta(\phi(t), \phi'(t))dt \\ &\le \int_{[0,T]\cap\{|\phi'(t)|>B\}} L_\delta(\phi(t), \phi'(t))dt + \int_{[0,T]\cap\tilde{E}} L_\delta(\phi(t), \phi'(t))dt \\ &\quad + \int_{[0,T]\cap(\{|\phi'(t)|\le B\}\cup\tilde{E}^c)} L_\delta(\phi(t), \phi'(t))dt \end{aligned}$$

$$\leq \frac{2\epsilon}{3} + \int_{[0,T]\cap(\{|\phi'(t)\leq B\}\cup\tilde{E}^c)} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt$$

$$= \frac{2\epsilon}{3} + \int_{[0,T]} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt - \int_{[0,T]\cap\tilde{E}} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt$$

$$- \int_{[0,T]\cap\{|\phi'(t)|>B\}\cap\tilde{E}^c} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt$$

$$\leq \epsilon + \int_{[0,T]} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t))dt.$$

Indeed $\theta(t) = 0$ on the set $\{|\phi'(t)| > B\} \cap \tilde{E}^c$, while (7.89) implies that the second integral in the next to last line is bounded by $\epsilon/3$.

We next prove.

Lemma 7.32 *Let* $u : [0, T] \to \mathbb{R}$ *be nonnegative and absolutely continuous and* $\delta > 0$*. Then there exists an* $\eta > 0$*, a Borel set* $E \subset [0, T]$ *with* $Leb(E) < \delta$ *and two finite collections* $(J_i)_{i\in\mathcal{I}_+}$ *and* $(H_j)_{j\in\mathcal{I}_0}$ *of subintervals of* $[0, T]$ *such that*

$$[0, T] = E \cup \bigcup_{i\in\mathcal{I}_+} J_i \cup \bigcup_{j\in\mathcal{I}_0} H_j$$

and for all $i \in \mathcal{I}_+$, $j \in \mathcal{I}_0$,

$$\inf_{t\in J_i} u(t) > \eta, \qquad u(t) = 0 \text{ on } H_j \cap E^c.$$

Proof Given $t \in [0, T]$ such that $u(t) > 0$, let $\mathcal{O}_t$ be the largest open interval containing t such that $u(s) > 0$ for all $s \in \mathcal{O}_t$. Let $m_t = \max\{u(t),\ t \in \mathcal{O}_t\}$. Since u is absolutely continuous, there is a finite number of intervals $\mathcal{O}_t$ such that $m_t > 1/m$, for each $m \geq 1$. Hence there are at most countably many open intervals $\{\mathcal{O}_i,\ i \geq 1\}$ where u is positive. Choose M large enough such that

$$\text{Leb}\left(\cup_{i=M+1}^{\infty}\mathcal{O}_i\right) \leq \frac{\delta}{2}.$$

For $1 \leq i \leq M$, let $J_i \subset \mathcal{O}_i$ be a closed interval such that

$$\text{Leb}\,(\mathcal{O}_i \setminus J_i) \leq \frac{\delta}{2M}.$$

Let

$$E = \left(\cup_{i=M+1}^{\infty}\mathcal{O}_i\right) \cup \left(\cup_{i=1}^{M}\mathcal{O}_i \setminus J_i\right).$$

Clearly $\mathrm{Leb}(E) \le \delta$. Let $\overline{M}$ be the number of connected components of $[0, T] \setminus \cup_{i=1}^{M} J_i$. For $1 \le j \le \overline{M}$, let H_j denote the closure of the j-th connected component of $[0, T] \setminus \cup_{i=1}^{M} J_i$. H_j is an interval. Moreover

$$\inf_{1 \le i \le M} \inf_{t \in J_i} u(t) = \eta > 0, \quad \text{and}$$

$$u(t) = 0, \quad \text{if } t \in H_j \cap E^c.$$

We require this result for the proof of Lemma 4.6 of [14]. This is a (more general) variant of Lemma 5.43 of [13] (cf. also Lemma 7.31).

Lemma 7.33 *Assume that β_j ($j = 1, \ldots, k$) is bounded and Lipschitz continuous. Then for all ϕ with $I_T(\phi) < \infty$ and $\epsilon > 0$, there exists a step function θ such that*

$$I_T^{\delta}(\phi, \theta) \ge I_T^{\delta}(\phi) - \epsilon.$$

Proof If none of the $\beta_j(\phi(t))$ vanishes on the interval $[0, T]$, then the proof of Lemma 7.31 applies. If that is not the case, we note that since ϕ is absolutely continuous and β_j is Lipschitz continuous, $t \to \beta_j(\phi(t))$ is absolutely continuous. Hence we can apply Lemma 7.32 to the function $u(t) := \beta_j(\phi(t))$, and associate to each $1 \le j \le k$ intervals $(J_i^j)_{i \in \mathcal{I}_+}$ and $(H_i^j)_{i \in \mathcal{I}_0}$. It is not hard to see that to each $\eta > 0$ one can associate a real $\eta > 0$, an integer M, a collection $(I_i)_{1 \le i \le M}$ of subintervals of $[0, T]$, with the following properties

$$[0, T] = E \cup \bigcup_{1 \le i \le M} I_i,$$

with $\mathrm{Leb}(E) \le \delta$, and moreover to each $1 \le i \le N$ we can associate a subset $\mathcal{A} \subset \{1, 2, \ldots, k\}$ such that

$$\beta_j(\phi(t)) > \eta, \quad \text{if } j \in \mathcal{A}, t \in I_i, \quad \text{and } \beta_j(\phi(t)) = 0, \quad \text{if } j \notin \mathcal{A}, t \in I_i \cap E^c.$$

Each interval I_i is an intersection of J_i^j's for $j \in \mathcal{A}$ and of H_i^j's for $j \notin \mathcal{A}$.

On each subinterval I_i, by considering the process with rates and jump directions $\{\beta_j, h_j, j \in \mathcal{A}\}$ only, we can deduce from Lemma 7.31 that there exists a step function θ such that

$$\int_{I_i} \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t)) \ge \int_{I_i} L_\delta(\phi(t), \phi'(t))dt - \frac{\epsilon}{2M}.$$

In fact there exists a unique stepfunction θ defined on $[0, T]$, such that each of the above inequality holds and moreover, by the same argument as in the proof of

Lemma 7.31, provided η is small enough,

$$\int_E \tilde{\ell}_\delta(\phi(t), \phi'(t), \theta(t)) \geq -\frac{\epsilon}{2}.$$

The result follows.

We now define for $M \in \mathbb{N}$

$$\mathcal{K}(M) := \bigcap_{m \geq M} \{\phi \in C([0, T]; A) | V_{2^{-m}}(\phi) \leq \tfrac{1}{\log m}\},$$

where V_δ is the modulus of continuity:

$$V_\delta(\phi) := \sup_{s,t \in [0,T],\ |s-t|<\delta} |\phi(s) - \phi(t)|.$$

We readily observe (Arzelà-Ascoli) that $\mathcal{K}(M)$ is compact in $C([0, T]; \mathbb{R}^d)$.

We next obtain exponential tightness for the sequence $Y^{N,x}$ defined in (7.70).

Lemma 7.34 *Assume that β_j ($j = 1, \ldots, k$) is bounded. There exists a positive constant a such that for all M large enough and for all $x \in A$,*

$$\limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}[Y^{N,x} \notin \mathcal{K}(M)] \leq -a \frac{M}{\log M}.$$

Proof Suppose that

$$V_{2^{-m}}(Y^N) \leq \frac{1}{\log m}, \text{ for } m = M, \ldots, M(N), \text{ where } M(N) = \left\lceil \frac{\log(N/T)}{\log 2} \right\rceil. \tag{7.90}$$

It is plain that $m \geq M(N)$ implies that $2^{-m} < T/N$, hence $V_{2^{-(m+1)}}(Y^N) = \frac{1}{2} V_{2^{-m}}(Y^N)$. Then, provided $N > 4T$, $M(N) \geq 2$, hence for any $m \geq M(N)$, $m + 1 \leq m^2$, and also $(2 \log m)^{-1} \leq (\log(m + 1))^{-1}$, and it follows that (7.90) implies that $Y^{N,x} \in \mathcal{K}(M)$.

Now if $M \leq m \leq M(N)$,

$$\left\{ V_{2^{-m}}(Y^N) > \frac{1}{\log m} \right\} \subset \bigcup_{j=0}^{N-1} \left\{ \sup_{t_j \leq s \leq t_j + 2^{1-m}} |Z_s^{N,x} - Z_{t_j}^{N,x}| > \frac{1}{2 \log m} \right\}.$$

Consequently, with the help of Lemma 7.6, for some $C > 0$ and provided M is large enough,

$$\mathbb{P}[Y^{N,x} \notin \mathcal{K}(M)] \leq \sum_{m=M}^{M(N)} \mathbb{P}\left(V_{2^{-m}}(Y^N) > \frac{1}{\log m}\right)$$
$$\leq N \sum_{m=M}^{M(N)} \exp\left(-\frac{CN}{\log m} \log\left(\frac{C2^m}{\log m}\right)\right)$$
$$\leq M(N)N \exp\left(-\frac{CN}{\log M} \log\left(\frac{C2^M}{\log M}\right)\right),$$

where the last inequality follows from the fact that for $x > 0$ large enough, the mapping $x \to (\log x)^{-1} \log\left(\frac{C2^x}{\log x}\right)$ is increasing. Consequently

$$\frac{1}{N} \log \mathbb{P}\left[Y^{N,x} \notin \mathcal{K}(M)\right] \leq \frac{\log M(N)}{N} + \frac{\log N}{N} - \frac{c}{\log M} \log\left(\frac{C2^M}{\log M}\right).$$

It remains to take the limsup as $N \to \infty$.

We now establish the main local estimate for Y^N.

Lemma 7.35 *Assume that β_j ($j = 1, \dots, k$) is bounded. For all $\delta > 0$, we have uniformly in $x \in A$ and θ in a bounded set,*

$$\limsup_{N\to\infty} \log \mathbb{E}\left[\exp\left(N\langle Y^{N,x}(\tfrac{T}{N}) - Y^{N,x}(0), \theta\rangle\right)\right]$$
$$\leq T \cdot \sum_{j=1}^{k} \sup_{z^j \in A,\, |z^j - x| \leq \delta} \beta_j(z^j)(\mathrm{e}^{\langle\theta, h_j\rangle} - 1).$$

Proof It is not hard to verify that for any $\theta \in \mathbb{R}^d$, the process

$$M_t^\theta := \exp\left(N\langle Z_t^{N,x} - x, \theta\rangle - N \sum_{j=1}^{k} (e^{\langle\theta, h_j\rangle} - 1) \int_0^t \beta_j(Z^{N,x}(s))ds\right)$$

is a martingale with $M_0^\theta = 1$, hence $\mathbb{E}[M_t^\theta] = 1$. Let

$$S_{N,\delta} := \{\sup_{0\leq t\leq T/N} |Z_t^{N,x} - x] \leq \delta\}.$$

Since $M_t^\theta > 0$, $\mathbb{E}[M_{T/N}^\theta \mathbf{1}_{S_{N,\delta}}] \le 1$. But on the event $S_{N,\delta}$,

$$M_{T/N}^\theta \ge \exp\left(N\langle Z^{N,x}(T/N) - x, \theta\rangle - T\sum_{j=1}^{k} \sup_{z^j\in A, |z^j - x|\le\delta} \beta_j(z^j)(e^{\langle\theta,h_j\rangle} - 1)\right),$$

hence

$$\mathbb{E}\left[\exp\left(N\langle Z^{N,x}(T/N) - x, \theta\rangle\right)\mathbf{1}_{S_{N,\delta}}\right] \le \exp\left(T\sum_{j=1}^{k} \sup_{z^j\in A, |z^j - x|\le\delta} \beta_j(z^j)(e^{\langle\theta,h_j\rangle} - 1)\right).$$

On the other hand, from Lemma 7.6, for some $C > 0$, whenever $|\theta| \le B$,

$$\begin{aligned}
\mathbb{E}&\left[\exp\left(N\langle Z^{N,x}(T/N) - x, \theta\rangle\right)\mathbf{1}_{S^c_{N,\delta}}\right] \\
&\le \sum_{\ell=1}^{\infty} e^{N(\ell+1)\delta|\theta|}\mathbb{P}\left(\ell\delta \le |Z^{N,x}(T/N) - x| \le (\ell+1)\delta\right) \\
&\le \sum_{\ell=1}^{\infty} \exp\left(N\delta\left[(\ell+1)B - C\ell\log(CN\ell\delta)\right]\right) \\
&\le \sum_{\ell=1}^{\infty} a(N,\delta)^\ell \\
&\le 2a(N,\delta),
\end{aligned}$$

provided N is large enough, such that

$$a(N,\delta) := \exp\left(N\delta\left[2B - C\log(CN\delta)\right]\right) \le 1/2.$$

Finally

$$\mathbb{E}\left[\exp\left(N\langle Z^{N,x}(T/N) - x, \theta\rangle\right)\right] \le \exp\left(T\sum_{j=1}^{k} \sup_{z^j\in A, |z^j - x|\le\delta} \beta_j(z^j)(e^{\langle\theta,h_j\rangle} - 1)\right) + 2a(N,\delta).$$

The result follows from the fact that $a(N,\delta) \to 0$ as $N \to \infty$, for any $\delta > 0$.

We next establish

Lemma 7.36 *Let β_j ($j = 1, \ldots, k$) be bounded and continuous. Let $\theta : [0, T] \to \mathbb{R}^d$ be a step function, $\delta > 0$ and $\mathcal{K} \subset \mathcal{K}(M)$ be a compact set, such that the subset $\mathcal{K}^{ac}$ consisting of those elements of $\mathcal{K}$ which are absolutely continuous is dense in $\mathcal{K}$. Then*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[Y^{N,x} \in \mathcal{K}] \le - \inf_{\phi\in\mathcal{K},\, \phi(0)=x} I^\delta(\phi, \theta)$$

uniformly in x.

Proof Let θ be a fixed step function from $[0, T]$ into $\mathbb{R}^d$, which we assume w.l.o.g. to be right continuous, and let $\mathcal{K}$ be a given compact subset of $C([0, T]; \mathbb{R}^d)$, which has the property that $\mathcal{K}^{ac}$ is dense in $\mathcal{K}$. We define for each $\delta > 0$ the mapping $g_\delta : \mathbb{R}^{2d} \to \mathbb{R}$ by

$$g_\delta(z, \theta) = \sum_{j=1}^{k} \sup_{|z_j - x| \le \delta} \beta_j(z_j)(e^{\langle\theta, h_j\rangle} - 1).$$

We let $t_\ell := \ell T/N$, and define for $z \in \mathcal{K}^{ac}$, the two quantities

$$\tilde{S}_N(z, \theta) = \sum_{\ell=1}^{N} \langle z(t_\ell) - z(t_{\ell-1}), \theta(t_{\ell-1})\rangle - \frac{T}{N} \sum_{\ell=1}^{N} g_\delta(z(t_{\ell-1}), \theta(t_{\ell-1})),$$

$$S(z, \theta) = \int_0^T \langle z'(t), \theta(t)\rangle dt - \frac{T}{N} \sum_{\ell=1}^{N} g_\delta(z(t_{\ell-1}), \theta(t_{\ell-1})).$$

Choose any $\eta > 0$. We can assume that N_0 has been chosen large enough, such that

$$\sup_{z\in\mathcal{K}^{ac}} |\tilde{S}_N(z, \theta) - S(z, \theta)| \le \eta.$$

Indeed, this difference is bounded by twice the number of jumps of θ times the sup of $|\theta(t)|$, times the maximal oscillation of z on intervals of length $1/N$ in $[0, T]$.

It follows from Lemma 7.35 and the Markov property that, provided N_0 has been chosen large enough, for any $N \ge N_0$,

$$\mathbb{E}\left[\exp\left(N\tilde{S}_N(Y^{N,x}, \theta)\right)\right] \le \exp(N\eta).$$

Clearly, on the event $Y^{N,x} \in \mathcal{K}$,

$$\exp\left[N\left(\tilde{S}_N(Y^N, \theta) - \inf_{z\in\mathcal{K}^{ac}} \tilde{S}_N(z, \theta)\right)\right] \ge 1,$$

and combining this fact with the previous inequalities, we deduce that

$$\begin{aligned}\mathbb{P}(Y^{N,x} \in \mathcal{K}) &\leq \mathbb{E}\exp\left[N\left(\tilde{S}_N(Y^{N,x},\theta) - \inf_{z\in\mathcal{K}^{ac}} \tilde{S}_N(z,\theta)\right)\right]\\ &\leq \exp(N\eta)\exp\left(-N\inf_{z\in\mathcal{K}^{ac}}\tilde{S}_N(z,\theta)\right)\\ &\leq \exp(2N\eta)\exp\left(-N\inf_{z\in\mathcal{K}^{ac}} S_N(z,\theta)\right)\end{aligned}$$

Now, uniformly in $z \in \mathcal{K}^{ac}$, $S_N(z,\theta) \to I^\delta(z,\theta)$, where

$$I^\delta(z,\theta) = \int_0^T \langle z'(t), \theta(t)\rangle dt - \int_0^T g_\delta(z(t),\theta(t))dt.$$

The result follows from the last two facts, since $\eta > 0$ can be chosen arbitrarily small, and $\mathcal{K}^{ac}$ is dense in $\mathcal{K}$.

We now have

Lemma 7.37 *Assume that β_j ($j = 1, \ldots, k$) is bounded and Lipschitz continuous. Then for all $K > 0$, $\delta > 0$ and $\epsilon > 0$,*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[\mathrm{d}(Y^{N,x}, \Phi_x^\delta(K)) > \epsilon] \leq -K + \epsilon$$

uniformly in $x \in A$.

Proof We fix $\epsilon, \delta, K > 0$ and choose $M \in \mathbb{N}$ such that $a\frac{M}{\log M} > K - \epsilon$, where a is the constant appearing in Lemma 7.35.

For absolute continuous $\phi : [0,T] \to A$ with $I^\delta(\phi) < \infty$, there exists a step function θ^ϕ such that

$$I^\delta(\phi,\theta^\phi) \geq I^\delta(\phi) - \tfrac{\epsilon}{2}$$

(cf. Lemma 7.31). It can easily be verified by elementary calculus that the function $I^\delta(\cdot,\theta^\phi)$ is continuous for the sup norm topology on the set of absolutely continuous functions. Hence there exists a number $0 < \eta^\phi < \frac{\epsilon}{2}$ such that for all absolutely continuous $\tilde{\phi}$ with $\|\phi - \tilde{\phi}\| < \eta^\phi$,

$$I^\delta(\tilde{\phi},\theta^\phi) \geq I^\delta(\phi) - \epsilon. \tag{7.91}$$

We consider the compact set

$$\mathcal{K}^x(M) := \{\phi \in \mathcal{K}(M) | \phi(0) = x\}$$

(cf. the definition preceding Lemma 7.34). By a compactness argument, there exist finitely many absolutely continuous functions $\{\phi_i,\ 1 \le i \le m\} \subset \mathcal{K}^x(M)$ with $I^\delta(\phi_i) < \infty$ (and corresponding $\theta_i := \theta^{\phi_i}$ and $\eta_i := \eta^{\phi_i}$) such that

$$\mathcal{K}^x(M) \subset \bigcup_{i=1}^{m} B_{\eta_i}(\phi_i).$$

For each $1 \le i \le m$, we define the compact set

$$\mathcal{K}_i^x(M) := \overline{B_{\eta_i}(\phi_i) \cap \mathcal{K}^x(M)}.$$

We now let

$$\mathcal{I} := \{1 \le i \le m \mid \mathrm{d}(\phi_i, \Phi_x^\delta(K)) \ge \eta_i\}.$$

Then $\mathrm{d}(Y^{x,N}, \Phi_x^\delta(K)) \ge \epsilon$ and $Y^{x,N} \in \mathcal{K}_i^x(M)$ imply $i \in \mathcal{I}$, since $\eta_i \le \epsilon/2$. Hence

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[\mathrm{d}(Y^{x,N}, \Phi_x^\delta(K)) \ge \epsilon] \\
&\quad \le \limsup_{N\to\infty} \frac{1}{N} \log \big\{\mathbb{P}[Y^{x,N} \notin \mathcal{K}^x(M)] + \sum_{i\in\mathcal{I}} \mathbb{P}[Y^{x,N} \in \mathcal{K}_i^x(M)]\big\}.
\end{aligned}$$

Applying first Lemma 7.36 and then (7.91), we obtain

$$\begin{aligned}
\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}[Y^{x,N} \in \mathcal{K}_i^x(M)] &\le - \inf_{\phi\in\mathcal{K}_i^x(M)} I^\delta(\phi, \theta_i) \\
&\le -I^\delta(\phi_i) + \epsilon \\
&< -K + \epsilon
\end{aligned}$$

as $I^\delta(\phi_i) > K$ (recall that $i \in \mathcal{I}$). The result now follows from the two last inequalities, Lemma 7.34 and the fact that $a\frac{M}{\log M} > K - \epsilon$.

7.6.4 Main Results

Theorem 7.11 *Assume that Assumption 7.2 is satisfied. For $F \subset D([0,T];A)$ closed and $x \in A$, we have*

$$\limsup_{y_N\in A^N, y_N\to x,\ N\to\infty} \frac{1}{N} \log \mathbb{P}[Z^{N,y_N} \in F] \le -I_x(F).$$

Proof We first let $I_x(F) =: K < \infty$ and $\epsilon > 0$. By Lemma 7.22, there exits a $\delta^\epsilon > 0$ such that for all $\delta \le \delta^\epsilon$,

$$y \in A, |x-y| < \delta \Rightarrow I_y(F) \ge I_x(F) - \epsilon = K - \epsilon.$$

For $\delta \le \delta^\epsilon$, we define

$$F^\delta := \{\phi \in F \,|\, |\phi(0) - x| \le \delta\},$$
$$S^\delta := \bigcup_{y \in A, |x-y| \le \delta} \Phi_y(K - 2\epsilon).$$

F^δ is closed in $D([0,T]; A; \mathrm{d}_D)$ and S^δ is compact in $D([0,T]; A; \mathrm{d}_D)$ by Proposition 7.1. Furthermore, the two sets have no common elements. Hence, by the Hahn-Banach Theorem,

$$\mathrm{d}(F^\delta, S^\delta) =: \eta^\delta > 0. \tag{7.92}$$

Note that η^δ is increasing as δ is decreasing, since the sets F^δ and S^δ are decreasing. We now let $|y-x| \le \delta$ and $\eta \le \eta^\delta$. Let Y^N be defined as in the paragraph preceding Lemma 7.27. We have

$$\begin{aligned}\mathbb{P}[Z^{N,y} \in F] &= \mathbb{P}[Z^{N,y} \in F^\delta] \\ &\le \mathbb{P}[\mathrm{d}(Y^{N,y}, F^\delta) < \tfrac{\eta}{2}] + \mathbb{P}[\|Y^{N,y} - Z^{N,y}\| \ge \tfrac{\eta}{2}]. \end{aligned} \tag{7.93}$$

Let now $\phi(0) = y$ with $\mathrm{d}(\phi, F^\delta) < \frac{\eta}{2}$, hence from (7.92)

$$\mathrm{d}(\phi, \Phi_y(K - 2\epsilon)) \ge \frac{\eta}{2}. \tag{7.94}$$

Let $\tilde{\delta}$ be such that Corollary 7.8 with K replaced by $K - 2\epsilon$ and ϵ by $\frac{\eta}{4}$ holds with δ replaced by $2\tilde{\delta}$. Hence (7.94) implies

$$\mathrm{d}(\phi, \Phi_y^{2\tilde{\delta}}(K - 2\epsilon - \tfrac{\eta}{4}) > \frac{\eta}{4}. \tag{7.95}$$

Indeed, if that is not the case, there exists a $\tilde{\phi} \in \Phi_y^{2\tilde{\delta}}(K - 2\epsilon - \frac{\eta}{4})$ with $\|\phi - \tilde{\phi}\| \le \frac{\eta}{4}$. Then Corollary 7.8 implies that there exists $\bar{\phi} \in \Phi_y(K - 2\epsilon)$ with $\|\bar{\phi} - \tilde{\phi}\| \le \frac{\eta}{4}$; consequently $\|\bar{\phi} - \phi\| \le \frac{\eta}{2}$, which contradicts (7.95). We hence obtain by

Lemma 7.37,

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}[\mathrm{d}(Y^{N,y},F^{\delta})<\tfrac{\eta}{2}]\\ \le\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}[\mathrm{d}(Y^{N,y},\varPhi_y^{2\tilde{\Delta}}(K-2\epsilon-\tfrac{\eta}{4})>\frac{\eta}{4}]\\ \le-(K-2\epsilon-\frac{\eta}{2})\end{aligned}\tag{7.96}$$

uniformly in $y\in A$ with $|y-x|\le\delta$.

Furthermore, Lemma 7.27 implies

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}[\|Y^{N,y}-Z^{N,y}\|\ge\tfrac{\eta}{2}]=-\infty\tag{7.97}$$

uniformly in $y\in A$.

Combining Inequalities (7.93), (7.96) and (7.97), we obtain

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}[Z^{N,y}\in F]\le-(K-2\epsilon-\frac{\eta}{2})$$

uniformly in $y\in A$, $|x-y|\le\delta$. The result now follows as ϵ and η can be chosen arbitrarily small.

The result in case $I_x(F)=\infty$ follows, since this implies that $I_x(F)>K$ for all $K>0$.

We will need the following stronger version. Recall the definition of A^N at the start of Sect. 7.2.

Theorem 7.12 *Assume that Assumption 7.2 is satisfied. For $F\subset D([0,T];A)$ closed and any compact subset $K\subset A$, we have*

$$\limsup_{N\to\infty}\frac{1}{N}\log\sup_{x\in K\cap A^N}\mathbb{P}[Z^{N,x}\in F]\le-\inf_{x\in K}I_x(F).$$

Proof We use the same argument as in the proof of Corollary 5.6.15 in [3]. From Theorem 7.11, for any $x\in A$, any $\delta>0$, there exists $\epsilon_{x,\delta}>0$ and $N_{x,\delta}\ge1$ such that whenever $N\ge N_{x,\delta}$, $y\in A_N$ with $|y-x|<\epsilon_{x,\delta}$,

$$\frac{1}{N}\log\mathbb{P}[Z^{N,y}\in F]\le-I_x^{\delta}(F),$$

where $I_x^{\delta}(F)=\min[I_x(F)-\delta,\delta^{-1}]$. Consider now a compact set $K\subset A$. There exists a finite set $\{x_i,\ 1\le i\le I\}$ such that $K\subset\cup_{i=1}B(x_i,\epsilon_{x_i})$, where $B(x,\epsilon)=$

$\{y;\ |y-x|<\epsilon\}$. Consequently, for $N \geq \sup_{1\leq i\leq I} N_{x_i,\delta}$, any $y \in A_N \cap K$,

$$\frac{1}{N}\log \mathbb{P}[Z^{N,y} \in F] \leq -\min_{1\leq i\leq I} I^{\delta}_{x_i}(F) \leq -\inf_{x\in K} I^{\delta}_{x}(F).$$

It remains to take the sup over $y \in K \cap A_N$ on the left, take the lim sup as $N \to \infty$, and finally let δ tend to 0 to deduce the result.

7.7 Exit Time from a Domain

In this section we establish the results for the exit time of the process from a domain. To this end, we follow the line of reasoning of [3, Section 5.7] and modify the arguments when necessary.

We let $O \subsetneq A$ be relatively open in A (with $O = \tilde{O} \cap A$ for $\tilde{O} \subset \mathbb{R}^d$ open) and $x^* \in O$ be a stable equilibrium of (7.2). By a slight abuse of notation, we say that

$$\widetilde{\partial O} := \partial \tilde{O} \cap A$$

is the *boundary* of O. For $y, z \in A$, we define the following functionals.

$$V(x,z,T) := \inf_{\phi\in D([0,T];A),\phi(0)=x,\phi(T)=z} I_{T,x}(\phi)$$

$$V(x,z) := \inf_{T>0} V(x,z)$$

$$\bar{V} := \inf_{z\in\widetilde{\partial O}} V(x^*,z).$$

In other words, $\bar{V}$ is the minimal energy required to leave the domain O when starting from x^*.

Assumption 7.13

(D1) *x^* is the only stable equilibrium point of (7.2) in O and the solution Y^x of (7.2) with $x = Y^x(0) \in O$ satisfies*

$$Y^x(t) \in O \text{ for all } t>0 \text{ and } \lim_{t\to\infty} Y^x(t) = x^*.$$

(D2) *For a solution Y^x of (7.2) with $x = Y^x(0) \in \widetilde{\partial O}$, we have*

$$\lim_{t\to\infty} Y^x(t) = x^*.$$

(D3) $\bar{V} < \infty$.

(D4) *For all* $\rho > 0$ *there exist constants* $T(\rho)$, $\epsilon(\rho) > 0$ *with* $T(\rho), \epsilon(\rho) \downarrow 0$ *as* $\rho \downarrow 0$ *such that for all* $z \in \widetilde{\partial O} \cup \{x^*\}$ *and all* $x, y \in \overline{B(z, \rho)} \cap A$ *there exists an*

$$\phi = \phi(\rho, x, y) : [0, T(\rho)] \to A$$
$$\text{with } \phi(0) = x, \phi(T(\rho)) = y \text{ and } I_{T(\rho)}(\phi) < \epsilon(\rho).$$

(D5) *For all* $z \in \widetilde{\partial O}$ *there exists an* $\eta_0 > 0$ *such that for all* $\eta < \eta_0$ *there exists a* $\tilde{z} = \tilde{z}(\eta) \in A \setminus \bar{O}$ *with* $|z - \tilde{z}| > \eta$.

Let us shortly comment on Assumption 7.13. By (D1), O is a subset of the domain of attraction of x^*. (D2) is violated by the applications we have in mind: we are interested in situations where $\widetilde{\partial O}$ is the *characteristic boundary* of O, i.e., the boundary separating two regions of attraction of equilibria of (7.2). In order to relax this assumption, we shall add an approximation argument in Sect. 7.7.3. By (D3), it is possible to reach the boundary with finite energy. This assumption is always satisfied for the epidemiological models we consider. For $z = x^*$, (D4) is also always satisfied in our models as the rates β_j are bounded from above and away from zero in small neighborhoods of x^*; hence, the function $\phi(x, y, \rho)$ can, e.g., be chosen to be linear with speed one (see, e.g., [13] Lemma 5.22). (D5) allows us to consider a trajectory which crosses the boundary $\widetilde{\partial O}$, in such a way that all paths in a sufficiently small tube around that trajectory do exit O.

We are interested in the following quantity:

$$\tau^{N,x} := \tau^N := \inf\{t > 0 | Z^{N,x}(t) \notin O\},$$

i.e., the first time that $Z^{N,x}$ exits O.

7.7.1 Auxiliary Results

Assumption 7.13 (A4) gives the following analogue of Lemma 5.7.8 of [3].

Lemma 7.38 *Assume that Assumption 7.13 holds. Then for any* $\delta > 0$, *there exists an* $\rho_0 > 0$ *such that for all* $\rho < \rho_0$,

$$\sup_{z \in \widetilde{\partial O} \cup x^*, x, y \in \overline{B(z,\rho)}} \inf_{T \in [0,1]} V(x, y, T) < \delta.$$

We can recover the analogue of Lemma 5.7.18 of [3] by using Lemma 7.38.

Lemma 7.39 *Assume that Assumptions 7.2 and 7.13 hold. Then, for any $\eta > 0$ there exists a ρ_0 such that for all $\rho < \rho_0$ there exists a $T_0 < \infty$ such that*

$$\liminf_{N\to\infty} \frac{1}{N} \log \inf_{x\in\overline{B(x^*,\rho)}} \mathbb{P}[\tau^{N,x} \leq T_0] > -(\bar{V} + \eta).$$

Proof We follow the same line of reasoning as in the proof of Lemma 5.7.18 in [3]. Let $x \in \overline{B(x^*, \rho)}$. We use Lemma 7.38 for $\delta = \eta/4$ (and we let ρ be small enough for Lemma 7.38 to hold). We construct a continuous path ψ^x with $\psi^x(0) = x$, $\psi^x(t_x) = x^*$ ($t_x \leq 1$) and $I_{t_x,x}(\psi^x) \leq \eta/4$. We then use Assumption 7.13 (D3). For $T_1 < \infty$, we can construct a path $\phi \in C[0, T_1]$ such that $\phi(0) = x^*$, $\phi(T_1) = z \in \partial\widetilde{O}$ and $I_{T_1,0}(\phi) \leq \bar{V} + \eta/4$. Subsequently, we use Lemma 7.38 and obtain a path $\tilde{\psi}$ with $\tilde{\psi}(0) = z$, $\tilde{\psi}(s_x) \notin O$ ($s \leq 1$), $I_{s,z}(\tilde{\psi}) \leq \eta/4$ and $\text{dist}(\bar{z}, O) =: \Delta > 0$.[10] We finally let θ^x be the solution of the ODE (7.2) with $\theta^x(0) = \bar{z}$ on $[0, 2 - t_x - s]$, consequently $I_{2-t_x-s,\bar{z}}(\theta^x) = 0$, see Lemma 7.7.

We concatenate the paths ψ^x, ϕ, $\tilde{\psi}$ and θ^x and obtain the path $\phi^x \in C[0, T_0]$ ($T_0 = T_1 + 2$ independent of x) with $I_{T_0,x}(\phi^x) \leq \bar{V} + \eta/2$.

Finally, we define

$$\Psi := \bigcup_{x\in\overline{B(x^*,\rho)}} \left\{\psi \in D([0, T_0]; A) \middle| \|\psi - \phi^x\| < \Delta/2\right\};$$

hence $\Psi \subset D([0, T_0]; A)$ is open, $(\phi^x)_{x\in\overline{B(x^*,\rho)}} \subset \Psi$ and $\{Z^{N,x} \in \Psi\} \subset \{\tau^{N,x} \leq T_0\}$. We now use Theorem 7.10.

$$\begin{aligned}
\liminf_{N\to\infty} \frac{1}{N} \log \inf_{x\in\overline{B(x^*,\rho)}} \mathbb{P}[Z^{N,x} \in \Psi] &\geq -\sup_{x\in\overline{B(x^*,\rho)}} \inf_{\phi\in\Psi} I_{T_0,x}(\phi) \\
&\geq -\sup_{x\in\overline{B(x^*,\rho)}} I_{T_0,x}(\phi^x) \\
&> -(\bar{V} + \eta).
\end{aligned}$$

We also require the following result (analogue of Lemma 5.7.19 of [3]).

Lemma 7.40 *Assume that Assumption 7.13 holds. Let $\rho > 0$ such that $\overline{B(x^*, \rho)} \subset O$ and*

$$\sigma_\rho^{N,x} := \inf\{t > 0 | Z_t^{N,x} \in \overline{B(x^*, \rho)} \text{ or } Z_t^{N,x} \notin O\}.$$

[10] Note that the Assumption (D5) is required here.

Then

$$\lim_{t\to\infty}\limsup_{N\to\infty}\frac{1}{N}\log\sup_{x\in O}\mathbb{P}[\sigma_\rho^{N,x}>t]=-\infty.$$

Proof We adapt the proof of [3] Lemma 5.7.19 to our case.

Note first that for $x\in\overline{B(x^*,\rho)}$, $\sigma_\rho^{N,x}=0$; we hence assume from now on that $x\notin\overline{B(x^*,\rho)}$. For $t>0$, we define the closed set $\Psi_t\subset D([0,t];A)$,

$$\Psi_t:=\{\phi\in D([0,t];A)|\phi(s)\in\overline{O\setminus B(x^*,\rho)}\text{ for all }s\in[0,t]\};$$

hence for all x, N,

$$\{\sigma_\rho^{N,x}>t\}\subset\{Z^{N,x}\in\Psi_t\}.$$

By Theorem 7.12, this implies for all $t>0$,

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\log\sup_{x\in\overline{O\setminus B(x^*,\rho)}}\mathbb{P}[\sigma_\rho^{N,x}>t]&\le\limsup_{N\to\infty}\frac{1}{N}\log\sup_{x\in\overline{O\setminus B(x^*,\rho)}}\mathbb{P}[Z^{\epsilon,x}\in\Psi_t]\\&\le-\inf_{\phi\in\Psi_t}I_{t,\phi(0)}(\phi).\end{aligned}$$

It hence suffices to show that

$$\lim_{t\to\infty}\inf_{\phi\in\Psi_t}I_{t,\phi(0)}(\phi)=\infty.\tag{7.98}$$

To this end, consider $x\in\overline{O\setminus B(x^*,\rho)}$ and recall that Y^x is the solution of (7.2) (on $[0,t]$ for all $t>0$). By Assumption 7.13 (D2), there exists a $T_x<\infty$ such that $Y^x(T_x)\in\overline{B(x^*,3\rho)}$. We have (here B denotes the Lipschitz constant of b),

$$\begin{aligned}|\phi^x(t)-\phi^y(t)|&\le|x-y|+\int_0^t|b(\phi^x(s))-b(\phi^y(s))|ds\\&\le+|x-y|+\int_0^t B|\phi^x(s)-\phi^y(s)|ds\end{aligned}$$

and therefore by Gronwall's inequality $|Y^x(T_x)-Y^y(T_x)|\le|x-y|e^{T_xB}$; consequently, there exists a neighborhood W_x of x such that for all $y\in W_x$, $Y^y(T_x)\in\overline{B(x^*,3\rho)}$. By the compactness of $\overline{O\setminus B(x^*,\rho)}$, there exists a finite open subcover $\cup_{i=1}^k W_{x_i}\supset\overline{O\setminus B(x^*,\rho)}$; for $T:=\max_{i=1,\dots,k}T_{x_i}$ and $y\in\overline{O\setminus B(x^*,\rho)}$ this implies that $Y^y(s)\in\overline{B(x^*,2/3\rho)}$ for some $s\le T$.

Assume now that (7.98) is false. Then there exits an $M < \infty$ such that for all $n \in \mathbb{N}$ there exists an $\phi_n \in \Psi_{nT}$ with $I_{nT}(\phi_n) \leq M$. The function ϕ_n is concatenated by functions $\phi_{n,k} \in \Psi_T$ and we obtain

$$M \geq I_{nT}(\phi_n) = \sum_{k=1}^{n} I_T(\phi_{n,k}) \geq n \min_{k=1,\ldots,n} I_T(\phi_{n,k}).$$

Hence there exists a sequence $(\psi_k)_k \subset \Psi_T$ with $\lim_{k\to\infty} I_T(\psi_k) = 0$. Note now that the set

$$\phi(t) := \{\phi \in C[0,T] | I_{T,\phi(0)}(\phi) \leq 1, \phi(s) \in \overline{O \setminus B(x^*, \rho)} \text{ for all } s \in [0,T]\} \subset \Psi_T$$

is compact (as a subset of $(C[0,T], \|\cdot\|_\infty)$); hence there exists a subsequence $(\psi_{k_l})_l$ of $(\psi_k)_k$ such that $\lim_{l\to\infty} \psi_{k_l} =: \psi^* \in \phi(t)$ in $(C[0,T], \|\cdot\|_\infty)$. By the lower semi-continuity of I_T (cf. Lemma 7.20) this implies

$$0 = \liminf_{l\to\infty} I_T(\psi_{n_l}) \geq I_T(\psi^*),$$

which in turn implies that ψ^* solves (7.2) for $x = \psi^*(0)$. But then, $\psi^*(s) \in \overline{B(x^*, 2/3\rho)}$ for some $s \leq T$, a contradiction to $\psi^* \in \Psi_T$.

The following lemma is the analogue of [3] Lemma 5.7.21.

Lemma 7.41 *Assume that Assumptions 7.2 and 7.13 hold. Let $C \subset A \setminus O$ be closed. Then*

$$\lim_{\rho\to 0} \limsup_{N\to\infty} \frac{1}{N} \log \sup_{x \in \overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} \mathbb{P}[Z^{N,x}_{\sigma_\rho} \in C] \leq -\inf_{z\in C} V(x^*, z).$$

Proof We adapt the proof of [3] Lemma 5.7.21 to our situation. We can assume without loss of generality that $\inf_{z\in C} V(x^*, z) > 0$ (else the assertion is trivial). For $\inf_{z\in C} V(x^*, z) > \delta > 0$, we define

$$V_C^\delta := (\inf_{z\in C} V(x^*, z) - \delta) \wedge 1/\delta > 0.$$

By Lemma 7.38, there exists a $\rho_0 = \rho_0(\delta) > 0$ such that for all $0 < \rho < \rho_0$,

$$\sup_{y \in \overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} V(x^*, y) < \delta;$$

hence

$$\inf_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)},\ z\in C} V(y,z) \geq \inf_{z\in C} V(x^*,z) - \sup_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} V(x^*,y) > V_C^\delta. \tag{7.99}$$

For $T > 0$, we define the closed set $\Phi^T \subset D([0,T];A)$ by

$$\Phi^T := \Phi := \{\phi \in D([0,T];A) | \phi(t) \in C \text{ for some } t \in [0,T]\}.$$

We then have for $y \in \overline{B(x^*, 3\rho) \setminus B(x^*, 2\rho)}$,

$$\mathbb{P}[Z^{N,y}_{\sigma_\rho} \in C] \leq \mathbb{P}[\sigma^{N,y}_\rho > T] + \mathbb{P}[Z^{N,y} \in \Phi^T]. \tag{7.100}$$

In the following, we bound the two parts in Inequality (7.100) from above.

For the second part, we note first that (cf. Inequality (7.99))

$$\inf_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)},\ \phi\in\Phi^T} I_{T,y}(\phi) \geq \inf_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)},\ z\in C} V(y,z) > V_C^\delta;$$

hence, we obtain by Theorem 7.12

$$\begin{aligned}
\limsup_{N\to\infty} \frac{1}{N} \log \sup_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} \mathbb{P}[Z^{N,y} \in \Phi^T] \\
\leq - \inf_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)},\ \phi\in\Phi^T} I_{T,y}(\phi) \\
< -V_C^\delta.
\end{aligned} \tag{7.101}$$

For the first part in Inequality (7.100), we use Lemma 7.40: There exists a $0 < T_0 < \infty$ such that for all $T \geq T_0$

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} \mathbb{P}[\sigma^{N,y} > T] < -V_C^\delta. \tag{7.102}$$

We let $T \geq T_0$ and $\rho < \rho_0$ and combine Inequalities (7.100)–(7.102). Hence there exists an $N_0 > 0$ such that for all $N > N_0$,

$$\begin{aligned}
\frac{1}{N} \log \sup_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} \mathbb{P}[Z^{N,y}_{\sigma_\rho} \in C] \\
\leq \frac{1}{N} \log \Big(\sup_{y\in\overline{B(x^*,3\rho)\setminus B(x^*,2\rho)}} \mathbb{P}[\sigma^{N,y}_\rho > T]
\end{aligned}$$

$$+ \sup_{y \in \overline{B(x^*,3\rho) \setminus B(x^*,2\rho)}} \mathbb{P}[Z^{N,y} \in \Phi^T]\Big)$$
$$< \frac{1}{N} \log\left(2e^{-NV_C^\delta}\right) = \frac{1}{N}\log 2 - V_C^\delta;$$

and

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{y \in \overline{B(x^*,3\rho) \setminus B(x^*,2\rho)}} \mathbb{P}[Z^{N,x}_{\sigma_\rho} \in C] \le -V_C^\delta.$$

Taking the limit $\delta \to 0$ finishes the proof.

The next lemma is the analogue of Lemma 5.7.22 of [3].

Lemma 7.42 *Assume that Assumption 7.13 holds. Then, for all* $\rho > 0$ *such that* $\overline{B(x^*, \rho)} \subset O$ *and for all* $x \in O$,

$$\lim_{N\to\infty} \mathbb{P}[Z^{N,x}_{\sigma_\rho} \in \overline{B(x^*, \rho)}] = 1.$$

Proof Let $x \in O \setminus \overline{B(x^*, \rho)}$ (the case $x \in \overline{B(x^*, \rho)}$ is clear). Let furthermore $T := \inf\{t \ge 0 | \phi(t) \in B(x^*, \rho/2)\}$. Since Y^x is continuous and never reaches $\partial\widetilde{O}$ (Assumption 7.13 (D1)), we have $\inf_{t\ge 0} \operatorname{dist}(Y^x(t), \partial\widetilde{O}) =: \Delta > 0$. Hence we have the following implication:

$$\sup_{t\in[0,T]} |Z^{N,x}_t - Y^x(t)| \le \frac{\Delta}{2} \Rightarrow Z^{N,x}_{\sigma_\rho} \in \overline{B(x^*, \rho)}.$$

In other words,

$$\mathbb{P}[Z^{N,x}_{\sigma_\rho} \notin \overline{B(x^*, \rho)}] \le \mathbb{P}\Big[\sup_{t\in[0,T]} |Z^{N,x}_t - Y^x(t)| > \frac{\Delta}{2}\Big],$$

and the right hand side of the last inequality converges to zero as $N \to \infty$ by Theorem 7.3.

The next lemma is the analogue of [3, Lemma 5.7.23].

Lemma 7.43 *Assume that Assumption 7.13 holds. Then, for all* $\rho, c > 0$, *there exists a constant* $T = T(c, \rho) < \infty$ *such that*

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in O} \mathbb{P}[\sup_{t\in[0,T]} |Z^{N,x}_t - x| \ge \rho] < -c.$$

Proof Let $\rho, c > 0$ be fixed. For $T, N > 0$ and $x \in O$ we have

$$\begin{aligned}\mathbb{P}[\sup_{t\in[0,T]} |Z_t^{N,x} - x| \geq \rho] &= \mathbb{P}\Big[\sup_{t\in[0,T]} \frac{1}{N}|\sum_j h_j P_j\Big(N\int_0^t \beta_j(Z_s^{N,x})ds\Big)| \geq \rho\Big] \\ &\leq \mathbb{P}\Big[\sum_j P_j(N\bar{\beta}T) \geq N\rho\bar{h}^{-1}\Big] \\ &\leq k\mathbb{P}\Big[P(N\bar{\beta}T) \geq N\rho\bar{h}^{-1}k^{-1}\Big] \qquad (7.103)\end{aligned}$$

for a standard Poisson process P. We now let, with $c_1(T) = \bar{\beta}T$ and $c_2 = \rho\bar{h}^{-1}k^{-1}$,

$$T < T_0 := \frac{e^{-1}c_2}{2\bar{\beta}} \wedge \frac{e^{-c/c_2-1}c_2}{\bar{\beta}} \quad \text{and} \quad N > N_0 := 1/c_2 \wedge \frac{\log 2k}{c_1(T)}. \qquad (7.104)$$

We then deduce from (7.11)

$$k\mathbb{P}\Big[P(Nc_1(T)) \geq Nc_2\Big] \leq k\left(\frac{e}{c_2}c_1(T)\right)^{Nc_2}. \qquad (7.105)$$

Finally

$$\big(\frac{e}{c_2}c_1(T)\big)^{Nc_2} = \Big(\big(\frac{e}{c_2}c_1(T)\big)^{-c_2}\Big)^{-N} < e^{-Nc} \qquad (7.106)$$

by (7.104). The assertion now follows by combining the Inequalities (7.103), (7.105) and (7.106).

7.7.2 Main Results

We can now deduce the analogue of Theorem 5.7.11 (a) in [3] . The proof of [3] carries over.

Theorem 7.14 *Assume that Assumption 7.13 holds. Then, for all $x \in O \cap AN$ and $\delta > 0$,*

$$\lim_{N\to\infty} \mathbb{P}\big[e^{(\bar{V}-\delta)N} < \tau^{N,x} < e^{(\bar{V}+\delta)N}\big] = 1.$$

Moreover, for all $x \in O$, as $N \to \infty$,

$$\frac{1}{N}\log\mathbb{E}(\tau^{N,x}) \to \bar{V}.$$

Proof Upper bound of exit time:

We fix $\delta > 0$ and apply Lemma 7.39 with $\eta := \delta/4$. Hence, for $\rho < \rho_0$ there exists a $T_0 < \infty$ and an $N_0 > 0$ such that for $N > N_0$,

$$\inf_{x \in \overline{B(x^*,\rho)}} \mathbb{P}[\tau^{N,x} \leq T_0] > e^{-N(\bar{V}+\eta)}.$$

Furthermore, by Lemma 7.40 there exists a $T_1 < \infty$ and $N_1 > 0$ such that for all $N > N_1$,

$$\inf_{x \in O} \mathbb{P}[\sigma_\rho^{N,x} \leq T_1] > 1 - e^{-2N\eta}.$$

For $T := T_0 + T_1$ and $N > N_0 \vee N_1 \vee 1/\eta$, we hence obtain

$$\begin{aligned} q^N := q &:= \inf_{x \in O} \mathbb{P}[\tau^{N,x} \leq T] \\ &\geq \inf_{x \in O} \mathbb{P}[\sigma_\rho^{N,x} \leq T_1] \inf_{y \in \overline{B(x^*,\rho)}} \mathbb{P}[\tau^{N,y} \leq T_0] \\ &> (1 - e^{-2N\eta}) e^{-N(\bar{V}+\eta)} \\ &\geq e^{-N(\bar{V}+2\eta)}. \end{aligned} \tag{7.107}$$

This yields for $k \in \mathbb{N}$

$$\begin{aligned} \mathbb{P}[\tau^{N,x} > (k+1)T] &= \big(1 - \mathbb{P}[\tau^{N,x} \leq (k+1)T | \tau^{N,x} > kT]\big)\mathbb{P}[\tau^{N,x} > kT] \\ &\leq (1-q)\mathbb{P}[\tau^{N,x} > kT] \end{aligned}$$

and hence inductively

$$\sup_{x \in O} \mathbb{P}[\tau^{N,x} > kT] \leq (1-q)^k.$$

This implies, exploiting (7.107) for the last inequality

$$\sup_{x \in O} \mathbb{E}[\tau^{N,x}] \leq T\big(1 + \sum_{k=1}^{\infty} \sup_{x \in O} \mathbb{P}[\tau^{N,x} > kT]\big) \leq T \sum_{k=0}^{\infty} (1-q)^k = \frac{T}{q} \leq T e^{N(\bar{V}+2\eta)}; \tag{7.108}$$

by Chebychev's Inequality we obtain

$$\mathbb{P}[\tau^{N,x} \geq e^{N(\bar{V}+\delta)}] \leq e^{-N(\bar{V}+\delta)} \mathbb{E}[\tau^{N,x}] \leq T e^{-\delta N/2}$$

which approaches zero as $N \to \infty$ as required.

<u>Lower bound of exit time:</u>

For $\rho > 0$ such that $\overline{B(x^*, 3\rho)} \subset O$, we define recursively $\theta_0 := 0$ and for $m \in \mathbb{N}_0$,

$$\tau_m^x := \tau_m := \inf\{t \geq \theta_m^x | Z_t^{N,x} \in \overline{B(x^*, \rho)} \text{ or } Z_t^{N,x} \notin O\},$$

$$\theta_{m+1}^x := \theta_{m+1} := \inf\{t \geq \tau_m^x | Z_t^{N,x} \in \overline{B(x^*, 3\rho) \setminus B(x^*, 2\rho)}\},$$

with the convention $\theta_{m+1} := \infty$ if $Z_{\tau_m}^N \notin O$. Note that we have $\tau^{N,x} = \tau_m^x$ for some $m \in \mathbb{N}_0$.

For fixed $T_0 > 0$ and $k \in \mathbb{N}$ we have the following implication: If for all $m = 0, \ldots, k$, $\tau_m \neq \tau^N$ and for all $m = 1 \ldots, k$, $\tau_m - \tau_{m-1} > T_0$, then

$$\tau^N > \tau_k = \sum_{m=1}^{k} (\tau_m - \tau_{m-1}) > kT_0.$$

In particular, we have for $k := \lfloor T_0^{-1} e^{N(\bar{V}-\delta)} \rfloor + 1$ (note that $\theta_m - \tau_{m-1} \leq \tau_m - \tau_{m-1}$),

$$\begin{aligned}
\mathbb{P}[\tau^{N,x} \leq e^{N(\bar{V}-\delta)}] &\leq \mathbb{P}[\tau^{N,x} \leq kT_0] \\
&\leq \sum_{m=0}^{k} \mathbb{P}[\tau^{N,x} = \tau_m^x] + \sum_{m=1}^{k} \mathbb{P}[\theta_m^x - \tau_{m-1}^x \leq T_0] \\
&= \mathbb{P}[\tau^{N,x} = \tau_0^x] + \sum_{m=1}^{k} \mathbb{P}[\tau^{N,x} = \tau_m^x] \\
&\quad + \sum_{m=1}^{k} \mathbb{P}[\theta_m^x - \tau_{m-1}^x \leq T_0].
\end{aligned} \tag{7.109}$$

In the following, we bound the three parts in (7.109) from above. To this end, we assume $\bar{V} > 0$ for now. The simpler case $\bar{V} = 0$ is treated below.

For the first part, we have

$$\mathbb{P}[\tau^{N,x} = \tau_0^x] = \mathbb{P}[Z_{\sigma_\rho}^{N,x} \notin O]. \tag{7.110}$$

For the second part, we use the fact that $Z^{N,x}$ is a strong Markov process and that the τ_m's are stopping times. We obtain for $m \geq 1$ and $x \in O$,

$$\mathbb{P}[\tau^{N,x} = \tau_m^x] \leq \sup_{y \in \overline{B(x^*, 3\rho) \setminus B(x^*, 2\rho)}} \mathbb{P}[Z_{\sigma_\rho}^{N,y} \notin O]. \tag{7.111}$$

Similarly, we obtain for the third part for $m \geq 1$ and $x \in O$,

$$\mathbb{P}[\theta_m^x - \tau_{m-1}^x \leq T_0] \leq \sup_{y \in O} \mathbb{P}[\sup_{t \in [0,T_0]} |Z_t^{N,y} - y| \geq \rho]. \tag{7.112}$$

The upper bounds in (7.111) and (7.112) can now be bounded by using the Lemmas 7.41 and 7.43, respectively. We fix $\delta > 0$. By Lemma 7.41 (for $C = A \setminus O$), there exists a $\rho = \rho(\delta) > 0$ and an $N_1 = N_1(\rho, \delta) > 0$ such that for all $N > N_1$,

$$\sup_{y \in \overline{B(x^*,3\rho) \setminus B(x^*,2\rho)}} \mathbb{P}[Z_{\sigma_\rho}^{N,y} \notin O] \leq \exp\big(-N(\bar{V} - \delta/2)\big). \tag{7.113}$$

By Lemma 7.43 (for $\rho = \rho(\delta)$ as above and $c = \bar{V}$), there exists a constant $T_0 = T(\rho, \bar{V}) < \infty$ and an $N_2 = N_2(\rho, \delta) > 0$ such that for all $N > N_2$,

$$\sup_{y \in O} \mathbb{P}[\sup_{t \in [0,T_0]} |Z_t^{N,y} - y| \geq \rho] \leq \exp\big(-N(\bar{V} - \delta/2)\big). \tag{7.114}$$

We now let $N > N_1 \vee N_2$ (and large enough for $T_0^{-1} \exp\big(N(\bar{V} - \delta)\big) > 1$ for the specific T_0 above). Then first by Inequalities (7.109)–(7.112), then by (7.113) and (7.114)

$$\begin{aligned} \mathbb{P}[\tau^{N,x} \leq e^{N(\bar{V}-\delta)}] &\leq \mathbb{P}[Z_{\sigma_\rho}^{N,x} \notin O] + k \sup_{y \in \overline{B(x^*,3\rho) \setminus B(x^*,2\rho)}} \mathbb{P}[Z_{\sigma_\rho}^{N,y} \notin O] \\ &\quad + k \sup_{y \in O} \mathbb{P}[\sup_{t \in [0,T_0]} |Z_t^{N,y} - y| \geq \rho] \\ &\leq \mathbb{P}[Z_{\sigma_\rho}^{N,x} \notin O] + 4T_0^{-1} \exp\big(-N\delta/2\big). \end{aligned}$$

The right-hand side of the last inequality tends to zero as $\epsilon \to 0$ by Lemma 7.42, finishing the proof for $\bar{V} > 0$.

Finally, let us assume that $\bar{V} = 0$ and that the assertion is false for a given $x \in O$. Then there exists a $\mu_0 \in (0, 1/2)$ and a $\delta_0 > 0$ such that for all $\bar{N} > 0$ there exists an $N > \bar{N}$ with

$$\mu_0 \leq \mathbb{P}[\tau^{N,x} \leq e^{-N\delta_0}].$$

We fix $\rho > 0$ such that $\overline{B(x^*, 2\rho)} \subset O$. Using the strong Markov property of Z and the fact that σ_ρ is a stopping time again, we have that for all $\bar{N} > 0$ there exists an $N > \bar{N}$ with

$$\begin{aligned} \mu_0 &\leq \mathbb{P}[\tau^{N,x} \leq e^{-N\delta_0}] \\ &\leq \mathbb{P}[Z_{\sigma_\rho}^{N,x} \notin \overline{B(x^*, \rho)}] + \sup_{y \in O} \mathbb{P}[\sup_{t \in [0,e^{-N\delta_0}]} |Z_t^{N,y} - y| \geq \rho]. \end{aligned} \tag{7.115}$$

By Lemma 7.42, there exists an N_0 such that for all $N > N_0$,

$$\mathbb{P}[Z^{N,x}_{\sigma_\rho} \notin \overline{B(x^*, \rho)}] \leq \frac{\mu_0}{2}. \tag{7.116}$$

We now set $c := -2\epsilon_0 \log \frac{\mu_0}{2}$. Then by Lemma 7.43, there exists a $T = T(c, \rho) > 0$ and an $N_1 > N_0$ such that for all $N > N_1$,

$$e^{-N\delta_0} < T \tag{7.117}$$

and

$$\sup_{y \in O} \mathbb{P}[\sup_{t \in [0,T]} |Z^{N,y}_t - y| \geq \rho] \leq e^{-Nc/2} < \frac{\mu_0}{2}. \tag{7.118}$$

Combining Inequalities (7.116)–(7.118) yields a contradiction to Inequality (7.115), finishing the proof.

Expected exit time:

We have shown in particular that $\mathbb{P}(\tau^{N,x} > e^{(\bar{V}-\delta)N}) \to 1$ as $N \to \infty$. Consequently, from Chebycheff,

$$\mathbb{E}(\tau^{N,x}) \geq e^{(\bar{V}-\delta)N}\mathbb{P}(\tau^{N,x} > e^{(\bar{V}-\delta)N}),$$

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{E}(\tau^{N,x}) \geq \bar{V} - \delta$$

for all $\delta > 0$. This together with (7.108) implies the second statement of the Theorem.

7.7.3 The Case of a Characteristic Boundary

Since we are mainly interested in studying the time of exit form the basin of attraction of one local equilibrium to that of another, we need to consider situations which do not satisfy the above assumptions. More precisely, we want to delete Assumption (D2), and keep Assumptions (D1), (D3)–(D5). We assume that there exists a collection of open sets $\{O_\rho,\ \rho > 0\}$ which is such that

- $\overline{O}_\rho \subset O$ for any $\rho > 0$.
- $d(O_\rho, \partial \tilde{O}) \to 0$, as $\rho \to 0$.
- O_ρ satisfies Assumptions (D1), (D2), (D3), (D4) and (D5) for any $\rho > 0$.

We can now establish

Corollary 7.9 *Let then O be a domain satisfying Assumptions (D1), (D3), (D4) and (D5), such that there exists a sequence $\{O_\rho,\ \rho > 0\}$ satisfying the three above conditions. Then the conclusion of Theorem 7.14 is still true.*

Proof If we define $\bar{V}_\rho$ as $\bar{V}$, but with O replaced by O_ρ, it follows from Lemma 7.38 that $\bar{V}_\rho \to \bar{V}$ as $\rho \to 0$. By an obvious monotonicity property and the continuity of the quasi-potential, the lower bound

$$\lim_{N\to\infty} \mathbb{P}\big[\tau^{N,x} > e^{(\bar{V}-\delta)N}\big] = 1$$

follows immediately from Theorem 7.14. The proof of the upper bound is done as in the proof of Theorem 7.14, once (7.107) is established. Let us now give the argument. Let $\tau_\rho^{N,x}$ denote the time of exit from O_ρ. The same argument used to establish (7.107) above leads to the statement that for any $\rho, \eta > 0$, there exists $N_{\rho,\eta}$ such that for all $N \geq N_{\rho,\eta}$,

$$\inf_{x\in O} \mathbb{P}[\tau^{N,x} \leq T] \geq e^{-N(\bar{V}+\eta)}.$$

Now utilizing (D4), (D5) and the compactness of $\bar{O}\backslash O_\rho$, it is not hard to deduce from Theorem 7.10 that for $\rho > 0$ small enough,

$$\liminf_{N\to\infty} \log \inf_{x\in(\bar{O}\backslash O_\rho)\cap A^N} \mathbb{P}_x(\tau^{N,x} \leq 1) > -\eta.$$

The wished result follows now from the last two lower bounds and the strong Markov property.

Finally the result for $\mathbb{E}(\tau^{N,x})$ now follows from the first part of the result, exactly as in the proof of Theorem 7.14.

7.8 Example: The SIRS Model

We finally show that Theorem 7.14 applies to the SIRS model from Example 7.1. Assumptions (A) and (B) have already been verified in Sect. 7.2. For (C), we note that major problems only occur for the balls centered at the "corner points" of the set A. Only for a corner point x (with corresponding vector v), we have $v \notin \mathcal{C}_x$ (recall that we define $\mathcal{C}_x$ corresponding to the modified rates β^δ). For simplicity of exposition, we concentrate on the ball B centered at $x = (1, 0)^\top$. The same argument applies to the balls centered at the other corners and in a simpler form to all other balls. For the balls B_i not centered at the corners, the vectors v_i can be represented by μ^i's for which $\mu^i_j > 0$ implies that $\beta^\delta_j(z) > \lambda > 0$ (for an appropriate constant λ which can be chosen independently of i) for all $z \in B_i$.

This simplifies the discussion below significantly. In particular, Assumption (C) is satisfied due to Theorem 7.3. We first note that for all $x \in A$, $y \in \mathbb{R}^d$, $L(x, y) = \tilde{L}(x, y)$, cf. Theorem 7.4 below. As before, we define the vector

$$v = (-1/2, 1/4)^\top \quad \text{and} \quad \mu_1 = 0, \quad \mu_2 = \frac{1}{2}, \quad \mu_3 = \frac{1}{4},$$

in particular $\mu \in \tilde{V}_v$ (but $V_{x,v} = \emptyset$!). In order to simplify the notation, we do not normalize v. We let $\eta < \eta_0 := 1/2$ and note that for $t \in [0, \eta]$,

$$\beta_2(\phi^x(t)) = \gamma\Big(1 - \frac{t}{2}\Big) \geq \frac{3}{4}\gamma, \quad \beta_3(\phi^x(t)) = \frac{\nu}{4}t.$$

Let us prove that Assumption (C) is satisfied. Let X^N be a Poisson random variable with mean μN. We note that by Theorem 7.3 for $\xi > 1$,

$$\mathbb{P}[X^N > \xi N\mu] \leq \bar{C}_1 \exp\big(-N\bar{C}_2(\xi)\big), \quad \mathbb{P}[X^N < \xi^{-1}N\mu] \leq \tilde{C}_1 \exp\big(-N\tilde{C}_2(\xi)\big) \tag{7.119}$$

for appropriate constants $\bar{C}_1$, $\tilde{C}_1$, $\bar{C}_2$, $\tilde{C}_2$ with $\bar{C}_2(\xi) = O((\xi - 1)^2)$ as $\xi \downarrow 1$ and $\tilde{C}_2(\xi) = O((1-\xi^{-1})^2)$ as $\xi \downarrow 1$. The first bound is obtained by applying Theorem 7.3 to $d = k = h_1 = 1$, $\beta_1 \equiv \mu$ and $x = 0$. We get

$$\begin{aligned}
\mathbb{P}[X^N > \xi N\mu] &= \mathbb{P}\Big[\frac{1}{N}X^N > \xi\mu\Big] \\
&= \mathbb{P}\Big[\frac{1}{N}X^N - \mu > (\xi - 1)\mu\Big] \\
&\leq \mathbb{P}\Big[\Big|\frac{1}{N}X^N - \mu\Big| > (\xi - 1)\mu\Big] \\
&= \mathbb{P}\Big[|Z^{N,0}(1) - \mu| > (\xi - 1)\mu\Big] \\
&\leq \mathbb{P}\Big[\sup_{t\in[0,1]} |Z^{N,0}(t) - \mu t| > (\xi - 1)\mu\Big].
\end{aligned}$$

Let us define the process $\hat{Z}^{N,x}$ as the solution of (7.3) with constant rates μ_j. For $\epsilon > 0$ small enough, we define

$$\begin{aligned}
X_2^{N,\epsilon} &:= \#\text{jumps of type } h_2 \text{ of } \tilde{Z}^{N,x} \text{ in } [0, \epsilon], \\
X_3^{N,\epsilon} &:= \#\text{jumps of type } h_3 \text{ of } \tilde{Z}^{N,x} \text{ in } [0, \epsilon]
\end{aligned}$$

and

$$F_2^{N,\epsilon} := \Big\{\frac{15}{32}N\epsilon < X_2^{N,\epsilon} < \frac{17}{32}N\epsilon\Big\}, \quad F_3^{N,\epsilon} := \Big\{\frac{7}{32}N\epsilon < X_3^{N,\epsilon} < \frac{9}{32}N\epsilon\Big\}.$$

hence

$$\tilde{Z}^{N,x}(\epsilon) \in \tilde{B} := \left\{ z \in A \middle| 1 - \frac{17}{32}\epsilon < z_1 < 1 - \frac{15}{32}\epsilon,\ \frac{3}{16}\epsilon < z_2 < \frac{5}{16}\epsilon \right\}$$

and

$$\sup_{t\in[0,\epsilon]} |\tilde{Z}^{N,x}(t) - \phi^x(t)| < \epsilon$$

on $F_2^{N,\epsilon} \cap F_3^{N,\epsilon}$. Furthermore, for $z \in \tilde{B}$ and $t \in [0, \eta - \epsilon]$,

$$\mathrm{dist}(\phi^z(t), \partial A) \geq \frac{3}{16}\epsilon$$

and

$$|\tilde{Z}^{N,z}(t) - \phi^z(t)| < \frac{3}{16}\epsilon \Rightarrow |\tilde{Z}^{N,z}(t) - \phi^{\tilde{x}}(t)| < \epsilon,$$

where $\tilde{x} = \phi^x(\epsilon) = (1 - \epsilon/2, \epsilon/4)^\top$. We compute by using the Markov property of Z^N,

$$\begin{aligned}
&\mathbb{P}\Big[\sup_{t\in[0,\eta]} |\tilde{Z}^{N,x}(t) - \phi^x(t)| < \epsilon\Big] \\
&\geq \mathbb{P}\Big[\sup_{t\in[0,\eta]} |\tilde{Z}^{N,x}(t) - \phi^x(t)| < \epsilon;\, F_2^{N,\epsilon} \cap F_3^{N,\epsilon}\Big] \\
&\geq \mathbb{P}\Big[F_2^{N,\epsilon} \cap F_3^{N,\epsilon}\Big] \cdot \inf_{z\in\tilde{B}} \mathbb{P}\Big[\sup_{t\in[0,\eta-\epsilon]} |\tilde{Z}^{N,z}(t) - \phi^z(t)| < \frac{3}{16}\epsilon\Big] \\
&\geq \mathbb{P}\Big[F_2^{N,\epsilon} \cap F_3^{N,\epsilon}\Big] \cdot \inf_{z\in\tilde{B}} \mathbb{P}\Big[\sup_{t\in[0,\eta-\epsilon]} |\hat{Z}^{N,z}(t) - \phi^z(t)| < \frac{3}{16}\epsilon\Big] \\
&\geq 1 - \hat{C}_1 \exp\big(-N\hat{C}_2(\epsilon)\big)
\end{aligned}$$

for appropriate constants $\hat{C}_1$, $\hat{C}_2$ with $\hat{C}_2(\epsilon) = O(\epsilon^2)$ as $\epsilon \downarrow 0$ by Theorem 7.3 and Inequalities (7.119) as required. As the rates are vanishing like polynomials, (7.6) is satisfied.

Our model has a disease-free equilibrium $(0, 0)$, and if $\beta > \gamma$ it has a stable endemic equilibrium (and then the disease-free equilibrium is unstable). We assume from now on that $\beta > \gamma$, and we seek to estimate the time it takes for the random perturbations to drive our system form the stable endemic equilibrium to the disease-free equilibrium. The characteristic boundary which we want to hit is the set $\{x = 0, 0 \leq y \leq 1\}$. We note, however, that not only the Assumption 7.13 (D2) but also the Assumption 7.13 (D5) fail to be satisfied here. Consequently we cannot apply Corollary 7.9. We will now show that if we denote by $\tau^{N,x} = \inf\{t > 0,\ Z^N(t) \in \tilde{O}\}$

where $\tilde{O} = \{z_1 = 0\}$, then Theorem 7.14 applies. All we have to show is that for any $\delta > 0$, $\bar{V}$ being defined as in Sect. 7.7,

$$\lim_{N\to\infty} \mathbb{P}\big[e^{(\bar{V}-\delta)N} < \tau^{N,x} < e^{(\bar{V}+\delta)N}\big] = 1.$$

For any $\eta > 0$, let $O_\eta = \{(z_1, z_2) \in [0, 1]^2, z_1 > \eta, z_1 + z_2 \le 1\}$, $\tilde{O}_\eta = \{z_1 = \eta, 0 \le z_2 \le 1 - \eta\}$, and $\tau_\eta^{N,x} = \inf\{t > 0,\ Z^N(t) \in \tilde{O}_\eta\}$. We note that all the above assumptions, including Assumption 7.13 (D1), (D2),..., (D5) are satisfied for this new exit problem, so that

$$\lim_{N\to\infty} \mathbb{P}\big[e^{(\bar{V}_\eta-\delta/2)N} < \tau_\eta^{N,x} < e^{(\bar{V}_\eta+\delta/2)N}\big] = 1,$$

where $\bar{V}_\eta = \inf_{z\in\tilde{O}_\eta} V(x^*, z)$. Now for $\eta_0 > 0$ such that whenever $\eta \le \eta_0$, $\bar{V} - \delta/2 \le \bar{V}_\eta < \bar{V}$. Moreover clearly $\tau_\eta^{N,x} \le \tau^{N,x}$. From these follows clearly the fact that $\mathbb{P}(\tau^{N,x} > e^{(\bar{V}-\delta)N}) \to 1$ as $N \to \infty$. It remains to establish the upper bound.

The crucial result which allows us to overcome the new difficulty is the

Lemma 7.44 *For any $\eta > 0$, $t > 0$,*

$$\liminf_{N\to\infty} \frac{1}{N} \log \inf_{x\in A^N\cap O_\eta^c} \mathbb{P}(\tau^{N,x} < t) \ge -\eta \log\left(\frac{\beta}{\gamma}\right).$$

Proof The first component of the process $Z^{N,x}(t)$ is dominated by the process

$$x_1 + \frac{1}{N} P_1\left(N\beta \int_0^t Z_1^{N,x}(s)ds\right) - \frac{1}{N} P_1\left(N\gamma \int_0^t Z_1^{N,x}(s)ds\right),$$

which is a continuous time binary branching process with birth rate β and death rate γ. This process goes extinct before time t with probability (see the formula in the middle of page 108 in [1])

$$\left(\frac{\gamma e^{N(\beta-\gamma)t} - \gamma}{\beta e^{N(\beta-\gamma)t} - \gamma}\right)^{Nx_1}.$$

The result follows readily, since $x \in O_\eta^c$ implies that $x_1 \le \eta$.

In order to adapt the proof of the upper bound in Theorem 7.14, all we have to do is to extend the proof of Lemma 7.39 to the time of extinction in the SIRS model, which we now do. Indeed, from Lemma 7.39, for any $\eta, \delta, \rho > 0$, there exists T_0

such that

$$\liminf_{N\to\infty}\frac{1}{N}\log\inf_{x\in\overline{B(x^*,\rho)}}\mathbb{P}(\tau_\eta^{N,x}\le T_0-1)>-(\bar{V}+\delta/2).$$

On the other hand, from Lemma 7.44, provided $\eta<\frac{\delta}{2\log(\beta/\gamma)}$,

$$\liminf_{N\to\infty}\frac{1}{N}\log\inf_{x\in A^N\cap O_\eta^c}\mathbb{P}(\tau^{N,x}<1)\ge-\delta/2.$$

The statement of Lemma 7.39 now follows from the strong Markov property and the last two estimates.

Acknowledgements The authors thank an anonymous Referee, whose careful reading and detailed remarks allowed us to improve an earlier version of this paper.

This research was supported by the ANR project MANEGE, the DAAD, and the Labex Archimède.

Appendix: Change of Measure

We assume that $Z^{N,x}=Z^N$ has rates $\{N\beta_j|j=1,\dots,k\}$ under $\mathbb{P}$ and rates $\{N\tilde{\beta}_j|j=1,\dots,k\}$ under $\tilde{\mathbb{P}}$. We furthermore assume that for $x\in A$,

$$\tilde{\beta}_j(x)>0 \text{ only if } \beta_j(x)>0.$$

Hence, $\tilde{\mathbb{P}}|_{\mathcal{F}_t}$ is absolutely continuous with respect to $\mathbb{P}|_{\mathcal{F}_t}$ but not necessarily vice versa.

We require Theorem B.6 of [13] which gives us an important change of measure formula.

Theorem 7.15 *For all $T>0$ and non-negative, $\mathcal{F}_T$-measurable random variables X, we have*

$$\mathbb{E}[\xi_T X]=\tilde{\mathbb{E}}[X],$$

where $\tilde{\mathbb{E}}$ denotes the expectation with respect to $\tilde{\mathbb{P}}$ and

$$\xi_T:=\exp\Big(\sum_\tau\Big[\log\big(\tilde{\beta}_{j(\tau)}(Z^N(\tau-))\big)-\log\big(\beta_{j(\tau)}(Z^N(\tau-)\big)\Big]\\ -N\sum_j\int_0^T\big(\tilde{\beta}_j(Z^N(t))-\beta_j(Z^N(t))\big)dt\Big);$$

here, we sum over the jump times $\tau \in [0, T]$ of Z^N; $j(\tau)$ denotes the corresponding type of the jump direction. In other words, we have

$$\xi_T = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\Big|_{\mathcal{F}_T}.$$

We observe that (under $\mathbb{P}$) $\xi_T = 0$ if and only if there exists a jump time $\tau \in [0, T]$ (with jump type $j(\tau)$) and $\tilde{\beta}_{j(\tau)}(Z^N(\tau-)) = 0$.

We deduce the following result. Note that since $\tilde{\mathbb{P}}[\xi_T = 0] = 0$, ξ_T^{-1} is well-defined $\tilde{\mathbb{P}}$-almost surely.

Corollary 7.10 *For every non-negative measurable function $X \geq 0$,*

$$\mathbb{E}[X] \geq \tilde{\mathbb{E}}[\xi_T^{-1} X].$$

Proof As $X \geq 0$, we have

$$\mathbb{E}[X] \geq \mathbb{E}[X \mathbb{1}_{\{\xi_T \neq 0\}}] = \tilde{\mathbb{E}}[X \mathbb{1}_{\{\xi_T \neq 0\}} \xi_T^{-1}] = \tilde{\mathbb{E}}[X \xi_T^{-1}].$$

References

1. K.B. Athreya, P.E. Ney, *Branching Processes* (Springer, New York, 1972)
2. P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1999)
3. A. Dembo, O. Zeitouni, *Large Deviations Techniques and Applications* (Springer, Berlin, 2009)
4. P. Dupuis, R.S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations* (Wiley, New York, 1997)
5. P. Dupuis, R.S. Ellis, A. Weiss, Large deviations for Markov processes with discontinuous statistics. I. General upper bounds. Ann. Probab. **19**(3), 1280–1297 (1991)
6. J. Feng, T.G. Kurtz, *Large Deviations for Stochastic Processes* (American Mathematical Society, Providence, 2006)
7. M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, Berlin, 2012)
8. P. Kratz, E. Pardoux, B.S. Kepgnou, Numerical methods in the context of compartmental models in epidemiology. ESAIM Proc. **48**, 169–189 (2015)
9. T.G. Kurtz, Strong approximation theorems for density dependent Markov chains. Stoch. Process. Appl. **6**(3), 223–240 (1978)
10. K. Pakdaman, M. Thieullen, G. Wainrib, Diffusion approximation of birth-death processes: comparison in terms of large deviations and exit points. Stat. Probab. Lett. **80**(13–14), 1121–1127 (2010)
11. D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion* (Springer, Berlin, 2005)
12. H. Roydon, *Real Analysis* (Collier-Macmillan, London, 1968)
13. A. Shwartz, A. Weiss, *Large Deviations for Performance Analysis* (Chapman Hall, London, 1995)
14. A. Shwartz, A. Weiss, Large deviations with diminishing rates. Math. Oper. Res. **30**(2), 281–310 (2005)
15. M. Sion, On general minimax theorems. Pac. J. Math. **8**, 171–176 (1958)

Chapter 8
The Girsanov Theorem Without (So Much) Stochastic Analysis

Antoine Lejay

Abstract In this pedagogical note, we construct the semi-group associated to a stochastic differential equation with a constant diffusion and a Lipschitz drift by composing over small times the semi-groups generated respectively by the Brownian motion and the drift part. Similarly to the interpretation of the Feynman-Kac formula through the Trotter-Kato-Lie formula in which the exponential term appears naturally, we construct by doing so an approximation of the exponential weight of the Girsanov theorem. As this approach only relies on the basic properties of the Gaussian distribution, it provides an alternative explanation of the form of the Girsanov weights without referring to a change of measure nor on stochastic calculus.

Keywords Girsanov theorem · Lie-Trotter-Kato formula · Feynman-Kac formula · Stochastic differential equation · Euler scheme · Splitting scheme · Flow · Heat equation · Cameron-Martin theorem

8.1 Introduction

This pedagogical paper aims at presenting the Girsanov theorem—a change of measure for the Brownian motion—using the point of view of operator analysis. We start from the sole knowledge of the Brownian distribution and its main properties. By doing so, stochastic calculus is avoided excepted for identifying the limit. Therefore, in a simplified context, we give an alternative proof of a result which is usually stated and proved using stochastic analysis and measure theory.

In 1944, R.H. Cameron and W.T. Martin proved the celebrated theorem on the change of the Wiener measure. They later extend it [10, 11].

A. Lejay (✉)
Université de Lorraine, CNRS, Inria, IECL, Nancy, France
e-mail: Antoine.Lejay@univ-lorraine.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_8

Theorem 8.1 (Cameron and Martin [9]) *Let F be a continuous functional on the space of continuous functions* $\mathrm{C}([0,1];\mathbb{R})$ *with respect to the uniform norm. Let b be a continuous function in* $\mathrm{C}([0,1];\mathbb{R})$ *whose derivative b' has bounded variation. Then for a Wiener process (or a Brownian motion)*[1] W,

$$\mathbb{E}[F(\mathsf{W})] = \mathbb{E}\left[F(\mathsf{W}+b)\exp\left(-\int_0^1 b'(s)\,\mathrm{d}\mathsf{W}_s - \frac{1}{2}\int_0^1 |b'(s)|^2\,\mathrm{d}s\right)\right].$$

Later in 1960, Girsanov states in [20] a variant of this theorem for solutions of stochastic differential equations. Even when the diffusivity σ is constant, the drift is itself a non-linear functional of the Brownian motion.

Theorem 8.2 (Girsanov [20]) *Let* W *be a n-dimensional Wiener process on a probability space* $(\Omega,\mathscr{F},\mathbb{P})$ *with respect to a filtration* $(\mathscr{F}_t)_{t\geq 0}$. *Let* X *be the solution on* $(\Omega,\mathbb{P})$ *to*

$$\mathsf{X}_t(\omega) = \mathsf{X}_0(\omega) + \int_0^t a(s,\omega)\,\mathrm{d}\mathsf{W}(s,\omega) + \int_0^t b(s,\omega)\,\mathrm{d}s$$

where a is matrix valued, b is vector valued, and

(A1) *The applications a and b are measurable with respect to* $(s,\omega)\in[0,1]\times\Omega$.
(A2) *For each $t\geq 0$, a is $\mathscr{F}_t$-measurable.*
(A3) *Almost everywhere,*[2] $\int_0^1 \|a(t,\omega)\|^2\,\mathrm{d}t < +\infty$.
(A4) *Almost everywhere,* $\int_0^1 |b(t,\omega)|\,\mathrm{d}t < +\infty$.

Let $\phi = (\phi^1,\dots,\phi^n)$ be a vector-valued function on $[0,1]\times\Omega$ *such that (A1)-(A3) are satisfied.*

Let us set $\widetilde{\mathbb{P}}[\mathrm{d}\omega] = \exp(\mathsf{Z}_0^1(\phi,\omega))\mathbb{P}[\mathrm{d}\omega]$ *where*

$$\mathsf{Z}_s^t(\phi,\omega) = \int_s^t \phi^i(u,\omega)\delta_{ij}\,\mathrm{d}\mathsf{W}^j(u,\omega) - \frac{1}{2}\int_0^t\left(\sum_{i=1}^n \phi^i(u,\omega)^2\right)\mathrm{d}u.$$

Let us also set $\widetilde{\mathsf{W}}(t,\omega) = \mathsf{W}(t,\omega) - \int_0^t \phi(s,\omega)\,\mathrm{d}s$. *If* $\widetilde{\mathbb{P}}[\Omega]=1$, *then* $\widetilde{\mathsf{W}}$ *is a Wiener process with respect to* $(\mathscr{F}_t)_{t\geq 0}$ *on* $(\Omega,\mathscr{F},\widetilde{\mathbb{P}})$ *and* $(\mathsf{X},\widetilde{\mathbb{P}})$ *is solution to*

$$\mathsf{X}_t(\omega) = \mathsf{X}_0(\omega) + \int_0^t a(s,\omega)\,\mathrm{d}\mathsf{W}_s(\omega) + \int_0^t (b(s,\omega) + a(s,\omega)\phi(s,\omega))\,\mathrm{d}s.$$

[1] In the original paper [9], the result is stated for $2^{-1/2}\mathsf{W}$ and $2^{1/2}b$.

[2] The norm of a matrix is $\|a\| = \left(\sum_{i,j=1}^n |a_{i,j}^2|\right)^{1/2}$ while the norm of a vector is $\|a\| = \left(\sum_{i=1}^n |a_i|^2\right)^{1/2}$.

Soon after, these results were extended in many directions, for example to deal with semi-martingales (see e.g. [51]). The study of the weights for the change of measures and their exponential nature gives rise to the theory of Doléans-Dade martingales [14, 35].

These theorems provide us with measures which are equivalent to the Wiener's one. The converse is also true [42]: Absolute continuity of Wiener or diffusions measures can only be reached by adding terms of bounded variation [42, 46].

The Cameron-Martin and Girsanov theorems have a profound meaning as well as a deep impact on modern stochastic calculus. For example, for example, they are one of the cornerstones of Malliavin calculus [6, 36], likewise a major tool in filtering, statistics of diffusion processes [33, 34], mathematical finance [26], ...

The Girsanov theorem has also been extended to some Gaussian processes, including the fractional Brownian motion [18, 24].

The Feynman-Kac formula related the Brownian motion with some PDEs. It involves a probabilistic representation with an exponential weight (see Sect. 8.9.1). The Feynman-Kac formula could be proved by many ways, including stochastic calculus from one side and the Trotter-Kato-Lie formula on the other side (see [17, Chap. 3, Sect. 5] or [12] for a nice introduction to this subject, and [21] for a proof of the Feynman-Kac formula with this procedure).

To illustrate the latter approach, let us consider three linear matrix-valued equations

$$\dot{X} = \mathfrak{A}X,\ \dot{Y} = \mathfrak{B}Y,\ \dot{R} = (\mathfrak{A} + \mathfrak{B})R \text{ with } X_0 = Y_0 = R_0 = \mathrm{Id},$$

where $\mathfrak{A}$ and $\mathfrak{B}$ are $d \times d$-matrices. These equations are easily solved by

$$X_t = \exp(t\mathfrak{A}),\ Y_t = \exp(t\mathfrak{B}) \text{ and } R_t = \exp(t(\mathfrak{A} + \mathfrak{B})).$$

The solutions X, Y and R satisfy the *semi-group property*: $X_{t+s} = X_t X_s$. If $\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}$, then $R_t = Y_t X_t$. This is no longer true in general. However, as shown first by Lie [32], the solution R could be constructed from X and Y by the following limit procedure:

$$R_t = \exp(t(\mathfrak{A}+\mathfrak{B})) = \lim_{n\to\infty} X_{t/n}Y_{t/n}X_{t/n}Y_{t/n}\cdots X_{t/n}Y_{t/n} = \lim_{n\to\infty}(e^{\mathfrak{A}t/n}e^{\mathfrak{B}t/n})^n.$$

Trotter [50] and Kato [28] have shown that this could be generalized for large families of linear unbounded operators $\mathfrak{A}$ and $\mathfrak{B}$. In this case, $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are families on linear operators with the semi-groups property.

On the space of continuous, bounded functions $\mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$, consider the (scaled) Laplace operator $\mathfrak{A} = \frac{1}{2}\triangle$ and let $\mathfrak{B}$ be defined by $\mathfrak{B}f(x) = U(x)f(x)$ for any $f \in \mathrm{C}(\mathbb{R}^d, \mathbb{R})$, where the continuous function U is called a *potential*. For a Brownian motion B, $X_t = \mathbb{E}[f(x + \mathsf{B}_t)]$ whereas $Y_t f(x) = \exp(tU(x))f(x)$ for any $f \in \mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$, $x \in \mathbb{R}^d$ and $t \geq 0$.

With the Trotter-Kato-Lie formula, we compose over short times the semi-group X of the Laplace operator with the one Y of the potential term. Using the Markov property of the Brownian motion, for any bounded, measurable function f,

$$R_t f(x) = \lim_{n\to\infty} \mathbb{E}\left[\exp\left(\frac{t}{n}\sum_{i=0}^{n-1} U(x+\mathsf{B}_{it/n})\right) f(x+\mathsf{B}_t)\right]. \tag{8.1}$$

With the right integrability conditions on U, we obtain in the limit an exponential representation of the Feynman-Kac formula [27], that is

$$R_t f(x) = \mathbb{E}\left[\exp\left(\int_0^t U(x+\mathsf{B}_s)\,\mathrm{d}s\right) f(x+\mathsf{B}_t)\right],$$

which gives a probabilistic representation to the PDE

$$\partial_t R_t f(x) = \frac{1}{2}\triangle R_t f(x) + U(x) R_t f(x) \text{ with } R_0 f(x) = f(x).$$

The seminal derivation of the Feynman-Kac formula by Kac in [27] used an approximation of the Brownian motion by a random walk, which leads to an expression close to (8.1).

What happens now if we use for $\mathfrak{B}$ the first-order differential operator $b\nabla\cdot$ for a function b? This means that we consider giving a probabilistic representation of the semi-group $(R_t)_{t\geq 0}$ related to the PDE

$$\partial_t R_t f(x) = \frac{1}{2}\triangle R_t f(x) + b(x)\nabla R_t f(x) \text{ with } R_0 f(x) = f(x).$$

Of course, a probabilistic representation is derived by letting X_t be the solution of the SDE $\mathsf{X}_t = x + \mathsf{B}_t + \int_0^t b(\mathsf{X}_s)\,\mathrm{d}s$ and $R_t f(x) = \mathbb{E}[f(\mathsf{X}_t)]$. With the Girsanov theorem (Theorem 8.2), $R_t f(x) = \mathbb{E}[\mathsf{Z}_t f(x+\mathsf{B}_t)]$ for the Girsanov weight Z given by (8.5) below.

The Feynman-Kac formula is commonly understood as a byproduct of the Trotter-Kato-Lie formula [22], at least in the community of mathematical physics in relation with the Schrödinger equation [40, 45]. The Trotter-Kato-Lie formula also offers a simple interpretation, leading to *splitting procedures* that provide explicit construction of numerical schemes (see e.g., [8, 23], among others).

Splitting schemes have been proposed to solve SDE.[3] Surprisingly enough, we found no trace where the Girsanov theorem is presented as a by-product of the

[3]In the domain of SDE, among others, the Ninomiya-Victoir scheme [41] relies on an astute way to compose the operators.

Trotter-Kato-Lie formula[4] Yet its probabilistic interpretation is very simple: we combine Brownian evolution over a short time with a transport equation through the ODE $\dot{\mathsf{X}} = b(\mathsf{X})$ over a short time, and so on...

A heuristic argument shows how the exponential weight appears. For a continuous, bounded function f, the semi-group solution $(P_t)_{t\geq 0}$ such that $\partial_t P_t f(x) = \frac{1}{2}\triangle P_t f(x)$ is given by $P_t f(x) = \mathbb{E}[f(x+\mathsf{B}_t)]$. The one $(Q_t)_{t\geq 0}$ that gives the solution to $\partial_t Q_t f(x) = b(x)\nabla Q_t f(x)$ is for $f \in \mathrm{C}^1_{\mathrm{b}}(\mathbb{R}^d, \mathbb{R})$

$$Q_t f(x) = f(\mathsf{Y}_t(x)) = f(x) + tb(x)\nabla f(x) + \mathrm{O}(t^2) \text{ with } \dot{\mathsf{Y}}_t = b(\mathsf{Y}_t) \text{ and } \mathsf{Y}_0 = x. \tag{8.2}$$

Thus, in short time,

$$\begin{aligned} P_t Q_t f(x) &= \mathbb{E}[f(x+\mathsf{B}_t) + tb(x+\mathsf{B}_t)\nabla f(x+\mathsf{B}_t)] + \mathrm{O}(t^2) \\ &= \mathbb{E}[f(x+\mathsf{B}_t) + tb(x)\nabla f(x+\mathsf{B}_t)] \\ &\quad +\mathbb{E}[t(b(x+\mathsf{B}_t) - b(x)\nabla f(x+\mathsf{B}_t))] + \mathrm{O}(t^2). \end{aligned} \tag{8.3}$$

Since the Gaussian density $p(t,x) = \exp(-x^2/2t)/\sqrt{2\pi t}$ satisfies $\partial_x p(t,x) = -\frac{x}{t}p(t,x)$, an integration by parts implies that[5]

$$\begin{aligned} \mathbb{E}[\nabla f(x+\mathsf{B}_t)] &= \int p(t,y)\partial_y f(x+y)\,\mathrm{d}y = \int \frac{y}{t}p(t,y)f(x+y)\,\mathrm{d}y \\ &= \mathbb{E}\left[\frac{\mathsf{B}_t}{t}f(x+\mathsf{B}_t)\right] \end{aligned}$$

so that

$$\mathbb{E}[tb(x)\nabla f(x+\mathsf{B}_t)] = \mathbb{E}\left[\mathsf{B}_t b(x) f(x+\mathsf{B}_t)\right].$$

Since b is Lipschitz, the term involving $b(x+\mathsf{B}_t) - b(x)$ in (8.3) is considered as a higher order term, roughly of order $\mathrm{O}(t^{3/2})$ since B_t is roughly of order $\mathrm{O}(\sqrt{t})$ and it is multiplied by t. Thus,

$$P_t Q_t f(x) = \mathbb{E}\left[f(x+\mathsf{B}_t) + \mathsf{B}_t b(x) f(x+\mathsf{B}_t)\right] + \mathrm{O}(t^{3/2}).$$

[4]In [31], R. Léandre gives an interpretation of the Girsanov formula and Malliavin calculus in terms of manipulation on semi-groups.

[5]This is one of the central ideas of Malliavin calculus to express the expectation involving the derivative of a function as the expectation involving the function multiplied by a weight.

Again owing to the regularity of the Brownian motion since $(\mathsf{B}_t^2 - t)_{t\geq 0}$ is a martingale, we get

$$P_t Q_t f(x) = \mathbb{E}\left[\exp\left(b(x)\mathsf{B}_t - \frac{t}{2}b(x)^2\right) f(x+\mathsf{B}_t)\right] + \mathrm{O}(t^{3/2})$$

to take into account the Taylor expansion of the exponential at order 2. The exponential weight is an approximation of the Girsanov weight in short time so that a similar analysis may be performed to show that $P_t Q_t f(x) = \mathbb{E}[\mathsf{Z}_t f(x + \mathsf{B}_t)] + \mathrm{O}(t^{3/2})$.

We also see that limit of our approach, as it requires both f to be $\mathrm{C}_\mathrm{b}^1(\mathbb{R}^d, \mathbb{R})$ and b to be Lipschitz. Such a restriction does not hold with the classical approach.

Using the above set of ideas, we then aim at proving rigorously the Girsanov theorem in a restricted setting, by considering only the solution of the stochastic differential equation

$$\mathsf{X}_t = x + \mathsf{B}_t + \int_0^t b(\mathsf{X}_s)\,\mathrm{d}s$$

for a drift $b \in \mathrm{C}_\mathrm{b}^1(\mathbb{R}^d, \mathbb{R}^d)$ and a d-dimensional Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. As a rule, we try to stuck on functional analysis arguments, and not on stochastic calculus to see how far we could go. When X is the canonical process,[6] we show that for any $f \in \mathrm{C}_\mathrm{b}^1(\mathbb{R}^d, \mathbb{R})$,

$$\widehat{\mathbb{E}}[f(\mathsf{X}_t)] = \mathbb{E}[\mathsf{Z}_t f(\mathsf{X}_t)] \tag{8.4}$$

$$\text{with } \mathsf{Z}_t = \exp\left(\int_0^t b(\mathsf{X}_s)\,\mathrm{d}\mathsf{B}_s - \frac{1}{2}\int_0^t |b(\mathsf{X}_s)|^2\,\mathrm{d}s\right) \tag{8.5}$$

$$= 1 + \int_0^t \mathsf{Z}_s b(\mathsf{X}_s)\,\mathrm{d}\mathsf{B}_s, \tag{8.6}$$

where $\widehat{\mathbb{E}}$ is the expectation of the distribution $\widehat{\mathbb{P}}$ of X, and $\mathbb{P}$ the one of the Brownian motion. It is only at the last stage that stochastic calculus is used, to give the expression for Z by passing to the limit. All the other arguments come from functional analysis or simple computations on the Gaussian density.

In the course of events, we obtain an upper bound for the weak rate of convergence of the Euler scheme with a drift coefficient of class $\mathrm{C}_\mathrm{b}^1(\mathbb{R}^d, \mathbb{R}^d)$, as well as some insight on the exponential nature of the Girsanov weight.

We actually gives two derivations. One is based on the Euler scheme and could be used without reference to the splitting procedure. It was used, with a different formulation, by G. Maruyama in 1954 to show the differentiability of the transition

[6] This is, $\mathsf{X}_t(\omega) = \omega(t)$ when the probability space Ω is $\mathrm{C}(\mathbb{R}_+, \mathbb{R})$.

densities of SDE [37]. This approach now seems to be part of the "folklore". The second one is an application of the Trotter-Kato-Lie using the heuristic given above. Both procedures complement each others and gives two alternative discretization of the Girsanov weights.

Our hypotheses are a priori more restrictive than the full Girsanov theorem. Actually,

- Using the Markov property of Z and X, it is possible to show that

  $$\mathbb{E}[\mathsf{Z}_{t_2} f(\mathsf{X}_{t_2}) f(\mathsf{X}_{t_1})] = \mathbb{E}\big[\mathbb{E}_{\mathsf{X}_{t_1}}[\mathsf{Z}_{t_2,t_1} f(\mathsf{X}_{t_2-t_1})]\mathsf{Z}_{t_1} f(\mathsf{X}_{t_1})\big] = \widehat{\mathbb{E}}[f(\mathsf{X}_{t_2}) f(\mathsf{X}_{t_1})]$$

  for $0 \leq t_1 \leq t_2$, and so on. A limiting argument shows that the transform is valid for functionals of the Brownian motion, not only for the marginal distribution.
- The restriction that $f \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$ in (8.4) is easily removed by using a sequence of smooth approximations of a function f which is only continuous.
- The Girsanov theorem assumes nothing on the regularity of the drift and requires only integrability condition. Here, the hypothesis that b is $\mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R}^d)$ (or at least Lipschitz continuous) is crucial for our analysis. Yet any drift could be reached by regularizing the drift both in the PDE and the expression of the change of measure. However, non-anticipative functional drifts cannot be treated by this approach. We comment this further in Sect. 8.10.
- We use the independence of the increments of the Brownian motion as well as the explicit expression of its density (the heat or Gaussian kernel) and its derivative. Considering a diffusion coefficient σ which is smooth enough, we could apply to approach to

  $$\mathsf{X}_t = x + \int_0^t \sigma(\mathsf{X}_s)\,\mathrm{d}\mathsf{B}_s,$$

  where the Markov property is used, as well as Gaussian controls over the density of the process and its derivatives, as well as expressions over the derivative of X_t using Malliavin calculus. It requires fine stochastic and analytic tools while our aim is to be as basic as possible, we use only the Brownian motion seen as a process with independent, Gaussian increments.
- Time could be added in b, provided that b is uniformly Lipschitz continuous in space with respect to the time and uniformly bounded.

Finally, let us mention an alternative derivation of the Girsanov theorem that relies mostly on algebraic manipulations on the exponential weight [1]. This approach also avoids stochastic calculus as much as they can. Yet more sophisticated tools than ours are used.

Outline

The semi-groups are introduced in Sect. 8.2. The main estimates on elementary steps are given in Sect. 8.3. The convergence of the semi-groups associated to the Euler scheme is shown in Sect. 8.4, while the ones for the splitting procedure is shown in Sect. 8.5. The limits of the weights are identified in Sect. 8.7, leading to our proof of the Girsanov theorem. The infinitesimal generator of the SDE with drift is identified in Sect. 8.8 by the sole use of functional analysis. At last, to complement our idea to avoid stochastic calculus, we present in Sect. 8.9 a representation of the Girsanov weights that does not involve stochastic integrals for some special form the of drift. Finally, we discuss in Sect. 8.10 the difference between SDEs and ODEs. The article ends with two appendices.

8.2 Brownian Motion and Flows

For a dimension d, let B be a d-dimensional Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The drift b we consider belongs to $\mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$, the space of bounded, continuous functions with a bounded, continuous first order derivatives.

We set for convenience $\mathscr{B}_{t,s}(x, \omega) = x + \mathsf{B}_t(\omega) - \mathsf{B}_s(\omega)$ for any $x \in \mathbb{R}^d$, $t \geq s \geq 0$ and $\omega \in \Omega$. This family $(\mathsf{B}_{t,x}(x, \omega))_{t \geq s \geq 0}$ is called a *random dynamical system*. Clearly,

$$\mathscr{B}_{t,s}(\mathscr{B}_{s,r}(x, \omega), \omega) = \mathscr{B}_{t,r}(x, \omega) \text{ for any } t \geq s \geq r \geq 0 \text{ and any } x \in \mathbb{R}^d. \tag{8.7}$$

This property (8.7) is called the *flow property*.

For any $s \geq 0$, let us define $(\mathscr{X}_{t,s}(x, \omega))_{t \geq s}$ as the unique solution to

$$\begin{aligned} \mathscr{X}_{t,s}(x, \omega) &= x + \mathsf{B}_t(\omega) - \mathsf{B}_s(\omega) + \int_s^t b(\mathscr{X}_{r,s}(x, \omega))\,\mathrm{d}r \\ &= \mathscr{B}_{t,s}(x, \omega) + \int_s^t b(\mathscr{X}_{r,s}(x, \omega))\,\mathrm{d}r,\ t \geq s. \end{aligned} \tag{8.8}$$

Actually, there is no need of a theory of stochastic differential equation for this, so that $\mathscr{X}_{\cdot,s}(x, \omega)$ could be defined pathwise for any Brownian path $\mathsf{B}(\omega)$ of the Brownian motion B.

Proposition 8.1 *For every ω in Ω and any starting point $x \in \mathbb{R}^d$, there exists a unique solution $\mathscr{X}(x, \omega)$ to (8.8) defined on $[s, +\infty)$.*

Proof Let us consider $\mathscr{U}_{t,s}(x, \omega) = \mathscr{X}_{t,s}(x, \omega) - \mathscr{B}_{t,s}(x, \omega)$ and set $\beta(t, y, \omega) = b(\mathscr{B}_{t,s}(x) + y)$ so that $b(\mathscr{X}_{t,s}(x, \omega)) = \beta(t, \mathscr{U}_{t,s}(x, \omega), \omega)$. This way, $\mathscr{X}_{\cdot,s}(x, \omega)$

is solution to (8.8) if and only if $\mathscr{U}_{\cdot,s}(x,\omega)$ is solution to

$$\mathscr{U}_{t,s}(x,\omega) = \int_s^t \beta(r, \mathscr{U}_{r,s}(x,\omega), \omega)\,\mathrm{d}r,\ t \geq s. \tag{8.9}$$

For any $s \geq 0$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$, Eq. (8.9) has a unique solution, since β is bounded in time and Lipschitz continuous in space. Then, (8.8) has necessarily a unique solution as the map which transforms $\mathscr{X}$ to $\mathscr{U}$ is one-to-one. □

To simplify our notations, we drop from now any reference to the event $\omega \in \Omega$, which is implicit.

From the uniqueness of the solution to (8.8), the family $(\mathscr{X}_{t,s})_{t\geq s\geq 0}$ satisfies the flow property:

$$\mathscr{X}_{t,s}(\mathscr{X}_{s,r}(x)) = \mathscr{X}_{t,s}(x) \text{ for } 0 \leq r \leq s \leq t.$$

For $x \in \mathbb{R}^d$, we also consider the solution to the one-dimensional ODE

$$\mathscr{Y}_{t,r}(x) = x + \int_r^t b(\mathscr{Y}_{s,r}(x))\,\mathrm{d}s,\ t \geq r.$$

This family also satisfies the flow property: $\mathscr{Y}_{t,r}(x) = \mathscr{Y}_{t,s}(\mathscr{Y}_{s,r}(x))$ for any $0 \leq r \leq s \leq t$ and any $x \in \mathbb{R}^d$. As b is time-homogeneous, $\mathscr{Y}_{t,s}(x) = \mathscr{Y}_{t-s}(x)$ for any $0 \leq s \leq t$.

Let us introduce several families of linear operators on $\mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$ to $\mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$ defined by

$$\begin{aligned}
X_{s,t}g(x) &= \mathbb{E}[g(\mathscr{X}_{t,s}(x))],\\
P_{s,t}g(x) &= \mathbb{E}[g(\mathscr{B}_{t,s}(x))],\\
Q_{s,t}g(x) &= g(\mathscr{Y}_{t,s}(x)),\\
V_{s,t}g(x) &= \mathbb{E}[(1 + b(x)(\mathsf{B}_t - \mathsf{B}_s))g(\mathscr{B}_{t,s}(x))],\\
E_{s,t}g(x) &= \mathbb{E}[g(\mathscr{B}_{t,s}(x) + (t-s)b(x))],
\end{aligned}$$

for $g \in \mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$ and any $0 \leq s \leq t$.

The time indices of X have been inverted with respect to the ones of $\mathscr{X}$. The same convention holds for the other operators. The reason is the following: as the increments $\mathsf{B}_t - \mathsf{B}_s$ and $\mathsf{B}_s - \mathsf{B}_r$ are independent,

$$\mathbb{E}[g(\mathscr{X}_{t,s}(\mathscr{X}_{r,s}(x))] = \mathbb{E}[X_{s,t}g(\mathscr{X}_{r,s}(x))] = X_{r,s}X_{s,t}g(x).$$

Acting the same for the other operators,

$$X_{r,s}X_{s,t} = X_{r,t},\ Q_{r,s}Q_{s,t} = Q_{r,t} \text{ and } P_{r,s}P_{s,t} = P_{r,t}. \tag{8.10}$$

Besides, $\mathscr{X}_{s,s}(x) = x$ so that $X_{s,s} = \mathrm{Id}$. The same holds for the pairs $(\mathscr{Y}, Q)$ and $(\mathscr{B}, P)$. This means that $(X_{r,t})_{0\leq r\leq t}$, $(Q_{r,t})_{0\leq r\leq t}$ and $(P_{r,t})_{0\leq r\leq t}$ are *non-homogeneous semi-groups*.[7] Their infinitesimal generators are

$$\mathfrak{B} := b\nabla \text{ for } (Q_{r,t})_{0\leq r\leq t},\ \mathfrak{L} := \frac{1}{2}\triangle \text{ for } (P_{r,t})_{0\leq r\leq t} \text{ and } \mathfrak{A} := \mathfrak{L} + \mathfrak{B} \text{ for } (X_{r,t})_{0\leq r\leq t}$$

with the appropriate domains (see Sect. 8.8).

8.3 Some Estimates

The next lemma is central for our analysis. Its proof is immediate so that we skip it.

Lemma 8.1 *For bounded, linear operators A_i and B_i, $i = 1, \dots, n$,*

$$\prod_{i=1}^{n}(A_i + B_i) - \prod_{i=1}^{n} A_i = \sum_{j=1}^{n}\left(\prod_{i=1}^{j-1}(A_i + B_i)\right) B_j \left(\prod_{i=j+1}^{n} A_i\right).$$

Actually, such a lemma just follows from algebraic manipulations which are valid in any Banach algebra. In the context of semi-group, it can be seen as a discrete version of the perturbation formula [17, 43] which is central for the analysis of semigroups. This perturbation is also related to the so-called *parametrix method* for constructing densities (see e.g., [4] for an interpretation in the stochastic context).

8.3.1 Differentiability of the Flows

We recall here a classical result about the differentiability of the flow. A formal proof is given for the sake of clarity.

To simplify the notations, we set

$$\beta = \|\nabla b\|_\infty.$$

Proposition 8.2 *For any $0 \leq s \leq t$, the maps $x \mapsto \mathscr{Y}_{t,s}(x)$ and $x \mapsto \mathscr{X}_{t,s}(x)$ are differentiable. Besides,*

$$\|\nabla \mathscr{Y}_{t,s}(x)\| \leq \exp((t-s)\beta) \text{ and } \|\nabla \mathscr{X}_{t,s}(x)\| \leq \exp((t-s)\beta),\ \forall t \geq s \geq 0. \tag{8.11}$$

[7]These semi-groups are actually time-homogeneous. We however found it more convenient to keep the time dependence for our purpose.

Proof For $i = 1, \dots, d$ and $\mathscr{Y}_{t,s}(x) = (\mathscr{Y}^1_{t,s}(x), \dots, \mathscr{Y}^d_{t,s}(x))$,

$$\partial_{x_i} \mathscr{Y}^j_{t,s}(x) = \delta_{i,j} + \int_s^t \sum_{k=1}^n \partial_{x_k} b_j(\mathscr{Y}_{s,r}(x)) \partial_i \mathscr{Y}^k_{s,r}(x)\, \mathrm{d}r.$$

After having considered formally this equation, it is then possible to show that $\nabla \mathscr{Y}_{t,s}(x)$ is really the derivative of $\mathscr{Y}_{t,s}(x)$.

By identifying $\nabla \mathscr{Y}_{t,s}(x)$ with a matrix $(\partial_{x_i} Y^j_{t,s}(x))_{i,j=1,\dots,d}$,

$$\|\nabla \mathscr{Y}_{t,s}(x)\| \leq 1 + \int_s^t \beta \|\nabla \mathscr{Y}_{r,s}(x)\|\, \mathrm{d}r$$

so that (8.11) for $\nabla \mathscr{Y}_{t,s}(x)$ follows from the Grönwall lemma. Similarly,

$$\partial_{x_i} \mathscr{X}^j_{t,s}(x) = \delta_{i,j} + \int_s^t \sum_{k=1}^n \partial_{x_k} b_j(\mathscr{X}_{r,s}(x)) \partial_i \mathscr{X}^k_{r,s}(x)\, \mathrm{d}r$$

so that a control similar to the one on $\nabla \mathscr{Y}_t$ holds for $\nabla \mathscr{X}_{t,s}$. □

The control over the derivative of the flow is then transferred as a control on the semi-groups.

Corollary 8.1 *For any $t \geq s \geq 0$ and $g \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$,*

$$\|\nabla X_{s,t} g\|_\infty \leq \exp(\beta(t-s)) \|\nabla g\|_\infty.$$

Proof For any $0 \leq s \leq t$, the chain rule implies that

$$\nabla X_{s,t} g(x) = \nabla \mathbb{E}[g(\mathscr{X}_{t,s}(x))] = \mathbb{E}[\nabla \mathscr{X}_{t,s}(x) \nabla g(\mathscr{X}_{t,s}(x))].$$

The result stems from (8.11). □

8.3.2 The Heat Semi-group

Being associated to the *heat equation*, we call $(P_{s,t})_{0 \leq s \leq t}$ the *heat semi-group* (see section "The Heat Semi-group" in Appendix 2).

Lemma 8.2 *For any $0 \leq s \leq t$,*

$$\begin{aligned} &\|P_{s,t} g\|_\infty \leq \|g\|_\infty,\ g \in \mathrm{C}_\mathrm{b}(\mathbb{R}^d, \mathbb{R}) \\ \text{and } &\|\nabla P_{s,t} g\|_\infty \leq \|\nabla g\|_\infty,\ g \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R}). \end{aligned} \tag{8.12}$$

In addition, for $g \in \mathrm{C_b}(\mathbb{R}^d, \mathbb{R})$ *and* $i = 1, \dots, d$,

$$\partial_{x_i} P_{s,t} g(x) = \partial_{x_i} \mathbb{E}[g(\mathscr{B}_{t,s}(x))] = \mathbb{E}\left[\frac{\mathsf{B}_t^i - \mathsf{B}_s^i}{t-s} g(\mathscr{B}_{t,s}(x))\right]. \tag{8.13}$$

For $\mu \in \mathbb{R}$, $\sigma > 0$, let us denote by $\mathscr{N}(\mu, \sigma)$ the Gaussian distribution of mean ν and variance σ.

Proof Inequalities (8.12) are immediate. For (8.13), since $x + \mathsf{B}_t - \mathsf{B}_s \sim \mathscr{N}(x, t-s)$,

$$P_{s,t} g(x) = \int_{\mathbb{R}^d} \frac{1}{(2\pi(t-s))^{d/2}} \exp\left(\frac{-|y-x|^2}{2(t-s)}\right) g(x+y)\, \mathrm{d}y.$$

Since

$$\partial_{x_i} \exp\left(\frac{-|y-x|^2}{2(t-s)}\right) = \frac{(y_i - x_i)}{t-s} \exp\left(\frac{-|y-x|^2}{2(t-s)}\right), \tag{8.14}$$

we obtain the integration by part formula (8.13). □

Remark 8.1 An immediate consequence of (8.13) is that

$$\|\nabla P_{s,t} g\|_\infty \le \frac{\|g\|_\infty}{\sqrt{t-s}},\ \forall 0 < s < t,\ \forall g \in \mathrm{C_b}(\mathbb{R}^d, \mathbb{R}). \tag{8.15}$$

8.3.3 The Transport Semi-group

The semi-group $(Q_{s,t})_{0 \le s \le t}$ is associated to a transport equation (see section "The Transport Semi-group" in Appendix 2).

Lemma 8.3 *For any* $0 \le s \le t$,

$$\|Q_{s,t} g\|_\infty \le \|g\|_\infty \text{ for } g \in \mathrm{C_b}(\mathbb{R}^d, \mathbb{R}). \tag{8.16}$$

In addition,

$$\|\nabla Q_{s,t} g\|_\infty \le \|\nabla g\|_\infty \exp((t-s)\beta),\ \forall t \ge 0 \text{ for } g \in \mathrm{C_b^1}(\mathbb{R}^d, \mathbb{R}). \tag{8.17}$$

Proof Inequality (8.16) is immediate and (8.17) follows from Proposition 8.2. □

For $g \in \mathrm{C_b^1}(\mathbb{R}^d, \mathbb{R})$, the Newton formula implies that

$$\begin{aligned} g(\mathscr{Y}_{t,r}(x)) &= g(x) + \int_r^t b(\mathscr{Y}_{s,r}(x)) \nabla g(\mathscr{Y}_{s,r}(x))\, \mathrm{d}s \\ &= g(x) + (t-r) b(x) \nabla g(x) + R(r,t,g,x) \end{aligned} \tag{8.18}$$

with a remainder term

$$R(r,t,g,x) = \int_r^t (b(\mathscr{Y}_{s,r}(x))\nabla g(\mathscr{Y}_{s,r}(x)) - b(x)\nabla g(x))\,\mathrm{d}s. \tag{8.19}$$

8.4 The Euler Scheme

We give a first convergence result which is related to the Euler scheme.

Let us set $\mathscr{E}_{t,s}(x) = x + \mathsf{B}_t - \mathsf{B}_s + b(x)(t-s) = \mathscr{B}_{t,s}(x) + b(x)(t-s)$. This is one step of the *Euler scheme*, in the sense that for $t_i = iT/n$, $i = 0,\dots,n$,

$$\xi_i(x) = \mathscr{E}_{t_i,t_{i-1}} \circ \cdots \circ \mathscr{E}_{t_1,0}(x)$$

and the $\xi_i(x)$'s are easily recursively computed by

$$\xi_0(x) = x \text{ and } \xi_{i+1}(x) = \mathscr{E}_{t_{i+1},t_i}(\xi_i(x)) = \xi_i + \mathsf{B}_{t_{i+1}} - \mathsf{B}_{t_i} + b(\xi_i(x))\frac{1}{n}$$

for $i = 0,\dots,n-1$.

The Euler scheme provides a simple way to approximate the flow $\mathscr{X}_{T,0}(x)$ as

$$\mathscr{E}_{T,(n-1)T/n} \circ \cdots \circ \mathscr{E}_{T/n,0}(x) \xrightarrow[n\to\infty]{} \mathscr{X}_{T,0}(x) \text{ almost surely.} \tag{8.20}$$

We give an insight of the proof in Appendix 1.

In Proposition 8.3 below, an immediate consequence of the next lemma, we provide a *weak rate* of convergence for the Euler scheme.

Lemma 8.4 *For* $g \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d,\mathbb{R})$,

$$\|E_{s,t}g - X_{s,t}g\|_\infty \le \|\nabla g\|_\infty K(t-s)^{3/2}.$$

Proof Actually,

$$\mathscr{X}_{t,s}(x) - \mathscr{E}_{s,t}(x) = \int_s^t (b(\mathscr{X}_{r,s}(x)) - b(x))\,\mathrm{d}r$$

so that with $\beta = \|\nabla b\|_\infty$,

$$|\mathscr{X}_{t,s}(x) - \mathscr{E}_{t,s}(x)| \le \beta \int_s^t \left| \mathsf{B}_r - \mathsf{B}_s - \int_s^r b(\mathscr{X}_{u,s}(x))\,\mathrm{d}u \right| \mathrm{d}r.$$

Then for $g \in \mathrm{C}^1(\mathbb{R}^d; \mathbb{R})$,

$$\|E_{s,t}g - X_{s,t}g\|_\infty \leq \|\nabla g\|_\infty \beta \int_s^t \mathbb{E}[|\mathsf{B}_r - \mathsf{B}_s|]\,\mathrm{d}r + \|\nabla g\|_\infty \frac{(t-s)^2}{2}\beta\|b\|_\infty.$$

But $\mathbb{E}[|\mathsf{B}_r - \mathsf{B}_s|] \leq \sqrt{r-s}$. Hence the result. □

For $g \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$, let us denote

$$\|g\|_\star = \max\{\|g\|_\infty, \|\nabla g\|_\infty\}.$$

For a bounded linear operator $A : \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R}) \to \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$, let us set

$$\|A\|_{\infty\to\infty} = \sup_{g\in\mathrm{C}_\mathrm{b}, \|g\|_\infty=1} \|Ag\|_\infty \text{ and } \|A\|_\star = \sup_{g\in\mathrm{C}^1_\mathrm{b}, \|g\|_\star=1} \|Ag\|_\star. \tag{8.21}$$

Proposition 8.3 (Weak Rate of Convergence of the Euler Scheme) *For any $g \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$,*

$$\left\| \left(\prod_{i=0}^{n-1} E_{iT/n,(i+1)T/n} \right) g - X_{0,T}g \right\|_\infty \leq \|g\|_\star \frac{C}{\sqrt{n}}, \tag{8.22}$$

for a constant C that depends only on $\|b\|_\star$ and T.

Proof Writing $D_{s,t} = E_{s,t} - X_{s,t}$ and using Lemma 8.1,

$$\begin{aligned}&E_{0,1/n}\cdots E_{(n-1)T/n,T}\,g - X_{0,1/n}\cdots X_{(n-1)T/n,T}\,g\\ &\quad= \sum_{j=0}^{n-1} \left(\prod_{i=j+1}^{n-1} E_{iT/n,(i+1)T/n} \right) D_{(j-1)T/n,jT/n} \left(\prod_{i=0}^{j-1} X_{iT/n,(i+1)T/n} \right) g.\end{aligned}$$

With Lemma 8.4, the product property (8.10) of $(X_{s,t})$ and Corollary 8.1,

$$\|D_{(j-1)T/n,jT/n}X_{0,(j-1)T/n}g\|_\infty \leq \|\nabla g\|_\infty K \frac{T^{3/2}}{n^{3/2}}.$$

In addition, since $\|E_{s,t}\|_{\infty\to\infty} \leq 1$ and

$$\|E_{r,s}E_{s,t}\|_{\infty\to\infty} \leq \|E_{r,s}\|_{\infty\to\infty}\|E_{s,t}\|_{\infty\to\infty},$$

on get easily (8.22) from the above inequality. □

Remark 8.2 The weak rate of convergence of the Euler scheme is generally established for smoother coefficients (e.g. $b \in \mathrm{C}^4(\mathbb{R}^d, \mathbb{R}^d)$) to achieve a rate 1 [49]. In [38, 39], it is shown that for α-Hölder continuous coefficients with $\alpha < 2$, the order of convergence is $\alpha/2$. This approach excludes the integer values of α, and the terminal condition is required to be $(2+\alpha)$-Hölder continuous. With our regularity condition on the drift, we complete this result for $\alpha = 1$ when the diffusivity is constant. This result can also be recovered with the results in [29].

Remark 8.3 For a closely related approach with semi-groups, the article [5] provides us with some conditions in a more general context to exhibit the rate of convergence of Euler schemes.

8.5 An Intermediary Convergence Result

We prove that the products of the operators $V_{s,t}$ over finer and finer partitions of $[0,T]$ converges to the operator $X_{0,T}$. This result is crucial to study the convergence of the products of $Q_{s,t}P_{s,t}$ in the next section.

Proposition 8.4 *For any $g \in \mathrm{C}^1_{\mathrm{b}}(\mathbb{R}^d, \mathbb{R})$,*

$$\left\| \left(\prod_{i=0}^{n-1} V_{iT/n,(i+1)T/n} \right) g - X_{0,T} g \right\|_\infty \leq \|g\|_\star \frac{C}{\sqrt{n}},$$

for a constant C that depends only on $\|b\|_\star$ and T.

Proof From the Newton formula,

$$g(\mathscr{E}_{t,s}(x)) = g(\mathscr{B}_{t,s}(x)) + (t-s)\int_0^1 b(x)\nabla g(\mathscr{B}_{t,s}(x) + \tau(t-s)b(x))\,\mathrm{d}\tau,$$

so that

$$\begin{aligned} &|E_{s,t}g(x) - V_{s,t}g(x)| \\ &\quad\leq \left|\mathbb{E}[g(\mathscr{E}_{t,s}(x)] - \mathbb{E}[g(\mathscr{B}_{t,s}(x))] - \mathbb{E}[(t-s)b(x)\nabla g(\mathscr{B}_{t,s}(x))]\right| \\ &\quad\leq (t-s)\int_0^1 \|b\|_\infty \left|\mathbb{E}[\nabla g(\mathscr{B}_{t,s}(x) + \tau(t-s)b(x)) - \nabla g(\mathscr{B}_{t,s}(x))]\,\mathrm{d}\tau\right|. \end{aligned}$$

Again with an integration by parts on the density of the normal distribution, for $G \sim \mathscr{N}(0,1)$,

$$\mathbb{E}[\nabla\phi(\mu+\sigma G)] = \frac{1}{\sqrt{2\pi}}\int \exp\left(\frac{-|y|^2}{2}\right)\nabla\phi(\mu+\sigma y)\,\mathrm{d}y = \mathbb{E}\left[\frac{G}{\sigma}\phi(\mu+\sigma G)\right],$$

from which we obtain for $\tau \in [0,1]$,

$$\mathbb{E}[\nabla g(\mathscr{B}_{t,s}(x) + \tau(t-s)b(x))] = \mathbb{E}\left[\frac{\mathsf{B}_t - \mathsf{B}_s}{t-s} g(\mathscr{B}_{t,x}(x) + \tau(t-s)b(x)\right].$$

Thus,

$$|E_{s,t}g(x) - V_{s,t}g(x)| \leq \frac{\|b\|_\infty^2}{2}\mathbb{E}[|\mathsf{B}_t - \mathsf{B}_s|]\|\nabla g\|_\infty(t-s) \leq C\|\nabla g\|_\infty(t-s)^{3/2}.$$

With Lemma 8.4,

$$|X_{s,t}g(x) - V_{s,t}g(x)| \leq C'(t-s)^{3/2}\|\nabla g\|_\infty.$$

The result follows from the same argument as in the proof of Proposition 8.3. □

8.6 The Splitting Procedure

The composition $Q_{s,t}P_{s,t}$ corresponds to a *splitting* (or composition): first, we follow the flow generated $\mathscr{B}$ starting from x and then the one generated by $\mathscr{Y}_{\cdot,s}$ starting at $\mathscr{B}_{t,s}(x)$. The convergence of the products of Q and P over finer and finer partitions is the spirit of the Trotter-Kato-Lie approach to construct the semigroup X (see [12, 17]).

The main point is that if g is $\mathrm{C}_\mathrm{b}^1(\mathbb{R}, \mathbb{R})$, then

$$Q_{s,t}g(x) = g(\mathscr{Y}_{t,s}(x)) = g(x) + \int_s^t b\nabla g(\mathscr{Y}_{r,s}(x))\,\mathrm{d}r = g(x) + (t-s)b\nabla g(x) + \cdots.$$

We use this Taylor development to obtain some control over the product $Q_{s,t}P_{s,t}$, together with an integration by parts on the Brownian density.

Proposition 8.5 *For any $g \in \mathrm{C}_\mathrm{b}^1(\mathbb{R}^d, \mathbb{R})$,*

$$\left\|\left(\prod_{i=0}^{n-1} Q_{iT/n,(i+1)T/n}P_{iT/n,(i+1)T/n}\right)g - X_{T,0}g\right\|_\infty \leq \|g\|_\star \frac{C}{\sqrt{n}},$$

for a constant C that depends only on $\|b\|_\star$ and T.

Proof With (8.12) and (8.17),

$$\|\nabla(Q_tP_t)g\|_\infty \leq \exp(t\|\nabla b\|_\infty)\|\nabla P_tg\|_\infty \leq \exp(t\beta)\|g\|_\infty.$$

From (8.18),

$$Q_{s,t}P_{s,t}g(x) = P_{s,t}g(x) + (t-s)b(x)\nabla P_{s,t}g(x) + R(s,t,P_{s,t}g,x)$$

where the remainder term $R(s, t, \cdot, x)$ is defined by (8.19). With (8.13),

$$Q_{s,t} P_{s,t} g(x) = V_{t,s} g(x) + R(s, t, P_{s,t} g, x).$$

Since $b \in \mathrm{C}^1_\mathrm{b}(\mathbb{R}^d, \mathbb{R})$,

$$|\mathscr{Y}_{t,s}(x) - x| = \left| \int_s^t b(\mathscr{Y}_{r,s}(x))\,\mathrm{d}r \right| \leq \|b\|_\infty (t-s). \tag{8.23}$$

Again from (8.13),

$$\begin{aligned} &R(s, t, P_{s,t} g, x) \\ &\quad = \int_s^t \mathbb{E}\left[\frac{\mathsf{B}_t - \mathsf{B}_s}{t-s} \cdot \Big(b(\mathscr{Y}_{r,s}(x)) g(\mathscr{Y}_{r,s}(x) + \mathsf{B}_t - \mathsf{B}_s) - b(x) g(\mathscr{B}_{t,s}(x)) \Big)\,\mathrm{d}r \right] \end{aligned}$$

so that with (8.23),

$$\begin{aligned} \|R(s, t, P_{s,t} g, \cdot)\|_\infty &\leq \mathbb{E}\left[\frac{|\mathsf{B}_t - \mathsf{B}_s|}{t-s} \right] \int_s^t (\|\nabla b\|_\infty \|g\|_\infty + \|b\|_\infty \|\nabla g\|_\infty) r\,\mathrm{d}r \\ &\leq 2\|b\|_\star \|g\|_\star (t-s)^{3/2}. \end{aligned}$$

This proves that

$$\left\| \left(\prod_{i=0}^{n-1} Q_{iT/n,(i+1)T/n} P_{iT/n,(i+1)T/n} \right) g - \left(\prod_{i=0}^{n-1} V_{iT/n,(i+1)T/n} \right) g \right\|_\infty \leq \|g\|_\star \frac{C}{\sqrt{n}},$$

and the result follows from Proposition 8.4. □

8.7 The Weights and Their Limits: The Girsanov Theorem

The expression of the weight will be obtained by two ways: The first one involves the expression of the Euler scheme. For this, we need however a change of measure which is easily deduced from the explicit expression of the Gaussian density. The second one involves the expressions of V in the splitting procedure. With the discrete time approximation, both expressions leads to different expressions for the weights, whose limit is however the same. The expression obtained by the Euler scheme is a discretization of the stochastic integral in the exponential weight given by (8.5). The ones obtained using the splitting is a discretization of the SDE (8.6) the exponential weights solve.

We have now all the elements to prove the Girsanov theorem with our approach.

Theorem 8.3 (Girsanov, Simplified Form) *There exists a stochastic process* $(\mathsf{Z}_t(x))_{t\in[0,T]}$ *given by*

$$\mathsf{Z}_t(x) = \exp\left(\int_0^t b(x+\mathsf{B}_s)\,\mathrm{d}\mathsf{B}_s - \frac{1}{2}\int_0^t |b(x+\mathsf{B}_s)|^2\,\mathrm{d}s\right) \tag{8.24}$$

$$\text{or } \mathsf{Z}_t(x) = 1 + \int_0^t \mathsf{Z}_s(x) b(x+\mathsf{B}_s)\,\mathrm{d}\mathsf{B}_s \tag{8.25}$$

such that

$$\mathbb{E}[g(\mathscr{X}_{T,0}(x))] = \mathbb{E}[\mathsf{Z}_T(x) g(\mathscr{B}_{T,0}(x))]$$

for any bounded, measurable function g.

The weight $(\mathsf{Z}_t(x))_{t\geq 0}$ will be identified as a limit, and stochastic analysis is needed for this. For this, we will rewrite the probabilistic representation of the operators given by both the Euler scheme and the splitting scheme as involving some exponential term, and then combine them. The expression (8.24) is related to the representation provided by the Euler scheme while (8.25) stems from the splitting scheme.

The proof of this theorem is then obtained by combining Propositions 8.7 and 8.8 below with Remark 8.5.

8.7.1 *Exponential Representation of the One-Step Euler Scheme*

The following exponential representation is a direct consequence of the expression of the density of the Gaussian distribution.

Lemma 8.5 *For any* $x \in \mathbb{R}^d$ *and* $0 \leq s \leq t$,

$$\mathbb{E}[g(\mathscr{E}_{t,s}(x))] = E_{t,s}g(x) = \mathbb{E}\left[\mathbf{E}_{s,t}(x) g(\mathscr{B}_{t,s}(x))\right]$$

with

$$\mathbf{E}_{s,t}(x) = \exp\left(b(x)\cdot(\mathsf{B}_t - \mathsf{B}_s) - \frac{(t-s)}{2}|b(x)|^2\right).$$

Proof When $G \sim \mathcal{N}(0, \sigma^2 \mathrm{Id}_{d\times d})$ and $H \sim \mathcal{N}(\mu, \sigma^2 \mathrm{Id}_{d\times d})$, then for any measurable, bounded function $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}[\phi(\mu + G)] = \mathbb{E}[\phi(H)] &= \int \phi(x) \frac{\exp\left(\frac{-|x-\mu|^2}{2\sigma^2}\right)}{(2\pi\sigma^2)^{d/2}}\,\mathrm{d}x \\
&= \int \phi(x) \exp\left(\frac{x\cdot\mu}{\sigma^2} - \frac{|\mu|^2}{2\sigma^2}\right) \frac{\exp\left(\frac{-|x|^2}{2\sigma^2}\right)}{(2\pi\sigma^2)^{d/2}}\,\mathrm{d}x \\
&= \mathbb{E}\left[\Psi(G,\mu)\phi(G)\right] \\
\text{with } \Psi(G,\mu) &= \exp\left(\frac{G\cdot\mu}{\sigma^2} - \frac{|\mu|^2}{2\sigma^2}\right).
\end{aligned}$$

With

$$\begin{aligned}
\phi(y) = g(x+y),\ G = \mathsf{B}_t - \mathsf{B}_s \sim \mathcal{N}(0, (t-s)\mathrm{Id}_{d\times d}) \\
\text{and } H = (t-s)b(x) + \mathsf{B}_t - \mathsf{B}_s \sim \mathcal{N}((t-s)b(x), (t-s)\mathrm{Id}_{d\times d}),
\end{aligned}$$

this yields the result. □

8.7.2 The Weights

Let us set for $x \in \mathbb{R}^d$ and $0 \le r \le t$,

$$\mathbf{V}_{r,t}(x) = 1 + b(x)(\mathsf{B}_t - \mathsf{B}_r) \text{ so that } V_{r,t}g(x) = \mathbb{E}[\mathbf{V}_{r,t}(x) g(\mathscr{B}_{t,r}(x))].$$

Proposition 8.6 *For any $x \in \mathbb{R}^d$, $T > 0$ and $n \in \mathbb{N}$,*

$$\prod_{i=0}^{n-1} E_{iT/n,(i+1)T/n} g(x) = \mathbb{E}\left[\mathsf{Z}^E_{n,n}(x) g(\mathscr{B}_{t,0}(x))\right] \tag{8.26}$$

$$\textit{and } \prod_{i=0}^{n-1} V_{iT/n,(i+1)T/n} g(x) = \mathbb{E}\left[\mathsf{Z}^V_{n,n}(x) g(\mathscr{B}_{t,0}(x))\right]$$

$$\textit{with } \mathsf{Z}^E_{k,n}(x) = \prod_{i=0}^{k-1} \mathbf{E}_{iT/n,(i+1)T/n}(\mathscr{B}_{iT/n,0}(x))$$

$$\textit{and } \mathsf{Z}^V_{k,n}(x) = \prod_{i=0}^{k-1} \mathbf{V}_{iT/n,(i+1)T/n}(\mathscr{B}_{iT/n,0}(x)).$$

Proof Let us work on $E_{t,s}$, the proof being similar for $V_{t,s}$.

Set $F(x) = \mathbb{E}[g(\mathscr{E}_{t,s}(x))]$ so that

$$\begin{aligned} E_{r,s}E_{s,t}g(x) &= \mathbb{E}[E_{s,t}g(\mathscr{E}_{s,r}(x))] = \mathbb{E}[F(\mathscr{E}_{s,r}(x))] \\ &= \mathbb{E}[\mathbf{E}_{r,s}(x)F(\mathscr{B}_{s,r}(x))] = \mathbb{E}[\mathbf{E}_{r,s}(x)\mathbf{E}_{s,t}(\mathscr{B}_{s,r}(x))g(\mathscr{B}_{t,s}(\mathscr{B}_{s,r}(x)))] \\ &= \mathbb{E}[\mathbf{E}_{r,s}(x)\mathbf{E}_{s,t}(\mathscr{B}_{s,r}(x))g(\mathscr{B}_{t,r}(x))]. \end{aligned}$$

By iterating this computation, this leads to (8.26). □

8.7.3 Uniform Integrability of the Weights

For $n \geq 1$ and $T > 0$ fixed, let us denote by $\Delta_i^n \mathsf{B} = \Delta_{iT/n,(i+1)T/n}^n \mathsf{B}$, the increments of the Brownian motion. For $i = 1, \dots, n$, set $\mathscr{F}_i^n = \sigma(\Delta_0^n \mathsf{B}, \dots, \Delta_{i-1}^n \mathsf{B})$, so that $(\mathscr{F}_i^n)_{i \geq 0}$ is the filtration generated by the increments of the Brownian motion.

We fix a starting point $x \in \mathbb{R}^d$.

Lemma 8.6 *For each n, both* $(\mathsf{Z}_{k,n}^E(x))_{k\geq 0}$ *and* $(\mathsf{Z}_{k,n}^V(x))_{n\geq 0}$ *are discrete martingales with respect to* $(\mathscr{F}_i^n)_{i=1,\dots,n}$. *Besides* $\mathbb{E}[\mathsf{Z}_{k,n}^E(x)] = \mathbb{E}[\mathsf{Z}_{k,n}^V(x)] = 1$.

Proof The proof is immediate using the independence of the increments of the Brownian motion. □

Remark 8.4 Although $\mathsf{Z}_{n,n}^E(x)$ and $\mathsf{Z}_{n,n}^V(x)$ are close, they are certainly not equal as $\mathsf{Z}_{n,n}^E(x)$ remains positive while $\mathsf{Z}_{n,n}^V(x)$ is negative with a positive probability.

To be able to pass to the limit, we prove that the family weights are uniformly integrable martingales.

Lemma 8.7 *Both* $(\mathsf{Z}_{k,n}^E(x))_{n\geq 1}$ *and* $(\mathsf{Z}_{k,n}^V(x))_{n\geq 1}$ *are uniformly integrable.*

Remark 8.5 Provided that $\mathsf{Z}_{n,n}^E(x)$ and $\mathsf{Z}_{n,n}^V(x)$ converge in distribution to some random variable Z_T, the uniform integrability of these random variables is sufficient to assert that

$$\lim_{n\to\infty} \mathbb{E}[\mathsf{Z}_{n,n}^E(x)g(\mathscr{B}_{T,0}(x))] = \lim_{n\to\infty} \mathbb{E}[\mathsf{Z}_{n,n}^V(x)g(\mathscr{B}_{T,0}(x))] = \mathbb{E}[\mathsf{Z}_T(x)g(\mathscr{B}_{T,0}(x))].$$

Proof Using the independence of the increment and the property of the Laplace transform of the Gaussian distribution,

$$\mathbb{E}\left[\left(\exp\left(b(x+\mathsf{B}_{iT/n})\Delta_i^n\mathsf{B} - \frac{T}{2n}|b(x+\mathsf{B}_{iT/n})|^2\right)\right)^2 \middle| \mathscr{F}_i^n\right] \leq \exp\left(\|b\|_\infty \frac{T}{n}\right),$$

so that, by computing iteratively the conditional expectations,

$$\sup_{\substack{k=1,\dots,n\\ n\geq 1}} \mathbb{E}[(\mathsf{Z}^E_{k,n}(x))^2] \leq \exp(\|b\|_\infty T).$$

Similarly,

$$\mathbb{E}\left[\left(1+b(x+\mathsf{B}_{iT/n})\Delta^n_i\mathsf{B}\right)^2\middle|\mathscr{F}^n_i\right] \leq 1+\frac{T}{n}\|b\|_\infty \leq \exp\left(\|b\|_\infty\frac{T}{n}\right),$$

so that

$$\sup_{\substack{k=1,\dots,n\\ n\geq 1}} \mathbb{E}[(\mathsf{Z}^V_{k,n}(x))^2] \leq \exp(\|b\|_\infty T).$$

This is sufficient to prove and uniform integrability. □

8.7.4 *Identification of the Limit of the Weights: Stochastic Calculus At last*

At this stage, we need some stochastic calculus in order to identify the limit of $\mathsf{Z}^E_{k,n}(x)$ and $\mathsf{Z}^V_{k,n}(x)$.

Let us rewrite $\mathsf{Z}^E_{k,n}(x)$ as the iterative family by setting $\mathsf{Z}^E_{0,n}(x)=1$ and

$$\begin{aligned} Z^E_{k+1,n}(x) &= \mathsf{Z}^E_{k,n}(x)\exp\left(b(x+\mathsf{B}_{t^n_k})\Delta^n_k\mathsf{B}-\frac{T}{2n}|b(x+\mathsf{B}_{t^n_k})|^2\right)\\ &= \exp\left(\sum_{i=0}^k b(x+\mathsf{B}_{t^n_i})\Delta^n_i\mathsf{B}-\frac{T}{2n}\sum_{i=0}^k |b(x+\mathsf{B}_{t^n_i})|^2\right) \end{aligned}$$

for $k=0,\dots,n-1$.

The next result is an immediate consequence of the definition of a stochastic integral. We recover the traditional expression for the Girsanov weight.

Proposition 8.7 *When $n\to\infty$ and $k_n\to\infty$ with $k_n/n \xrightarrow[n\to\infty]{} t$ for some $t\in[0,T]$, $\mathsf{Z}^E_{k_n,n}(x)$ converges in probability to $\mathsf{Z}_t(x)_{t\in[0,T]}$ given by (8.24).*

An application of the Itô formula yields the well known fact that Z_t given by (8.24) is solution to the Stochastic Differential Equation (8.25).

On the other hand, writing $\mathsf{Z}^V_{k,n}(x)$ as an iterative family yields that

$$Z^V_{k+1,n}(x) = \mathsf{Z}^V_{k,n}(1+b(x+\mathsf{B}_{t^n_k})\Delta^n_k\mathsf{B}) \text{ for } k=0,\dots,n-1$$

with $\mathsf{Z}^V_{0,n} = 1$. This could immediately be rewritten as the discrete analogue of (8.25) by

$$Z^V_{k+1,n}(x) = 1 + \sum_{i=0}^{k} \mathsf{Z}^V_{i,n}(x) b(x + \mathsf{B}_{t_i^n}) \Delta_i^n \mathsf{B} \text{ for } k = 0, \dots, n-1.$$

The convergence of $\mathsf{Z}^V_{k,n}(x)$ is less immediate than the one of $\mathsf{Z}^E_{k,n}(x)$. Hopefully, it can be dealt with the results of Duffie and Protter [16] (see also [30, Example 8.7, p. 33]).

Proposition 8.8 *When* $n \to \infty$ *and* $k_n \to \infty$ *with* $k_n/n \xrightarrow[n\to\infty]{} t$ *for some* $t \in [0, T]$, $\mathsf{Z}^V_{k_n,n}(x)$ *converges in distribution to* $\mathsf{Z}_t(x)$ *given by* (8.25) *or equivalently* (8.24).

8.8 The Infinitesimal Generator of the Semi-group $(X_t)_{t\geq 0}$

We compute now the infinitesimal generator of $(X_t)_{t\geq 0}$, still using only functional analysis. For the details, we refer for example to the book [12, 17].

The semi-groups $(X_{s,t})_{0\leq s\leq t}$, $(P_{s,t})_{0\leq s\leq t}$ and $(Q_{s,t})_{0\leq s\leq t}$ are indeed homogeneous in time, as actually $X_{s,t} = X_{0,t-s}$, $P_{s,t} = P_{0,t-s}$ and $Q_{s,t} = Q_{0,t-s}$. We now set $X_t = X_{0,t-s}$, $P_t = P_{0,t-s}$ and $Q_t = Q_{0,t-s}$.

Definition 8.1 (Strongly Continuous Semi-group) Let B be a Banach space with a norm $\|\cdot\|$. A semi-group $(F_t)_{t\geq 0}$ on B is said to be *strongly continuous* if for any $f \in \mathrm{B}$, $T_f f$ is continuous.

Definition 8.2 (Infinitesimal Generator) The *infinitesimal generator* of a semi-group $(F_t)_{t\geq 0}$ on B is a linear operator $\mathfrak{F} : \mathrm{Dom}(\mathfrak{F}) \subset \mathrm{B} \to \mathrm{B}$ such that

$$\mathrm{Dom}(\mathfrak{F}) = \left\{ f \in \mathrm{B} \,\middle|\, \lim_{h\to 0} \frac{F_h f - f}{h} \text{ exists} \right\}$$

and $\mathfrak{F} f = \lim_{h\to 0}(F_h f - f)/h$.

When $(\mathfrak{F}, \mathrm{Dom}(\mathfrak{F}))$ is closed for the graph norm $\|\cdot\| + \|\mathfrak{F}\cdot\|$ and densely defined in B, then it determines the semi-group uniquely [17, Proposition 1.4, p. 51].

We consider that the underlying Banach space is $\mathrm{B} = \mathrm{C_z}(\mathbb{R}^d, \mathbb{R})$, the space of continuous, bounded functions that vanish at infinity, and $\|\cdot\| = \|\cdot\|_\infty$. We denote by $\mathrm{C}^k_\mathrm{c}(\mathbb{R}^d, \mathbb{R}) \subset \mathrm{C_z}(\mathbb{R}^d, \mathbb{R})$ the space of functions of class $\mathrm{C}^k(\mathbb{R}^d, \mathbb{R})$ with compact support.

On B, we consider the heat operator $\mathfrak{L} = \frac{1}{2}\triangle$, whose domain is the closure of $\mathrm{C}^2_\mathrm{c}(\mathbb{R}^d, \mathbb{R})$ for the graph norm. This is the infinitesimal generator of $(P_t)_{t\geq 0}$ (see section "The Heat Semi-group" in Appendix 2).

We also consider $\mathfrak{B} = b\nabla$ whose domain is the closure of $C^1_c(\mathbb{R}^d, \mathbb{R})$ for the graph norm. The operator $(\mathfrak{B}, \mathrm{Dom}(\mathfrak{B}))$ is the infinitesimal generator of $(Q_t)_{t\geq 0}$ (See section "The Transport Semi-group" in Appendix 2).

The next result is an "almost" direct consequence of the Trotter-Kato-Lie formula.

Proposition 8.9 *The infinitesimal generator of the semi-group* $(X_t)_{t\geq 0}$ *is* $\mathfrak{A} = \mathfrak{L} + \mathfrak{B}$ *with* $\mathrm{Dom}(\mathfrak{A}) = \mathrm{Dom}(\mathfrak{L})$.

Proof To construct the semi-group $(X_t)_{t\geq 0}$, we apply Corollary 5.8 in [17, p. 227], which follows from the Chernoff product formula [17, Theorem 5.2, p. 220].

We show first that $\mathrm{Dom}(\mathfrak{L}) \subset \mathrm{Dom}(\mathfrak{B})$.

For any $\lambda > 0$, $(\lambda - \mathfrak{L})$ is invertible on B and $\mathrm{Dom}(\mathfrak{L}) = (\lambda - \mathfrak{L})^{-1}(\mathrm{B})$. Moreover,

$$(\lambda - \mathfrak{L})^{-1} f(x) = \int_0^{+\infty} e^{-\lambda t} P_t f(x)\,\mathrm{d}t. \tag{8.27}$$

Using the control (8.15), for any $f \in \mathrm{B}$,

$$\|\nabla(\lambda - \mathfrak{L})^{-1} f(x)\|_\infty \leq C(\lambda)\|f\|_\infty \tag{8.28}$$

with $C(\lambda) = \int_0^{+\infty} t^{-1/2} e^{-\lambda t}\,\mathrm{d}t$. Thus,

$$\|\mathfrak{B}(\lambda - \mathfrak{L})^{-1} f\|_\infty \leq \|b\|_\infty C(\lambda)\|f\|_\infty.$$

It follows that $\mathfrak{B}$ is well defined for any $\mathrm{Dom}(\mathfrak{L})$ since for any $g \in \mathrm{Dom}(\mathfrak{L})$, there exists $f \in \mathrm{B}$ such that $g = (\lambda - \mathfrak{L})^{-1} f$.

Both $(Q_t)_{t\geq 0}$ and $(P_t)_{t\geq 0}$ are *contraction* semi-groups,[8] since $\|P_t g\|_\infty \leq \|g\|_\infty$ and $\|Q_t g\|_\infty \leq \|g\|_\infty$ for any $g \in \mathrm{B} = C_z(\mathbb{R}^d, \mathbb{R})$. Thus, for any $t \geq 0$, $P_t Q_t$ is a bounded operator on B with norm 1.

It remains to show that $(\lambda - \mathfrak{L} - \mathfrak{B})(\mathrm{Dom}(\mathfrak{L}))$ is dense in $\mathrm{B} = C_z(\mathbb{R}^d, \mathbb{R})$ for some $\lambda > 0$. We actually show that $(\lambda - \mathfrak{L} - \mathfrak{B})$ is one-to-one between $\mathrm{Dom}(\mathfrak{L})$ to B.

For this, we consider finding the pairs $(f, g) \in \mathrm{Dom}(\mathfrak{L}) \times \mathrm{B}$ such that

$$(\lambda - \mathfrak{L} - \mathfrak{B}) f = g. \tag{8.29}$$

We rewrite (8.29) as

$$f - (\lambda - \mathfrak{L})^{-1}\mathfrak{B} f = (\lambda - \mathfrak{L})^{-1} g,$$

[8] These semi-groups satisfies the far more finer properties of being Feller, as for $(X_t)_{t\geq 0}$, but we do not use it here.

so that f is sought as

$$f = \mathfrak{K}g = \lim_{n\to\infty} \mathfrak{K}_n g \text{ with } \mathfrak{K}_n g = \sum_{k=0}^{n} \big((\lambda - \mathfrak{L})^{-1}\mathfrak{B}\big)^k (\lambda - \mathfrak{L})^{-1} g. \tag{8.30}$$

For any $k \geq 1$,

$$\big((\lambda - \mathfrak{L})^{-1}\mathfrak{B}\big)^k (\lambda - \mathfrak{L})^{-1} = (\lambda - \mathfrak{L})^{-1}\big(\mathfrak{B}(\lambda - \mathfrak{L})^{-1}\big)^k.$$

Inequality (8.8) proves that $\mathfrak{B}(\lambda - \mathfrak{L})^{-1}$ is a bounded operator on B with constant $\|b\|_\infty C(\lambda)$.

When $\lambda \to \infty$, $C(\lambda)$ decreases to 0. We choose λ large enough so that $\|b\|_\infty C(\lambda) < 1$. An immediate consequence of (8.27) is that $(\lambda - \mathfrak{L})^{-1}$ is bounded by $1/\lambda$. Thus, for any $k \geq 1$,

$$\|\big((\lambda - \mathfrak{L})^{-1}\mathfrak{B}\big)^k (\lambda - \mathfrak{L})^{-1} g\|_\infty \leq \frac{1}{\lambda}\big(\|b\|_\infty C(\lambda)\big)^k \|g\|_\infty.$$

This means that the series $\mathfrak{K}g$ defined in (8.30) converges in B. Moreover, it is easily checked that

$$(\lambda - \mathfrak{L} - \mathfrak{B})\mathfrak{K}_n (\lambda - \mathfrak{L})^{-1} g = -(\mathfrak{B}(\lambda - \mathfrak{L})^{-1})^{n+1} g + g.$$

With (8.28), under the condition that $C(\lambda)\|b\|_\infty < 1$, $f = \mathfrak{K}g$ solves $(\lambda - \mathfrak{L} - \mathfrak{B})f = g$ in B.

Since

$$\mathfrak{B}\mathfrak{K}_n g = \sum_{k=0}^{n} \mathfrak{B}[(\lambda - \mathfrak{L})^{-1}\mathfrak{B}]^k (\lambda - \mathfrak{L})^{-1} g = \sum_{k=1}^{n+1} [\mathfrak{B}(\lambda - \mathfrak{L})^{-1}]^k g,$$

it follows from (8.8) that when $C(\lambda)\|b\|_\infty < 1$, $(\mathfrak{B}\mathfrak{K}_n g)_{n\geq 1}$ forms a Cauchy sequence. Since $(\mathfrak{B}, \mathrm{Dom}(\mathfrak{B}))$ is a closed operator, its limit is necessarily $\mathfrak{B}\mathfrak{K}g$. Thus,

$$(\lambda - \mathfrak{L})\mathfrak{K}_n g = g - (\mathfrak{B}(\lambda - \mathfrak{L})^{-1})^{n+1} g + \mathfrak{B}\mathfrak{K}_n g.$$

Clearly, $\mathfrak{K}_n g \in \mathrm{Dom}(\mathfrak{L})$. Since $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is also a closed operator, we obtain that by passing to the limit, $\mathfrak{K}g \in \mathrm{Dom}(\mathfrak{L})$ and

$$(\lambda - \mathfrak{L})\mathfrak{K}g = g + \mathfrak{B}\mathfrak{K}g$$

so that $\lambda - \mathfrak{L} - \mathfrak{B}$ is invertible from $\mathrm{Dom}(\mathfrak{L})$ to B with inverse $\mathfrak{K}$ which is bounded on B. This implies that $(\mathfrak{L} + \mathfrak{B})$ is itself a closed operator, with domain $\mathrm{Dom}(\mathfrak{L})$. The latter domain is dense in B.

The Chernoff product formula [17, Theorem 5.2, p. 220] then proves that the infinitesimal generator of $(X_t)_{t\geq 0}$ is $(\mathfrak{L}+\mathfrak{B}, \mathrm{Dom}(\mathfrak{L}+\mathfrak{B}))$. □

8.9 A Case Where Itô Formula Could be Avoided (or Feynman, Kac, Girsanov and Doob Meet Together)

Still to stuck to our rule to avoid stochastic calculus for the sake of play, we show that for special form of the drift, an already known probabilistic representation can be obtained by combining the previous results. The representation is the starting point of the so-called *exact simulation method* from Beskos and Roberts [7] for simulating a Brownian motion with drift by performing an acceptance/rejection scheme on paths from the Brownian bridge.

8.9.1 The Feynman-Kac Formula

The Feynman-Kac formula provides a probabilistic representation to the solution to

$$\begin{cases} \partial_t u(t,x) = \frac{1}{2}\triangle u(t,x) + U(x)u(t,x), \\ u(0,x) = f(x). \end{cases} \tag{8.31}$$

Theorem 8.4 (Kac [27]) *Let U be bounded.*[9] *Then the solution of* (8.31) *is solution to*

$$u(t,x) = \mathbb{E}\left[\exp\left(\int_0^t U(x+\mathsf{B}_s)\,\mathrm{d}s\right) f(x+\mathsf{B}_t)\right]. \tag{8.32}$$

Proved first by M. Kac, it is also related to the Feynman path integral for solving the Schrödinger equation (See e.g. [40, 45]). This formula has many applications, in mathematical physics of course, but also in analysis. For example, it provides an effective way to compute Laplace transforms of functionals of the Brownian motion—the original goal of Kac [27] (for applications, see e.g. [25])—or to perform some change of measures [44].

The Feynman-Kac formula could also be proved naturally through the Trotter-Kato-Lie formula (see e.g. [21, 48]).

[9] This condition is stronger than the one given in the original article of M. Kac on that subject.

With $\mathfrak{L} = \frac{1}{2}\triangle$ and $\mathfrak{U}$ be defined on the space of bounded functions by $\mathfrak{U} f(x) = U(x) f(x)$, the semi-group $(U_t)_{t\geq 0}$ of $\mathfrak{U}$ is simply given by

$$U_t f(x) = \exp(U(x)t) f(x),\ t \geq 0,$$

a fact which is easily verified. Thus, by a computation similar to the one of Proposition 8.6, for any $t \geq 0$,

$$(U_{t/n} P_{t/n})^n = \mathbb{E}\left[\exp\left(\sum_{i=0}^{n-1} U(x + \mathsf{B}_{it/n})\frac{t}{n}\right) f(\mathsf{B}_t)\right] \xrightarrow[n\to\infty]{} u(t, x) \text{ given by (8.32).}$$

On the other hand, the Trotter-Kato-Lie formula implies that

$$\partial_t v(t, x) = (\mathfrak{L} + \mathfrak{U}) v(t, x) \text{ with } v(0, x) = f(x) \text{ when } v(t, x) = \lim_{n\to\infty} (U_{t/n} P_{t/n})^n f(x).$$

8.9.2 The Infinitesimal Generator of a Semi-group Under a h-Transform

Given a semi-group $(P_t)_{t>0}$ and a positive function ϕ, one could naturally set

$$P_t^\phi f = \frac{1}{\phi} P_t(\phi f),\ t \geq 0.$$

The rationale is that for any measurable, bounded function f,

$$P_t^\phi P_s^\phi f = \frac{1}{\phi} P_t\left(\frac{\phi}{\phi} P_s(\phi f)\right) = \frac{1}{\phi} P_{t+s}(\phi f) = P_{t+s}^\phi f \text{ for any } s, t \geq 0$$

so that $(P_t^\phi)_{t\geq 0}$ is still a semi-group.

Proposition 8.10 *Let ϕ be a positive function of class $\mathrm{C}_\mathrm{b}^2(\mathbb{R}^d, \mathbb{R}^d)$. Consider the semi-group $(P_t)_{t\geq 0}$ generated by $\mathfrak{P} = \mathfrak{L} + \mathfrak{V}$. Then the infinitesimal generator $\mathfrak{P}^\phi$ of $(P_t^\phi)_{t\geq 0}$ is given by*

$$\mathfrak{P}^\phi f = \mathfrak{P} f + \frac{1}{\phi} \nabla\phi \cdot \nabla f + \frac{1}{2\phi}(\triangle\phi) f.$$

Since ϕ is positive, let us write $\Phi = \ln(\phi)$ so that for $f \in \mathrm{C}_\mathrm{b}^2(\mathbb{R}^d, \mathbb{R}^d)$,

$$\mathfrak{P}^\phi f = \mathfrak{L} f + \mathfrak{V} f + \nabla\Phi \cdot \nabla f + \frac{1}{2}(\triangle\Phi) f + \frac{1}{2}(\nabla\Phi \cdot \nabla\Phi) f.$$

Remark 8.6 If ϕ is a harmonic function, then $\mathfrak{P}^{\phi}$ is the infinitesimal generator of a diffusion process with drift $\nabla\phi/\phi$. This is the spirit of the *h-transform* or *Doob's transform* introduced by Doob [15] for solving problem in potential analysis. This implies a large class of processes with conditioning (for example, to construct a Brownian bridge where the value at a given time is fixed) could be obtained through a process with a drift (Actually, the same computations hold when ϕ is also time dependent).

8.9.3 An Alternative Formulation for the Girsanov Weights for Some Special Form of the Drift

Given a bounded function ϕ, it is tempting but hopeless to look for a function ϕ such that $\nabla\phi = \phi b$ and $\triangle\phi = 0$. We are however free to choose the potential U.

Let $(S_t)_{t\geq 0}$ be the semi-group generated by $\mathfrak{P} = \mathfrak{L} + \mathfrak{U}$ with $U = -\frac{1}{2}\nabla\Phi\cdot\nabla\Phi + \frac{1}{2}\triangle\Phi$ with $\Phi = \log\phi$. Hence,

$$S_t^{\phi} f(x) = \frac{1}{e^{\Phi(x)}}\mathbb{E}_x\left[\exp\left(-\frac{1}{2}\int_0^t \nabla\Phi(\mathsf{B}_s)\cdot\nabla\Phi(\mathsf{B}_s)\,\mathrm{d}s + \int_0^t \frac{1}{2}\triangle\Phi(\mathsf{B}_s)\,\mathrm{d}s + \Phi(\mathsf{B}_t)\right) f(\mathsf{B}_t)\right]. \tag{8.33}$$

According to the above rules, the infinitesimal generator $\mathfrak{S}^{\phi}$ of S_t^{ϕ} is

$$\mathfrak{S}^{\phi} f = \mathfrak{L} f + \nabla\Phi\cdot\nabla f.$$

Proposition 8.11 *Assume that $b \in \mathrm{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ is a potential vector field, i.e., there exists $\Phi \in \mathrm{C}^2(\mathbb{R}^d, \mathbb{R})$ such that $b = \nabla\Phi$. Then for a bounded, measurable function f, $\mathbb{E}[f(X_t)] = S_t^{\phi} f(x)$ where S_t^{ϕ} is given by (8.33).*

Actually, (8.33) is not surprising. With the Itô's formula,

$$\Phi(\mathsf{B}_t) - \Phi(\mathsf{B}_0) = \int_0^t \nabla\Phi(\mathsf{B}_s)\,\mathrm{d}\mathsf{B}_s - \frac{1}{2}\int_0^t \triangle\Phi(\mathsf{B}_s)\,\mathrm{d}s,$$

from which the classical representation of the exponential weight of the Girsanov theorem is easily obtained.

8.10 Complement: On the Regularity of the Drift and the Difference Between Classical and Stochastic Analysis

With our approach, the drift was assumed to be Lipschitz continuous. The main reason for this condition is that from the very construction, the flow associated to the ODE $\dot{\mathsf{X}} = b(\mathsf{X})$ needs to be defined to construct the semi-group $(Q_t)_{t>0}$.

If b is not Lipschitz continuous, the equation $\dot{\mathsf{X}} = b(\mathsf{X})$ may have several solutions. A classical example is $b(x) = \sqrt{x}$.

The regularity of b could be weakened to define a flow or to consider a particular solution to $\dot{\mathsf{X}} = b(\mathsf{X})$ but still a minimal regularity of b should be enforced. For example, b should belong to some Sobolev space [2, 13].

On the other hand, the situation changes when SDE are considered. Striking results from Zvonkin [53] and afterwards Veretennikov [52] prove the existence of a unique *strong* solution to the SDE

$$\mathrm{d}\mathsf{X}_t = \sigma(\mathsf{X}_t)\,\mathrm{d}\mathsf{B}_t + b(\mathsf{X}_t)\,\mathrm{d}t \tag{8.34}$$

whatever the regularity of b provided that

$$\mathrm{d}\mathsf{X}_t = \sigma(\mathsf{X}_t)\,\mathrm{d}\mathsf{B}_t,\ t \geq 0 \tag{8.35}$$

has a strong solution (of course, with suitable integrability conditions on b). The solution to (8.34) defines a flow of diffeomorphisms when σ is a positive constant and b is Hölder continuous.

Taking $\sigma(z) = \epsilon z$ for $\epsilon > 0$ which is taken as small as possible proved the uniqueness of a solution to the SDE $\mathrm{d}\mathsf{X}_t = \epsilon\,\mathrm{d}\mathsf{B}_t + b(\mathsf{X}_t)\,\mathrm{d}t$ seen as a noisy perturbation of the ODE $\dot{\mathsf{X}} = b(\mathsf{X})$ even when the latter ODE has no single solution. This fact is discussed among other references in [3, 19].

For the Girsanov theorem seen as a change of measure, the regularity of b plays no role. One has only to ensure that the exponential super-martingale $\mathsf{Z}_t = \exp\left(\int_0^t b(\mathsf{X}_s)\,\mathrm{d}\mathsf{B}_s + \frac{1}{2}\int_0^t b(\mathsf{X}_s)\,\mathrm{d}s\right)$ defining the weight is actually a martingale, and numerous conditions have been given on b (See e.g. [20, 33]). For this, the drift appears only in an integrated form. An immediate application is the existence of a *weak* solution to the (8.34) when (8.35) has one, whatever the regularity of b (of course, with the suitable integrability conditions on b, see e.g. [47]).

We then see that a difference holds by considering (8.35) and $\dot{\mathsf{X}} = b(\mathsf{X})$ separately through the Trotter-Kato-Lie, or by considering directly SDE of type (8.34), which draws a line of separation between stochastic and ordinary differential equations. This explains why our conditions on b are stronger than the one required usually when invoking the Girsanov theorem.

Acknowledgements I wish to thank K. Coulibaly-Pasquier, V. Bally and A. Kohatsu-Higa for some motivating and interesting discussions on this approach. This article is a follow-up of a talk given at the "groupe de travail" of the Probability and Statistics teams of Institut Élie Cartan de Lorraine (Nancy) on the link between the Trotter-Kato-Lie and the Feynman-Kac formula, and I am grateful to the audience for his/her patience.

Appendix 1: Almost Sure Convergence of the Euler Scheme

The convergence property (8.20) of the Euler scheme follows from the recursive inequality

$$|\xi_{i+1}(x) - \mathscr{X}_{t_{i+1},0}(x)| \leq |\xi_i(x) - \mathscr{X}_{t_i,0}(x)| + \rho_i + \frac{T}{n}|b(\xi_i(x)) - b(\mathscr{X}_{t_i,0}(x))|$$
$$\leq \exp\left(T\|\nabla b\|_\infty\right) \sum_{j=1}^{i} \rho_j \tag{8.36}$$

with

$$\rho_i = \int_{t_i}^{t_{i+1}} |b(\mathscr{X}_{r,0}(x)) - b(\mathscr{X}_{t_i,0}(x))|\,\mathrm{d}r \leq \frac{T}{n}\|\nabla b\|_\infty \sup_{r\in[t_i,t_{i+1}]} |\mathscr{X}_{r,0}(x) - \mathscr{X}_{t_i,0}(x)|.$$

For this, we have used the fact that b is Lipschitz continuous and $1 + \frac{T}{n}\|\nabla b\|_\infty \leq \exp(T\|\nabla b\|_\infty/n)$.

It can be proved that similarly to the Brownian path, each path of $\mathscr{X}_{t,0}(x)$ is α-Hölder continuous for any $\alpha < 1/2$. This is a direct consequence of the Kolmogorov lemma on the regularity of stochastic processes. This proves that the right hand side of (8.36) converges to 0 at rate $\alpha < 1/2$.

Appendix 2: The Heat and the Transport Semi-group

The underlying Banach space is $\mathrm{B} = \mathrm{C}_{\mathrm{z}}(\mathbb{R}^d, \mathbb{R})$, the space of continuous, bounded functions that vanish at infinity. The norm on B is $\|f\| = \sup_{x\in\mathbb{R}^d} |f(x)|$.

The Heat Semi-group

The heat semi-group is

$$P_t f(x) = \int \frac{1}{(2\pi t)^{d/2}} \exp\left(\frac{-|x-y|^2}{2t}\right) f(y)\,\mathrm{d}y$$

for a measurable function g which is bounded or in $\mathrm{L}^2(\mathbb{R}^d)$, the space of square integrable functions.

Since the marginal distribution of the Brownian motion B at any time t is normal one with mean 0 and variance t, $P_t f(x) = \mathbb{E}[f(x + B_t)]$.

Using Fourier transform or computing derivatives,

$$\partial_t P_t f(x) = \frac{1}{2}\triangle P_t f(x),\ x \in \mathbb{R}^d,\ t > 0.$$

Multiplying the above equation by $g(x) \in \mathrm{C}^2_\mathrm{c}(\mathbb{R}^d, \mathbb{R})$, performing an integration by parts then integrating between 0 and t lead to

$$\int_{\mathbb{R}^d} (P_t g(x) - g(x)) f(x)\,\mathrm{d}x = \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} P_t f(x) \triangle g(x)\,\mathrm{d}x.$$

Then, passing to the limit and since f is freely chosen,

$$\lim_{t\to 0} \frac{P_t g(x) - g(x)}{t} = \frac{1}{2}\triangle g(x),\ \forall x \in \mathbb{R}^d,\ g \in \mathrm{C}^2_\mathrm{c}(\mathbb{R}^d, \mathbb{R}). \tag{8.37}$$

Thus, if $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ is the infinitesimal generator of $(P_t)_{t\geq 0}$ (this is necessarily a close operator), $\mathfrak{L} = \frac{1}{2}\triangle$ on $\mathscr{C}^2_\mathrm{c}(\mathbb{R}^d, \mathbb{R}) \subset \mathrm{Dom}(\mathfrak{L})$. The latter space being dense in the underlying space $\mathrm{C}_\mathrm{z}(\mathbb{R}^d, \mathbb{R})$ with respect to $\|\cdot\|_\infty$ and $\mathrm{Dom}(\mathfrak{L})$ with respect to the graph norm $\|\cdot\|_\infty + \|\mathfrak{L}\cdot\|_\infty$, (8.37) characterizes $(\mathfrak{L}, \mathrm{Dom}(\mathfrak{L}))$ when $\mathrm{C}_\mathrm{z}(\mathbb{R}^d, \mathbb{R})$ is the ambient Banach space.

In other words, we recover that the infinitesimal generator of the Brownian motion is $\mathfrak{L} = \frac{1}{2}\triangle$ with a suitable domain. This could of course be easily obtained from the Itô's formula. Also, we see the link between the heat equation (8.10), the Brownian motion and the heat semi-group.

The Transport Semi-group

Let us consider now the flow $(\mathscr{Y}_t)$. Since $\mathscr{Y}(x)$ is solution to $\mathscr{Y}_t(x) = x + \int_0^t b(\mathscr{Y}_s(x))\,\mathrm{d}s$, the Newton formula for $f \in \mathrm{C}^1_\mathrm{c}(\mathbb{R}^d, \mathbb{R})$ implies that

$$f(\mathscr{Y}_t(x)) - f(x) = \int_0^t b(\mathscr{Y}_s(x)) \nabla f(\mathscr{Y}_s(s))\,\mathrm{d}s.$$

Hence, the infinitesimal generator of $(Q_t)_{t\geq 0}$ is $\mathfrak{B} = b\nabla\cdot$ whose domain $\mathrm{Dom}(\mathfrak{B})$ is the closure of $\mathrm{C}^1_\mathrm{c}(\mathbb{R}^d, \mathbb{R})$ to the graph norm (see e.g. [17, § II.3.28, p. 91]).

The semi-group $(Q_t)_{t>0}$ is also linked to a PDE, called the *transport equation*. We have seen in Proposition 8.2 that $x \mapsto \mathscr{Y}_t(x)$ is differentiable. It is actually of class $\mathrm{C}^1(\mathbb{R}^d, \mathbb{R})$. Thus, $Q_t f(x) = f \circ \mathscr{Y}_t(x)$ is also differentiable. Applying the

Newton formula to $f \circ \mathscr{Y}_t = Q_t f$ and using the flow property of $\mathscr{Y}_t$ leads to

$$f(\mathscr{Y}_{t+\epsilon}(x)) - f(\mathscr{Y}_t(x)) = Q_t f(\mathscr{Y}_\epsilon(x)) - Q_t f(x) = \int_0^\epsilon b(\mathscr{Y}_s(x)) \nabla Q_t f(\mathscr{Y}_s(x))\, \mathrm{d}s.$$

Dividing each side by ϵ and passing to the limit implies that

$$\partial_t Q_t f(x) = b(x) \nabla Q_t f(x).$$

Conversely, it is also possible to start from the transport equation $\partial_t u(t, x) = b(x)\nabla u(t, x)$ to construct the flow $\mathscr{Y}$ through the so-called *method of characteristics*, that is to find the paths $\mathscr{Z} : \mathbb{R}_+ \to \mathbb{R}^d$ such that $u(t, \mathscr{Z}_t)$ is constant over the time t.

References

1. J. Akahori, T. Amaba, S. Uraguchi, An algebraic approach to the Cameron-Martin-Maruyama-Girsanov formula. Math. J. Okayama Univ. **55**, 167–190 (2013)
2. L. Ambrosio, D. Trevisan, Lecture notes on the DiPerna-Lions theory in abstract measure spaces (2015). arXiv: 1505.05292
3. R. Bafico, P. Baldi, Small random perturbations of Peano phenomena. Stochastics **6**(3–4), 279–292 (1981/1982). https://doi.org/10.1080/17442508208833208
4. V. Bally, A. Kohatsu-Higa, A probabilistic interpretation of the parametrix method. Ann. Appl. Probab. **25**(6), 3095–3138 (2015). https://doi.org/10.1214/14-AAP1068
5. V. Bally, C. Rey, Approximation of Markov semigroups in total variation distance. Electron. J. Probab. **21**, 44 (2016). Paper No. 12. https://doi.org/10.1214/16-EJP4079
6. D.R. Bell, *The Malliavin Calculus*. Pitman Monographs and Surveys in Pure and Applied Mathematics (Longman Scientific & Technical/Wiley, Harlow/New York, 1987)
7. A. Beskos, G.O. Roberts, Exact simulation of diffusions. Ann. Appl. Probab. **15**(4), 2422–2444 (2005). https://doi.org/10.1214/105051605000000485
8. S. Blanes, F. Casas, A. Murua, Splitting and composition methods in the numerical integration of differential equations. Bol. Soc. Esp. Mat. Apl. S^e MA **45**, 89–145 (2008)
9. R.H. Cameron, W.T. Martin, Transformations of Wiener integrals under translations. Ann. Math. (2) **45**, 386–396 (1944)
10. R.H. Cameron, W.T. Martin, Transformations of Wiener integrals under a general class of linear transformations. Trans. Am. Math. Soc. **58**, 184–219 (1945)
11. R.H. Cameron, W.T. Martin, The transformation of Wiener integrals by nonlinear transformations. Trans. Am. Math. Soc. **66**, 253–283 (1949). https://doi.org/10.109/S0002-9947-1949-0031196-6
12. P.R. Chernoff, *Product Formulas, Nonlinear Semigroups, and Addition of Unbounded Operators*. Memoirs of the American Mathematical Society, vol. 140 (American Mathematical Society, Providence, 1974)
13. R.J. DiPerna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. Invent. Math. **98**(3), 511–547 (1989)
14. C. Doléans-Dade, Quelques applications de la formule de changement de variables pour les semimartingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **16**, 181–194 (1970)
15. J.L. Doob, Conditional Brownian motion and the boundary limits of harmonic functions. Bull. Soc. Math. France **85**, 431–458 (1957)

16. D. Duffie, P. Protter, From discrete- to continuous-time finance: weak convergence of the financial gain process. J. Math. Finance **2**, 1–15 (1992)
17. K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics, vol. 194. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli, R. Schnaubelt (Springer, New York, 2000)
18. M. Erraoui, E.H. Essaky, Canonical representation for Gaussian processes, in *Séminaire de Probabilités XLII*. Lecture Notes in Mathematics, vol. 1979 (Springer, Berlin, 2009), pp. 365–381. https://doi.org/10.1007/978-3-642-01763-6_13.
19. F. Flandoli, Remarks on uniqueness and strong solutions to deterministic and stochastic differential equations. Metrika **69**(2–3), 101–123 (2009). https://doi.org/10.1007/s00184-008-0210-7
20. I.V. Girsanov, On transforming a class of stochastic processes by absolutely continuous substitution of measures. Theor. Probab. Appl. **5**, 285–301 (1960). https://doi.org/10.1137/1105027
21. J.A. Goldstein, Remarks on the Feynman-Kac formula, in *Partial Differential Equations and Dynamical Systems*. Research Notes in Mathematics, vol. 101 (Pitman, Boston, 1984), pp. 164–180
22. H. Gzyl, The Feynman-Kac formula and the Hamilton-Jacobi equation. J. Math. Anal. Appl. **142**(1), 74–82 (1989). https://doi.org/10.1016/0022-247X(89)90165-0
23. E. Hairer, C. Lubich, G. Wanner, Structure-preserving algorithms for ordinary differential equations, in *Geometric Numerical Integration*. Springer Series in Computational Mathematics, 2nd edn., vol. 31 (Springer, Berlin, 2006)
24. H. Hult, Approximating some Volterra type stochastic integrals with applications to parameter estimation. Stoch. Process. Appl. **105**(1), 1–32 (2003). https://doi.org/10.1016/S0304-4149(02)00250-8
25. M. Jeanblanc, J. Pitman, M. Yor, The Feynman-Kac formula and decomposition of Brownian paths. Mat. Appl. Comput. **16**(1), 27–52 (1997)
26. M. Jeanblanc, M. Yor, M. Chesney, *Mathematical Methods for Financial Markets*. Springer Finance (Springer, London, 2009). https://doi.org/10.1007/978-1-84628-737-4
27. M. Kac, On distributions of certain Wiener functionals. Trans. Am. Math. Soc. **65**, 1–13 (1949)
28. T. Kato, Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups, in *Topics in Functional Analysis (Essays Dedicated to M. G. Kreĭn on the Occasion of his 70th Birthday)*. Advances in Mathematics Supplied Studies, vol. 3 (Academic, New York, 1978), pp. 185–195
29. A. Kohatsu-Higa, A. Lejay, K. Yasuda, Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift. J. Comput. Appl. Math. **326C**, 138–158 (2017). https://doi.org/10.1016/j.cam.2017.05.015
30. T.G. Kurtz, P.E. Protter, Weak convergence of stochastic integrals and differential equations, in *Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995)*. Lecture Notes in Mathematics, vol. 1627 (Springer, Berlin, 1996), pp. 1–41. https://doi.org/10.1007/BFb0093176
31. R. Leandre, Applications of the Malliavin calculus of Bismut type without probability. WSEAS Trans. Math. **5**(11), 1205–1210 (2006)
32. S. Lie, F. Engel, *Theorie der Transformationsgruppen* (Teubner, Leipzig, 1888)
33. R.S. Liptser, A.N. Shiryaev, *Statistics of Random Processes. I. General Theory*. Applications of Mathematics (New York), 2nd edn., vol. 5 (Springer, Berlin, 2001). Translated from the 1974 Russian original by A.B. Aries. https://doi.org/10.1007/978-3-662-13043-8
34. R.S. Liptser, A.N. Shiryaev, *Statistics of Random Processes. II. Applications*. Applications of Mathematics (New York), 2nd edn., vol. 6 (Springer, Berlin, 2001). Translated from the 1974 Russian original by A. B. Aries. https://doi.org/10.1007/978-3-662-10028-8
35. B. Maisonneuve, Quelques martingales remarquables associées à une martingale continue. Publ. Inst. Statist. Univ. Paris 17.fasc. **3**, 13–27 (1968)

36. P. Malliavin, *Stochastic Analysis*. Grundlehren der Mathematischen Wissenschaften, vol. 313 (Springer, Berlin, 1997). https://doi.org/10.1007/978-3-642-15074-6
37. G. Maruyama, On the transition probability functions of the Markov process. Nat. Sci. Rep. Ochanomizu Univ. **5**, 10–20 (1954)
38. R. Mikulevičius, E. Platen, Rate of convergence of the Euler approximation for diffusion processes. Math. Nachr. **151**, 233–239 (1991). https://doi.org/10.1002/mana.19911510114
39. R. Mikulevičius, C. Zhang, Weak Euler approximation for Itô diffusion and jump processes. Stoch. Anal. Appl. **33**(3), 549–571 (2015). https://doi.org/10.1080/07362994.2015.1014102
40. E. Nelson, Feynman integrals and the Schrödinger equation. J. Math. Phys. **5**, 332–343 (1964)
41. S. Ninomiya, N. Victoir, Weak approximation of stochastic differential equations and application to derivative pricing. Appl. Math. Finance **15**(1–2), 107–121 (2008). https://doi.org/1.108/13504860701413958
42. S. Orey, Conditions for the absolute continuity of two diffusions. Trans. Am. Math. Soc. **193**, 413–426 (1974)
43. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences, vol. 44 (Springer, New York, 1983). https://doi.org/10.1007/978-1-4612-5561-1
44. B. Roynette, M. Yor, *Penalising Brownian Paths*. Lecture Notes in Mathematics, vol. 1969 (Springer, Berlin, 2009). https://doi.org/10.1007/978-3-540-89699-9
45. B. Simon, Schrödinger semigroups. Bull. Am. Math. Soc. (N.S.) **7**(3), 447–526 (1982). https://doi.org/10.1090/S0273-0979-1982-15041-8
46. A.V. Skorohod, On the densities of probability measures in functional spaces, in *Proceedings of. Fifth Berkeley Symposium on Mathematical Statistics and Probability (Berkeley, CA, 1965/1966)* (University of California Press, Berkeley, 1967); vol. II: Contributions to Probability Theory, Part 1
47. D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes*. Classics in Mathematics (Springer, Berlin, 2006). Reprint of the 1997 edn.
48. S. Takanobu, On the error estimate of the integral kernel for the Trotter product formula for Schrödinger operators. Ann. Probab. **25**(4), 1895–1952 (1997). https://doi.org/10.1214/aop/1023481116
49. D. Talay, Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. RAIRO Modél. Math. Anal. Numér. **20**(1), 141–179 (1986)
50. H.F. Trotter, On the product of semi-groups of operators. Proc. Am. Math. Soc. **10**, 545–551 (1959)
51. J. H. Van Schuppen, E. Wong, Transformation of local martingales under a change of law. Ann. Probab. **2**, 879–888 (1974)
52. A.J. Veretennikov, On strong solutions and explicit formulas for solutions of stochastic integral equations. Math. USSR Sb. **39**(3), 387–403 (1981). https://doi.org/10.1070/SM1981v039n03ABEH001522
53. A. Zvonkin, A transformation of the phase space of a diffusion process that removes the drift. Math. USSR Sb. **22**, 129–149 (1975). https://doi.org/10.1070/SM1974v022n01ABEH001689

Chapter 9
On Drifting Brownian Motion Made Periodic

Paul McGill

Abstract The Brownian reference measure on periodic functions provides a framework for investigating more general circular processes. These include a significant class of periodic diffusions. We illustrate by proposing simple analytic criteria for finiteness and absolute continuity of the intrinsic circular measure associated to drifting Brownian motion. Our approach exploits a property of approximate bridges.

Keywords Drifting Brownian motion · Approximate bridge · Circular measure · Girsanov formula

AMS Classification 2010 Primary: 60J65; Secondary: 60H10, 28C20

9.1 Introduction

Notation Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\mathcal{M}_\phi = \phi^2 - \phi'$. Write $\mathbb{P}_x[z_t = y]$ for the semigroup density of the real diffusion (z_t). We reserve F to represent a generic bounded Borel functional on $\mathbb{C}([0, \infty))$, $\mathbb{C}([0, 1])$ or $\mathbb{C}(\mathbb{T})$ as appropriate. Processes are indexed by $[0, 1]$, unless otherwise indicated, and $C > 0$ is constant.

Any regular real diffusion (z_t) specifies a unique measure $\check{z}$ on $\mathbb{C}(\mathbb{T})$ via the relation

$$\check{\mathbb{E}}[F(\check{z})] = \int_{\mathbb{C}(\mathbb{T})} F(\gamma)\check{z}(d\gamma)$$
$$= \int_{\mathbb{R}} \mathbb{E}_x[F(z)|z_1 = x]\,\mathbb{P}_x[z_1 = x]dx \qquad (9.1)$$

P. McGill (✉)
Département de Mathématiques, Université Claude Bernard Lyon 1, Villeurbanne, France
e-mail: pmcgill@protonmail.com

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_9

where $F \geq 0$. This mixture of periodic bridges is intrinsic to the (z_t) semigroup. We drafted $\check{\mathbb{E}}$ to signal the possibility of infinite mass; a finite $\check{z}$ determines the law of a periodic process—aka (z_t) made periodic.

For the case of Brownian motion (β_t), translation invariance means that

$$\check{\mathbb{E}}[F(\check{\beta})] = \int_{\mathbb{R}} \mathbb{E}_x[F(x+\xi)] \frac{dx}{\sqrt{2\pi}} \tag{9.2}$$

using the standard bridge $(\xi_t) \overset{\text{law}}{=} (\beta_t - t\beta_1)$ for $\beta_0 = 0$. Compare [3] but beware their normalization. The circular Brownian measure furnishes a convenient framework for doing analysis on $\mathbb{C}(\mathbb{T})$. Applications include the study of periodic diffusions with explicit Radon-Nikodym derivative.

We illustrate by proposing simple analytic criteria for absolute continuity $\check{z} \ll \check{\beta}$ and $\check{z} < \infty$ in the case of drifting Brownian motion, namely $\exists\, \phi : \mathbb{R} \to \mathbb{R}$ such that

$$dz_t = d\beta_t - \phi(z_t)dt. \tag{9.3}$$

Restricting to solutions of (9.3) facilitates calculation. It also opens a fresh perspective on [2, 3]. Example 9.3 in section four describes the Ricatti circular measure that features, albeit implicitly, in [3] and [9].

The approach adopted below will start from the absolute continuity of individual bridges. We therefore broaden our remit to embrace

$$\mathbb{P}_x[z \in .\,] \ll \mathbb{P}_x[\beta \in .\,] \quad ; \quad \mathbb{P}_{x\to y}[z \in .\,] \ll \mathbb{P}_{x\to y}[\beta \in .\,] \quad ; \quad \check{z} \ll \check{\beta}. \tag{9.4}$$

Conditions that guarantee the first are well-understood (Novikov, Kazamaki,..). We will establish the other parts under decidedly weaker assumptions.

Our underpinning hypothesis stipulates that

$$\phi \textit{ has locally bounded variation and } \mathcal{M}_\phi(da) \geq -C da \tag{9.5}$$

noting how the Radon measure $\mathcal{M}_\phi(da)$ features in the exponential local martingale

$$\begin{aligned} \mathcal{E}_t(\phi, \beta) &= \exp\left(-\textstyle\int_0^t \phi(\beta_s)d\beta_s - \frac{1}{2}\int_0^t \phi^2(\beta_s)ds\right) \\ &= \exp\left(\Phi(\beta_0) - \Phi(\beta_t) - \tfrac{1}{2}\textstyle\int_0^t \mathcal{M}_\phi(\beta_s)ds\right). \end{aligned} \tag{9.6}$$

Condition (9.5), while effective and widely applicable, is far from optimal. We deploy it here in order to shorten the proofs. Recall also, e.g. [11], that $\Phi := \int_0^{\cdot} \phi$ satisfies $\Phi' \overset{\text{a.e.}}{=} \phi$.

Remark 9.1

(1) The circular measure $\check{z}$ determines all periodic bridge laws and the mixing density $\mathbb{P}_x[z_1 = x]$.

(2) Condition (9.5) says that $h \geq -C$ a.e. and $\mu^\perp \geq 0$ in the Lebesgue decomposition $\mathcal{M}_\phi(da) = h(a)da + \mu^\perp(da)$.
(3) $\mathcal{M}_\phi$ is Miura's map of mKdV/KdV renown [9].

The main result of this paper sets out sufficient conditions for (9.4). In the statement we conflate derivative and Radon-Nikodym factor.

Proposition 9.1 *Assume condition (9.5).*

(P1) If $\Phi \geq -C$ *then* $\mathbb{P}_x[z \in .] \ll \mathbb{P}_x[\beta \in .]$ *with derivative* $\mathcal{E}_1(\phi, \beta)$.
(P2) The derivative for $\mathbb{P}_{x\to y}[\, z \in .\,] \ll \mathbb{P}_{x\to y}[\, \beta \in .\,]$ *equals*

$$e^{\Phi(x)-\Phi(y)}e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(\beta)}\mathbb{P}_x[\beta_1 = y]/\mathbb{P}_x[z_1 = y].$$

(P3) $\check{z}(d\gamma) = \exp\{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(\gamma)\}\check{\beta}(d\gamma)$.
(P4) $\check{z}$ *has finite total mass iff* $\int dx\, \mathbb{E}[\,\exp\{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(x+\xi)\}\,]$ *converges.*

Remark 9.2

(1) Well-known (P1) handles an ultimately restoring drift. We use it to introduce the proof of (P2).
(2) Proposition 9.1, together with Corollary 9.1 below, refines [2, Section 3].
(3) Subject to existence, formula (P2) befits the continuous version of $y \to \mathbb{E}_x[F(z)|z_1 = y]$.
(4) The measure $\check{z}$ is finite iff $\int \mathbb{P}_x[z_1 = x]dx$ converges but direct verification is fraught. The practical test set forth in Corollary 9.1 wields the criterion (P4).
(5) From (9.6) we see that $\psi = e^{-\Phi}$ solves $\psi''/\psi = \mathcal{M}_\phi$. In [7, Section 8.6], they apply the converse. Given a suitable k, the logarithmic derivative of the groundstate for $\psi''/\psi = k$ defines ϕ satisfying $k = \mathcal{M}_\phi$.

9.2 Proof of Proposition 9.1

The key result of this section, Lemma 9.1, deals with absolute continuity of the approximate bridges. Our rationale for allowing the endpoint to vary in a non-trivial bounded interval is that this expedites the deployment of martingale calculus. We begin however by recalling some consequences of condition (9.5).

Consider first the implications for (9.3) which describes a diffusion having scale and speed measure given by

$$s'(x) = e^{2\Phi(x)} \quad ; \quad m(dx) = 2e^{-2\Phi(x)}dx. \tag{9.7}$$

Applying Itô's formula to $u = s(z)$ we get

$$du_t = s'(z_t)d\beta_t + [\tfrac{1}{2}s'' - \phi s'](z_t)dt$$
$$= s' \circ s^{-1}(u_t)d\beta_t \tag{9.8}$$

where, as noted in [12], local boundedness of ϕ implies $s' \circ s^{-1}$ is locally Lipschitz. So for $z_0 = x$ we conclude (e.g. [10, Ch.IX §2]) that Eq. (9.8), and consequently (9.3), admits a unique strong solution on its interval of non-explosion $[0, \zeta)$.

Next we consider the ramifications of (9.5) for the local martingale (9.6). Since the distributional derivative ϕ' is a Radon measure we may specify

$$\int_0^t \mathcal{M}_\phi(\beta_s)ds = \int \mathcal{M}_\phi(da)L_t^a$$

by fixing a bicontinuous version of Brownian local time $(a, t) \to L_t^a$. Moreover, noting the lower bound

$$\int_0^t \mathcal{M}_\phi(\beta_s)ds = \int L_t^a \mathcal{M}_\phi(da)$$
$$\geq -C \int L_t^a da$$
$$= -C\, t, \tag{9.9}$$

it follows from $\Phi \geq -C$ that (9.6) is uniformly bounded when $\beta_0 = x$.

Proof of (P1) By the preceding $\mathbb{Q}_x[A] = \mathbb{E}_x[\mathcal{E}_1(\phi, \beta); A]$ defines a probability law on the Borel sets of $\mathbb{C}[0, 1]$ such that, *vide* [10, Chapter VIII],

$$\beta^{\mathbb{Q}} = \beta + \int_0^{\cdot} \phi(\beta_s)ds$$

is a $\mathbb{Q}$-Brownian motion. Uniqueness in (9.3) then validates

$$\mathbb{E}_x[F(z)] = \mathbb{E}_x^{\mathbb{Q}}[F(\beta)]$$
$$= \mathbb{E}_x[F(\beta)\mathcal{E}_1(\phi, \beta)]. \tag{9.10}$$

Proof (P2) ⇒ (P3–4) Rearranging then integrating (P2) with $x = y$, we apply (9.1) on the left and (9.2) on the right to obtain

$$\check{\mathbb{E}}[F(\check{z})] = \int_{\mathbb{R}} \mathbb{E}\left[F(x+\xi)e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(x+\xi)} \right] \frac{dx}{\sqrt{2\pi}}.$$

This is the formal expression for (P3) and assigning $F \equiv 1$ yields (P4).

To complete the proof of Proposition 9.1 it remains to justify (P2). There are two steps. First we verify absolute continuity when the endpoint lies in a bounded interval. Invoking semigroup regularity then lets us secure same for the bridge.

For the next lemma we temporarily lift the restriction that all processes are defined on [0, 1] and introduce $(\mathcal{F}_t)_{t\geq 0}$ to represent the underlying filtration. Recall the notation $\beta^T := (\beta_{t\wedge T})$ and standard convention whereby F vanishes on the explosion set $(\zeta \leq 1)$.

Lemma 9.1 *Under condition (9.5) the relation*

$$\mathbb{E}_x[F(z_{\cdot\wedge 1})g(z_1)] = \mathbb{E}_x\left[F(\beta_{\cdot\wedge 1})g(\beta_1)\mathcal{E}_1(\phi,\beta)\right]$$

holds for all $g \in \mathbb{C}_c(\mathbb{R})$.

Proof Introducing $T_n = \inf\{t > 0 : |\beta_t| \geq n\}$, and writing $S_n = T_n \wedge n$, then (9.9) ensures uniform boundedness of $(\mathcal{E}_t^{S_n}(\phi,\beta))$ when $\beta_0 = x$. The argument justifying (9.10) now yields

$$\mathbb{E}_x[F(z^{S_n}_{\cdot\wedge 1})g(z^{S_n}_1); A] = \mathbb{E}_x[F(\beta^{S_n}_{\cdot\wedge 1})g(\beta^{S_n}_1)\mathcal{E}_{S_n}(\phi,\beta); A]$$

for arbitrary $A \in \mathcal{F}_{S_n}$. Fixing $A = (S_n > 1) \in \mathcal{F}_1$ the equation simplifies

$$\mathbb{E}_x[F(z_{\cdot\wedge 1})g(z_1); S_n > 1] = \mathbb{E}_x[F(\beta_{\cdot\wedge 1})g(\beta_1)\mathcal{E}_1(\phi,\beta); S_n > 1]$$

by Doob's theorem. However both integrands are uniformly bounded: clearly for each term on the left while the right side has bounded factors

$$F(\beta_{\cdot\wedge 1}) \quad ; \quad e^{\Phi(x)-\Phi(\beta_1)}g(\beta_1) \quad ; \quad e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(\beta)}.$$

Passing to limit we obtain

$$\mathbb{E}_x[F(z_{\cdot\wedge 1})g(z_1); \zeta > 1] = \mathbb{E}_x[F(\beta_{\cdot\wedge 1})g(\beta_1)\mathcal{E}_1(\phi,\beta)]$$

and the explosion set does not contribute.

Proof of (P2) Defining $F_n(\beta) = F(\beta^{t_n})$ where $t_n = (n-1)/n$, Lemma 9.1 implies that for $\mu(da) := \frac{1}{2}\mathcal{M}_\phi(a)da + \frac{1}{2}C da \geq 0$ and almost every y

$$\begin{aligned}
&\mathbb{E}_x[F_n(z)|z_1 = y]\mathbb{P}_x[z_1 = y] \\
&\quad = e^{\Phi(x)-\Phi(y)+\frac{1}{2}C}\mathbb{E}_x[F_n(\beta)e^{-\int L_1^a \mu(da)}|\beta_1 = y]\mathbb{P}_x[\beta_1 = y]. \qquad (9.11)
\end{aligned}$$

We claim that each term in (9.11) has a y-continuous version. This ensures equality and hence, since the F_n constitute a determining family, identity of the measures in question. To verify our claim we recall first from [4, §4.11], that the semigroup density of β killed at rate μ

$$(x, y, t) \to p_t^{(\mu)}(x, y) = \mathbb{E}_x\left[e^{-\int L_t^a \mu(da)} \mid \beta_t = y\right]\mathbb{P}_x[\beta_t = y] > 0 \tag{9.12}$$

is jointly continuous in all three variables when $t > 0$. Next, for $g \in \mathbb{C}_c(\mathbb{R})$, we apply the Markov property to decompose

$$\begin{aligned}\mathbb{E}_x\left[F_n(\beta)e^{-\int L_1^a \mu(da)} g(\beta_1)\right] &= \mathbb{E}_x^{(\mu)}\left[F_n(\beta)\mathbb{E}_{\beta_{t_n}}^{(\mu)}\left[g(\beta_{1/n})\right]\right] \\ &= \int \mathbb{E}_x^{(\mu)}\left[h(\beta_{t_n})p_{1/n}^{(\mu)}(\beta_{t_n}, y)\right] g(y)dy\end{aligned}$$

using $h(\beta_{t_n}) := \mathbb{E}_x^{(\mu)}\left[F_n(\beta)|\beta_{t_n}\right]$. It follows that for almost every y

$$\begin{aligned}&\mathbb{E}_x\left[F_n(\beta)e^{-\int L_1^a \mu(da)}|\beta_1 = y\right]\mathbb{P}_x[\beta_1 = y] \\ &= \mathbb{E}_x^{(\mu)}\left[h(\beta_{t_n})p_{1/n}^{(\mu)}(\beta_{t_n}, y)\right].\end{aligned}$$

However the last is continuous—inherited from (9.12) using h bounded and the Gaussian domination

$$p_{1/n}^{(\mu)}(x, y) \le \mathbb{P}_x[\beta_{1/n} = y] \le \sqrt{n/2\pi}.$$

This implies continuity of the conditional expectation and hence, Φ being continuous, continuity for all terms on the right side of (9.11). Since the same holds on the left, by continuity of $\mathbb{P}_x[z_1 = y] > 0$, our claim is vindicated.

This completes the proof of Proposition 9.1. The result suggests a wide range of applications, notably in the area of semigroup perturbation [2, 3], but we confine our immediate focus to the repercussions for $\check{z}$.

9.3 Applications

The combination of (P3) with (9.2) provides a general method for establishing properties of $\check{z}$. Besides their wider import, the corollaries below were selected for their relevance to the examples of the next section.

The first concerns a test for finiteness. Starting from (P4), we seek an approximating function $m_\phi(x)$ for $\int_0^1 \mathcal{M}_\phi(x+\xi)$ such that $\check{z} < \infty$ holds whenever $\exp\{-\frac{1}{2}m_\phi\} \in L_1(\mathbb{R})$. Our options for m_ϕ are dictated by the choice of approximation estimate.

As a practical demonstration we apply Kuiper's [6] oscillation bound

$$\mathbb{P}[\xi_1^\bullet - \xi_1^\circ > x] = 2\sum_{k\geq 1}(4x^2k^2 - 1)e^{-2x^2k^2},$$

where $(\xi_t^\bullet)$ (resp. (ξ_t°)) denotes the maximum (resp. minimum), to control the error in the case $m_\phi^\Delta(x) = \inf_{|2u|\leq|x|} \mathcal{M}_\phi(x+u)$. Our proof employs the maximal process $\xi^* = \sup(\xi^\bullet, -\xi^\circ)$.

Corollary 9.1 *Assume ϕ' exists for $|x| \geq N$ sufficiently large. Then integrability at infinity of* $\exp\{-\frac{1}{2}m_\phi^\Delta\}$ *guarantees $\check{z} < \infty$.*

Proof To verify (P4) we note first that (9.9) lets us assume $|x|$ large. Then by the same token

$$\mathbb{E}\left[e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(x+\xi)}; \xi_1^* > |x|/2\right] \leq e^{\frac{1}{2}C}\mathbb{P}\left[\xi^* > |x|/2\right]$$
$$\leq e^{\frac{1}{2}C}\mathbb{P}\left[\xi_1^\bullet - \xi_1^\circ > |x|/2\right],$$

which has Gaussian decay, while for $|x| > 2N$

$$\mathbb{E}\left[e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(x+\xi)}; \xi_1^* \leq |x|/2\right] \leq e^{-\frac{1}{2}m_\phi^\Delta(x)}$$

and our hypothesis applies.

Remark 9.3

(1) We can replace m_ϕ^Δ by $\mathcal{M}_\phi$ if this is monotone far out. Thus polynomial drift conforms.
(2) Applying $\mathbb{P}[\xi^\bullet > x] = e^{-2x^2}$ (see [1, IV.26]) on each half-line separately yields a better, if ultimately futile, Gaussian bound.
(3) Jeulin [5] studied $\int dx\, \mathbb{E}[\exp\{-\frac{1}{2}\int \mu(da)\ell_1^{a-x}\}]$ for μ Radon and (ℓ_t^a) the bridge local time. While (P4) is a special case, the method proves too cumbersome for the simple examples of the next section.
(4) Condition (9.5), and hence Corollary 9.1, fails for $\phi(x) = |x|$. Nevertheless, a different argument comprising an estimate with Brownian local time confirms that $\check{z} < \infty$ here too.
(5) Given Borel $\mathcal{R} : \mathbb{C}(\mathbb{T}) \to [0, \infty)$ then applying Corollary 9.1 to $\mathcal{M}_\phi - 2\mathcal{R}$ provides a test for existence of $\check{\mathbb{E}}[\exp\{\int_0^1 \mathcal{R}(\check{z})\}]$.

Our second application of Proposition 9.1 uncovers a general reflection formula. The proof invokes Brownian symmetry $\check{\mathbb{E}}[F(-\check{\beta})] = \check{\mathbb{E}}[F(\check{\beta})]$ itself a consequence of $(\xi_t) \stackrel{\text{law}}{=} (-\xi_t)$ in definition (9.2).

Corollary 9.2 *Defining* $\Delta\mathcal{M}_\phi(dx) = \mathcal{M}_\phi(dx) - \mathcal{M}_\phi(-dx)$ *then*

$$\check{\mathbb{E}}[F(-\check{z})] = \check{\mathbb{E}}\left[F(\check{z})e^{\frac{1}{2}\int_0^1 \Delta\mathcal{M}_\phi(\check{z})}\right]$$

holds $\forall\, F \geq 0$. *In particular,* $\check{z}$ *and* $\mathcal{M}_\phi$ *are symmetric together.*

Proof Combining (P3) with Brownian symmetry we find that

$$\begin{aligned}\check{\mathbb{E}}\big[F(-\check{z})\big] &= \check{\mathbb{E}}\big[F(-\check{\beta})e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(\check{\beta})}\big] \\ &= \check{\mathbb{E}}\big[F(\check{\beta})e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(-\check{\beta})}\big].\end{aligned}$$

It remains to identify the latter. Using (P3), which applies to bounded F, we deduce

$$\check{\mathbb{E}}\big[F(\check{z})e^{\frac{1}{2}\int_0^1 \Delta\mathcal{M}_\phi(\check{z})};\, A_n(\check{z})\big] = \check{\mathbb{E}}\big[F(\check{\beta})e^{-\frac{1}{2}\int_0^1 \mathcal{M}_\phi(-\check{\beta})};\, A_n(\check{\beta})\big]$$

for $A_n(z) = \{\int_0^1 \Delta\mathcal{M}_\phi(z) \leq n\}$ where it suffices to take the (monotone) limit.

There are other ways to investigate $\check{z}$. Given an explicit semigroup one can work directly from definition (9.1). E.g. in [3] they disintegrate $\check{\beta}$ wrt $\int_0^1 \check{\beta}$: for centred bridge $\bar{\xi} = \xi - \int_0^1 \xi$ and bounded Borel $g \geq 0$ then

$$\check{\mathbb{E}}\left[F(\check{\beta})g(\textstyle\int_0^1 \check{\beta})\right] = \displaystyle\int_{\mathbb{R}} \mathbb{E}\left[F(\bar{\xi} + y)\right] g(y)\frac{dy}{\sqrt{2\pi}}.$$

This formula offers an alternative to definition (9.2), which describes the disintegration over the initial = final value.

In the reverse direction, Remark 9.1(1) hints that periodic z-bridges may inherit certain attributes of $\check{z}$. See Lemma 9.2 for an interesting addendum.

9.4 Examples

Our examples are simplest possible; they have polynomial drift of degree at most two. So (9.5) applies automatically and Corollary 9.1 implies $\check{z} < \infty$ iff $\deg(\phi) \geq 1$.

Example 9.1 Brownian motion with constant drift $\lambda \in \mathbb{R}$ satisfies (9.4). In fact Proposition 9.1 handles all bar $\mathbb{P}_x[z \in .] \ll \mathbb{P}_x[\beta \in .]$ where we guarantee uniform integrability of (9.6) on $[0, 1]$ by $\mathbb{E}_x\left[\mathcal{E}_1^2(-\lambda, \beta)\right] < \infty$. Note also $\check{z} = e^{-\frac{1}{2}\lambda^2}\check{\beta}$ from (P3).

Example 9.2 For $\alpha \in \mathbb{R}_*$ the Ornstein-Uhlenbeck process $q = q(\alpha)$ solves

$$dq_t = d\beta_t - \alpha q_t dt.$$

Then (P1) gives $\mathbb{P}_x[q \in .] \ll \mathbb{P}_x[\beta \in .]$ when $\alpha > 0$. Remark however that the result extends to $\alpha > -\frac{1}{2}\pi$ via Novikov's criterion for (9.6), as verified by analytic continuation in the Cameron-Martin formula [10, XI.1.8],

$$\mathbb{E}_x\left[e^{-\frac{1}{2}\alpha^2\int_0^1\beta^2}\right] = (\cosh x)^{-1/2}e^{-\frac{1}{2}\alpha x \tanh\alpha}.$$

The other parts of (9.4) are immediate and we obtain the total mass for $\check{q}$ as in [3], using $\sigma^2 = (1-e^{-2\alpha})/2\alpha > 0$ to evaluate

$$\begin{aligned}\int_{\mathbb{R}} \mathbb{P}[q_0 = x = q_1]dx &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} e^{-(x-e^{-\alpha}x)^2/2\sigma^2}dx \\ &= \frac{e^{\alpha/2}}{2\sinh\frac{1}{2}|\alpha|}.\end{aligned}$$

Symmetry follows by Corollary 9.2, since $\mathcal{M}_\phi$ is even, while from (P3) we discover the drift-reversal relation $e^{\alpha/2}\check{q}(-\alpha) = e^{-\alpha/2}\check{q}(\alpha)$.

Example 9.3 The Ricatti process with real parameter λ satisfies

$$dp_t = d\beta_t - (\lambda + p_t^2)dt$$

on the interval $0 < t < \zeta = \inf\{t > 0 : p_t = -\infty\}$. Using (9.7) we identify $-\infty$ as an exit boundary in the sense of [4], 4.1. Thus $\mathbb{P}_x[p_t^* = \infty] > 0$ for all $t > 0$ meaning $\mathbb{P}_x[p \in .] \not\ll \mathbb{P}_x[\beta \in .]$. Nevertheless, the other assertions of (9.4) follow from Proposition 9.1. Lastly, if Λ_0 denotes the periodic groundstate eigenvalue for $\psi'' = -\psi(\lambda + \beta')$ on $[0, 1]$ then (P3) lets us rewrite the leading formula of [3] in Ricatti style

$$\begin{aligned}\mathbb{P}[\lambda < \Lambda_0] &= \check{\mathbb{E}}\left[1 - e^{-2\check{P}}\,;\ \check{P} > 0\right] \\ &= \check{\mathbb{P}}[\check{P} > 0] - \check{\mathbb{P}}[\check{P} < 0]\end{aligned}$$

with $\check{P} := \int_0^1 \check{p}$ and invoking Corollary 9.2 for the last.

We finish with a result concerning the reverse-drift process characterized by

$$dw_t = d\beta_t + \phi(w_t)dt.$$

Assuming (9.5), then (P3) shows that $\check{z} = C\check{w}$ iff $\mathcal{M}_\phi - \mathcal{M}_{-\phi} = 2\phi'$ is constant a.e. wrt $\check{\beta}$. So for linear drift Remark 9.1(1) suggests that (w_t) and (z_t) share periodic bridges. The confirmation, by substituting the explicit semigroup in formula (P2), is emphatic.

Lemma 9.2 *For ϕ linear then $\mathbb{E}_x[F(z)|z_1 = y] = \mathbb{E}_x[F(w)|w_1 = y]$.*

Remark 9.4

(1) Drift-reversal in the Ricatti SDE of Example 9.3 maps (p_t) to $(-p_t)$. So their respective bridges start, and finish, at opposing points.
(2) Another proof of Lemma 9.2 uses the SDE for (z_t) conditioned by z_1. When $\phi(x) = \alpha x + \lambda$ there exists a Brownian motion (b_t) satisfying

$$dz_t = db_t + \lambda \tanh \tfrac{1}{2}\alpha(1-t)dt$$
$$-\alpha \left[\frac{z_1 - z_t \cosh \alpha(1-t)}{\sinh \alpha(1-t)} \right] dt$$

and this is invariant for $(\alpha, \lambda) \to (-\alpha, -\lambda)$.
(3) Examples 9.1–9.2 suggest that dispersion diminishes the circular measure in the ordering defined by formula (P3).
(4) In [8] the authors consider absolute continuity of the complex Brownian bridge under a phase change depending only on the modulus. Their result exploits independence properties of the skew-product decomposition and is unrelated to Proposition 9.1.

References

1. A. Borodin, P. Salminen, *Handbook of Brownian Motion – Facts and Formulae* (Birkhaüser, Basel, 2002)
2. S. Cambronero, Some Ornstein-Uhlenbeck potentials for the one-dimensional Schrödinger operator part II: position-dependent drift. Revista de Mathemática: Teoria y Applicationes **9**(2), 31–38 (2002)
3. S. Cambronero, H.P. McKean, The ground state eigenvalue of Hill's equation with white noise potential. Commun. Pure Appl. Math. **52**(10), 1277–1294 (1999)
4. K. Itô, H.P. McKean, *Diffusion Processes and Their Sample Paths*, 2nd edn. (Springer, Berlin, 1974)
5. Th. Jeulin, Application de la théorie du grossissement à l'étude des temps locaux browniens, in *Grossissement de filtrations: exemples et applications*. Lecture Notes in Mathematics, vol. 1118 (Springer, Berlin, 1985), pp. 197–304
6. N.H. Kuiper, Tests concerning random points on a circle. Indag. Math. **22**, 32–37, 38–47 (1960)
7. H.P. McKean, K.T. Vaninsky, Statistical mechanics of nonlinear wave equations, in *Trends and Perspectives in Applied Mathematics*, ed. by L. Sirovich (Springer, Berlin, 1994), pp. 239–264
8. A. Nahmod, L. Rey-Bellet, S. Sheffield, G. Staffilani, Absolute continuity of Brownian bridges under certain gauge transformations. Math. Res. Lett. **18**(5), 875–887 (2011)
9. J. Quastel, B. Valko, KdV preserves white noise. Commun. Math. Phys. **277**(3), 707–714 (2008)

10. D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 3rd edn. (Springer, Berlin, 1999)
11. W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 1966)
12. A.K. Zvonkin, A transformation of the phase space of a process that removes the drift. Math. USSR Sb. **22**, 129–149 (1974)

Chapter 10
On the Markovian Similarity

Laurent Miclo

Abstract Two finite Markov generators L and $\widetilde{L}$ are said to be intertwined if there exists a Markov kernel Λ such that $L\Lambda = \Lambda\widetilde{L}$. The goal of this paper is to investigate the equivalence relation between finite Markov generators obtained by imposing mutual intertwinings through invertible Markov kernels, in particular its links with the traditional similarity relation. Some consequences on the comparison of speeds of convergence to equilibrium for finite irreducible Markov processes are deduced. The situation of infinite state spaces is also quickly mentioned, by showing that the Laplacians of isospectral compact Riemannian manifolds are weakly Markov-similar.

Keywords Markov generators · Markov intertwinings · Similarity relation · Isospectrality · Convergence to equilibrium · φ-Entropies

MSC 2010 Primary: 60J27, Secondary: 60J35, 60J25, 05C50, 37A30, 58J53

10.1 Introduction

Intertwining of Markov processes is an old subject, coming back to Rogers and Pitman [17] (or even to the book of Dynkin [8] for the deterministic version), which lately has attracted a renewed interest, see for instance the paper of Pal and Shkolnikov [15] and the references therein. Very recently, Patie and Savov [16] have used intertwinings between reversible Laguerre diffusions and certain non-local and non-reversible Markov processes to get information on the spectral decompositions of the latters. This arises a natural question: when are two Markov

L. Miclo (✉)
Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse Cedex, France
e-mail: miclo@math.univ-toulouse.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_10

processes intertwined? To avoid a trivial answer, we will introduce a notion of Markov-similarity, where a Markovian requirement is imposed on the relation of similitude. Indeed, we will begin by investigating its links with the usual similarity in the framework of general finite Markov processes. Next we will discuss the consequences for the comparison of mixing speeds of Markov-similar finite ergodic Markov processes. Then we will only scratch the surface of the corresponding question in the non-finite setting, in particular by checking that a weak Markov-similarity holds for isospectral Riemannian manifolds.

As announced, we first study the finite state space situation. Let V be a finite set, endowed with a **Markov generator** L: it is a $V \times V$ matrix $(L(x, y))_{x,y\in V}$ whose off-diagonal entries are non-negative and whose row sums vanish:

$$\forall x \in V, \qquad \sum_{y\in V} L(x, y) = 0.$$

Consider $\widetilde{L}$ another Markov generator on a finite set $\widetilde{V}$ (more generally, all objects associated to $\widetilde{L}$ will receive a tilde). A **Markov kernel** Λ from V to $\widetilde{V}$ is a $V \times \widetilde{V}$ matrix $(\Lambda(x, \widetilde{x}))_{(x,\widetilde{x})\in V\times\widetilde{V}}$ whose entries are non-negative and whose row sums are all equal to 1. We say that L is **intertwined** with $\widetilde{L}$, if there exists a Markov kernel Λ from V to $\widetilde{V}$ such that $L\Lambda = \Lambda\widetilde{L}$. If furthermore there exists a Markov kernel $\widetilde{\Lambda}$ from $\widetilde{V}$ to V such that $\widetilde{L}\widetilde{\Lambda} = \widetilde{\Lambda}L$, then L and $\widetilde{L}$ are said to be **mutually intertwined**. This notion is not very interesting, because any finite Markov generators L and $\widetilde{L}$ are always mutually intertwined. Indeed, any finite Markov generator L admits an **invariant probability measure** μ, namely satisfying $\mu[L[f]] = 0$ for all functions f defined on V (where we used the traditional matrix notations: any measure is seen as a row vector and any function as a column vector). Let $\widetilde{\mu}$ be an invariant measure for $\widetilde{L}$ and define two Markov kernels Λ and $\widetilde{\Lambda}$ by

$$\begin{aligned} \forall (x, \widetilde{x}) \in V \times \widetilde{V}, \qquad & \Lambda(x, \widetilde{x}) := \widetilde{\mu}(\widetilde{x}), \\ \forall (\widetilde{x}, x) \in \widetilde{V} \times V, \qquad & \widetilde{\Lambda}(\widetilde{x}, x) := \mu(x). \end{aligned}$$

By using these Markov kernels, it is immediate to check that L and $\widetilde{L}$ are mutually intertwined.

So let us add a more stringent requirement. A Markov kernel Λ from V to $\widetilde{V}$ is said to be a **link**, if it is invertible (as a matrix). Here we depart from the terminology introduced by Diaconis and Fill [5], since for them a link is just a Markov kernel. In particular, V and $\widetilde{V}$ have the same cardinality, which will be denoted $|V|$.

Definition 10.1 The Markov generators L and $\widetilde{L}$ are said to be **Markov-similar** if there exist two links Λ and $\widetilde{\Lambda}$, respectively from V to $\widetilde{V}$ and from $\widetilde{V}$ to V such that

$$L\Lambda = \Lambda\widetilde{L} \text{ and } \widetilde{L}\widetilde{\Lambda} = \widetilde{\Lambda}L \tag{10.1}$$

The first motivation of this paper stems from the natural question: when are two finite Markov generators Markov-similar?

Of course, two finite Markov-similar Markov generators are linked by a similitude relation, so they are similar in the usual sense, namely they have the same eigenvalues (in $\mathbb{C}$) and the corresponding Jordan blocks have the same dimensions. But despite the results presented in this introduction, the reverse implication is not always true, as we will see in Sect. 10.3.

Recall the usual notion of transience for the points of V relatively to L. Let $x, y \in V$, we say that x **leads to** y, if there exists a finite sequence $x = x_0, x_1, x_2, \ldots, x_l = y$, with $l \in \mathbb{Z}_+$, such that $L(x_{k-1}, x_k) > 0$ for all $k \in [\![l]\!]:=\{1, 2, \ldots, l\}$. A point $x \in V$ is said to be **transient**, if there exists $y \in S$ such that x leads to y but y does not lead to x. The finite Markov generator L is said to be **non-transient**, if there is no transient point. In particular, if L is **irreducible** (namely, any point $x \in V$ leads to any point $y \in S$), then L is non-transient.

Theorem 10.1 *Two non-transient Markov generators L and $\widetilde{L}$ are Markov-similar if and only if they are similar.*

It is well-known that the number of irreducible classes (whose definition will be recalled in the beginning of Sect. 10.3) of a non-transient Markov generator is an information included into the spectrum of L, since it is the multiplicity of the eigenvalue 0. So according to the above result, two finite Markov-similar non-transient Markov generators have the same number of irreducible classes. Nevertheless the cardinalities of these classes can be different. This may first sound strange (but this is the deep reason behind the aggregation (10.8) considered in the transient setting, see Sect. 10.3) and is illustrated by the example below.

Example 10.1 Assume that the finite set V is partitioned into $V = \sqcup_{n\in[\![\mathfrak{n}]\!]} C_n$, with $\mathfrak{n} \in \mathbb{N}$. For $n \in [\![\mathfrak{n}]\!]$, let be given μ_n a probability measure whose support is C_n. On each C_n, consider the generator $L_n:=\mu_n - I_{C_n}$ where I_{C_n} is the $C_n \times C_n$ identity matrix and where μ_n stands for the matrix whose rows are all equal to μ_n. The spectrum of $-L_n$ consists of the simple eigenvalue 0 and of the eigenvalue 1 with (geometric) multiplicity $|C_n| - 1$. Next, define the generator L on V which acts as L_n on C_n for all $n \in [\![\mathfrak{n}]\!]$, namely $L:= \oplus_{n\in[\![\mathfrak{n}]\!]} L_n$. Then the spectrum of $-L$ has the eigenvalue 0 with multiplicity $\mathfrak{n}$ and the eigenvalue 1 with multiplicity $|V| - \mathfrak{n}$. Thus L is diagonalizable and its similarity class is the set of diagonalizable matrices which are isospectral to L. In particular a generator $\widetilde{L}$ defined in a similar fashion will be Markov-similar to L if and only if $|\widetilde{V}| = |V|$ and $\widetilde{\mathfrak{n}} = \mathfrak{n}$. It follows that $\{|C_n| : n \in [\![\mathfrak{n}]\!]\}$ can be different from $\{|\widetilde{C}_n| : n \in [\![\mathfrak{n}]\!]\}$ (as multisets), for instance we can have $\mathfrak{n} = 2$, $|C_1| = 1$, $|C_2| = 3$, $|\widetilde{C}_1| = 2$ and $|\widetilde{C}_2| = 2$.

Proposition 10.5 of Sect. 10.3 gives an extension of Theorem 10.1 to subMarkov generators, which corresponds to Markov processes which can be absorbed.

Remark 10.1

(a) A more stringent requirement in (10.1) would impose that the links Λ and $\widetilde{\Lambda}$ are inverse of each other: $\Lambda\widetilde{\Lambda} = I$, the identity kernel, as in the usual similarity relation. But this implies that Λ is a **deterministic kernel**, in the sense there exists a bijection $\sigma : V \to \widetilde{V}$ such that

$$\forall (x, \widetilde{x}) \in V \times \widetilde{V}, \qquad \Lambda(x, \widetilde{x}) = \delta_{\sigma(x)}(\widetilde{x})$$

(see for instance [14]). The link $\widetilde{\Lambda}$ is then the deterministic kernel associated to σ^{-1}. It follows that

$$\forall (x, y) \in V^2, \qquad L(x, y) = \widetilde{L}(\sigma(x), \sigma(y)).$$

namely, L can be identified with $\widetilde{L}$, up to a permutation of the state space.

Under this form, it would correspond to a discrete and non symmetric version of the question "can one hear the shape of a drum?" popularized by Kac in [12], where Laplace operators on two-dimensional compact domains with Dirichlet condition on the boundary (assumed to be smooth or polygonal), should be replaced by finite Markovian generators.

(b) Consider links Λ, $\widetilde{\Lambda}$ such that (10.1) is satisfied with respect to some Markov generators L, $\widetilde{L}$. Then $\Lambda\widetilde{\Lambda}$ is an invertible Markov kernel from V to V which commutes with L:

$$L\Lambda\widetilde{\Lambda} \;=\; \Lambda\widetilde{L}\widetilde{\Lambda} \;=\; \Lambda\widetilde{\Lambda}L$$

and symmetrically, $\widetilde{\Lambda}\Lambda$ is an invertible Markov kernel from $\widetilde{V}$ to $\widetilde{V}$ commuting with $\widetilde{L}$. In [14], the convex set of Markov kernels commuting with a given Markov generator was studied, in particular in correlation with the notion of weak hypergroup, on which we will come back at the end of Sect. 10.4.

One of the interest of Markov-similarity of two generators is that it should enable the comparison between their speeds of convergence to equilibrium or to absorption. Assume that L is a finite irreducible Markov generator and let μ be its unique invariant probability. If m_0 is a given initial probability on V, define for any $t > 0$, $m_t := m_0 \exp(tL)$, the distribution at time t of the Markov process starting from m_0 and whose evolution is dictated by L. As t goes to infinity, m_t converges toward μ and there are several ways to measure the discrepancy between m_t and μ. Let $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function such that $\psi(1) = 0$. The set of such functions will be denoted Ψ. For $\psi \in \Psi$, the ψ-entropy of a probability measure m with respect to μ is given by

$$E_\psi[m|\mu] := \sum_{x \in V} \psi\left(\frac{m(x)}{\mu(x)}\right) \mu(x).$$

Consider the worst cases over the initial conditions, namely

$$\forall \psi \in \Psi,\ \forall t \geq 0, \qquad E(\psi, t) := \sup\{E_\psi[m_0 \exp(tL)|\mu] \;:\; m_0 \in \mathscr{P}(V)\}$$

where $\mathscr{P}(V)$ stands for the set of all probability measures on V. Then we have:

Proposition 10.1 *Let L and $\widetilde{L}$ be two Markov-similar generators. Then there exists a constant $T \geq 0$ such that*

$$\forall \psi \in \Psi,\ \forall t \geq 0, \qquad E(\psi, T+t) \ \leq\ \widetilde{E}(\psi, t) \text{ and } \widetilde{E}(\psi, T+t) \ \leq\ E(\psi, t)$$

where $\widetilde{E}(\psi, t)$ is defined as $E(\psi, t)$, but with L replaced by $\widetilde{L}$ and μ by $\widetilde{\mu}$.

So in some sense, after the warming up time T, the convergences to equilibrium are similar for the Markov processes generated by L and $\widetilde{L}$. More precise results in this direction will be given in Sect. 10.4, in particular for some initial distributions no warming up period is necessary, but the crucial quantitative estimation of T will remain to be investigated.

To extend the previous considerations to infinite state spaces, one must begin by choosing an appropriate notion of "non degeneracy" of the links. Recall that in general, a **Markov kernel** Λ from a measurable space $(V, \mathscr{V})$ to a measurable space $(\widetilde{V}, \widetilde{\mathscr{V}})$, is a mapping from $V \times \widetilde{\mathscr{V}}$ such that:

- for any $x \in V$, $\Lambda(x, \cdot)$ is a probability measure on $(\widetilde{V}, \widetilde{\mathscr{V}})$,
- for any $A \in \widetilde{\mathscr{V}}$, $\Lambda(\cdot, A)$ is a $(V, \mathscr{V})$-measurable mapping.

When V and $\widetilde{V}$ are finite (it is then understood that they are endowed with their full σ-algebras), one recovers the above definition, namely Λ can be identified with a $V \times \widetilde{V}$ matrix whose entries are non-negative and whose row sums are all equal to 1.

Let $\mathscr{B}$ (respectively $\widetilde{\mathscr{B}}$) the vector space of bounded measurable functions on $(V, \mathscr{V})$ (resp. $(\widetilde{V}, \widetilde{\mathscr{V}})$). A Markov kernel Λ from $(V, \mathscr{V})$, to $(\widetilde{V}, \widetilde{\mathscr{V}})$ induces an operator from $\widetilde{\mathscr{B}}$ from $\mathscr{B}$ via

$$\forall x \in V,\ \forall \widetilde{f} \in \widetilde{\mathscr{B}}, \qquad \Lambda[\widetilde{f}](x) := \int \widetilde{f}(\widetilde{x})\, \Lambda(x, d\widetilde{x}) \tag{10.2}$$

The Markov kernel is then said to be a **weak link** if it is a one-to-one operator.

Let L be a **Markov generator** on $(V, \mathscr{V})$, in the sense that it is defined on a subspace $\mathscr{D}(L)$ of $\mathscr{B}$ such that the corresponding martingale problems are well-posed for any initial condition (for a thorough exposition of these concepts, see e.g. the book of Ethier and Kurtz [9]). If V is finite, it corresponds to Definition 10.1 given in the beginning of this introduction.

Definition 10.2 The Markovian generators L and $\widetilde{L}$ are said to be **weakly Markov-similar**, if there exist two links, Λ from $(V, \mathscr{V})$ to $(\widetilde{V}, \widetilde{\mathscr{V}})$ and $\widetilde{\Lambda}$ from $(\widetilde{V}, \widetilde{\mathscr{V}})$ to $(V, \mathscr{V})$, such that

$$L\Lambda \ =\ \Lambda\widetilde{L} \text{ and } \widetilde{L}\widetilde{\Lambda} \ =\ \widetilde{\Lambda}L$$

In particular, these relations require that $\Lambda(\mathscr{D}(\widetilde{L})) \subset \mathscr{D}(L)$ and $\Lambda(\mathscr{D}(L)) \subset \mathscr{D}(\widetilde{L})$.

In these general definitions, we did not mention invariant probabilities, since when V is infinite, they may not exist. Nevertheless, if we are given a probability μ invariant for L, in the sense that

$$\forall f \in \mathscr{D}(L), \qquad \mu[L[f]] = 0$$

then the above notions can be slightly modified to be given a $\mathbb{L}^2$ flavor: the Markov operator defined in (10.2) can be extended into an operator from $\mathbb{L}^2(\mu)$ to $\mathbb{L}^2(\widetilde{\mu})$, with $\widetilde{\mu}:=\mu\Lambda$ and $(\mathscr{D}(L), L)$ can be replaced by its $\mathbb{L}^2(\mu)$-closure. The operator Λ is then Markovian, in the abstract sense that it preserves non-negativity and the functions taking only the value 1 (respectively μ- and $\widetilde{\mu}$-a.s.). Conversely, if the measurable spaces $(V, \mathscr{V})$ and $(\widetilde{V}, \widetilde{\mathscr{V}})$ are the Borelian spaces associated to Polish topological spaces, then any abstract Markovian operator from $\mathbb{L}^2(\widetilde{\mu})$ to $\mathbb{L}^2(\mu)$ corresponds to a Markov kernel. We are thus led naturally to the notions of weak (abstract) $\mathbb{L}^2(\mu)$-link and of weak (abstract) $\mathbb{L}^2$-Markov-similarity between L and $\widetilde{L}$, when $\widetilde{\mu}$ is left invariant by $\widetilde{L}$. Despite the fact that this subject would deserve a general investigation, here we restrict our attention to a very particular situation. We say that the Markov generator L with invariant probability μ is **nice**, if:

- The measurable space $(V, \mathscr{V})$ is the Borelian space associated to a Polish topological space.
- The operator L admits a unique invariant probability μ, which is in fact reversible, in the sense that

$$\forall f, g \in \mathscr{D}(L), \qquad \mu[fL[g]] = \mu[gL[f]] \tag{10.3}$$

  This assumption enables to consider the (Friedrich) minimal extension of L as a self-adjoint operator on $\mathbb{L}^2(\mu)$, with $\mathscr{D}(L)$ as new domain.
- The spectral decomposition of $-L$ only consists of eigenvalues, say $(\lambda_l)_{l\in\mathbb{Z}_+}$, with multiplicities.
- It is possible to choose a family $(\varphi_l)_{l\in\mathbb{Z}_+}$ of eigenvectors associated to the eigenvalues $(\lambda_l)_{l\in\mathbb{Z}_+}$, such that for any $l \in \mathbb{Z}_+$, the function φ_l is bounded (this is always true for the eigenvalue 0, since by the preceding point, its eigenspace is the set of the constant functions).

The interest of this notion is:

Proposition 10.2 *Two nice Markov generators L and $\widetilde{L}$ are weakly $\mathbb{L}^2$-Markov-similar if and only if they are isospectral.*

A typical example of a nice Markov generator is that of a reversible elliptic diffusion with regular coefficients on a compact manifold V. In this situation, one can endow V with a Riemannian structure and find a smooth function U, such that the underlying Markov generator L has the following form (known as Witten Laplacian)

$$L\cdot = \triangle \cdot - \langle \nabla U, \nabla \cdot \rangle$$

where $\triangle$ is the Laplace-Beltrami operator, $\langle\cdot,\cdot\rangle$ is the scalar product and ∇ is the gradient operator (see e.g. the book of Ikeda and Watanabe [11]). The corresponding reversible probability μ admits as density with respect to the Riemannian measure the one proportional to $\exp(-U)$. The compactness of V implies that the spectrum of $-L$ consists only of non-negative eigenvalues with finite multiplicities and without accumulation point. Denote them by

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \ldots$$

As solutions to elliptic equations, the corresponding eigenvectors are smooth and thus bounded.

Let $\widetilde{L}$ be another diffusion generator of the same kind (i.e. associated to a compact Riemannian manifold $\widetilde{V}$ and to a potential $\widetilde{U}$). Let $0 = \widetilde{\lambda}_0 < \widetilde{\lambda}_1 \leq \widetilde{\lambda}_2 \leq \ldots$ be its eigenvalues. As a consequence of Proposition 10.2, L and $\widetilde{L}$ are $\mathbb{L}^2$-Markov-similar if and only if we have $\lambda_n = \widetilde{\lambda}_n$ for all $n \in \mathbb{Z}_+$. In particular, the Laplace-Beltrami operators corresponding to isospectral compact Riemannian manifolds are $\mathbb{L}^2$-Markov-similar. This result enables the coupling by intertwining (generalizing the coupling constructed by Diaconis and Fill [5] for discrete time finite Markov chains) of the Brownian motions on such manifolds, suggesting that the question of isospectrality for compact Riemannian manifolds (see e.g. the review of Bérard [1] and references therein) could be revisited from a probabilistic point of view. The study of the links between the mixing speed of such Brownian motions, as in Proposition 10.2, is out of the scope of this paper.

The paper is organized as follows: the next section contains the proof of Theorem 10.1. Section 10.3 investigates the transient situation, where the characterization of Markov-similarity is not complete. The subMarkovian case will also be dealt with there. Section 10.4 collects the considerations about mixing speeds. The proof of Proposition 10.2 is given in the final section.

10.2 The Finite Non-transient Setting

This section is devoted to the proof of Theorem 10.1. It will be shown gradually, starting with the case of irreducible and reversible generators and ending with the general non-transient case.

Recall that a finite Markov generator L is said to be **reversible** with respect to a probability measure μ on V, if

$$\forall x, y \in V, \qquad \mu(x)L(x, y) = \mu(y)L(y, x)$$

This property is equivalent to the symmetry of L in $\mathbb{L}^2(\mu)$ mentioned in (10.3) for the general case. We begin by assuming that the generator L is irreducible and reversible. By irreducibility, the invariant measure μ is unique and positive on V.

The reversibility of L with respect to μ implies that L is diagonalizable. Denote the eigenvalues (with multiplicities) of $-L$ by

$$0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_{|V|} \tag{10.4}$$

(the strict inequality comes from irreducibility).

Consider another irreducible and reversible Markov generator $\widetilde{L}$. In this case, the similarity of L and $\widetilde{L}$ reduces the fact that L and $\widetilde{L}$ are isospectral, i.e. V and $\widetilde{V}$ have the same cardinality and

$$\forall k \in [\![|V|]\!], \qquad \widetilde{\lambda}_k = \lambda_k$$

Here is the first step in the direction of Theorem 10.1, it corresponds to Proposition 10.2 in the finite case.

Lemma 10.1 *Two finite, irreducible and reversible Markov generators are Markov-similar if and only if they are similar.*

Proof Let L and $\widetilde{L}$ be finite, irreducible and reversible Markov generators. If they are Markov-similar, there is a similarity relation between them, for instance $\widetilde{L} = \Lambda^{-1}L\Lambda$, so they are similar.

Conversely, assume that L and $\widetilde{L}$ are similar. Denote by (10.4) the common spectrum of $-L$ and $-\widetilde{L}$. Let $(\varphi_k)_{k\in[\![|V|]\!]}$ and $(\widetilde{\varphi}_k)_{k\in[\![|V|]\!]}$ be orthonormal bases of $\mathbb{L}^2(\mu)$ and $\mathbb{L}^2(\widetilde{\mu})$ consisting of corresponding eigenvectors. Without loss of generality, we can assume that $\widetilde{V} = V$ and that $\varphi_1 = \widetilde{\varphi}_1 = \mathbb{1}$ (the function always taking the value 1). To construct an invertible Markov kernel Λ from V to V such that $L\Lambda = \Lambda\widetilde{L}$, consider the operator A defined by

$$\forall k \in [\![|V|]\!], \qquad A[\widetilde{\varphi}_k] := \begin{cases} \varphi_k & \text{, if } k \geq 2 \\ 0 & \text{, if } k = 1 \end{cases}$$

For $\epsilon \in \mathbb{R}$, we are interested in the operator

$$\Lambda := \widetilde{\mu} + \epsilon A \tag{10.5}$$

where $\widetilde{\mu}$ is again interpreted as the matrix whose rows are all equal to the probability $\widetilde{\mu}$. It is immediately checked that

$$\forall k \in [\![|V|]\!], \qquad \Lambda[\widetilde{\varphi}_k] := \begin{cases} \epsilon\varphi_k & \text{, if } k \geq 2 \\ \varphi_1 & \text{, if } k = 1 \end{cases}$$

since by orthogonality, $\widetilde{\mu}[\widetilde{\varphi}_k] = \widetilde{\mu}[\widetilde{\varphi}_1\widetilde{\varphi}_k] = 0$. It implies the relation $L\Lambda = \Lambda\widetilde{L}$ and that Λ is invertible as soon as $\epsilon \neq 0$.

From the relation $\Lambda[\mathbb{1}] = \mathbb{1}$, it appears that the row sums of Λ are all equal to 1. Thus it remains to find $\epsilon \neq 0$ such that all the entries of Λ are non-negative. It is sufficient to take

$$0 < |\epsilon| \leq \min_{x,y\in V} \frac{|A(x,y)|}{\widetilde{\mu}(y)} \tag{10.6}$$

By exchanging the roles of L and $\widetilde{L}$, one constructs an invertible Markov kernel $\widetilde{\Lambda}$ such that $\widetilde{L}\widetilde{\Lambda} = \widetilde{\Lambda}L$ and this ends the proof of the lemma.

To extend the above lemma to all finite irreducible Markov generators, we need to recall more precisely the characteristic invariants for the similarity relation and to introduce the corresponding notation. Let R be a $N \times N$ real finite matrix. Seen as a complex matrix, it is similar to a block matrix, whose blocks are of Jordan type (λ_1, n_1), (λ_2, n_2), ..., (λ_r, n_r), where $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}$ are the eigenvalues of R (with geometric multiplicities) and $r \in \mathbb{N}$, $n_1, n_2, \dots, n_r \in \mathbb{N}$ with $n_1 + n_2 + \cdots + n_r = N$. Recall that a Jordan block of type (λ, n) is a $n \times n$ matrix whose diagonal entries are equal to λ, whose first above diagonal entries are equal to 1 and whose other entries vanish. The set $\{(\lambda_k, n_k) : k \in [\![r]\!]\}$ is a characteristic invariant for the similarity class of R and will be called the **characteristic set** of R. Note that this characteristic set of R is equal to $\{(\lambda_k, n_k) : k \in [\![r]\!]\}$, if and only if one can find a (complex) basis $(\varphi_{(k,l)})_{(k,l)\in S}$ of $\mathbb{C}^N$, where $S{:=}\{(k,l) : k \in [\![r]\!] \text{ and } l \in [\![n_k]\!]\}$, such that

$$\forall (k,l) \in S, \qquad R[\varphi_{(k,l)}] = \lambda_k \varphi_{(k,l)} + \varphi_{(k,l-1)}$$

where by convention, $\varphi_{(k,0)} = 0$ for all $k \in [\![r]\!]$. Such a basis will be said to be **adapted** to R.

Lemma 10.1 extends into:

Lemma 10.2 *Two finite and irreducible Markov generators are Markov-similar if and only if they are similar.*

Proof It is sufficient to adapt the arguments given in the reversible situation. Again we just need to show the direct implication. Let L and $\widetilde{L}$ be two finite and irreducible Markov generators which are similar. Up to a permutation, we identify the index set $\widetilde{S}$ with S in the above notation (with $R = -\widetilde{L}$ or $R = L$). Let $(\varphi_{(k,l)})_{(k,l)\in S}$ and $(\widetilde{\varphi}_{(k,l)})_{(k,l)\in S}$ be adapted bases associated to $-L$ and $-\widetilde{L}$. By irreducibility, 0 is an eigenvalue of multiplicity 1, so we can assume that $(\lambda_1, n_1) = (0, 1)$ and $\varphi_{(1,1)} = \mathbb{1}$. We begin by proving that

$$\forall (k,l) \in S \setminus \{(1,1)\}, \qquad \mu[\varphi_{(k,l)}] = 0$$

Indeed, for any $k \in [\![r]\!]$, we have $L[\varphi_{(k,1)}] = -\lambda_k \varphi_{(k,1)}$ with $\lambda_k \neq 0$. Integrating the previous relation with respect to μ, we obtain

$$\lambda_k \mu[\varphi_{(k,1)}] = 0$$

so that $\mu[\varphi_{(k,1)}] = 0$. Next we show that

$$\mu[\varphi_{(k,l)}] = 0 \tag{10.7}$$

by iteration on l, with $k \in [\![r]\!]$ fixed. If (10.7) is true for some $l \in [\![n_k - 1]\!]$, then integrating with respect to μ the relation

$$L[\varphi_{(k,l+1)}] = -\lambda_k \varphi_{(k,l+1)} + \varphi_{(k,l)}$$

we get $\lambda_k \mu[\varphi_{k,l+1}] = 0$, namely (10.7) with l replaced by $l+1$.

Let $\mathscr{F}$ be the vector space generated by the family $(\varphi_{(k,l)})_{(k,l)\in S\setminus\{(1,1)\}}$, i.e. the vector space of functions f defined on V such that $\mu[f] = 0$. Define similarly $\widetilde{\mathscr{F}}$ and an operator B from $\widetilde{\mathscr{F}}$ to $\mathscr{F}$ by

$$\forall (k,l) \in S \setminus \{(1,1)\}, \qquad B[\widetilde{\varphi}_{(k,l)}] := \varphi_{(k,l)}$$

Consider two bases of $\mathscr{F}$ and $\widetilde{\mathscr{F}}$ made up of real functions, a priori the entries of the matrix (still denoted B) associated to B in these bases are complex numbers. But we have that B is invertible and that on $\widetilde{\mathscr{F}}$, $LB = B\widetilde{L}$. Since the entries of L and $\widetilde{L}$ are real numbers, it follows that $L\Re(B) = \Re(B)\widetilde{L}$ and $L\Im(B) = \Im(B)\widetilde{L}$, where $\Re$ and $\Im$ stands for the real and imaginary parts. Furthermore, there exists a real number s such that the rank of $A := \Re(B) + s\Im(B)$ is $|V| - 1$ (use e.g. the polynomial mapping $\mathbb{C} \ni z \mapsto \det(\Re(B) + z\Im(B))$). Extend A into an operator from $\mathbb{L}^2(\widetilde{\mu})$ to $\mathbb{L}^2(\mu)$ by imposing $A[\mathbb{1}] = 0$ and note that $LA = A\widetilde{L}$ and that in the usual basis $(\mathbb{1}_x)_{x\in V}$ formed of the indicator functions of the points, the entries of A are real numbers. For $\epsilon \neq 0$, we consider again the operator Λ given by (10.5). The proof goes on as before, Λ being an invertible Markov kernel if (10.6) is satisfied.

It remains to relax the irreducibility assumption to prove Theorem 10.1. Recall that a finite Markov generator is non-transient, if and only if it admits a invariant probability measure which gives positive weights to all the points of V. This state space can then be partitioned into parts which ignore themselves and are of the type considered in Lemma 10.2. Nevertheless, Example 10.1 suggests that Theorem 10.1 cannot be a direct consequence of Lemma 10.2.

Proof (Proof of Theorem 10.1) The difference with the proof of Lemma 10.2 is that the eigenvalue 0 of a finite and non-transient Markov generator is no longer necessarily simple. Its multiplicity is the number $\mathfrak{n} \in \mathbb{N}$ of irreducible classes and the dimension of the Jordan blocks associated to each of the eigenvalue(s) 0 is 1. The arguments of the proof of Lemma 10.2 can be adapted by doing the following. If L and $\widetilde{L}$ are two finite and non-transient Markov generators, begin by choosing corresponding positive invariant probabilities measures μ and $\widetilde{\mu}$. Next choose an orthonormal (in $\mathbb{L}^2(\mu)$) basis $(\varphi_1, \varphi_2, \ldots, \varphi_{\mathfrak{n}})$ of the kernel of L, with $\varphi_1 = \mathbb{1}$ and similarly an orthonormal (in $\mathbb{L}^2(\widetilde{\mu})$) basis $(\widetilde{\varphi}_1, \widetilde{\varphi}_2, \ldots, \widetilde{\varphi}_{\mathfrak{n}})$ of the kernel of $\widetilde{L}$ with $\widetilde{\varphi}_1 = \mathbb{1}$. Complete these families of vectors into adapted bases for L and $\widetilde{L}$, with the convention that the index associated to the eigenvectors φ_1 and $\widetilde{\varphi}_1$ is $(1, 1)$. The argument goes on as before, since both φ_1 and $\widetilde{\varphi}_1$ are orthogonal to all the other eigenvectors in their respective bases.

10.3 On the Finite Transient Setting

As alluded to in the introduction, in general similarity does not imply Markov-similarity. Remaining in the finite state space framework, we investigate here in more detail the transient situation where this phenomenon appears, by obtaining a necessary condition, of spatial-spectral nature, for Markov-similarity. On a example, we will check that this condition is not sufficient. Thus the problem of finding characterizing invariants for Markov-similarity is still open and seems quite challenging.

We begin by recalling a traditional classification of the points of V according to L, refining the notion of transience defined in the introduction. Consider $x, y \in V$, if x leads to y and y to x, we say that x and y **communicate**. This defines an equivalence (**irreducibility**) relation.

Let $C_1, C_2, \ldots, C_{\mathfrak{n}}$, with $\mathfrak{n} \in \mathbb{N}$, be the associated equivalence classes. For $k, l \in [\![\mathfrak{n}]\!]$, we write $C_k \preceq_0 C_l$ if there exist $x \in C_l$ and $y \in C_k$ such that x leads to y. The relation $\preceq_0$ defines a partial order on $\mathscr{C}_0{:=}\{C_1, \ldots, C_{\mathfrak{n}}\}$. Consider

$$\begin{aligned} A_1 &:= \{a \in [\![\mathfrak{n}]\!] \;:\; C_a \text{ is minimal for } \preceq_0\} \\ B_1 &:= [\![\mathfrak{n}]\!] \setminus A_1 \\ V_1 &:= \bigsqcup_{a \in A_1} C_a \end{aligned}$$

From a probabilistic point of view, V_1 is the maximal subset of V supporting an invariant probability for L and $V \setminus V_1$ is the set of transient points. If $V_1 = V$, the construction stops here and L is non-transient. Otherwise, consider $\mathscr{C}_1{:=}\{C_a \;:\; a \in B_1\}$ and denote by $\preceq_1$ the restriction of the partial order $\preceq_0$ to $\mathscr{C}_1$. Define

$$\begin{aligned} A_2 &:= \{a \in B_1 \;:\; C_a \text{ is minimal for } \preceq_1\} \\ B_2 &:= B_1 \setminus A_2 \end{aligned}$$

For each $a \in A_2$, consider L_{C_a}, the $C_a \times C_a$ matrix extracted from L (also named the Dirichlet restriction of L to C_a), it is a **subMarkovian generator**, in the sense that the off-diagonal entries are non-negative and the row sums are non-positive. The Perron-Frobenius' theorem can be applied to show that $-L_{C_a}$ admits a smallest eigenvalue (in modulus), $\lambda_1(C_a) \geq 0$, called the **first Dirichlet eigenvalue** of L_{C_a}. Order the elements of the set $\{\lambda_1(C_a) \;:\; a \in A_2\}$ into $\lambda_{2,1} < \lambda_{2,2} < \cdots < \lambda_{2,\kappa_2}$, with $\kappa_2 \in [\![|A_2|]\!]$. For $l \in [\![2, 1+\kappa_2]\!]:=\{2, 3, \ldots, 1+\kappa_2\}$, we denote

$$V_l := \bigcup_{a \in A_2 \,:\, \lambda_1(C_a)=\lambda_{2,l-1}} C_a \tag{10.8}$$

so that $V_2 \sqcup V_3 \sqcup \cdots \sqcup V_{\kappa_2}$ forms a partition of $\bigsqcup_{a\in A_2} C_a$. It was not necessary to explicit such a partition in our first step (defining V_1), because for any $a \in A_1$, one has that L_{C_a} is still a Markov generator, so that $\lambda_1(C_a) = 0$ and we would have end up with the unique set V_1 (i.e. $\kappa_1 = 1$). The procedure goes on by iteration, in the second step (if B_2 is not empty), we construct some disjoint non-empty subsets $V_{\kappa_2+1}, V_{\kappa_2+2}, \ldots, V_{\kappa_2+\kappa_3}$, for some $\kappa_3 \in [\![|A_3|]\!]$, where A_3 is the set of minimal elements of B_2 for the restriction of $\preceq_1$, and so on. At the end of the construction, we have disjoint non-empty subsets $V_1, V_2, \ldots, V_{\mathfrak{m}}$, with $\mathfrak{m} \in [\![\mathfrak{n}]\!]$, such that $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_{\mathfrak{m}}$, as well as a finite sequence of positive integers $\kappa := (\kappa_1, \kappa_2, \ldots, \kappa_\eta)$, where η is the number of iterations of the previous procedure, in particular $\kappa_1 + \kappa_2 + \cdots + \kappa_\eta = \mathfrak{m}$. For $m \in [\![\mathfrak{m}]\!]$, let L_m be the $V_m \times V_m$ submatrix of L, it is a subMarkovian generator.

Consider another generator $\widetilde{L}$ and construct as above the subMarkovian generators $\widetilde{L}_1, \widetilde{L}_2, \ldots, \widetilde{L}_{\widetilde{\mathfrak{m}}}$, as well as the finite sequence $\widetilde{\kappa} := (\widetilde{\kappa}_k)_{k\in[\![\widetilde{\eta}]\!]}$. We say that L and $\widetilde{L}$ satisfy **Condition (C)** if $\widetilde{\kappa} = \kappa$ (in particular $\widetilde{\eta} = \eta$ and $\widetilde{\mathfrak{m}} = \mathfrak{m}$) and if for any $l \in [\![\mathfrak{m}]\!]$, L_l and $\widetilde{L}_l$ are similar. Example 10.4 at the end of this section shows that Condition (C) does not imply similarity in general. So let us call **Hypothesis (H)** the conjunction of (C) with similarity.

Proposition 10.3 *If two Markov generators L and $\widetilde{L}$ are Markov-similar then they satisfy (H).*

The following very simple example (which nevertheless played an important role in the study of certain Markov intertwinings in [6] and [13]) illustrates the above construction and the difference between Hypothesis (H) and the similarity relation.

Example 10.2 Consider on $V := [\![|V|]\!]$ a finite Markov generator $L := (L(x, y))_{x,y\in V}$ which is lower diagonal, whose diagonal entries are all different and such that the first lower diagonal is positive (namely, for all $k \in [\![2, |V|]\!]$, $L(k, k-1) > 0$). Then in the above decomposition, we have $\mathfrak{n} = \mathfrak{m} = |V|$ and for all $k \in [\![|V|]\!]$, $V_k = \{k\}$ and L_k is reduced to the real number $L(k, k)$, which is also its unique eigenvalue (note furthermore that necessarily, $L_1 = 0$). We have $\eta = |V|$ and $\kappa_k = 1$ for all $k \in [\![|V|]\!]$. Consider another Markov generator $\widetilde{L}$. The generators L and $\widetilde{L}$ satisfy Condition (C) if and only if $\widetilde{L}$ is of the same type (up to a permutation of the state space) and if for any $k \in [\![|V|]\!]$, $\widetilde{L}(k, k) = L(k, k)$. As a consequence L and $\widetilde{L}$ are similar, since they are diagonalizable, the $L(k, k)$, for $k \in [\![|V|]\!]$, being their distinct eigenvalues. Nevertheless, the mere similarity of L and $\widetilde{L}$ is a much weaker requirement, it does not imply that $\widetilde{L}$ is of the same type, and even if it is, it only asks for the equality of the spectra, i.e. of the sets $\{L(k, k) : k \in [\![|V|]\!]\} = \{\widetilde{L}(k, k) : k \in [\![|V|]\!]\}$. Another example in the same spirit is obtained by considering a finite

Markov generator L satisfying

$$\forall x \in [\![|V|]\!], \qquad L(x, x-1) = -L(x, x) > 0$$

(with $L(1,0) = 0 = L(1,1)$), so that all the entries outside the main and first lower diagonals vanish. Any eigenvalue λ of $-L$ is geometrically simple, because a corresponding eigenvector φ is completely determined by λ and by the value $\varphi(|V|)$ (by iteration on $k \in [\![|V|]\!]$, one computes $\varphi(|V|-k+1)$ via the relation $L(|V|-k+1, |V|-k)(\varphi(|V|-k) - \varphi(|V|-k+1)) = -\lambda\varphi(|V|-k+1)$). The dimension of the corresponding Jordan block is the cardinal of the set $\{x \in [\![|V|]\!] : L(x,x) = \lambda\}$. As above, another Markov generator $\widetilde{L}$ and L satisfy Condition (C) if and only if $\widetilde{L}$ is of the same type (up to a permutation of the state space) and if for any $k \in [\![|V|]\!]$, $\widetilde{L}(k,k) = L(k,k)$. Again (C) implies (H), due to the previous observation on the dimension of the Jordan blocks. The same remark about the mere similarity is equally valid, except that the last equality $\{L(k,k) : k \in [\![|V|]\!]\} = \{\widetilde{L}(k,k) : k \in [\![|V|]\!]\}$ must be understood in the sense of multi-sets.

In the definition of Condition (C), it is important not to forget the equality of the finite sequences $\kappa = \widetilde{\kappa}$, as shown by

Example 10.3 On V:=$\{1,2,3\}$, consider the two generators

$$L := \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & -2 \end{pmatrix} \text{ and } \widetilde{L} := \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 0 & -2 \end{pmatrix}$$

With the notation of the above decomposition, we have, for all $k \in [\![3]\!]$, $V_k = \{k\} = \widetilde{V}_k$ and $L_k = (k-1) = \widetilde{L}_k$. So the fact that L and $\widetilde{L}$ are not satisfying Condition (C) comes from $\kappa = (1,1) \neq (2) = \widetilde{\kappa}$. This also provides a very simple example of Markov generators which are similar but not Markov-similar.

Here is a simple consequence of Proposition 10.3:

Corollary 10.1 *Let the two Markov generators L and $\widetilde{L}$ be Markov-similar. If L is non-transient, then the same is true for $\widetilde{L}$.*

Proof Indeed, the non transience of a Markov generator L is equivalent to the fact that $\kappa = (\kappa_1) = (1)$.

As an extension of the observation made after the statement of Theorem 10.1, note that two finite Markov-similar Markov generators have the same number of irreducible classes. Indeed, for a general Markov generator L, this number is the sum of the multiplicities of the first Dirichlet eigenvalues of the subMarkovian generators $L_1, L_2, \ldots, L_{\mathrm{m}}$, with the notation of the above decomposition (which will be enforced for the remaining part of this section).

The proof of Proposition 10.3 asks for several steps. We start with

Lemma 10.3 *Let two Markov generators L, $\widetilde{L}$ and a link Λ be such that $L\Lambda = \Lambda\widetilde{L}$. Then we have $|V_1| \geq |\widetilde{V}_1|$ and $\Lambda_{V_1\times\widetilde{V}_1}$ (the submatrix of Λ indexed by $V_1 \times \widetilde{V}_1$) is a Markov kernel.*

Proof Consider μ an invariant probability for L whose support is V_1 (constructed as a mixture with positive weights of the invariant probabilities associated to the irreducibility classes forming V_1). The intertwining relation implies that $\mu\Lambda\widetilde{L} = 0$, namely $\widetilde{\mu}:=\mu\Lambda$ is an invariant probability for $\widetilde{L}$. The support of $\widetilde{\mu}$ is included into $\widetilde{V}_1$, since it is the largest subset of $\widetilde{V}$ supporting an invariant probability for $\widetilde{L}$. The equality

$$\forall \widetilde{x} \notin \widetilde{V}_1, \qquad \sum_{x\in V} \mu(x)\Lambda(x,\widetilde{x}) \; = \; \widetilde{\mu}(\widetilde{x}) \; = \; 0$$

implies that $\Lambda_{V_1\times(\widetilde{V}\setminus\widetilde{V}_1)} = 0$, namely $\Lambda_{V_1\times\widetilde{V}_1}$ is a Markov kernel. Another consequence of the fact that $\Lambda_{V_1\times(\widetilde{V}\setminus\widetilde{V}_1)}$ vanishes is that $|V_1| \geq |\widetilde{V}_1|$, otherwise Λ could not be invertible.

In particular, we get

Corollary 10.2 *Assume that the two Markov generators L and $\widetilde{L}$ are Markov-similar. Then the two Markov generators L_1 and $\widetilde{L}_1$ are Markov-similar.*

Proof Applying the previous lemma to the two intertwining relations (10.1), we get that $|V_1| = |\widetilde{V}_1|$, namely $\Lambda_{V_1,\widetilde{V}_1}$ can be seen as a square matrix. Denote $W_1:=V \setminus V_1$ and $\widetilde{W}_1:=V \setminus \widetilde{V}_1$. Since $L_{V_1,W_1} = 0$, $\widetilde{L}_{\widetilde{V}_1,\widetilde{W}_1} = 0$ and $\Lambda_{V_1,\widetilde{W}_1} = 0$, we deduce from the intertwining $L\Lambda = \Lambda\widetilde{L}$ that $L_1\Lambda_{V_1,\widetilde{V}_1} = L_{V_1,V_1}\Lambda_{V_1,\widetilde{V}_1} = \Lambda_{V_1,\widetilde{V}_1}\widetilde{L}_{\widetilde{V}_1,\widetilde{V}_1} = \Lambda_{V_1,\widetilde{V}_1}\widetilde{L}_1$. Furthermore $\Lambda_{V_1,\widetilde{V}_1}$ must be invertible, if we want Λ to be invertible. Applying the same considerations to the intertwining $\widetilde{L}\widetilde{\Lambda} = \widetilde{\Lambda}L$, it follows that L_1 and $\widetilde{L}_1$ are Markov-similar.

To extend by iteration the above result to all the subMarkov generators L_l and $\widetilde{L}_l$, for $l \in [\![\mathfrak{m}]\!]$, we must adapt the arguments to the subMarkovian setting. First note that the decomposition of the state space into the partition $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_{\mathfrak{m}}$ can be applied verbatim to a subMarkovian generator L (with the difference that the first step can already produce several subset V_1, V_2, ..., V_{κ_1}, with $\kappa_1 \in \mathbb{N}$). The probabilistic interpretation of V_1 has to be slightly modified, with respect to the strict Markovian case:

Lemma 10.4 *Consider L a subMarkovian generator and let $\mathscr{L}$ be the set of real numbers λ such that there exists a probability measure μ with $\mu L = -\lambda\mu$ (then λ is necessarily non-negative). We have $\mathscr{L} = \{\lambda_1(C_a) : a \in A_1\}$. Denote $\lambda_1 < \lambda_2 < \cdots < \lambda_{\kappa_1}$ the elements of $\mathscr{L}$. For any $k \in [\![\kappa_1]\!]$, V_k is the largest subset of V supporting a probability measure satisfying $\mu L = -\lambda_k\mu$.*

Proof Since the classes $C_1, C_2, \ldots, C_n$ are irreducible, we can apply to each of them the Perron-Frobenius' theorem, to get for $a \in [\![m]\!]$, a probability measure μ_a (called the quasi-stationary measure associated to L_{C_a}) whose support is C_a and which is such that $\mu_a L_{C_a} = -\lambda_1(C_a)\mu_a$, where $\lambda_1(C_a) \geq 0$ is the first Dirichlet eigenvalue of L_{C_a}. The particularity of the set of indices A_1 is that for each $a \in A_1$ and for any probability measure ν whose support is included in C_a, we have $\nu L = \nu L_{C_a}$ (with the identification of a measure whose support is included into C_a with a measure on C_a). It follows that for $a \in A_1$, we have $\mu_a L = -\lambda_1(C_a)\mu_a$, so that

$$\{\lambda_1(C_a) \; : \; a \in A_1\} \subset \mathscr{L}$$

Conversely, consider $\lambda \in \mathscr{L}$ and a probability measure μ satisfying $\mu L = -\lambda\mu$. Let us first check that $\text{supp}(\mu) \subset \sqcup_{a \in A_1} C_a$. We begin by remarking that if $x, y \in V$ are such that $\mu(x) > 0$ and $L(x, y) > 0$, then $\mu(y) > 0$. Indeed, otherwise in the equality

$$\sum_{z \in V \setminus \{y\}} \mu(z)L(z, y) = -\mu(y)L(y, y) - \lambda\mu(y)$$

the l.h.s. would be positive and the r.h.s. would vanish. It follows by iteration that if $\mu(x) > 0$ and if x leads to y, then $\mu(y) > 0$. In particular, the support of μ is an union of irreducibility classes and at least one of them is included into $\sqcup_{a \in A_1} C_a$. If all the irreducibility classes forming $\text{supp}(\mu)$ are included into $\sqcup_{a \in A_1} C_a$, we get that $\text{supp}(\mu) \subset \sqcup_{a \in A_1} C_a$. Otherwise, we can find $a \in A_1$ and $b \notin A_1$, with $C_a \sqcup C_b \subset \text{supp}(\mu)$ and there exist $x_0 \in C_b$ and $y_0 \in C_a$ with $L(x_0, y_0) > 0$. The restriction to C_a of $\mu L = -\lambda\mu$ writes down $\mu_{C_a} L_{C_a} = -\lambda\mu_{C_a}$, where μ_{C_a} is the restriction of μ to C_a. Since μ_{C_a} is positive, it follows from the uniqueness statement in Perron-Frobenius' theorem, that μ_{C_a} is proportional to the quasi-stationary measure μ_a associated to L_{C_a} and $\lambda = \lambda_1(C_a)$. Due to the property satisfied by x_0, y_0, we have that $\mu(x_0) > 0$ and $((\mu - \mu_{C_a})L)(y_0) > 0$. We deduce that

$$\begin{aligned}
-\lambda\mu(y_0) &= (\mu L)(y_0) \\
&= ((\mu - \mu_{C_a} + \mu_{C_a})L)(y_0) \\
&= ((\mu - \mu_{C_a})L)(y_0) + (\mu_{C_a} L)(y_0) \\
&> (\mu_{C_a} L)(y_0) \\
&= -\lambda\mu(y_0)
\end{aligned}$$

which is a contradiction.

The above arguments also show that μ is a mixture of the quasi-stationary measures associated to the irreducible classes included into $\sqcup_{a \in A_1} C_a$. Furthermore, the classes C_a, with $a \in A_1$, which are such that $\mu(C_a) > 0$ must satisfy $\lambda_1(C_a) = \lambda$. It follows that if $\mu L = -\lambda_1\mu$, then the support of μ is included into V_k, where $k \in [\![\kappa_1]\!]$ is such that $\lambda = \lambda_k$, and is equal to V_k if μ is chosen to be a non-degenerate convex combination of the quasi-stationary measures associated to the C_a included into V_k.

This result allows us to adapt the proof of Lemma 10.3 and Corollary 10.2 to get the following generalization, where a sublink stands for an invertible subMarkov kernel (i.e. a matrix with non-negative entries whose row sums are bounded above by 1). We also say that two subMarkovian generators L and $\widetilde{L}$ are subMarkov-similar if there exist two sublinks Λ and $\widetilde{\Lambda}$ such that (10.1) is valid.

Lemma 10.5 *Let L, $\widetilde{L}$ be two subMarkov generators and Λ a sublink such that $L\Lambda = \Lambda\widetilde{L}$. Then we have $\mathscr{L} \subset \widetilde{\mathscr{L}}$. Assume furthermore that L and $\widetilde{L}$ are subMarkov-similar. Then $\mathscr{L} = \widetilde{\mathscr{L}}$ and the subMarkov generators L_m and $\widetilde{L}_m$ are subMarkov-similar, for $m \in [\![\kappa_1]\!]$, as well as the subMarkov generators $L_W := L_{W,W}$ and $\widetilde{L}_{\widetilde{W}} := \widetilde{L}_{\widetilde{W},\widetilde{W}}$, where $W := V \setminus (V_1 \sqcup \cdots \sqcup V_{\kappa_1})$ and $\widetilde{W} := \widetilde{V} \setminus (\widetilde{V}_1 \sqcup \cdots \sqcup \widetilde{V}_{\kappa_1})$.*

Proof With the notation of Lemma 10.4, consider $\lambda \in \mathscr{L}$ and a probability measure μ on V such that $\mu L = -\lambda\mu$. The measure $\mu\Lambda$ is non-negative and cannot be 0, because Λ is invertible. We can thus define the probability measure $\widetilde{\mu} := \mu\Lambda/\mu\Lambda(\widetilde{V})$. By multiplying on the left the relation $L\Lambda = \Lambda\widetilde{L}$ by μ, we get that $\widetilde{\mu}\widetilde{L} = -\lambda\widetilde{\mu}$, so that $\lambda \in \widetilde{\mathscr{L}}$.

So if L and $\widetilde{L}$ are subMarkov-similar, we get $\mathscr{L} = \widetilde{\mathscr{L}}$. The arguments of the proofs of Corollary 10.2 and Lemma 10.3 can now be repeated, with the notion of invariant measure replaced by that of eigen-probability measure associated to $\lambda \in \mathscr{L}$ (with respect to $-L$ and $-\widetilde{L}$). Indeed, the subMarkov-similarity of the subMarkov generators L_W and $\widetilde{L}_{\widetilde{W}}$ is also valid in Corollary 10.2, using the sublinks $\Lambda_{W,\widetilde{W}}$ and $\widetilde{\Lambda}_{\widetilde{W},W}$. It was not asserted there, just because the subMarkov-similarity between subMarkov generators had not yet been defined.

Remark 10.2 From the above proof, it also follows that for all $m \in [\![\mathfrak{m}]\!]$, we have $\Lambda_{V_m,\bar{V}_m} = 0$, where

$$\bar{V}_m := \bigsqcup_{m \in J_m} V_m$$

$$J_m := [\![\kappa_1 + \cdots + \kappa_{j_m} + 1, \mathfrak{m}]\!] \setminus \{m\}$$

where $j_m \in [\![0, \eta - 1]\!]$ is such that $m \in [\![\kappa_1 + \cdots + \kappa_{j_m} + 1, \kappa_1 + \cdots + \kappa_{j_m+1}]\!]$.

Proposition 10.3 is now a simple consequence of the previous lemma. Indeed, extending naturally Conditions (C) and (H) to subMarkovian generators, we get:

Proposition 10.4 *Consider two subMarkov generators L and $\widetilde{L}$. If they are subMarkov-similar, then they satisfy Hypothesis (H).*

Proof Applying iteratively Lemma 10.5, we end up with the conclusion that for $l \in [\![\mathfrak{m}]\!]$, L_l is subMarkov-similar to $\widetilde{L}_l$ and $\mathscr{L}_l = \widetilde{\mathscr{L}}_l$. SubMarkov-similarity implying similarity, we conclude to the announced validity of Hypotheses (C) and (H).

Let us now mention an extension of Theorem 10.1 to the present subMarkov framework. In some sense, the following result is the "Dirichlet condition" analogue of Theorem 10.1 (whose "Neumann condition" corresponds to the fact that Markov processes are conservative). We say that a subMarkov generator L is isotransient, if $L = L_1$ (this appellation amounts to non-transcience for Markov generators). Note in particular that for any subMarkov generator L, L_m is isotransient for all $m \in [\![\mathfrak{m}]\!]$.

Proposition 10.5 *Two isotransient subMarkov generators L and $\widetilde{L}$ are subMarkov-similar if and only if they are similar.*

Proof As usual, the direct implication is obvious. We begin by showing the subMarkovian extension of Lemma 10.2, namely that two finite and irreducible subMarkov generators are similar if and only if they are subMarkov-similar. Consider two similar and irreducible subMarkov generators L and $\widetilde{L}$. By Perron-Frobenius' theorem, there a exists positive eigenvector ψ associated to the first Dirichlet eigenvalue λ_1 of L. The operator $L^\dagger \cdot :=\psi^{-1}(L - \lambda_1 I)[\psi \cdot]$ is a Markov generator (sometimes called the Doob transform of L, see e.g. [7]). Its spectrum is the spectrum of L shifted by λ_1. If $\widetilde{L}^\dagger$ is constructed similarly for $\widetilde{L}$, it appears that $L^\dagger$ and $\widetilde{L}^\dagger$ are similar irreducible Markov generators, so from Lemma 10.2, there exist two links $\Lambda^\dagger$ and $\widetilde{\Lambda}^\dagger$ such that

$$L^\dagger \Lambda^\dagger \ = \ \Lambda^\dagger \widetilde{L}^\dagger \text{ and } \widetilde{L}^\dagger \widetilde{\Lambda}^\dagger \ = \ \widetilde{\Lambda}^\dagger L^\dagger$$

It remains to define the non-negative kernels

$$\Lambda[\cdot]:=\psi \Lambda^\dagger \left[\frac{1}{\widetilde{\psi}} \cdot\right]$$

$$\widetilde{\Lambda}[\cdot]:=\widetilde{\psi} \widetilde{\Lambda}^\dagger \left[\frac{1}{\psi} \cdot\right]$$

to ensure that (10.1) is satisfied. To get sublinks, divide Λ and $\widetilde{\Lambda}$ by a sufficiently large constant.

When L and $\widetilde{L}$ are isotransient subMarkov generators, let $C_1, C_2, \ldots, C_{\mathfrak{n}}$, with $m \in \mathbb{N}$, be the irreducibility class(es) of L and let λ_1 be the common first Dirichlet eigenvalue of the corresponding restrictions L_{C_l}. For $l \in [\![\mathfrak{n}]\!]$, let ψ_l be a positive eigenvector on C_l associated to λ_1. Let ψ the function on V coinciding with these eigenvectors on each of the C_l, for $l \in [\![\mathfrak{n}]\!]$. Do the same with $\widetilde{L}$, remarking that $\widetilde{\lambda}_1 = \lambda_1$ by similarity. The previous arguments are then still valid.

We end this section with an example on four points illustrating that Hypothesis (H) is not sufficient for Markov-similarity.

Example 10.4 On $V:=[\![4]\!]$, consider for any $p \in [0, 1)$, the Markov generator

$$L^{(p)} := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 2p & 2(1-p) & -2 \end{pmatrix}$$

and we denote $L:=L^{(0)}$.

For all $p \in [0, 1)$, $-L^{(p)}$ has three eigenvalues: $\lambda_1:=0$, $\lambda_2:=1$ and $\lambda_3:=2$. The similarity class of $L^{(p)}$ depends on the geometric multiplicity of λ_3, which is either 2 or 1 (with then a Jordan block of dimension 2 associated to λ_3). Computing the eigenspace of λ_3, it appears that there is a Jordan block of dimension 2 associated to λ_3 if and only if $p \neq 1/2$.

Moreover the spatial decomposition of $L^{(p)}$ is immediate to obtain for all $p \in [0, 1)$: for all $k \in [\![4]\!]$, we have $V_k = \{k\}$, $L_k^{(p)} = L(k, k)$ and $\kappa^{(p)} = (1, 1, 1, 1)$. It follows that if $\widetilde{L}:=L^{(p)}$, with $p \in [0, 1)$, then L and $\widetilde{L}$ satisfy Condition (C). In particular (C) does not imply similarity for $p = 1/2$ and Hypothesis (H) is true if and only if $p \neq 1/2$. From now on, we assume that $\widetilde{L}:=L^{(p)}$, with a fixed $p \in [0, 1) \setminus \{1/2\}$ and we are wondering if L and $\widetilde{L}$ are Markov-similar. We show below that this is the case if and only if $p \in [0, 1/2)$.

Denote $A:=(0, 0, 2)$, $\widetilde{A}:=(0, 2p, 2(1-p))$ and

$$K := \begin{pmatrix} 0 & 0 & 0 \\ 2 & -2 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

so we can write

$$L = \begin{pmatrix} K & 0 \\ A & -2 \end{pmatrix} \text{ and } \widetilde{L} = \begin{pmatrix} K & 0 \\ \widetilde{A} & -2 \end{pmatrix}$$

Let Λ a link such that $L\Lambda = \Lambda\widetilde{L}$, then necessarily it can be written under the form

$$\Lambda = \begin{pmatrix} Q & 0 \\ B & d \end{pmatrix}$$

where Q is a link, $B = (a, b, c)$ and $d = 1 - a - b - c$, with $a, b, c, d \in [0, 1]$. Indeed, $\widetilde{L}\mathbb{1}_{\{4\}} = -2\mathbb{1}_{\{4\}}$, so that $\Lambda\mathbb{1}_{\{4\}}$ is an eigenfunction associated to the eigenvalue -2 of L. Since $p \neq 1/2$, such an eigenfunction is proportional to $\mathbb{1}_{\{4\}}$, which amounts to the above form of Λ. This form can also be deduced from Remark 10.2, which enables to see a priori that Q must be lower diagonal.

The intertwining relation $L\Lambda = \Lambda\widetilde{L}$ is then equivalent to

$$KQ = QK \text{ and } AQ - 2B = BK + d\widetilde{A} \tag{10.9}$$

Define

$$\varphi_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \varphi_2 := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \qquad \varphi_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

which are eigenvectors of K associated respectively to the eigenvalues $0, -2, -1$. The intertwining relation is equivalent to the existence of $x, y, z \in \mathbb{R}$ such that

$$Q\varphi_1 = x\varphi_1, \qquad Q\varphi_2 = y\varphi_2, \qquad Q\varphi_3 = z\varphi_3$$

and this means that

$$Q = \begin{pmatrix} x & 0 & 0 \\ x - y & y & 0 \\ x + y - 2z & z - y & z \end{pmatrix}$$

The fact that Λ is required to be an invertible Markov kernel is then equivalent to the constraints

$$x = 1, \qquad 0 < y \leq 1, \qquad y \leq z \leq (1 + y)/2$$

It follows that Condition (10.9) is equivalent to

$$\begin{cases} 1 + y - 2z = 2a \\ 2(z - y) = c + 2dp \\ 2z - c = 2(1 - p)d \end{cases}$$

itself equivalent to

$$\begin{cases} a = (1 + y - 2z)/2 \\ b = (1 + y - 2z)/2 \\ c = \frac{2}{2p-1}((1 - p)y + (2p - 1)z) \end{cases}$$

Summing these equations, we get

$$a + b + c = 1 + \frac{1}{2p - 1}y$$

so the requirement $a + b + c \leq 1$ is equivalent to $p < 1/2$ (recall that $y > 0$). Conversely, if $p \in [0, 1/2)$, taking e.g. $z = 1/4$ and $y > 0$ small enough leads to a solution for the link Λ.

Similar considerations show that there exists a link $\widetilde{\Lambda}$ such that $\widetilde{\Lambda}L = \widetilde{L}\widetilde{\Lambda}$ if and only if $p \in [0, 1/2)$.

Remark 10.3 In addition to Remark 10.2, in general the link Λ is not such that Λ_{V_m,V_m} is itself a link, for $m \in [\![\mathfrak{m}]\!]$. Indeed, in the above example under this restriction, we would have end up with $\Lambda = I$, the identity matrix, which does not enable to intertwine L and $L^{(p)}$ for $p \in (0, 1/2)$.

10.4 Comparisons of Mixing Speeds

The goal of this section is to discuss the consequences of Markov-similarity on speeds of convergence to equilibrium, and especially to prove Proposition 10.1. We introduce some sets of Markov kernels and probability measures associated to finite Markov-similar generators which play a crucial role.

More precisely, if L and $\widetilde{L}$ are finite Markov-similar generators as in the introduction, denote $\mathscr{K}(L, \widetilde{L})$ the set of Markov kernels Λ from V to $\widetilde{V}$ such that $L\Lambda = \Lambda\widetilde{L}$ and $\mathscr{P}(L, \widetilde{L}) := \mathscr{P}(V)\mathscr{K}(L, \widetilde{L})$, namely the set of probability measures $\widetilde{m}$ on $\widetilde{V}$ such that there exists $m \in \mathscr{P}(V)$ (recall that $\mathscr{P}(V)$ stands for the set of all probability measures on V) and $\Lambda \in \mathscr{K}(L, \widetilde{L})$ such that $\widetilde{m} = m\Lambda$. The sets $\mathscr{K}(\widetilde{L}, L)$ and $\mathscr{P}(\widetilde{L}, L)$ are defined symmetrically, by inverting the roles of L and $\widetilde{L}$. Here is the advantage of Markov-similarity:

Lemma 10.6 *Assume furthermore that L is irreducible. Then the set $\mathscr{P}(L, \widetilde{L})$ is a neighborhood of $\widetilde{\mu}$ the invariant measure of $\widetilde{L}$, which is necessarily also irreducible.*

Remark 10.4 We don't know whether $\mathscr{P}(L, \widetilde{L})$ is always convex or not.

Proof (Proof of Lemma 10.6) Since L and $\widetilde{L}$ are Markov-similar, we deduce from Corollary 10.1 that the non-transience of L implies that of $\widetilde{L}$. Furthermore the number of irreducible class(es) of $\widetilde{L}$ is that of L, so $\widetilde{L}$ is irreducible. It follows that $\widetilde{L}$ has a unique invariant measure and it is given by $\widetilde{\mu} = \mu\Lambda$, for all $\Lambda \in \mathscr{K}(L, \widetilde{L})$. Note that μ belongs to the interior of $\mathscr{P}(V)$ (as a subset of the space of all signed measures on V), because it gives a positive weight to all the points. By Markov-similarity, we can find a link Λ in $\mathscr{K}(L, \widetilde{L})$. Its invertibility implies that it transforms a neighborhood of μ into a neighborhood of $\widetilde{\mu}$. So $\mathscr{P}(L, \widetilde{L})$ is a neighborhood of $\widetilde{\mu}$.

Remark that $\mathscr{P}(L, \widetilde{L})$ is left invariant by the Markov semi-group $(\exp(tL))_{t\geq 0}$ generated by L. Indeed, for any initial conditions $m_0 \in \mathscr{P}(V)$, $\widetilde{m}_0 \in \mathscr{P}(\widetilde{V})$ and any time $t \geq 0$, denote

$$m_t := m_0 \exp(tL) \text{ and } \widetilde{m}_t := \widetilde{m}_0 \exp(t\widetilde{L})$$

the marginal distribution at time t obtained respectively through the evolutions generated by L and $\widetilde{L}$. Assume $\widetilde{m}_0 \in \mathscr{P}(L, \widetilde{L})$, so there exist $m_0 \in \mathscr{P}(V)$ and

$\Lambda \in \mathscr{K}(L, \widetilde{L})$ such that $\widetilde{m}_0 = m_0\Lambda$. For any $t \geq 0$, we get

$$\begin{aligned}\widetilde{m}_t &= \widetilde{m}_0 \exp(t\widetilde{L})\\ &= m_0\Lambda \exp(t\widetilde{L})\\ &= m_0 \exp(tL)\Lambda\\ &= m_t\Lambda\end{aligned}$$

This implies that $\widetilde{m}_t \in \mathscr{P}(L, \widetilde{L})$, as announced. Since we also have $\widetilde{\mu} = \mu\Lambda$, it follows that

$$\begin{aligned}E_\psi(\widetilde{m}_t|\widetilde{\mu}) &= E_\psi(m_t\Lambda|\mu\Lambda)\\ &\leq E_\psi(m_t|\mu)\end{aligned} \tag{10.10}$$

because ψ-entropies decrease under the action of Markov kernels, for any $\psi \in \Psi$. This well-known property, which holds on general measurable spaces (see e.g. Proposition 1.1 of [4]), is an important reason behind our interest in considering intertwining kernels which are Markovian. Thus, seen through the ψ-entropy, the convergence of $\widetilde{m}_t$ toward $\widetilde{\mu}$ for large $t \geq 0$ is dominated by that of m_t toward μ. In particular we deduce that

$$\widetilde{m}_0 \in \mathscr{P}(L, \widetilde{L}) \Rightarrow \forall t \geq 0,\ \forall \psi \in \Psi, \quad E_\psi(\widetilde{m}_t|\widetilde{\mu}) \ \leq \ E(\psi, t)$$

Proposition 10.1 is now a simple consequence of

Lemma 10.7 *Under the assumption of Lemma 10.6, there exists $T \geq 0$, such that for any $\widetilde{m}_0 \in \mathscr{P}(\widetilde{V})$, $\widetilde{m}_T \in \mathscr{P}(L, \widetilde{L})$.*

Of course in practice, the problem will be to describe $\mathscr{P}(L, \widetilde{L})$ and to estimate T.

Proof As in the proof of Lemma 10.6, consider a link $\Lambda \in \mathscr{K}(L, \widetilde{L})$. The set $\mathscr{P}(V)\Lambda \subset \mathscr{P}(L, \widetilde{L})$ is convex and left invariant by the semigroup $(\exp(t\widetilde{L}))_{t\geq 0}$. Thus it is sufficient to see that for any $\widetilde{x} \in \widetilde{V}$, there exists $T_{\widetilde{x}} \geq 0$ such that $\delta_{\widetilde{x}} \exp(T_{\widetilde{x}}\widetilde{L}) \in \mathscr{P}(V)\Lambda$. Indeed, by stability of $\mathscr{P}(L, \widetilde{L})$ by the semi-group generated by $\widetilde{L}$, we get

$$\forall \widetilde{x} \in \widetilde{V}, \qquad \delta_{\widetilde{x}} \exp(T\widetilde{L}) \in \mathscr{P}(L, \widetilde{L})$$

with $T := \max\{T_{\widetilde{x}} : \widetilde{x} \in \widetilde{V}\}$. By convexity of the mapping $\mathscr{P}(\widetilde{V}) \ni \widetilde{m}_0 \mapsto \widetilde{m}_T$ and of the set $\mathscr{P}(V)\Lambda$, it appears then that

$$\forall \widetilde{m}_0 \in \mathscr{P}(\widetilde{V}), \qquad \widetilde{m}_0 \exp(T\widetilde{L}) \in \mathscr{P}(V)\Lambda \ \subset \ \mathscr{P}(L, \widetilde{L})$$

But for any fixed $\widetilde{x} \in \widetilde{V}$, we have that $\delta_{\widetilde{x}} \exp(t\widetilde{L})$ converges toward $\widetilde{\mu}$ for large t, so for large enough $T_{\widetilde{x}} \geq 0$, $\delta_{\widetilde{x}} \exp(T_{\widetilde{x}}\widetilde{L})$ belongs to the neighborhood $\mathscr{P}(L, \widetilde{L})$ of $\widetilde{\mu}$.

In this paper, we adopted an equivalence relation point of view on Markov intertwinings, through the Markov-similarity. But the order relation aspect of the Markov intertwinings is also very interesting and maybe more relevant for applications. Such considerations can be found in [14], but let us slightly modify the definitions given there by saying that the Markov generator $\widetilde{L}$ on the finite set $\widetilde{V}$ is dominated by the Markov generator L on the finite set V (written $\widetilde{L} \prec L$), if there exists a injective Markov kernel Λ from V to $\widetilde{V}$ such that $L\Lambda = \Lambda\widetilde{L}$. The requirement that Λ is one-to-one (with respect to the functional interpretation (10.2)) means that $|V| \geq |\widetilde{V}|$ and that Λ has maximal rank as a matrix. Note that two Markov generators L and $\widetilde{L}$ are Markov-similar if and only if $\widetilde{L} \prec L$ and $L \prec \widetilde{L}$. Most of the results presented up to now have variants for the domination relation $\prec$. In this spirit, Lemma 10.6 can be strengthened into

Lemma 10.8 *Assume that the two Markov generators L and $\widetilde{L}$ are such that $\widetilde{L} \prec L$ and L is irreducible. Then $\widetilde{L}$ is irreducible and if $\widetilde{\mu}$ is its invariant probability, $\mathscr{P}(L, \widetilde{L})$ is a neighborhood of $\widetilde{\mu}$.*

Proof We begin by proving that $\widetilde{L}$ is irreducible. Let $\Lambda \in \mathscr{K}(L, \widetilde{L})$ be injective. Let $\widetilde{f}$ be a function on $\widetilde{V}$ such that $\widetilde{L}[\widetilde{f}] = 0$. By the intertwining relation, we get that $L[\Lambda[\widetilde{f}]] = 0$, so that by irreducibility of L, $\Lambda[\widetilde{f}]$ is constant and by injectivity of Λ, $\widetilde{f}$ is constant (since $\Lambda[\mathbb{1}] = \mathbb{1}$). This property implies that if $\widetilde{V}$ is decomposed into irreducible classes with respect to $\widetilde{L}$, then there is only one terminal class (namely $\widetilde{A}_1$ is a singleton, with the notation introduced at the beginning of Sect. 10.3). So to prove that $\widetilde{L}$ is irreducible, it is sufficient that to show that $\widetilde{L}$ admits an invariant probability whose support is $\widetilde{V}$. By the intertwining relation, we get that $\widetilde{\mu} := \mu\Lambda$ is an invariant probability of $\widetilde{L}$, if μ is the invariant probability of L. It remains to see that $\widetilde{\mu}$ gives a positive weight to all the elements of $\widetilde{V}$. Let $\mathscr{M}(V)$ be the set of signed measures on V. The Markov kernel Λ can be seen as an operator from $\mathscr{M}(V)$ to $\mathscr{M}(\widetilde{V})$ via:

$$\forall m \in \mathscr{M}(V),\ \forall \widetilde{x} \in \widetilde{V}, \qquad m\Lambda(\widetilde{x}) := \sum_{x \in V} m(x)\Lambda(x, \widetilde{x})$$

It corresponds to the dual operator of Λ seen as an operator on functions, through the natural duality between functions and signed measures on V. In particular, seen as an operator on signed measures, Λ is onto. As a consequence, for any $\widetilde{x} \in \widetilde{V}$, we can find $m_{\widetilde{x}} \in \mathscr{M}(V)$ such that $m_{\widetilde{x}}\Lambda = \delta_{\widetilde{x}}$. Since μ gives a positive weight to all elements of V, we can find $\epsilon > 0$ small enough so that for all real numbers $(a_x)_{x \in V}$ with $|a_x| \leq \epsilon$ for all $x \in V$, $\mu + \sum_{x \in V} a_x m_x$ is a non negative measure. It follows that

$$\widetilde{\mu} + \sum_{x \in V} a_x \delta_x = (\mu + \sum_{x \in V} a_x m_x)\Lambda$$

is a non-negative measure. This is only possible, for all $(a_x)_{x \in V}$ as above, if and only if $\widetilde{\mu}$ gives a positive weight to all the elements of $\widetilde{V}$.

The same argument shows that Λ transforms neighborhoods of μ into neighborhoods of $\widetilde{\mu}$, so $\mathscr{P}(L, \widetilde{L})$ is a neighborhood of $\widetilde{\mu}$.

The proof of Lemma 10.7 can now be adapted (replacing the link Λ by an injective Markov kernel) to show:

Proposition 10.6 *Under the assumption of Lemma 10.7, there exists $T \geq 0$ such that $\widetilde{m}_T \in \mathscr{P}(L, \widetilde{L})$ for all $\widetilde{m}_0 \in \mathscr{P}(\widetilde{V})$. It follows that*

$$\forall \psi \in \Psi, \forall t \geq 0, \qquad \widetilde{E}(\psi, T+t) \leq E(\psi, t)$$

with the notation introduced in Proposition 10.1, because for any $\psi \in \Psi$, $\widetilde{m}_0 \in \mathscr{P}(L, \widetilde{L})$ and $t \geq 0$, we have

$$E_\psi(\widetilde{m}_t|\widetilde{\mu}) \leq E(\psi, t)$$

according to (10.10).

Let us illustrate the previous considerations on the simplest example.

Example 10.5 Consider the two points set $V{:=}\{0, 1\}$. Any generator L on V can be written under the form $L = l(\mu - \mathrm{Id})$, where Id is the 2×2-identity matrix, $l \geq 0$ and μ is "the" invariant measure of L (note that except if $L = 0$, which corresponds to $l = 0$, L has a unique invariant measure μ). It appears that $-L$ is diagonalizable and its eigenvalues are 0 and l. The generator L is non-transient if and only if $a > 0$ and $\mu > 0$ (in the sense that $\mu(0) > 0$ and $\mu(1) > 0$). The generator $L \neq 0$ is transient if and only if $l > 0$ and μ is a Dirac mass. The left case is $L = 0$. Consider another generator $\widetilde{L} = \widetilde{l}(\widetilde{\mu} - \mathrm{Id})$ on $\{0, 1\}$. According to Corollary 10.1 and Theorem 10.1, it is Markov-similar to the non-transient L if and only if $l = \widetilde{l}$ and $\widetilde{\mu} > 0$. From Corollary 10.1, we also deduce that the generator $\widetilde{L} \neq 0$ is Markov-similar to the transient $L \neq 0$ if and only if $\widetilde{l} = l$ and $\widetilde{\mu}$ is a Dirac mass. Finally the unique generator Markov similar to $L = 0$ is 0 itself.

From now on, we assume that L and $\widetilde{L}$, as above, are non-transient and Markov-similar. Let Λ be a Markov kernel on $\{0, 1\}$ such that $L\Lambda = \Lambda\widetilde{L}$. This amounts to $\mu\Lambda = \widetilde{\mu}$, namely $\mathscr{K}(L, \widetilde{L})$ is the set of Markov kernels transporting μ on $\widetilde{\mu}$ (in general, it is only a subset of those Markov kernels). Since $\Lambda - \mathrm{Id}$ is a Markov generator, we can find $a \geq 0$ and a probability measure ν on $\{0, 1\}$ such that

$$\Lambda = (1 - a)\,\mathrm{Id} + a\nu$$

This is not sufficient to insure that Λ is a Markov kernel: to get that the entries are non-negative, we need furthermore that $a \in [0, 1/(1 - \min(\nu))]$, but it will not be convenient to work directly with this condition. The relation $\mu\Lambda = \widetilde{\mu}$ is equivalent to

$$a\nu = \widetilde{\mu} - (1 - a)\mu$$

For the l.h.s. to be non-negative, we must have $a \geq 1 - \min(\widetilde{\mu}/\mu)$. The kernel Λ can be written under the form

$$\Lambda_a := (1-a)(\mathrm{Id} - \mu) + \widetilde{\mu}$$

For the entries of this matrix to be non-negative, we must have:

- For $a > 1$: for all $x \in \{0, 1\}$,

  $$(1-a)(1-\mu(x)) + \widetilde{\mu}(x) \geq 0$$

  i.e.

  $$a \leq 1 + \frac{\widetilde{\mu}(x)}{1-\mu(x)}$$

  Let $a_+ := 1 + \min(\widetilde{\mu}/(1-\mu))$, this condition is $a \leq a_+$.
- For $a = 1$, $\Lambda_1 = \widetilde{\mu}$ has non-negative entries.
- For $a < 1$: for all $x \in \{0, 1\}$,

  $$(a-1)\mu(x) + \widetilde{\mu}(x) \geq 0$$

  and we recover the condition $a \geq a_- := 1 - \min(\widetilde{\mu}/\mu)$.

  Thus we get

  $$\mathscr{K}(L, \widetilde{L}) = \{\Lambda_a \,:\, a \in [a_-, a_+]\}$$

Since the mapping $a \mapsto \Lambda_a$ is affine and that the set of probability measures on $\{0, 1\}$ is of dimension 1, it appears that $\mathscr{P}(L, \widetilde{L})$ is the segment generated by the four probabilities $\eta_{\star,y} := \delta_y \Lambda_{a_\star}$, with $y \in \{0, 1\}$ and $\star \in \{-, +\}$. Let $x_0, x_1 \in \{0, 1\}$ be respectively such that

$$\frac{\widetilde{\mu}}{\mu}(x_0) = \min\left(\frac{\widetilde{\mu}}{\mu}\right)$$
$$\frac{\widetilde{\mu}}{1-\mu}(x_1) = \min\left(\frac{\widetilde{\mu}}{1-\mu}\right)$$

We have, for any $y \in \{0, 1\}$,

$$\eta_{-,y} = \frac{\widetilde{\mu}(x_0)}{\mu(x_0)}(\delta_y - \mu) + \widetilde{\mu}$$
$$\eta_{+,y} = \frac{\widetilde{\mu}(x_1)}{1-\mu(x_1)}(\mu - \delta_y) + \widetilde{\mu}$$

So, denoting $\bar{x}:=1-x$ for all $x \in \{0,1\}$, we compute that $\eta_{-,\bar{x}_0} = \delta_{\bar{x}_0}$ and $\eta_{+,x_1} = \delta_{\bar{x}_1}$. We also get

$$\eta_{-,x_0}(x_0) = \frac{\widetilde{\mu}(x_0)}{\mu(x_0)} \text{ and } \eta_{+,\bar{x}_1}(x_1) = \frac{\widetilde{\mu}(x_1)}{\mu(\bar{x}_1)} \tag{10.11}$$

So η_{-,x_0} is a Dirac mass if and only if $\widetilde{\mu} = \mu$ and $\eta_{+,\bar{x}_1}$ is a Dirac mass if and only if $\widetilde{\mu}$ is the image of μ by the involution of $\{0,1\}$, $x \mapsto \bar{x}$. Without loss of generality, assume that $\mu(1) \geq \mu(0)$ and $\widetilde{\mu}(1) \geq \widetilde{\mu}(0)$. In particular $\widetilde{\mu}$ is the image of μ by the involution of $\{0,1\}$, $x \mapsto \bar{x}$ if and only if μ and $\widetilde{\mu}$ are the uniform measure. Next let us dismiss the cases where $\mu = \widetilde{\mu}$, i.e. $L = \widetilde{L}$, because it is clear then that $\mathscr{P}(L, L) = \mathscr{P}(V)$. From the above considerations, it follows that for $L \neq \widetilde{L}$, we will have $\mathscr{P}(L, \widetilde{L}) = \mathscr{P}(V)$ if and only if the convex hull of $\{\delta_{\bar{x}_0}, \delta_{\bar{x}_1}\}$ is $\mathscr{P}(V)$, i.e. if $x_0 \neq x_1$. But we just assumed that μ and $\widetilde{\mu}$ are non-decreasing, we have $x_1 = 0$. Thus $\mathscr{P}(L, \widetilde{L}) = \mathscr{P}(V)$ if and only if $\frac{\widetilde{\mu}(1)}{\mu(1)} \geq \frac{\widetilde{\mu}(0)}{\mu(0)}$, i.e. $\widetilde{\mu}(1) \leq \mu(1)$. Note that $\mathscr{P}(L, \widetilde{L}) = \mathscr{P}(V) = \mathscr{P}(\widetilde{L}, L)$ is in fact equivalent to $L = \widetilde{L}$. A first conclusion is that if $\widetilde{\mu}(1) \leq \mu(1)$, then

$$\forall \psi \in \Psi, \ \forall t \geq 0, \qquad \widetilde{E}(\psi, t) \leq E(\psi, t)$$

Assume next that $\widetilde{\mu}(1) > \mu(1)$. Then we have $x_0 = x_1 = 0$ and from (10.11) (taking into account that $\widetilde{\mu}(0)/\mu(0) \geq \widetilde{\mu}(0)/\mu(1)$) we deduce that

$$\mathscr{P}(L, \widetilde{L}) = \left[\frac{\widetilde{\mu}(0)}{\mu(0)}\delta_0 + \left(1 - \frac{\widetilde{\mu}(0)}{\mu(0)}\right)\delta_1, \delta_1\right]$$

It follows that we can take in Lemma 10.7,

$$\begin{aligned} T &= \min\left\{t \geq 0 \,:\, \forall \widetilde{m}_0 \in \mathscr{P}(\{0,1\}), \widetilde{m}_t(1) \geq 1 - \frac{\widetilde{\mu}(0)}{\mu(0)}\right\} \\ &= \min\left\{t \geq 0 \,:\, \delta_0 \exp(t\widetilde{L})(1) \geq 1 - \frac{\widetilde{\mu}(0)}{\mu(0)}\right\} \\ &= \min\left\{t \geq 0 \,:\, (1 - \exp(-lt))\widetilde{\mu}(1) \geq 1 - \frac{\widetilde{\mu}(0)}{\mu(0)}\right\} \\ &= -\frac{1}{l}\ln\left(1 - \frac{\widetilde{\mu}(1) - \mu(1)}{\widetilde{\mu}(1)(1 - \mu(1))}\right) \end{aligned}$$

We end this section by pointing out the links between the objects introduced above with the notion of weak hypergroup. Let be given a Markov generator L on the finite set V. The set $\mathscr{K}(L, L)$ was called the Markov commutator of L in [14], since it consists of the Markov kernels commuting with L (as already mentioned in Remark 10.1). Following this previous paper, the generator L is said to be a **weak hypergroup** with respect to $x_0 \in V$ if for any $m \in \mathscr{P}(V)$, there exists

$K \in \mathscr{K}(L, L)$ such that $K(x_0, \cdot) = m(\cdot)$. Taking advantage of the fact that for any Markov generators L, $\widetilde{L}$ and $\widehat{L}$, we have the inclusion $\mathscr{K}(L, \widetilde{L})\mathscr{K}(\widetilde{L}, \widehat{L}) \subset \mathscr{K}(L, \widehat{L})$, we deduce the following criterion:

Proposition 10.7 *Assume that L and $\widetilde{L}$ are two Markov generators on V and $\widetilde{V}$ respectively, such that $\widetilde{L}$ is a weak hypergroup with respect to $\widetilde{x}_0$ and there exists $x_0 \in V$ and $\Lambda \in \mathscr{K}(L, \widetilde{L})$ with $\Lambda(x_0, \cdot) = \delta_{\widetilde{x}_0}$. Then we have $\mathscr{P}(L, \widetilde{L}) = \mathscr{P}(\widetilde{V})$ and by consequence,*

$$\forall \psi \in \Psi,\ \forall t \geq 0, \qquad \widetilde{E}(\psi, t) \leq E(\psi, t)$$

This condition generalizes the deduction of $\mathscr{P}(L, \widetilde{L}) = \mathscr{P}(\widetilde{V})$ given in Example 10.5, which is continued below:

Example 10.6 We come back to the two point case, with the notation introduced in Example 10.5. Consider $L = l(\mu - \mathrm{Id})$ a non-transient generator, where $\mu(0) \leq \mu(1)$. Let us check that L is a weak hypergroup with respect to 0. We begin by computing the commutator $\mathscr{K}(L, L)$. We have seen that any Markov kernel K on $\{0, 1\}$ can be written under the form $(1 - k)\,\mathrm{Id} + k\nu$, where $\nu \in \mathscr{P}(\{0, 1\})$ and $k \in [0, 1/(1 - \min(\nu))]$. It appears that K commutes with L if and only if $k\nu$ commutes with μ, namely if $k = 0$ or $\nu = \mu$. So we get

$$\mathscr{K}(L, L) = \{K_k := (1 - k)\,\mathrm{Id} + k\mu \;:\; k \in [0, 1/(1 - \min(\mu))]\}$$

Since $K_0(0, \cdot) = \delta_0$ and $K_{1/(1-\mu(0))}(0, \cdot) = \delta_1$, we get that L is a weak hypergroup.

Consider another non-transient Markov generator $\widetilde{L}$, to fulfill the assumptions of Proposition 10.7, we are wondering if we can find $x_0 \in \{0, 1\}$ and $\Lambda \in \mathscr{K}(L, \widetilde{L})$ such that $\Lambda(x_0, \cdot) = \delta_0$. As we have already deduced it from (10.11), this is equivalent to $\widetilde{\mu}(1) \leq \mu(1)$.

10.5 On Infinite State Spaces

The goal of this short section is to prove Proposition 10.2 and to suggest that the infinite state space situation would deserve to be investigated further.

Let L and $\widetilde{L}$ be two Markov generators, respectively on the measurable spaces $(V, \mathscr{V})$ and $(\widetilde{V}, \widetilde{\mathscr{V}})$. The simple implication in Proposition 10.2 holds under weaker assumptions than L and $\widetilde{L}$ being nice:

Lemma 10.9 *Assume that L and $\widetilde{L}$ admit unique invariant probabilities μ and $\widetilde{\mu}$ which are reversible, and that their spectra consist of eigenvalues, respectively in $\mathbb{L}^2(\mu)$ and $\mathbb{L}^2(\widetilde{\mu})$. If L and $\widetilde{L}$ are weakly Markov-similar in the abstract sense, then they are isospectral.*

Proof Let Λ be an abstract weak $\mathbb{L}^2$-link such that $L\Lambda = \Lambda\widetilde{L}$. As the operator Λ is Markovian, it has norm 1. Denote by $(\widetilde{\lambda}_l)_{l\in\mathbb{Z}_+}$ and $(\widetilde{\varphi}_l)_{l\in\mathbb{Z}_+}$ the eigenvalues and respective orthonormal eigenvectors of $\widetilde{L}$. For any $l \in \mathbb{Z}_+$, $\Lambda[\widetilde{\varphi}_l]$ belongs to $\mathbb{L}^2(\mu)$ (its norm is less than or equal to 1) and from the intertwining relation we deduce that it is an eigenfunction of L associated to the eigenvalue $\widetilde{\lambda}_l$. Taking into account that Λ is one-to-one, we deduce that the spectrum of $\widetilde{L}$ is included into that of L. Conversely, considering $\widetilde{\Lambda}$ an abstract weak $\mathbb{L}^2$-link such that $\widetilde{L}\widetilde{\Lambda} = \widetilde{\Lambda}L$, we get the reverse inclusion.

The proof of the reciprocal implication is an extension of that of Lemma 10.1.

Lemma 10.10 *If the two nice generators L and $\widetilde{L}$ are isospectral, then they are weakly Markov-similar.*

Proof Let $(\varphi_l)_{l\in\mathbb{Z}_+}$ and $(\widetilde{\varphi}_l)_{l\in\mathbb{Z}_+}$ be bounded orthonormal eigenvectors of L and $\widetilde{L}$, respectively, associated to the same family of eigenvalues $(\lambda_l)_{l\in\mathbb{Z}_+}$. We can and will assume that $\varphi_0 = \mathbb{1}$ and $\widetilde{\varphi}_0 = \mathbb{1}$ are the constant eigenvectors associated to the eigenvalue 0. We will construct an operator Λ such that $L\Lambda = \Lambda\widetilde{L}$ by requiring that

$$\forall l \in \mathbb{Z}_+, \qquad \Lambda[\widetilde{\varphi}_l] = a_l\varphi_l$$

for a conveniently chosen sequence $(a_l)_{l\in\mathbb{Z}_+}$. First we impose that $a_0 = 1$, so that $\Lambda[\mathbb{1}] = \mathbb{1}$. Next we choose the remaining coefficients positive and satisfying

$$\sum_{l\in\mathbb{N}} a_l \,\|\varphi_l\|_\infty \,\|\widetilde{\varphi}_l\|_\infty \leq 1 \tag{10.12}$$

This is possible, since the eigenvectors are bounded. Let us check that such an operator Λ preserves non-negativity. It is sufficient to show that if $\widetilde{f}$ is a measurable function defined on $\widetilde{V}$ and taking values in $[0, 1]$, then $\Lambda[\widetilde{f}] \geq 0$ μ-a.s. Since $\widetilde{f} \in \mathbb{L}^2(\widetilde{\mu})$, we can decompose it on the orthonormal basis $(\widetilde{\varphi}_l)_{l\in\mathbb{Z}_+}$:

$$\widetilde{f} = \sum_{l\in\mathbb{Z}_+} b_l\widetilde{\varphi}_l$$

where a priori the coefficients $(b_l)_{l\in\mathbb{Z}_+}$ belong to $\mathbb{R}$. We have $b_0 = \widetilde{\mu}[\mathbb{1}\widetilde{f}] \in [0, 1]$ and

$$\begin{aligned}\forall l \in \mathbb{N}, \qquad |b_l| &= |\widetilde{\mu}[\widetilde{\varphi}_l\widetilde{f}]| \\ &\leq \|\widetilde{\varphi}_l\|_\infty \,\widetilde{\mu}[\mathbb{1}\widetilde{f}] \\ &= \|\widetilde{\varphi}_l\|_\infty \, b_0\end{aligned}$$

Thus we get

$$\begin{aligned}\Lambda[\widetilde{f}] &= \Lambda\left[\sum_{l\in\mathbb{Z}_+} b_l\widetilde{\varphi}_l\right]\\ &= \sum_{l\in\mathbb{Z}_+} a_l b_l \varphi_l\\ &= b_0\mathbb{1} + \sum_{l\in\mathbb{N}} a_l b_l \varphi_l\\ &\geq b_0 - \sum_{l\in\mathbb{N}} a_l |b_l|\, \|\varphi_l\|_\infty\\ &\geq \left(1 - \sum_{l\in\mathbb{N}} a_l\, \|\varphi_l\|_\infty\, \|\widetilde{\varphi}_l\|_\infty\right) b_0\\ &\geq 0\end{aligned}$$

according to (10.12). It follows that Λ is a Markov operator in the abstract sense. It comes from a Markov kernel, due to the assumption on the state spaces.

It remains to show that $\mu\Lambda$ is equal to the invariant probability $\widetilde{\mu}$ of $\widetilde{L}$. By the intertwining relation, $\mu\Lambda$ is an invariant probability of $\widetilde{L}$, thus by uniqueness of the latter, we have $\mu\Lambda = \widetilde{\Lambda}$.

It is natural to imagine a strong version of Proposition 10.2. The Markov operator $\Lambda : \mathbb{L}^2(\widetilde{\mu}) \to \mathbb{L}^2(\mu)$ is said to be a **strong link** if it is invertible and its inverse is bounded. This notion leads to the definition: two Markov generators L and $\widetilde{L}$ are **strongly Markov-similar** if they can be mutually intertwined through strong links. We are wondering if two nice isospectral Markov generators would not be strongly Markov-similar if their eigenvectors are uniformly bounded, namely with the above notation, if

$$\sup_{l\in\mathbb{Z}_+} \|\varphi_l\|_\infty < +\infty \text{ and } \sup_{l\in\mathbb{Z}_+} \|\widetilde{\varphi}_l\|_\infty < +\infty$$

Note that the examples of isospectral flat manifolds presented in the review of Gordon [10] can be shown to be strongly Markov-similar, by transforming the transplantation maps (i.e. unitary instead of Markovian intertwining maps, see the papers of Bérard [2, 3]) into strong links. More precisely, it is sufficient to take $a, b > 0$ such that $4a + 3b = 1$ in the matrix T displayed page 763 of Gordon [10].

Acknowledgements This paper was motivated by the presentation of Pierre Patie of his paper with Mladen Savov [16], I'm grateful to him for all the explanations he gave me. I'm also thankful to the ANR STAB (Stabilité du comportement asymptotique d'EDP, de processus stochastiques et de leurs discrètisations) for its support.

References

1. P. Bérard, Variétés riemanniennes isospectrales non isométriques. Astérisque (177–178), Exp. No. 705, 127–154 (1989). Séminaire Bourbaki, vol. 1988/89
2. P. Bérard, Transplantation et isospectralité. I. Math. Ann. **292**(3), 547–559 (1992). https://doi.org/10.1007/BF01444635
3. P. Bérard, Transplantation et isospectralité. II. J. Lond. Math. Soc. (2) **48**(3), 565–576 (1993). https://doi.org/10.1112/jlms/s2-48.3.565
4. P. Del Moral, M. Ledoux, L. Miclo, On contraction properties of Markov kernels. Probab. Theory Relat. Fields **126**(3), 395–420 (2003). https://doi.org/10.1007/s00440-003-0270-6
5. P. Diaconis, J.A. Fill, Strong stationary times via a new form of duality. Ann. Probab. **18**(4), 1483–1522 (1990)
6. P. Diaconis, L. Miclo, On times to quasi-stationarity for birth and death processes. J. Theor. Probab. **22**(3), 558–586 (2009). https://doi.org/10.1007/s10959-009-0234-6
7. P. Diaconis, L. Miclo, On quantitative convergence to quasi-stationarity (2014). Available at http://hal.archives-ouvertes.fr/hal-01002622
8. E.B. Dynkin, Markov processes, in *Translated with the Authorization and Assistance of the Author by J. Fabius, V. Greenberg, A. Maitra, G. Majone*, vols. I, II. Die Grundlehren der Mathematischen Wissenschaften, Bände 121, vol. 122 (Academic, New York, 1965)
9. S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (Wiley, New York, 1986)
10. C.S. Gordon, Survey of isospectral manifolds, in *Handbook of Differential Geometry*, vol. I (North-Holland, Amsterdam, 2000), pp. 747–778. https://doi.org/10.1016/S1874-5741(00)80009-6
11. N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library, vol. 24, 2nd edn. (North-Holland, Amsterdam, 1989)
12. M. Kac, Can one hear the shape of a drum? Am. Math. Mon. **73**(4, Part II), 1–23 (1966)
13. L. Miclo, On absorption times and Dirichlet eigenvalues. ESAIM Probab. Stat. **14**, 117–150 (2010). https://doi.org/10.1051/ps:2008037
14. L. Miclo, On the Markov commutator (2015). Preprint available at https://hal.archives-ouvertes.fr/hal-01143511
15. S. Pal, M. Shkolnikov, Intertwining diffusions and wave equations (2013). ArXiv e-prints
16. P. Patie, M. Savov, Spectral expansions of non-self-adjoint generalized Laguerre semigroups (2015). ArXiv e-prints
17. L.C.G. Rogers, J.W. Pitman, Markov functions. Ann. Probab. **9**(4), 573–582 (1981)

Chapter 11
Sharp Rate for the Dual Quantization Problem

Gilles Pagès and Benedikt Wilbertz

Abstract In this paper we establish the sharp rate of the optimal dual quantization problem. The notion of dual quantization was introduced in Pagès and Wilbertz (SIAM J Numer Anal 50(2):747–780, 2012). Dual quantizers, at least in a Euclidean setting, are based on a Delaunay triangulation, the dual counterpart of the Voronoi tessellation on which "regular" quantization relies. This new approach to quantization shares an intrinsic stationarity property, which makes it very valuable for numerical applications.

We establish in this paper the counterpart for dual quantization of the celebrated Zador theorem, which describes the sharp asymptotics for the quantization error when the quantizer size tends to infinity. On the way we establish an extension of the so-called Pierce Lemma by a random quantization argument. Numerical results confirm our choices.

Keywords Dual quantization · Delaunay triangulation · Zador Theorem · Pierce Lemma · random quantization

11.1 Introduction

A new notion of vector quantization called *dual quantization* (or *Delaunay quantization* in a Euclidean framework) has been introduced in [12] after a first one dimensional try in [11]. Some applications were developed in [10], devoted to the

G. Pagès (✉)
Laboratoire de Probabilités, Statistique et Modélisation, UMR 8002, Sorbonne-Université, Paris, France
e-mail: gilles.pages@sorbonne-universite.fr

B. Wilbertz
Trendiction S.A., Luxembourg, Luxembourg

Laboratoire de Probabilités, Statistique et Modélisation, UMR 8002, Sorbonne-Université, Paris, France
e-mail: benedikt.wilbertz@gmx.de

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_11

design of new quantization based numerical schemes for multi-dimensional optimal stopping and stochastic control problems arising in Finance (see also [1]). The general principle of dual quantization consists of mapping an $\mathbb{R}^d$-valued random vector (r.v.) onto a non-empty finite subset (or *grid*) $\Gamma \subset \mathbb{R}^d$ using an appropriate random *splitting operator* $\mathscr{J}_\Gamma : \Omega_0 \times \mathbb{R}^d \to \Gamma$ (defined on an exogenous probability space $(\Omega_0, \mathscr{S}_0, \mathbb{P}_0)$) which satisfies the intrinsic *stationarity property*

$$\forall\, \xi \in \text{conv}(\Gamma), \qquad \mathbb{E}_{\mathbb{P}_0}\big(\mathscr{J}_\Gamma(\xi)\big) = \int_{\Omega_0} \mathscr{J}_\Gamma(\omega_0, \xi)\,\mathbb{P}_0(d\omega_0) = \xi, \tag{11.1}$$

where conv(Γ) denotes the convex hull of Γ in $\mathbb{R}^d$. Every r.v. $X : (\Omega, \mathscr{S}, \mathbb{P}) \to \text{conv}(\Gamma)$ defined on any probability space can be canonically extended to $(\Omega_0 \times \Omega, \mathscr{S}_0 \otimes \mathscr{S}, \mathbb{P}_0 \otimes \mathbb{P})$ in order to define the *dual quantization* induced by Γ as

$$\widehat{X}^{\Gamma,dual}(\omega_0, \omega) = \mathscr{J}_\Gamma\big(\omega_0, X(\omega)\big).$$

As a specific feature inherited from (11.1), it always satisfies the *dual* or *reverse stationary property*

$$\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}(\mathscr{J}_\Gamma(X)\,|\,X) = X.$$

This can be compared to the more classical Voronoi framework where the Γ-quantization of X is defined from a Borel nearest neighbour projection Proj_Γ by

$$\widehat{X}^{\Gamma,vor}(\omega) = \text{Proj}_\Gamma(X(\omega)).$$

The stationary property then reads: $\mathbb{E}\big(X\,|\,\widehat{X}^{\Gamma,vor}\big) = \widehat{X}^{\Gamma,vor}$, except that it holds only for grids which are critical points (typically local minima) of the so-called distortion function (see e.g. [5]) in a Euclidean framework.

To each quantization corresponds a functional approximation operator: Voronoi quantization is related to the *stepwise constant functional approximation operator* $f \circ \text{Proj}_\Gamma$ whereas dual quantization leads to an operator defined for every $\xi \in \text{conv}(\Gamma)$ by

$$\mathbb{J}_\Gamma(f)(\xi) = \mathbb{E}_{\mathbb{P}_0}\big(f(\mathscr{J}_\Gamma(\omega_0, \xi))\big) = \sum_{x\in\Gamma} f(x)\lambda_x(\xi), \tag{11.2}$$

where $\lambda_x(\xi) = \mathbb{P}_0(\mathscr{J}_\Gamma(.,\xi) = x)$, $x \in \Gamma$, are barycentric "pseudo-coordinates" of ξ in Γ satisfying $\lambda_x(\xi) \in [0, 1]$, $\sum_{x\in\Gamma} \lambda_x(\xi) = 1$ and $\sum_{x\in\Gamma} \lambda_x(\xi)x = \xi$. The operator $\mathbb{J}_\Gamma$ is an *interpolation* operator which turns out, under appropriate conditions, to be more regular (continuous and stepwise affine, see [10]) than its "Voronoi" counterpart. It is shown in [10–12] how we can take advantage of this intrinsic stationary property to produce more accurate error bounds for the resulting *cubature formula*

$$\mathbb{E}_{\mathbb{P}}\big(f(\widetilde{X}^{\Gamma,dual})\big) = \mathbb{E}_{\mathbb{P}}\big(\mathbb{J}_{\Gamma}(f)(X)\big) = \mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(f(\mathscr{J}_{\Gamma}(\omega_0,\xi))\big) = \sum_{x\in\Gamma} w_x^{dual} f(x) \tag{11.3}$$

where $w_x^{dual} = \mathbb{E}_{\mathbb{P}}(\lambda_x(X)) = \mathbb{P}\otimes\mathbb{P}_0(\mathscr{J}_{\Gamma}(\omega_0, X) = x)$, $x \in \Gamma$, regardless of any optimality property of Γ with respect to $\mathbb{P}_X$. Typically, if $f \in \mathrm{Lip}(\mathbb{R}^d, \mathbb{R})$ (Lipschitz continuous function) with coefficient $[f]_{\mathrm{Lip}}$,

$$\begin{aligned}\Big|\mathbb{E}_{\mathbb{P}} f(X) - \mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0} f(\widetilde{X}^{\Gamma,dual})\Big| &\le [f]_{\mathrm{Lip}}\big\|X - \widehat{X}^{\Gamma,dual}\big\|_{L^1(\mathbb{P}\otimes\mathbb{P}_0)}\\ &= [f]_{\mathrm{Lip}}\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\|\big)\\ &= [f]_{\mathrm{Lip}}\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\| \mid X)\big)\end{aligned}$$

whereas, if f has a Lipschitz continuous differential Df (the norm on $\mathbb{R}^d$ is denoted $\|\,.\,\|$), a second order Taylor expansion yields

$$\begin{aligned}\Big|\mathbb{E}_{\mathbb{P}} f(X) - \mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0} f(\widetilde{X}^{\Gamma,dual})\Big| &\le \Big\|f(X) - \mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(f(\mathscr{J}_{\Gamma}(\omega_0, X)) \mid X\big)\Big\|_{L^1(\mathbb{P}\otimes\mathbb{P}_0)}\\ &\le [Df]_{\mathrm{Lip}}\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\|^2\big)\\ &\le [Df]_{\mathrm{Lip}}\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\|^2 \mid X)\big)\end{aligned} \tag{11.4}$$

where $\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\|^p \mid X) = \sum_{x\in\Gamma} \lambda_x(X)\|X - x\|^p = \mathbb{J}_{\Gamma}(\|.\|^2)(X)$, $p = 1, 2$ (see Sect. 11.2.2 for more details).

More generally, if one aims at approximating $\mathbb{E}\big(f(X) \mid g(Y)\big)$ by its dually quantized counterpart $\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0\otimes\mathbb{P}_1}\big(f(\mathscr{J}_{\Gamma_X}(\omega_0, X)) \mid \mathscr{J}_{\Gamma_Y}(\omega_1, Y)\big)$ (with obvious notations), it is also possible, under natural additional assumptions, to get error bounds based on both dual quantization error moduli related to the quantizations of X and Y respectively, see e.g. the proof (Step 2) of Proposition 2.1 in [10].

This suggests to investigate the properties and the asymptotic behavior of the (Γ, L^p)-mean dual quantization error, $p \in (0, \infty)$, defined by

$$\begin{aligned}\Big\|X - \widehat{X}^{\Gamma,dual}\Big\|_{L^p(\mathbb{P}\otimes\mathbb{P}_0)} &= \Big\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\Big\|_{L^p(\mathbb{P}\otimes\mathbb{P}_0)}\\ &= \Big[\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\Big(\mathbb{E}_{\mathbb{P}\otimes\mathbb{P}_0}\big(\|X - \mathscr{J}_{\Gamma}(\omega_0, X)\|^p \mid X\big)\Big)\Big]^{\frac{1}{p}}\end{aligned}$$

so as to make it as small as possible. This program can be summed up in four phases:

– The first step is to minimize the above conditional expectation, i.e. $\mathbb{E}\,(\|\xi - \mathscr{J}_{\Gamma}(\omega_0, \xi)\|^p)$ for every $\xi \in \mathrm{conv}(\Gamma)$, for a fixed grid Γ i.e. to determine the *best* random splitting operator $\mathcal{J}_\Gamma$. In a regular quantization setting, this phase corresponds to showing that the nearest neighbour projection on Γ is the best projection on Γ.

– The second step is "optional" . It aims at finding grids which minimize the mean dual quantization error $\left\| X - \mathcal{J}_\Gamma(\omega_0, X) \right\|_{L^p(\mathbb{P}\otimes\mathbb{P}_0)}$ *among all grids* Γ *whose convex hull contains the support of the distribution of* X or equivalently such that $\mathbb{P}\big(X \in \mathrm{conv}(\Gamma)\big) = 1$.
– The third step is to extend dual quantization to r.v.s X with unbounded support with in mind that, if the stationarity can no longer hold outside conv(Γ), one still benefits from it inside conv(Γ) (see e.g. (11.4)). A balance between the quantization errors induced by the outside and the inside of conv(Γ) will lead to the appropriate unfolding of the grid Γ in the optimization phase.

  The first two steps have been already solved in [12]. We discuss in-depth the third one in Sect. 11.2.2. The aim of this paper is to solve the last (and fourth) step:

– the fourth step aims at elucidating the rate of decay to 0 of the *optimal* L^p-mean dual quantization error modulus, i.e. minimized over all grids Γ of size at most N—as N grows to infinity.

This means establishing in a dual quantization framework the counterpart of Zador's Theorem—recalled below—which rules the convergence rate of optimal "regular" (Voronoi) quantization.

To be more precise, we will establish such a theorem, for L^∞-bounded r.v.s but also, once mean dual quantization error will have been extended in an appropriate way, following [12], to L^p-integrable r.v.s.

Let us now introduce in more formal way the (local and mean) dual quantization error moduli, following [12]. For a grid $\Gamma \subset \mathbb{R}^d$, we define the L^p*-mean dual quantization error* of X induced by the grid Γ by

$$d_p(X; \Gamma) = \| F_p(X; \Gamma) \|_{L^p(\mathbb{P})} \tag{11.5}$$

where F_p denotes the *local dual quantization error* function defined by

$$F_p(\xi; \Gamma) = \inf \left\{ \Big(\sum_{x\in\Gamma} \lambda_x \|\xi - x\|^p \Big)^{\frac{1}{p}},\ \lambda_x \in [0,1],\ \sum_{x\in\Gamma} \lambda_x\, x = \xi,\ \sum_{x\in\Gamma} \lambda_x = 1 \right\}. \tag{11.6}$$

Note that $F_p(\xi; \Gamma) < +\infty$ if and only if $\xi \in \mathrm{conv}(\Gamma)$ so that $d_p(X; \Gamma) < +\infty$ if and only if $X \in \mathrm{conv}(\Gamma)$ $\mathbb{P}$-*a.s.*. and that $d_p(X; \Gamma) = \left\| X - \widehat{X}^{\Gamma, dual} \right\|^p_{L^p(\mathbb{P}\otimes\mathbb{P}_0)}$. Hence, this notion only makes sense for compactly supported r.v.s. In particular, if the support of $\mathbb{P}_X$ is compact and contains $d+1$ affinely independent points, $d_{n,p}(X, \Gamma) = +\infty$ as long as $n \le d$. This new quantization modulus leads to an optimal dual quantization problem at *level* n,

$$d_{n,p}(X) = \inf\Big\{d_{n,p}(X, \Gamma),\ \Gamma \subset \mathbb{R}^d,\ |\Gamma| \le n\Big\}$$
$$= \inf\Big\{\|F_p(X; \Gamma)\|_p,\ \Gamma \subset \mathbb{R}^d,\ |\Gamma| \le n\Big\}$$

where $|\Gamma|$ denotes the cardinality of Γ.

One important application of quantization in general is the use of quantization grids to devise cubature formulas for numerical integration (see (11.3)) or conditional expectation approximation (see [10]). The main feature here is the stationarity which allows to derive a second order formula for the error (11.3) (see (11.4)). As, by construction, dual quantization can achieve stationarity only for compactly supported r.v.s/distributions, we explain in Sect. 11.2.2 how the extension of dual quantization to non-compactly supported random variables, as defined in [12], still preserves in practice this second order rate. We therefore define the splitting operator $\mathscr{J}_\Gamma$ outside conv(Γ) by setting

$$\forall\, \xi \in \mathbb{R}^d \setminus \mathrm{conv}(\Gamma), \qquad \mathscr{J}_\Gamma(\omega_0, \xi) = \mathrm{Proj}_{\mathrm{conv}\,(\Gamma)\cap\partial\Gamma}(\xi)$$

where $\mathrm{Proj}_{\mathrm{conv}\,(\Gamma)\cap\partial\Gamma}$ is a Borel nearest neighbour projection on Γ. This extension is not canonical: an alternative choice could be to set $\mathscr{J}_\Gamma(\omega_0, \xi) = \mathrm{Proj}_{\partial\Gamma\cap\mathrm{conv}(\Gamma)}(\xi)$ where $\partial\Gamma$ denotes the boundary of conv(Γ). A posteriori, after the grid optimization phase, this alternative extension often coincides with ours which turns out to be more tractable in terms of simulation. We will also prove that it does not deteriorate the resulting mean quantization error when the grid size $|\Gamma| \to +\infty$. Though the stationary property is lost as expected, we point out in Sect. 11.2.2 that this operator remains as performing as $\mathscr{J}_\Gamma$ is for bounded r.v.s when implementing cubature formulas for possibly unbounded L^p-integrable r.v.s.

Hence, we define the *extended local dual quantization error* function as

$$\bar{F}_p(\xi; \Gamma) := F_p(\xi; \Gamma)\, \mathbf{1}_{\mathrm{conv}(\Gamma)}(\xi) + \mathrm{dist}(X, \Gamma)\, \mathbf{1}_{\mathrm{conv}(\Gamma)^c}(\xi) \tag{11.7}$$

and the *extended L^p-mean dual quantization error* of X induced by Γ as

$$\bar{d}_p(X; \Gamma) = \|\bar{F}_p(X; \Gamma)\|_{L^p(\mathbb{P})}. \tag{11.8}$$

Finally, we define the *extended L^p-mean dual quantization error* at level n by

$$\bar{d}_{n,p}(X) = \inf\Big\{\bar{d}_p(X, \Gamma),\ \Gamma \subset \mathbb{R}^d,\ |\Gamma| \le n\Big\}. \tag{11.9}$$

At this stage, we also briefly recall a few facts about the (regular) *Voronoi optimal quantization problem* at level n associated to the nearest neighbour projection Proj_Γ:

it reads

$$e_{n,p}(X) = \inf\left\{\|\mathrm{dist}(X,\Gamma)\|_{L^p(\mathbb{P})},\ \Gamma \subset \mathbb{R}^d,\ |\Gamma| \le n\right\} \tag{11.10}$$

(where $\mathrm{dist}(x, A) = \inf_{a\in A} \|x-a\|$). It is well-known that, if $X \in L^p(\mathbb{P})$, $e_{n,p}(X) \downarrow 0$ as $n \to +\infty$. Moreover, the rate of convergence to 0 of $e_{n,p}(X)$ is ruled by Zador's Theorem (see Theorem 6.2 in [5]) and an extended version (see [7]) of the so-called Pierce Lemma (see Lemma 6.6, p. 82, in [5]). We recall below these results for the reader's convenience.

Theorem 11.1

(a) Sharp rate (Zador's Theorem). *Let* $p \in (0,+\infty)$. *Let* $X \in L^{p+\delta}_{\mathbb{R}^d}(\mathbb{P})$ *for some* $\delta > 0$ *with distribution* $\mathbb{P}_X = h.\lambda_d + \nu$, $\nu \perp \lambda_d$, *where* λ_d *denotes the Lebesgue measure on* $(\mathbb{R}^d, \mathscr{B}or(\mathbb{R}^d))$. *Then*

$$\lim_n n^{\frac{1}{d}} e_{n,p}(X) = Q^{vq}_{\|\cdot\|,p,d}\, \|h\|^{\frac{1}{p}}_{\frac{d}{p+d}}$$

where $\|h\|_{\frac{d}{p+d}} = \left(\int_{\mathbb{R}^d} h(\xi)^{\frac{d}{p+d}} d\xi\right)^{1+\frac{p}{d}}$ *and* $Q^{vq}_{\|\cdot\|,p,d} = \inf_n n^{\frac{1}{d}}\, e_{n,p}\big(\mathscr{U}([0,1]^d)\big)$ $\in (0,\infty)$.

(b) Non asymptotic upper bound (Extended Pierce's Lemma) *(see [7]). Let* $p, \delta > 0$. *There exists a real constant* $C_{d,p,\delta} \in (0,+\infty)$ *such that, for every* $\mathbb{R}^d$*-valued random vector* X *defined on a probability space* $(\Omega, \mathscr{A}, \mathbb{P})$,

$$\forall n \ge 1, \quad e_{n,p}(X) \le C_{d,p,\delta}\, \sigma_{p+\delta}(X) n^{-\frac{1}{d}}.$$

where, for any $r \in (0,+\infty)$, $\sigma_r(X) = \min_{a\in\mathbb{R}^d} \|X-a\|_{L^r} \le +\infty$.

Remark In fact, what is precisely proved in [7] (Lemma 1) is

$$\forall n \ge 1, \quad e_{n,p}(X) \le C_{d,\delta}\|X\|_{L^{p+\delta}} n^{-\frac{1}{d}}.$$

To get the above conclusion, just note that the L^p-mean quantization error is invariant by translation since, for every $a \in \mathbb{R}^d$ and $\Gamma \subset \mathbb{R}^d$, finite,

$$\mathrm{dist}(X,\Gamma) = \inf_{b\in\Gamma} \|X-b\| = \inf_{b\in\Gamma} \|(X-a)-(b-a)\| = \mathrm{dist}(X-a, \Gamma-a)$$

where $\Gamma - a = \{b-a,\ b \in \Gamma\}$ so that $e_{n,p}(X) = e_{n,p}(X-a)$ which in turn implies

$$\forall a \in \mathbb{R}^d, \quad e_{n,p}(X) \le C_{d,\delta}\|X-a\|_{L^{p+\delta}}\, n^{-\frac{1}{d}}.$$

Minimizing over a yields the announced result.

The above rate depends on d and is known as the *curse of dimensionality*. Its statement and proof goes back to Zador (PhD, 1954) for the uniform distributions on hypercubes, its extension to absolutely continuous distributions is due to Bucklew and Wise in [2]. A first general rigorous proof (according to mathematical standards) was provided in [5] in 2000 (see also [6] for a survey of the history of quantization).

It should be noted that $d_{n,p}(X)$ and $\bar{d}_{n,p}(X)$ do not coincide even for bounded r.v.s. We will extensively use (see [12]) that

$$d_{n,p}(X) \geq \bar{d}_{n,p}(X) \geq e_{n,p}(X).$$

This paper is entirely devoted to establishing the sharp asymptotics of the optimal dual quantization error moduli $d_{n,p}(X)$ and $\bar{d}_{n,p}(X)$ as n goes to infinity. The main result is stated in Theorem 11.2 (Zador's like theorem) (see Sect. 11.2.1 below). Theorem 11.3 (a Pierce like Lemma) is a companion result which provides a non-asymptotic upper bound for the exact rate simply involving moments of the r.v. X higher than p. Our proof follows the same approach as that in [5] (the first completely rigorous one to our knowledge), except that the splitting operator $\mathcal{J}_\Gamma$ is much more demanding to handle than the plain nearest neighbour projection: it requires non trivial arguments borrowed from convex analysis (including dual primal/methods) and geometry, both in a probabilistic framework. In one dimension, the exact rate $O(n^{-1})$ for $d_{n,p}(X)$ and $\bar{d}_{n,p}(X)$ follows from a random quantization argument detailed in Sect. 11.4 (extended Pierce Lemma for $d_{n,p}(X)$). This rate can be transferred in a d-dimensional framework to $O(n^{-\frac{1}{d}})$ using product (dual) quantization (see Proposition 11.1 below and Sect. 11.3.2), that is a tensorisation argument. Finally, the sharp upper bound is obtained in Sect. 11.5 by successive approximation procedures of the density of X, whereas the lower bound relies on a new "firewall" Lemma.

Notations

- $\|\cdot\|$ denotes a norm on $\mathbb{R}^d$.
- $\mathrm{conv}(A)$ stands for the convex hull of $A \subset \mathbb{R}^d$, $|A|$ for its cardinality, $\mathrm{diam}_{\|\cdot\|}(A) = \sup_{x,y\in A}\|x-y\|$ for its diameter and $\mathrm{aff.dim}(A)$ for the dimension of the affine subspace of $\mathbb{R}^d$ spanned by A.
- We denote $\binom{n}{i} := \frac{n!}{i!(n-i)!}$, $n,\ i \in \{0,\dots,n\}, n \in \mathbb{N} = \{1,2,\dots\}$.
- $\lfloor x \rfloor$ and $\lceil x \rceil$ will denote the lower and the upper integer part of the real number x respectively; set likewise $x_\pm = \max(\pm x, 0)$. For two sequences of real numbers (a_n) and (b_n), $a_n \sim b_n$ if $a_n = u_n b_n$ with $\lim_n u_n = 1$.
- For every $x = (x^1,\dots,x^d) \in \mathbb{R}^d$, $|x|_{\ell^r} = (|x^1|^r + \cdots |x^d|^r)^{1/r}$ denotes the ℓ^r-norm or pseudo-norm, $0 < r < +\infty$ and $|x|_{\ell^\infty} = \max_{1\leq i\leq d}|x_i|$ denotes the ℓ^∞-norm. A general norm on $\mathbb{R}^d$ will be denoted $\|\cdot\|$.
- λ_d denotes the Lebesgue measure on $(\mathbb{R}^d, \mathscr{B}or(\mathbb{R}^d))$ and $\|f\|_p = \left(\int_{\mathbb{R}^d}|f|^p d\lambda_d\right)^{\frac{1}{p}}$ denotes the L^p-(pseudo-)norm with respect to λ_d of the Borel function $f : \mathbb{R}^d \to \mathbb{R}$ for $p > 0$.

- supp(μ) denotes the support of a distribution μ on $(\mathbb{R}^d, \mathscr{B}or(\mathbb{R}^d))$.
- $\|X\|_{L^p}$ or $\|X\|_{L^p(\mathbb{P})}$ denotes the $L^p(\mathbb{P})$-(pseudo-)norm $\big(\mathbb{E}\|X\|^p\big)^{\frac{1}{p}}$ of a random vector $X : (\Omega, \mathscr{S}, \mathbb{P}) \to \mathbb{R}^d$ and $L^p_{\mathbb{R}^d}(\mathbb{P})$ the space of r.v.s for which this quantity is finite.

11.2 Main Results and Motivation for Extended Dual Quantization

11.2.1 Main Results

The theorem below establishes for any $p > 0$ and any norm on $\mathbb{R}^d$ the counterpart of Zador's Theorem in the framework of dual quantization for both $d_{n,p}$ and $\bar{d}_{n,p}$ error moduli.

Theorem 11.2

(*a*) *Let* $X \in L^\infty_{\mathbb{R}^d}(\mathbb{P})$. *Assume the distribution* $\mathbb{P}_X$ *of* X *reads* $\mathbb{P}_X = h.\lambda_d + \nu$, $\nu \perp \lambda_d$. *Then*

$$\lim_n \; n^{\frac{1}{d}}\, d_{n,p}(X) = \lim_n n^{\frac{1}{d}}\, \bar{d}_{n,p}(X) = Q^{dq}_{\|\cdot\|,p,d}\, \|h\|^{\frac{1}{p}}_{\frac{d}{p+d}}$$

where $Q^{dq}_{\|\cdot\|,p,d} = \inf_{n\geq 1} n^{\frac{1}{d}}\, d_{n,p}\big(\mathscr{U}([0,1]^d)\big) \in (0,\infty)$.

(*b*) *Let* $X \in L^{p'}_{\mathbb{R}^d}(\mathbb{P})$, $p' > p$*with* $\mathbb{P}_X$ *like in* (*a*). *Then*

$$\lim_n n^{\frac{1}{d}}\, \bar{d}_{n,p}(X) = Q^{dq}_{\|\cdot\|,p,d}\, \|h\|^{\frac{1}{p}}_{\frac{d}{p+d}}.$$

(*c*) *If* $d = 1$, *then*

$$d_{n,p}\big(\mathscr{U}([0,1])\big) = \left(\frac{2}{(p+1)(p+2)}\right)^{\frac{1}{p}} \frac{1}{n-1},$$

which implies $Q^{dq}_{|\,.\,|,p,1} = \left(\frac{2^{p+1}}{p+2}\right)^{\frac{1}{p}} Q^{vq}_{|\,.\,|,p,1}$.

Moreover, we will also establish in Sect. 11.5 an upper bound for the dual quantization coefficient $Q^{\mathrm{dq}}_{\|\cdot\|,p,d}$ when $\|\cdot\| = |\cdot|_{\ell^r}$.

Proposition 11.1 (Product Quantization) *Let $r, p \in [1, \infty)$ with $r \leq p$. Then it holds for every $d \in \mathbb{N}$*

$$Q^{dq}_{|\cdot|_{\ell^r},p,d} \leq d^{\frac{1}{r}} \cdot Q^{dq}_{|\cdot|,p,1}$$

where $|\,.\,|$ denotes standard absolute value on $\mathbb{R}$.

Since this upper bound achieves the same asymptotic rate as in the case of regular quantization (cf. Corollary 9.4 in [5]), this suggests the rate $O(d^{\frac{1}{r}})$ to be also the true one for $Q^{\mathrm{dq}}_{\|\cdot\|,p,d}$ as $d \to +\infty$.

As a step towards the above sharp rate theorem, we also need to establish a counterpart of the so-called Pierce Lemma (as stated in an operating form e.g. in [7]). In practice, it turns out to be quite useful for applications since it provides non-asymptotic error bounds which only depend on slightly higher moments of the r.v. X and the quantization level n as emphasized in [10] (see Sect. 11.4.2 for the proof).

Theorem 11.3 (d-Dimensional Extended Pierce Lemma)

(a) *Let $p, \eta > 0$. There exists a real constant $C_{d,p,\eta} > 0$ such that, for every $n \geq 1$ and every r.v. $X \in L^{p+\eta}_{\mathbb{R}^d}(\Omega, \mathscr{A}, \mathbb{P})$,*

$$\bar{d}_{n,p}(X) \leq C_{d,p,\eta}\sigma_{p+\eta,\|.\|}(X)\, n^{-1/d}$$

where $\sigma_{p+\eta,\|.\|}(X) = \inf_{a\in\mathbb{R}^d} \|X - a\|_{L^{p+\eta}}$ denotes the $L^{p+\eta}$-pseudo-standard deviation of X.

(b) *If $\mathrm{supp}(\mathbb{P}_X)$ is compact then, there exists a real constant $C'_{d,p,\eta} > 0$ such that, for every $n \geq 1$*

$$d_{n,p}(X) \leq C'_{d,p,\eta}\mathrm{diam}_{\|.\|}\big(\mathrm{supp}(\mathbb{P}_X)\big)\, n^{-1/d}.$$

11.2.2 How to Use the Extended L^p-Dual Quantization Error Modulus?

We briefly explain why the extended dual quantization error modulus, already introduced in [12] for non-compactly supported distributions, is the right tool to *perform automatically an optimized truncation* of non-compactly supported distributions. Basically, it uses its additional "outer Voronoi projection" (see (11.7)) as a *penalization term* which expands automatically the convex hull of the dually optimal grid at its appropriate "magnitude", making altogether the distribution outside of its convex hull "negligible" and sharing an optimal rate of decay $n^{-\frac{1}{d}}$ as its size n

goes to infinity. The specific choice of a Voronoi quantization among other possible solutions for this penalization is motivated by both its theoretical tractability and its simple implementability in stochastic grid optimization algorithms. This feature is of the highest importance for numerical integration or conditional expectation approximation. This is the main motivation to introduce and deeply investigate the sharp asymptotics of this L^p-mean extended dual quantization error modulus $\bar{d}_{n,p}(X)$.

We saw in [12] that *Euclidean dual quantization* of a compactly supported distribution produces *stationary* (dual) quantizers, namely r.v.s $\widehat{X}^{dual}$ satisfying $\mathbb{E}(\widehat{X}^{dual} \mid X) = X$. Hence, see Proposition 9 in [12], dual quantization based cubature formulas induce on functions $f \in \mathscr{C}^1_{\mathrm{Lip}}(\mathbb{R}^d, \mathbb{R})$ (Lipschitz functions with Lipschitz continuous gradient) an error at most equal to $[Df]_{\mathrm{Lip}} d_{2,n}(X)^2$. Taking into account the rate established in Theorem 11.2(a), this yields a $O(n^{-\frac{2}{d}})$ error rate.

There is no way to extend dual quantization to unbounded r.v.s so that it preserves the above stationarity property. However, with the choice we made (nearest neighbor projection on the grid outside its convex hull), natural heuristic arguments strongly suggest that the above order $O(n^{-\frac{2}{d}})$ is still satisfied for functions in $\mathscr{C}^1_{\mathrm{Lip}}(\mathbb{R}^d, \mathbb{R})$.

We consider an unbounded Borel distribution $\mu = \mathbb{P}_X$ of an $\mathbb{R}^d$-valued r.v. X. Let Γ_n be an *Euclidean L^2-optimal extended* dual quantization grid of size n for μ (see [12] or Theorem 11.5) and $\widehat{X}^{dual}$ the resulting Γ_n-valued extended dual quantization of X. Let $C_n = \mathrm{conv}\big(\Gamma_n\big)$. It is clear by construction of $\widehat{X}^{dual}$ that $\widehat{X}^{dual} = \widetilde{X}^{dual} + \widetilde{X}^{vor}$ where, with obvious notations,

$$\mathbf{1}_{\{X \in C_n\}} \mathbb{E}\left(\widetilde{X}^{dual} | X\right) = \mathbf{1}_{\{X \in C_n\}} X \ \text{ (dual stationarity) and } \ \widetilde{X}^{vor} = \mathrm{Proj}_{\Gamma_n \cap C_n}(X).$$

Hence, if $f \in \mathscr{C}^1_{\mathrm{Lip}}(\mathbb{R}^d, \mathbb{R})$, $\mathbb{E}\left((Df(X)|X - \widetilde{X}^{dual})|X \in C_n\right) = 0$ and

$$\begin{aligned}
&\left|\mathbb{E}\left(f\big(\widetilde{X}^{dual}\big)|X \in C_n\right) - \mathbb{E}\left(f\big(X\big)|X \in C_n\right)\right| \\
&\quad = \left|\mathbb{E}\left(f\big(\widetilde{X}^{dual}\big) - f\big(X\big) - Df(X).(X - \widetilde{X}^{dual})|X \in C_n\right)\right| \\
&\quad \le [Df]_{\mathrm{Lip}} d_{n,2}\big(\Gamma_n, \widetilde{X}^{dual} \mid X \in C_n\big)^2.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|\mathbb{E}\left(f\big(\widetilde{X}^{dual}\big)\mathbf{1}_{\{X \in C_n\}}\right) - \mathbb{E}\left(f\big(X\big)\mathbf{1}_{\{X \in C_n\}}\right)\right| &\le [Df]_{\mathrm{Lip}} d_{n,2}(\widetilde{X}^{dual}, \Gamma_n)^2 / \mathbb{P}(X \in C_n) \\
&\le [Df]_{\mathrm{Lip}} \bar{d}_{n,2}(X, \Gamma_n)^2 / \mathbb{P}(X \in C_n).
\end{aligned}$$

On the other hand,

$$\left|\mathbb{E}\left(f\big(\widetilde{X}^{vor}\big)\mathbf{1}_{\{X\notin C_n\}}\right)-\mathbb{E}\left(f(X)\mathbf{1}_{\{X\notin C_n\}}\right)\right| \le [f]_{\rm Lip}e_{n,2}\big(X,\Gamma_n\big)\mathbb{P}\big(X\notin C_n\big)^{\frac{1}{2}}$$
$$\le [f]_{\rm Lip}\bar{d}_{n,2}(X)\mathbb{P}\big(X\notin C_n\big)^{\frac{1}{2}}.$$

Relying on Theorem 11.2(b), we know that, if $\mu = h.\lambda_d \overset{\perp}{+} \nu$, then

$$\bar{d}_{n,2}(X) \sim Q^{dq}_{2,|.|_{eucl}}\, \|h\|^{\frac{1}{p}}_{\frac{d}{2+d}}\, n^{-\frac{1}{d}}.$$

The "outside" contribution will be negligible compared to the "inside" one as soon as

$$\mathbb{P}\big(X\notin C_n\big) = o\Big(\bar{d}_{n,2}(X,\Gamma_n)^2\Big) = o\Big(n^{-\frac{2}{d}}\Big). \tag{11.11}$$

This condition turns out to be not very demanding and can be checked, at least heuristically, as illustrated below: if $X \overset{d}{=} \mathcal{N}(0; I_d)$, one may conjecture, taking advantage of the spherical symmetries of the normal distribution, that C_n is approximately a sphere centered at 0 with radius $\rho_n = \max_{a\in\Gamma_n}|a|$. As

$$\mathbb{P}(|X|\ge\xi) \sim V_d\,\xi^{d-2}e^{-\frac{\xi^2}{2}} \;\text{ as }\; \xi\to+\infty \quad \big(\text{with } V_d=\lambda_{d-1}\big(S_d(0,1)\big)\big).$$

Condition (11.11) is satisfied as soon as $\liminf_n \frac{\rho_n}{\sqrt{\log n}} > \frac{2}{\sqrt{d}}$ ($\ge$ if $d = 1, 2$). As an example, one must have in mind that, for optimal *Voronoi* quantization, this inequality is satisfied since (see [9]) $\lim_n \dfrac{\rho_n}{\sqrt{\log n}} = \sqrt{2(1+2/d)} > \dfrac{2}{\sqrt{d}}$. More precisely, we have

$$\mathbb{P}\big(X\notin C_n\big) \sim \kappa_d(\log n)^{\frac{d}{2}-1}n^{-1-\frac{2}{d}}$$

so that

$$\bar{d}_{n,2}(X)\mathbb{P}\big(X\notin C_n\big)^{\frac{1}{2}} = O\Big(n^{-\frac{2}{d}-\frac{1}{2}}(\log n)^{\frac{d-2}{4}}\Big).$$

Numerical experiments, not reproduced here, carried out with the above $\mathcal{N}(0; I_d)$ distribution confirm that the radius of optimal dual quantizers always achieves this asymptotics which makes the above partially heuristic reasoning very likely. Moreover, we also tested the two rates of convergence of $\mathbb{P}(X \in C_n)$ and $\bar{d}_{n,2}(X)^2$, this time on the joint distribution of the $(W_1, \sup_{t\in[0,1]} W_t)$, W standard Brownian motion which has less symmetries (see Appendix). They also confirm that the above partially heuristic reasoning is most likely true.

11.3 Dual Quantization: Background and Basic Properties

11.3.1 Definition and First Properties

In the introduction, the definitions related to Voronoi (or regular) and dual quantizations of a r.v. X defined on a probability space $(\Omega, \mathcal{S}, \mathbb{P})$ have been recalled (see (11.7)–(11.9)). The aim of this section is to come back briefly to the origin and the motivations which led us to introduce dual quantization in [12]. On the way, we will also recall several basic results on dual quantization established in [12]. First, we will assume throughout the paper that the r.v. of interest, X, is truly d-dimensional in the sense that

$$\text{aff.dim}(\text{supp}(\mathbb{P}_X)) = d.$$

Let us start by a few practical points. First note that, although all the definitions below are related to a r.v. X, in fact the error moduli of interest only depend on the distribution $P = \mathbb{P}_X$, so we will also often write $d_p(P\Gamma)$ for $d_p(X, \Gamma)$ and $d_{n,p}(P)$. Furthermore, to alleviate notations, we will denote from now on F^p, d^p and $\bar{d}^p, \ldots$ instead of $(F_p)^p$, $(d_p)^p$ and $(\bar{d}_p)^p, \ldots$

Let us come back to the terminology *dual quantization*: it refers to a canonical example of the intrinsic stationary splitting operator: the dual quantization operator.

To be more precise, let $p \in [1, +\infty)$ and let $\Gamma = \{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ be a grid of size $n \geq d+1$ such that aff.dim$(\Gamma) = d$ i.e. Γ contains at least one $d+1$-tuple of affinely independent points.

The underlying idea is to "split" $\xi \in \text{conv}(\Gamma)$ across at most $d+1$ affinely independent points in Γ proportionally to its barycentric coordinates of ξ. There are usually many possible choices of such a Γ-valued $(d+1)$-tuple of affinely independent points, so we introduced in [12] a minimal inertia based criterion to select the most appropriate one, namely the function $F_p(\xi; \Gamma)$ defined for every ξ as the value of the minimization problem

$$F_p(\xi; \Gamma) = \inf_{(\lambda_1,\ldots,\lambda_n)} \left\{ \Big(\sum_{i=1}^{n} \lambda_i \|\xi - x_i\|^p \Big)^{\frac{1}{p}},\ \lambda_i \in [0,1],\ \sum_i \lambda_i \begin{bmatrix} x_i \\ 1 \end{bmatrix} = \begin{bmatrix} \xi \\ 1 \end{bmatrix} \right\}. \tag{11.12}$$

Owing to the compactness of the constraint set ($\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, $\sum_i \lambda_i x_i = \xi$), there exists at least one solution $\lambda^*(\xi)$ to the above minimization problem. Moreover, for any such solution, one shows using convex extremality arguments, that the set $I^*(\xi) := \big\{ i \in \{1, \ldots, n\} \text{ s.t. } \lambda_i^*(\xi) > 0 \big\}$ defines an affinely independent subset $\{x_i,\ i \in I^*(\xi)\}$ of Γ.

If, for every $\xi \in \mathrm{conv}(\Gamma)$, this solution is unique, the *dual quantization operator* is simply defined on $\mathrm{conv}(\Gamma)$ by

$$\forall\, \xi \in \mathrm{conv}(\Gamma),\ \forall\, \omega_0 \in \Omega_0,\ \mathscr{J}^*_\Gamma(\omega_0, \xi) = \sum_{i \in I(\xi)^*} x_i \mathbf{1}_{\{\sum_{j=1}^{i-1} \lambda^*_j(\xi) \le U(\omega_0) < \sum_{j=1}^{i} \lambda^*_j(\xi)\}} \tag{11.13}$$

where U denotes a random variable uniformly distributed over [0, 1], defined on an exogenous probability space $(\Omega_0, \mathscr{S}_0, \mathbb{P}_0)$. This operator $\mathscr{J}^*_\Gamma$ is then measurable (see [12]).

The above uniqueness assumption is not so stringent, especially for applications. Thus, in a purely Euclidean quadratic framework: $\|\,.\,\| = |\,.\,|_{\ell^2}$ (canonical Euclidean norm) and $p = 2$ and if Γ is said in "general position",[1] then $\Big\{\{\xi \text{ s.t. } I^*(\xi) = I\},\ |I| \le d+1\Big\}$ makes up a Borel partition of $\mathrm{conv}(\Gamma)$ (with possibly empty elements), known in 2-dimension as the *Delaunay triangulation* of Γ (see [14] for a precise connection with Delaunay triangulations).

In a more general framework, we refer to [12] for a construction of dual quantization operators. Such operators are splitting operators since, by construction, they satisfy the stationarity property (11.1).

The dual quantization operators $\mathscr{J}^*_\Gamma(\omega_0, \xi)$ plays the role of the nearest neighbour projections for regular Voronoi quantization. One checks that, by construction,

$$\forall\, \xi \in \mathrm{conv}(\Gamma), \quad \|\mathscr{J}^*_\Gamma(\xi) - \xi\|_{L^p(\mathbb{P}_0)} = \|F_p(\xi; \Gamma)\|_{L^p(\mathbb{P}_0)}$$

so that, as soon as $\mathrm{supp}(\mathbb{P}_X) \subset \Gamma$ (or equivalently $\mathbb{P}(X \in \mathrm{conv}(\Gamma)) = 1$),

$$d_p(X; \Gamma) = \|\mathscr{J}^*_\Gamma(X) - X\|_{L^p(\mathbb{P}_0 \otimes \mathbb{P})} = \|F_p(X; \Gamma)\|_{L^p(\mathbb{P}_0 \otimes \mathbb{P})}.$$

At this stage, it appears naturally that the second step of the optimization process consists of finding (at least) one grid which optimally "fits" (the distribution of) X for this criterion i.e. which is the solution to the second step of the optimization procedure

$$d_{n,p}(X) = \inf\Big\{\|\mathscr{J}^*_\Gamma(X) - X\|_{L^p(\mathbb{P}_0 \otimes \mathbb{P})},\ \mathscr{J}^*_\Gamma : \Omega_0 \times \mathrm{conv}(\Gamma) \to \Gamma,$$
$$\mathrm{conv}(\Gamma) \supset \mathrm{supp}(\mathbb{P}_X),\ |\Gamma| \le n\Big\}.$$

Note that if $X \in L^\infty_{\mathbb{R}^d}(\mathbb{P})$, $d_{n,p}(X) < +\infty$ if and only if $n \ge d+1$ (whereas it is identically infinite if X is not essentially bounded). The existence of an optimal grid (or dual quantizer) has been established in [12] (see Theorem 11.5 further on).

[1] No $d+2$ points of Γ lie on a sphere in $\mathbb{R}^d$.

The error modulus $d_{n,p}(X)$ can also be characterized as the *lowest L^p-mean approximation error by a r.v. having at most n values and satisfying the intrinsic stationarity property* as established in [12] (Theorem 2, precisely recalled in Theorem 11.4 below). It should be compared to the well-known property satisfied by the mean (regular) quantization error modulus $e_{n,p}(X)$, namely

$$e_{n,p}(X) = \inf\Big\{\|X - \widehat{X}\|_{L^p(\mathbb{P})},\ |\widehat{X}(\Omega)| \le n\Big\}.$$

A stochastic optimization procedure based on a stochastic gradient approach has been devised in [12] to compute optimal dual quantization grids w.r.t. various distributions (so far, uniform over $[0,1]^2$, normal, $(W_1, \sup_{t\in[0,1]} W_t)$, W standard Brownian motion, all in a quadratic Euclidean framework).

Let us conclude by recalling two results established in [12]. The first one is the characterization of dual quantization operator in terms in terms of best L^p-approximation (see [12], Theorem 2).

Theorem 11.4 *Let $X : \Omega, \mathcal{S}, \mathbb{P}) \to \mathbb{R}^d$ be a r.v. such that* $\text{aff.dim}(\text{supp}(\mathbb{P}_X)) = d$ *and let $n \in \mathbb{N}$, $n \ge d+1$. Then*

$$\begin{aligned}
d_{n,p}(X) = \inf\Big\{&\mathbb{E}\,\|X - \mathcal{J}_\Gamma(X)\|_{L^p} : \ \mathcal{J}_\Gamma : \Omega_0 \times \mathbb{R}^d \to \Gamma,\ \textit{intrinsic stationary},\\
&\text{supp}(\mathbb{P}_X) \subset \text{conv}(\Gamma),\ |\Gamma| \le n\Big\}\\
= \inf\Big\{&\mathbb{E}\,\|X - \widehat{X}\|_{L^p} : \widehat{X} : (\Omega_0 \times \Omega, \mathcal{S}_0 \otimes \mathcal{S}, \mathbb{P}_0 \otimes \mathbb{P}) \to \mathbb{R}^d,\\
&|\widehat{X}(\Omega_0 \times \Omega)| \le n,\ \mathbb{E}\,(\widehat{X}|X) = X\Big\} \le +\infty.
\end{aligned}$$

This quantity is finite if and only if $X \in L^\infty(\Omega, \mathcal{S}, \mathbb{P})$.

Finally, the following existence result for optimal dual quantizers *at level $n \in \mathbb{N}$* and the L^p-norm with $p \in (1,\infty)$ is established in [12]. Although we will not use it in our proofs, this result is recalled for the reader's convenience.

Theorem 11.5 (Existence of Optimal Quantizers) *Let $X \in L^p(\mathbb{P})$ for some $p \in (1,+\infty)$.*

(a) If $\text{supp}(\mathbb{P}_X)$ *is compact, then there exists for every $n \in \mathbb{N}$ a grid $\Gamma_n^* \subset \mathbb{R}^d$, $|\Gamma_n^*| \le n$ such that $d_p(X;\Gamma_n^*) = d_{n,p}(X)$.*
(b) If $\mathbb{P}_X$ is strongly continuous in the sense that it assigns no mass to hyperplanes of $\mathbb{R}^d$, then there exists for every $n \in \mathbb{N}$ a grid $\Gamma_n^* \subset \mathbb{R}^d$, $|\Gamma_n^*| \le n$ such that $\bar{d}_p(X;\Gamma_n^*) = \bar{d}_{n,p}(X)$.

If, furthermore $|\text{supp}(\mathbb{P}_X)| \ge n$*, then $|\Gamma_n^*| = n$.*

11.3.2 Local Properties of the Dual Quantization Functional

We establish or recall in this paragraph some first general properties of the local L^p-dual quantization functional F^p, which will be needed for the final proof of Theorem 11.2.

Proposition 11.2 *Let $\Gamma_1, \Gamma_2 \subset \mathbb{R}^d$ be finite grids and let $\xi \in \mathbb{R}^d$. Then*

$$\Gamma_1 \subset \Gamma_2 \Longrightarrow F_p(\xi;\Gamma_2) \le F_p(\xi;\Gamma_1).$$

Proof First note that the set $\left\{\lambda \in \mathbb{R}^n \mid \left[\begin{smallmatrix} x_1 & \dots & x_m \\ 1 & \dots & 1 \end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix} \xi \\ 1 \end{smallmatrix}\right]\right\}$ is clearly a compact set on which the continuous function $\lambda \mapsto \sum_{i=1}^n \lambda_i \|\xi - x_i\|^p$ attains a minimum. Assume $\Gamma_1 = \{x_1, \dots, x_m\}$ and $\Gamma_2 = \{x_1, \dots, x_m, x_{m+1}, \dots, x_n\}$. Then

$$\begin{aligned} F^p(\xi;\Gamma_2) = \min_{\substack{\lambda\in\mathbb{R}^n \\ \text{s.t.} \left[\begin{smallmatrix} x_1 & \dots & x_n \\ 1 & \dots & 1 \end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix} \xi \\ 1 \end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^n \lambda_i \|\xi - x_i\|^p &= \min_{\substack{\lambda\in\mathbb{R}^n, \lambda_{m+1}=\dots=\lambda_n=0 \\ \text{s.t.} \left[\begin{smallmatrix} x_1 & \dots & x_m \\ 1 & \dots & 1 \end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix} \xi \\ 1 \end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^m \lambda_i \|\xi - x_i\|^p \\ &\le \min_{\substack{\lambda\in\mathbb{R}^m \\ \text{s.t.} \left[\begin{smallmatrix} x_1 & \dots & x_m \\ 1 & \dots & 1 \end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix} \xi \\ 1 \end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^n \lambda_i \|\xi - x_i\|^p = F^p(\xi;\Gamma_1). \end{aligned}$$

□

We will also make use of the following three properties established in [12] (Propositions 11, 12, 13 respectively). In particular, the third claim yields a first upper bound for the asymptotics of the local L^p-dual quantization error when the size n of the grid goes to infinity.

Proposition 11.3

(a) Scalar bound: *Let $\Gamma = \{x_1, \dots, x_n\} \subset \mathbb{R}$ with $x_1 \le \dots \le x_n$. Then*

$$\forall \xi \in [x_1, x_n], \quad F^p(\xi;\Gamma) \le \max_{1\le i\le n-1} \left(\frac{x_{i+1}-x_i}{2}\right)^p.$$

(b) Local Product Quantization: *Let $\|\cdot\| = |\cdot|_{\ell^p}$ and let $\Gamma = \prod_{1\le j\le d} \Gamma_j$ for some $\Gamma_j \subset \mathbb{R}$. Then*

$$\forall\, \xi \in \mathbb{R}^d, \quad F_{p,|.|_{\ell^p}}(\xi;\Gamma) = \left(\sum_{j=1}^d F^p(\xi^j;\Gamma_j)\right)^{\frac{1}{p}}$$

and the same holds true with $\bar{F}_{p,\ell^p}$ on $\mathbb{R}^d$.

(c) Product Quantization: *Let* $C = a + L\,[0,1]^d$, $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$, $L > 0$, *be a hypercube, with edges parallel to the coordinate axis with common edge-length L. Let* Γ *be the product quantizer of size* $(m+1)^d$ *defined by*

$$\Gamma = \prod_{k=1}^{d} \Big\{ a_j + \frac{iL}{m},\ i = 0, \ldots, m \Big\}.$$

There exists a positive real constant $C_{\|.\|,p} = \sup_{|x|_{\ell^p}=1} \|x\|^p > 0$ *such that*

$$\forall\, \xi \in C, \quad F^p(\xi;\Gamma) \le d\, C_{\|\cdot\|,p} \cdot \Big(\frac{L}{2}\Big)^p \cdot m^{-p}. \tag{11.14}$$

11.4 Extended Pierce Lemma and Applications

The aim of this section is to provide a non-asymptotic "universal" upper-bound for the optimal (extended) L^p-mean dual quantization error in the spirit of [7] (see also [13]): it achieves nevertheless the optimal rate of convergence when the size n goes to infinity. Like for Voronoi quantization this upper-bound deeply relies on a random quantization argument and will be a key in the proof of the sharp rate (step 2 of the proof of Theorem 11.2). We begin with an extension of Pierce's lemma. This is one of the two main results of the paper.

11.4.1 Extended Pierce Lemma

For every integer $n \ge 1$, we define the set of "non-decreasing" n-tuples of $\mathbb{R}^n$ by

$$\mathscr{I}_n := \Big\{ (x_1, \ldots, x_n) \in \mathbb{R}^n,\ -\infty < x_1 \le x_2 \le \cdots \le x_n < +\infty \Big\}.$$

Let $(x_1, \ldots, x_n) \in \mathscr{I}_n$ (so that $\Gamma = \{x_1, \ldots, x_n\}$ has at most n elements) and let $\xi \in \mathbb{R}$. When $d = 1$, it is clear that the minimization problem (11.6) always has a unique solution when $\xi \in [x_1, x_n]$ so that, for every $\omega_0 \in \Omega_0 = [0,1]$, one has

$$\bar{\mathscr{J}}^*_{(x_1,\ldots,x_n)}(\omega_0, \xi) = \sum_{i=1}^{n-1} \Big(x_i \mathbf{1}_{\{\omega_0 \le \frac{x_{i+1}-\xi}{x_{i+1}-x_i}\}} + x_{i+1} \mathbf{1}_{\{\omega_0 \ge \frac{x_{i+1}-\xi}{x_{i+1}-x_i}\}} \Big) \mathbf{1}_{[x_i, x_{i+1})}(\xi)$$
$$+ x_1 \mathbf{1}_{(-\infty, x_1)}(\xi) + x_n \mathbf{1}_{[x_n, +\infty)}(\xi).$$

It follows from (11.7) that

$$\begin{aligned}\bar{F}_n^p(\xi, x_1, \ldots, x_n) &= \mathbb{E}_{\mathbb{P}_0}\left|\xi - \bar{\mathscr{J}}^*_{(x_1,\ldots,x_n)}(\omega_0, \xi)\right|^p \\ &= \sum_{i=1}^{n-1}\left(\frac{(x_{i+1}-\xi)^p(\xi - x_i)}{x_{i+1}-x_i} + \frac{(x_{i+1}-\xi)(\xi - x_i)^p}{x_{i+1}-x_i}\right)\mathbf{1}_{[x_i,x_{i+1})}(\xi) \\ &\quad + (x_1-\xi)^p\mathbf{1}_{(-\infty,x_1)}(\xi) + (\xi - x_n)^p\mathbf{1}_{[x_n,+\infty)}(\xi) \end{aligned} \tag{11.15}$$

(the subscript $_n$ is temporarily added to the functional $\bar{F}^p$, $\bar{F}_p$, etc, to emphasize that they are defined on $\mathscr{I}_n \times \mathbb{R}$). The functionals $\bar{F}_n^p$ share three important properties extensively used in what follows:

- *Additivity*: Let $(x_1, \ldots, x_{i_0}, \ldots, x_n) \in \mathscr{I}_n$. Then for every $\xi \in \mathbb{R}$

$$\begin{aligned}\bar{F}_n^p(\xi, x_1, \ldots, x_n) &= \bar{F}_{i_0}^p(\xi, x_1, \ldots, x_{i_0})\mathbf{1}_{(-\infty,x_{i_0})}(\xi) \\ &\quad + \bar{F}^p_{n-i_0+1}(\xi, x_{i_0}, \ldots, x_n)\mathbf{1}_{[x_{i_0},+\infty)}(\xi).\end{aligned}$$

- *Consistency and monotony*: Let $(x_1, \ldots, x_n) \in \mathscr{I}_n$ and $\widetilde{x}_i \in [x_i, x_{i+1}]$ for an $i \in \{1, \ldots, n-1\}$. For every $\xi \in \mathbb{R}$,

$$\bar{F}^p_{n+1}(\xi, x_1, \ldots, x_{i-1}, x_i, \widetilde{x}_i, x_{i+1}, \ldots, x_n) \le \bar{F}_n^p(\xi, x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n). \tag{11.16}$$

When $\xi \in [x_1, x_n]$, $\bar{F}_n^p(\xi; x_1, \ldots, x_n)$ coincides with $F^p(\xi, \{x_1, \ldots, x_n\})$ and (11.16) is a consequence of the definition of F_p as the value function of the minimization problem (11.6).

Outside, (11.16) holds as an equality since it amounts to the nearest distance of ξ to $\{x_1, x_n\}$ (or $[x_1, x_n]$). As a consequence,

$$n \longmapsto \bar{d}_{n,p}(X) = \inf_{(x_1,\ldots,x_n)\in\mathscr{I}_n} \left\|\bar{F}_{p,n}(X, x_1, \ldots, x_n)\right\|_{L^p} \text{ is non-increasing.} \tag{11.17}$$

More generally, for every fixed $x^0 \in \mathbb{R}$, both

$$n \longmapsto \inf_{(x^0,x_2,\ldots,x_n)\in\mathscr{I}_n} \left\|\bar{F}_{p,n}(X, x^0, x_2, \ldots, x_n)\right\|_{L^p} \tag{11.18}$$

and

$$n \longmapsto \inf_{(x_1,x_2,\ldots,x_{n-1},x^0)\in\mathscr{I}_n} \left\|\bar{F}_{p,n}(X, x_1, \ldots, x_{n-1}, x^0)\right\|_{L^p} \tag{11.19}$$

are non-increasing.

- *Scaling*: $\forall\,\omega \in \Omega_0, \forall\,(x_1,\dots,x_n) \in \mathcal{I}_n, \forall\,\xi \in \mathbb{R}, \forall\,\alpha \in \mathbb{R}_+, \forall\,\beta \in \mathbb{R}$,

$$(i)\quad \bar{F}_n^p(\alpha\,\xi+\beta, \alpha\,x_1+\beta,\dots,\alpha\,x_n+\beta) = \alpha\,\bar{F}_n^p(\xi, x_1,\dots,x_n),$$

$$(ii)\quad \bar{F}_n^p(\xi, x_1,\dots,x_n) = \bar{F}_n^p(-\xi, -x_n,\dots,-x_1).$$

The theorem below is the one dimensional version of Theorem 11.3(a) and a crucial step to its proof.

Theorem 11.6 *Let* $p,\ \eta > 0$. *There exists a real constant* $C_{p,\eta} > 0$ *such that, for every random variable* $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$,

$$\forall\,n \geq 1, \qquad \inf_{(x_1,\dots,x_n)\in\mathcal{I}_n} \|\bar{F}_{p,n}(X, x_1,\dots,x_n)\|_{L^p} \leq C_{p,\eta}\sigma_{p+\eta}(X)n^{-1}.$$

The proof that follows relies on a random quantization argument involving an n-sample of the Pareto(δ)-distribution on $[1,+\infty)$.

We will extensively make use of the Γ and B functions defined by $\Gamma(a) = \int_0^{+\infty} u^{a-1}e^{-u}du,\ a > 0$, and $B(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1}du,\ a,b > 0$, respectively, and satisfying $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Proof Step 1. We first assume that X is $[1,+\infty)$-valued and $n \geq 2$. Let $(Y_n)_{n\geq 1}$ be a sequence of i.i.d. Pareto(δ)-distributed random variables (with probability density $f_Y(y) = \delta y^{-\delta-1}\mathbf{1}_{\{y\geq 1\}}$) defined on a probability space $(\Omega', \mathcal{A}', \mathbb{P}')$.

Let $\delta = \delta(\eta, p) \in (0, \frac{\eta}{\lceil p \rceil})$ be chosen so that $\ell = \ell(p,\eta) = \frac{p}{\delta}$ is an integer and $\ell \geq 2$. For every $n \geq \ell(p,\eta)$, set $\widetilde{n} = n - \ell + 2 \in \mathbb{N}$, $\widetilde{n} \leq n$. It follows from the monotony properties (11.18) and (11.19) that

$$\begin{aligned}
&\inf_{(1,x_2,\dots,x_n)\in\mathcal{I}_n} \|\bar{F}_{p,n}(X, 1, x_2,\dots,x_n)\|_{L^p}\\
&\leq \inf_{(1,x_2,\dots,x_{\widetilde{n}})\in\mathcal{I}_{\widetilde{n}}} \|\bar{F}_{p,\widetilde{n}}(X, 1, x_2,\dots,x_{\widetilde{n}})\|_{L^p}\\
&\leq \|\bar{F}_{p,\widetilde{n}}(X, Y_0^{(n)}, Y_1^{(n)},\dots,Y_{\widetilde{n}-1}^{(n)})\|_{L^p(\Omega\times\Omega',\mathbb{P}\otimes\mathbb{P}')}
\end{aligned}$$

where, for every $n \geq 1$, $Y^{(n)} = (Y_1^{(n)},\dots,Y_n^{(n)})$ denotes the standard order statistics of the first n terms of the sequence $(Y_k)_{k\geq 1}$ and $Y_0^{(n)} = 1$. On the other hand, we recall (see e.g. [3]) that the joint distribution of $(Y_i^{(n)}, Y_{i+1}^{(n)})$, $1 \leq i \leq n-1$, is given by

$$\mathbb{P}'_{(Y_i^{(n)},Y_{i+1}^{(n)})}(du,dv) = \delta^2 \frac{n!}{(i-1)!(n-i-1)!}(1-u^{-\delta})^{i-1}v^{-\delta(n-i-1)}(uv)^{-\delta-1}\,du\,dv.$$

Step 2. Assume that $n \geq 3$. Since X and $(Y_1, \ldots, Y_0)$ are independent and $X \geq 1$

$$\|\bar{F}_{p,\widetilde{n}}(X, Y_0^{(n)}, Y_1^{(n)}, \ldots, Y_{\widetilde{n}-1}^{(n)})\|^p_{L^p(\Omega\times\Omega', \mathbb{P}\otimes\mathbb{P}')}$$
$$= \int_{[1,+\infty)} \|\bar{F}_{p,\widetilde{n}}(\xi, Y_0^{(n)}, Y_1^{(n)}, \ldots, Y_{\widetilde{n}-1}^{(n)})\|^p_{L^p(\Omega', \mathbb{P}')} \mathbb{P}_X(d\xi).$$

Relying on the expression (11.15) of the functional $\bar{F}_n^p$, we set for every $i = 0, \ldots, n-\ell$ and $\xi \geq 1$

$$(a)_i := \mathbb{E}\left(\frac{(Y_{i+1}^{(n)} - \xi)^p(\xi - Y_i^{(n)})}{Y_{i+1}^{(n)} - Y_i^{(n)}} \mathbf{1}_{\{Y_i^{(n)} < \xi \leq Y_{i+1}^{(n)}\}}\right),$$
$$(b)_i := \mathbb{E}\left(\frac{(Y_{i+1}^{(n)} - \xi)(\xi - Y_i^{(n)})^p}{Y_{i+1}^{(n)} - Y_i^{(n)}} \mathbf{1}_{\{Y_i^{(n)} < \xi \leq Y_{i+1}^{(n)}\}}\right)$$

and
$$(c)_{\widetilde{n}-1} := \mathbb{E}\left(\left(\xi - Y_{n-\ell+1}^{(n)}\right)^p \mathbf{1}_{\{\xi \geq Y_{n-\ell+1}^{(n)}\}}\right).$$

We will first inspect the sum $\sum_{i=0}^{n-\ell} (\star)_i$, where $\star = a, b$ successively.

Let $i \in \{1, \ldots, \tilde{n} - 1\}$. It follows from the above expression of the distribution of $(Y_i^{(n)}, Y_{i+1}^{(n)})$ that

$$(a)_i = \delta^2 \frac{n!}{(i-1)!(n-i-1)!}$$
$$\times \iint_{1 \leq u \leq \xi \leq v} \frac{(v-\xi)^p(\xi - u)}{v - u}(1 - u^{-\delta})^{i-1} v^{-\delta(n-i-1)}(uv)^{-\delta-1}\, du\, dv.$$

The change of variable $v = \xi(w + 1)$ yields

$$(a)_i = n(n-1)\binom{n-2}{i-1}\delta^2 \xi^{p-\delta(n-i)} \int_1^\xi (\xi - u)(1 - u^{-\delta})^{i-1} u^{-\delta-1} du$$
$$\times \int_0^{+\infty} \frac{w^p}{\xi(w+1) - u}(w+1)^{-\delta(n-i)-1} dw.$$

Noting that $\frac{\xi - u}{\xi(w+1)-u} \leq \frac{1}{w+1}$ then leads to

$$(a)_i \leq n(n-1)\binom{n-2}{i-1}\delta^2 n(n-1)\xi^{p-\delta(n-i)} \int_1^\xi (1 - u^{-\delta})^{i-1} u^{-\delta-1} du$$
$$\times \int_0^{+\infty} w^p (1+w)^{-\delta(n-i)-2} dw.$$

The change of variable $w = \frac{1}{y} - 1$ shows that

$$\int_0^{+\infty} w^p (1+w)^{-\delta(n-i)-2} dw = B\big(\delta(n-i) - p + 1, p+1\big)$$

whereas $\displaystyle\int_1^{\xi} (1-u^{-\delta})^{i-1} u^{-\delta-1} du = \frac{(1-\xi^{-\delta})^i}{\delta i}$. Hence

$$(a)_i \leq \delta n \binom{n-1}{i} (1-\xi^{-\delta})^i \xi^{p-\delta(n-i)} \frac{\Gamma(p+1)\Gamma(\delta(n-i)-p+1)}{\Gamma(\delta(n-i)+2)}$$

where we used the standard identity $\binom{n-1}{i} = \frac{n-1}{i}\binom{n-2}{i-1}$.

When $i = 0$, noting that the density of $Y_1^{(n)} = \min_{1\leq i\leq n} Y_i$ is $\delta n y^{-\delta n-1}\mathbf{1}_{\{y\geq 1\}}$, we get

$$\begin{aligned}
(a)_0 &= \mathbb{E}\left(\frac{(Y_1^{(n)}-\xi)^p(\xi-1)}{Y_1^{(n)}-1}\mathbf{1}_{\{1\leq\xi\leq Y_1^{(n)}\}}\right)\\
&= \delta n \int_{\xi}^{+\infty} (\xi-1)\frac{(v-\xi)^p}{v-1} v^{-\delta n-1} dv\\
&= \delta n \xi^{p-\delta n} \int_0^{+\infty} \frac{(\xi-1)}{\xi(w+1)-1} w^p (w+1)^{-\delta n-1} dw \;\; \text{(where we set } v = \xi(w+1))\\
&\leq \delta n \xi^{p-\delta n} B(\delta n - p + 1, p+1)
\end{aligned}$$

where we used in the last line that $\frac{\xi-1}{\xi(w+1)-1} \leq \frac{1}{w+1}$. As a consequence

$$\begin{aligned}
\sum_{i=0}^{n-\ell} (a)_i &\leq \delta\, n\, \Gamma(p+1) \sum_{i=0}^{n-\ell} \binom{n-1}{i} \xi^{p-\delta(n-i)} (1-\xi^{-\delta})^i \frac{\Gamma(\delta(n-i)-p+1)}{\Gamma(\delta(n-i)+2)}\\
&\leq \delta\, n\, \Gamma(p+1)\xi^p (1-\xi^{-\delta})^n \sum_{j=\ell}^{n} \binom{n-1}{j-1} (\xi^{\delta}-1)^{-j} \frac{\Gamma(\delta j - p + 1)}{\Gamma(\delta j + 2)}.
\end{aligned}$$

Now, using that for every $a > 0$, $\frac{\Gamma(x+a)}{\Gamma(x)} \sim x^a$ as $x \to +\infty$, we derive the existence of real constants $\tilde{\kappa}_{p,\delta}^{(0)}$, $\kappa_{p,\delta}^{(0)} > 0$ such that

$$\forall\, j \geq 0, \quad \frac{\Gamma(\delta j - p + 1)}{\Gamma(\delta j + 2)} \leq \tilde{\kappa}_{p,\delta}^{(0)}\, j^{-(p+1)} \leq \kappa_{p,\delta}^{(0)} \frac{j^{\lceil p\rceil - p}}{j(j+1)\cdots(j+\lceil p\rceil)}.$$

In turn, using that

$$\binom{n+\lceil p\rceil}{j+\lceil p\rceil} = \frac{(n+\lceil p\rceil)\cdots n}{(j+\lceil p\rceil)\cdots j}\binom{n-1}{j-1},$$

we finally obtain

$$\begin{aligned}\sum_{i=0}^{n-\ell}(a)_i &\le \kappa^{(0)}_{p,\delta}\, n\Gamma(p+1)\frac{\xi^p\delta(1-\xi^{-\delta})^n}{(n+\lceil p\rceil)\cdots(n+1)n}\sum_{j=\ell}^{n}\binom{n+\lceil p\rceil}{j+\lceil p\rceil}(\xi^\delta-1)^{-j}j^{\lceil p\rceil-p}\\ &\le \kappa^{(0)}_{p,\delta}\,\Gamma(p+1)\xi^p\delta(1-\xi^{-\delta})^n\frac{n^{\lceil p\rceil-p}}{(n+\lceil p\rceil)\cdots(n+1)}(\xi^\delta-1)^{\lceil p\rceil}\\ &\quad\times\Big(1+(\xi^\delta-1)^{-1}\Big)^{n+\lceil p\rceil}.\end{aligned}$$

Now, as $\xi \ge 1$,

$$(1-\xi^{-\delta})^n\xi^p(\xi^\delta-1)^{\lceil p\rceil}\Big(1+(\xi^\delta-1)^{-1}\Big)^{n+\lceil p\rceil} = (1-\xi^{-\delta})^2\xi^{p+\delta\lceil p\rceil} \le \xi^{p+\delta\lceil p\rceil}$$

so that, using again that $\xi \ge 1$ and $\delta < \frac{\eta}{\lceil p\rceil}$, we get $\xi^{p+\delta\lceil p\rceil} \le \xi^{p+\eta}$ which in turn implies

$$\sum_{i=0}^{n-\ell}(a)_i \le \kappa^{(0)}_{p,\delta}\,\delta\,\Gamma(p+1)\xi^{p+\eta}\frac{1}{n^p}.$$

Let us pass now to the second sum involving $(b)_i$. First note that, on the event $\Big\{Y_i^{(n)} \le \xi \le \dfrac{Y_i^{(n)}+Y_{i+1}^{(n)}}{2}\Big\}$ (which is clearly included in $\{Y_i^{(n)} \le \xi \le Y_{i+1}^{(n)}\}$), one has $(\xi-Y_i^{(n)})^p(Y_{i+1}^{(n)}-\xi) \le (\xi-Y_i^{(n)})(Y_{i+1}^{(n)}-\xi)^p$ so that, owing to what precedes, we can focus on $\sum_{i=0}^{n-\ell}(\widetilde{b})_i$ where

$$\begin{aligned}(\widetilde{b})_i &:= \mathbb{E}\left((\xi-Y_i^{(n)})^p\mathbf{1}_{\left\{\frac{Y_i^{(n)}+Y_{i+1}^{(n)}}{2}\le\xi\le Y_{i+1}^{(n)}\right\}}\right)\\ &\ge \mathbb{E}\left(\frac{(Y_{i+1}^{(n)}-\xi)(\xi-Y_i^{(n)})^p}{Y_{i+1}^{(n)}-Y_i^{(n)}}\mathbf{1}_{\left\{\frac{Y_i^{(n)}+Y_{i+1}^{(n)}}{2}\le\xi\le Y_{i+1}^{(n)}\right\}}\right).\end{aligned}$$

This time we will analyze successively the sum over $i = 1,\dots,n-\ell$ and the case $i = 0$.

$$\begin{aligned}
\sum_{i=1}^{n-\ell}(\widetilde{b})_i &= \delta^2 n(n-1)\iint_{\{1\le u\le\xi\le v\le 2\xi-u\}} du\,dv\,(uv)^{-\delta-1}(\xi-u)^p\sum_{i=1}^{n-\ell}\binom{n-2}{i-1}v^{-\delta(n-2-(i-1))}(1-u^{-\delta})^{i-1}\\
&\le \delta^2 n(n-1)\iint_{\{1\le u\le\xi\le v\le 2\xi-u\}} du\,dv(uv)^{-\delta-1}(\xi-u)^p(1-u^{-\delta}+v^{-\delta})^{n-2}\\
&\le \delta^2 n(n-1)\int_1^{\xi} du\,u^{-\delta-1}(\xi-u)^p\int_{\xi}^{2\xi-u} dv\,v^{-\delta-1}e^{-(n-2)(u^{-\delta}-v^{-\delta})}\\
&= \delta^2 n(n-1)\int_1^{\xi} du\,u^{-\delta-1}(\xi-u)^p e^{-(n-2)u^{-\delta}}\int_{\xi}^{2\xi-u} dv\,v^{-\delta-1}e^{(n-2)v^{-\delta}}
\end{aligned}$$

where we used in the in the second line that $n-\ell-1\le n-2$ since $\ell\ge 1$. Setting $v=y^{-\frac{1}{\delta}}$ yields

$$\begin{aligned}
\int_{\xi}^{2\xi-u} v^{-\delta-1}e^{(n-2)v^{-\delta}}\,dv &= \frac{1}{\delta}\int_{(2\xi-u)^{-\delta}}^{\xi^{-\delta}} e^{(n-2)y}dy\\
&\le \frac{1}{\delta}\big(\xi^{-\delta}-(2\xi-u)^{-\delta}\big)e^{(n-2)\xi^{-\delta}}\\
&\le (\xi-u)\xi^{-\delta-1}e^{(n-2)\xi^{-\delta}}
\end{aligned}$$

where we used in the last line the fundamental formula of Calculus. Consequently,

$$\begin{aligned}
\sum_{i=1}^{n-\ell}(\widetilde{b})_i &\le n(n-1)\delta^2\xi^{-\delta-1}\int_1^{\xi} u^{-\delta-1}(\xi-u)^{p+1}e^{-(n-2)(u^{-\delta}-\xi^{-\delta})}du\\
&= n(n-1)\xi^{-\delta-1}\delta\int_0^{(n-2)(1-\xi^{-\delta})}\Big(\xi-\big(\frac{x}{n-2}+\xi^{-\delta}\big)^{-\frac{1}{\delta}}\Big)^{p+1}e^{-x}\frac{dx}{n-2}
\end{aligned}$$

where we put $u=\big(\frac{x}{n-2}+\xi^{-\delta}\big)^{-\frac{1}{\delta}}$. Now, applying again the fundamental formula of Calculus to the function $z^{-\frac{1}{\delta}}$ yields,

$$\xi-\Big(\frac{x}{n-2}+\xi^{-\delta}\Big)^{-\frac{1}{\delta}} = (\xi^{-\delta})^{-\frac{1}{\delta}}-\Big(\frac{x}{n-2}+\xi^{-\delta}\Big)^{-\frac{1}{\delta}} \le \frac{x}{\delta(n-2)}\xi^{\delta+1}$$

so that
$$\begin{aligned}
\sum_{i=1}^{n-\ell}(\widetilde{b})_i &\le \frac{n(n-1)}{(n-2)^{p+2}}\delta^{-p}\xi^{(p+1)(\delta+1)-(\delta+1)}\int_0^{(n-2)(1-\xi^{-\delta})} x^{p+1}e^{-x}dx\\
&\le \kappa_{p,\delta}^{(1)}\Gamma(p+2)n^{-p}\xi^{p(\delta+1)}
\end{aligned}$$

for some constant $\kappa_{p,\delta}^{(1)}>0$.

When $i = 0$, keeping in mind that $Y_1^{(n)} = \min_{1\le i\le n} Y_i$, we get

$$\begin{aligned}(\widetilde{b})_0 &\le (\xi-1)^p \mathbb{P}(\xi \le Y_1^{(n)} \le 2\xi - 1) = (\xi-1)^p\big(\xi^{-n\delta} - (2\xi-1)^{-n\delta}\big)\\ &\le n\delta(\xi-1)^{p+1}\xi^{-n\delta-1} = n\delta\xi^{p(1+\delta)}g(1/\xi)\end{aligned}$$

where $g(u) = (1-u)^{p+1}u^{(n+p)\delta}$, $u \in (0,1)$. One checks that g attains its maximum over $(0,1]$ at $u^* = \frac{(n+p)\delta}{(n+p)\delta+p+1}$ so that

$$\sup_{u\in(0,1]} g(u) = g(u^*) = \left(\frac{p+1}{(n+p)\delta+p+1}\right)^{p+1}(u^*)^{(n+p)\delta} \le \left(\frac{1}{1+\frac{n+p}{p+1}\delta}\right)^{p+1}.$$

Finally, there exists a real constant $\kappa_{p,\delta}^{(2)} > 0$ such that

$$(\widetilde{b})_0 \le \xi^{p(\delta+1)}\frac{\delta n}{(1+\frac{n+p}{p+1}\delta)^{p+1}} \le \kappa_{p,\delta}^{(2)}\xi^{p(\delta+1)}n^{-p}.$$

As concerns the $(c)_{n-\ell+1}$ term, we proceed as follows.

$$\begin{aligned}\mathbb{E}\left(\big(\xi - Y_{n-\ell+1}^{(n)}\big)^p \mathbf{1}_{\{\xi\ge Y_{n-\ell+1}^{(n)}\}}\right) &\le \xi^p\mathbb{P}(\xi \ge Y_{n-\ell+1}^{(n)})\\ &\le \xi^{p(1+\delta)}\mathbb{E}\,(Y_{n-\ell+1}^{(n)})^{-p\delta}.\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}\,(Y_{n-\ell+1}^{(n)})^{-p\delta} &= \frac{\Gamma(n+1)}{\Gamma(n-\ell+1)\Gamma(\ell)}\int_0^1 (1-v)^{n-\ell}v^{\ell+p-1}dv\\ &= \frac{\Gamma(n+1)}{\Gamma(\ell)}\frac{\Gamma(\ell+p)}{\Gamma(n+p+1)}\\ &\sim \frac{\Gamma(\ell+p)}{\Gamma(\ell)}n^{-p} = O(n^{-p}).\end{aligned}$$

Finally, for every $\xi \ge 1$,

$$(c)_{n-\ell+1} \le \kappa_{p,\delta}^{(3)}\xi^{p(1+\delta)}n^{-p}.$$

Consequently, there exists a real constant $\kappa_{p,\eta} = \max_{j=0,\dots,3}\kappa_{p,\delta}^{(j)} > 0$ such that, for every $n \ge n_{p,\eta} = \ell(\eta,p)\vee 3$,

$$\forall\,\xi \ge 1, \qquad n^p\,\mathbb{E}\,\bar{F}_n^p(\xi, Y_0^{(n)},\dots,Y_{\tilde{n}+1}^{(n)}) \le \kappa_{p,\eta}\,\xi^{p+\eta}$$

since $p\,\delta \le \eta$. Hence for every r.v. X, we derive by integrating in $\xi \in [1, +\infty)$ with respect to $\mathbb{P}_X(d\xi)$:

$$n^p \inf_{(1,x_2,\ldots,x_n)\in\mathscr{I}_n} \mathbb{E}\,\bar{F}_n^p(X, 1, x_2, \ldots, x_n) \le n^p \mathbb{E}\,\bar{F}_n^p(X, Y_0^{(n)}, \ldots, Y_{\widetilde{n}+1}^{(n)})$$
$$\le \kappa_{p,\eta}\,\mathbb{E}\;X^{p+\eta}.$$

Step 3. If X is a non-negative random variable, applying the second step to $X+1$ and using the scaling property (i) satisfied by $F_{p,n}$ yields for $n \ge n_{p,\eta}$ (as defined in Step 2),

$$\begin{aligned}\inf_{(0,x_2,\ldots,x_n)\in\mathscr{I}_n} \|\bar{F}_{p,n}(X, 0, x_2, \ldots, x_n)\|_{L^p} &= \inf_{(1,x_2,\ldots,x_n)\in\mathscr{I}_n} \|\bar{F}_{p,n}(X+1, 1, \ldots, x_n)\|_{L^p}\\ &\le \kappa_{p,\eta}^{1/p} \frac{\|1+X\|_{L^{p+\eta}}^{1+\frac{\eta}{p}}}{n}\\ &\le C_{p,\eta}^{(0)} \frac{1+\|X\|_{L^{p+\eta}}^{1+\frac{\eta}{p}}}{n}\end{aligned}$$

with $C_{p,\eta}^{(0)} = (2^{1+\eta}\kappa_{p,\eta})^{\frac{1}{p}}$.

We may assume that $\|X\|_{L^{p+\eta}} \in (0, \infty)$. Then, applying the above bound to the non-negative random variable $\widetilde{X} = \frac{X}{\|X\|_{L^{p+\eta}}}$ and taking again advantage of the scaling property (i), we obtain

$$\begin{aligned}&\inf_{(0,x_2,\ldots,x_n)\in\mathscr{I}_n} \left\|\bar{F}_{p,n}(X, 0, x_2, \ldots, x_n)\right\|_{L^p}\\ &= \|X\|_{L^{p+\eta}} \inf_{(0,x_2,\ldots,x_n)\in\mathscr{I}_n} \left\|\bar{F}_{p,n}(\widetilde{X}, 0, x_2, \ldots, x_n)\right\|_{L^p}\\ &\le \|X\|_{L^{p+\eta}} C_{p,\eta}^{(0)} \frac{1+1}{n} = 2\,C_{p,\eta}^{(0)}\,\|X\|_{L^{p+\eta}} \frac{1}{n}.\end{aligned}$$

Step 4. Let X be a real-valued random variable and let, for every integer $n \ge 1$, $x_1, \ldots, x_n \in (-\infty, 0)$, $x_{n+1} = 0$ and $x_{n+2}, \ldots, x_{2n+1} \in (0, +\infty)$. It follows from the additivity property that

$$\begin{aligned}\bar{F}_{2n+1}^p(X, x_1, \ldots, x_{2n+1}) &= \bar{F}_{n+1}^p(X_+, x_{n+1}, \ldots, x_{2n+1})\mathbf{1}_{\{X\ge 0\}}\\ &\quad + \bar{F}_{n+1}^p(-X_-, x_1, \ldots, x_{n+1})\mathbf{1}_{\{X<0\}}\\ &= \bar{F}_{n+1}^p(X_+, x_1, \ldots, x_{n+1})\mathbf{1}_{\{X\ge 0\}}\\ &\quad + \bar{F}_{n+1}^p(X_-, -x_{n+1}, \ldots, -x_1)\mathbf{1}_{\{X<0\}}.\end{aligned}$$

Consequently, using that $X_+ \times X_- \equiv 0$ and that $x_{n+1} = 0$, we get

$$\inf_{\substack{(x_1,\ldots,x_{2n+1})\in\mathscr{I}_{2n+1}\\ x_{n+1}=0}} \big\|\bar{F}_{p,2n+1}(X, x_1, \ldots, x_{2n+1})\big\|_{L^p}^p$$

$$\le \inf_{(0,y_2,\ldots,y_{n+1})\in\mathscr{I}_{n+1}} \big\|\bar{F}_{p,n+1}(X_+, 0, y_2, \ldots, y_{n+1})\big\|_{L^p}^p$$

$$+ \inf_{(0,y_2,\ldots,y_{n+1})\in\mathscr{I}_{n+1}} \big\|\bar{F}_{p,n}(X_-, 0, y_2, \ldots, y_{n+1})\big\|_{L^p}^p.$$

Hence, it follows from Step 2 that, for every $n \ge n_{p,\eta} - 1$,

$$\inf_{(x_1,\ldots,x_{2n+1})\in\mathscr{I}_{2n+1}} \big\|\bar{F}_{p,2n+1}(X, x_1, \ldots, x_{2n+1})\big\|_{L^p}^p$$

$$\le \left(\|X_-\|_{L^{p+\eta}}^p + \|X_+\|_{L^{p+\eta}}^p\right)\left(\frac{2C_{p,\eta}^{(0)}}{n+1}\right)^p.$$

Now, using that $(a+b) \le 2^{1-\frac{1}{q}}(a^q + b^q)^{\frac{1}{q}}$, $a, b \ge 0$, with $q = 1 + \frac{\eta}{p} \ge 1$, we derive that

$$\|X_-\|_{L^{p+\eta}}^p + \|X_+\|_{L^{p+\eta}}^p \le 2^{\frac{\eta}{p+\eta}}\left(\|X_-\|_{L^{p+\eta}}^{p+\eta} + \|X_+\|_{L^{p+\eta}}^{p+\eta}\right)^{\frac{p}{p+\eta}} = 2^{\frac{\eta}{p+\eta}}\|X\|_{L^{p+\eta}}^p$$

since $X_- \times X_+ \equiv 0$. Now, the monotonicity property (11.17) implies that, for every $n \ge 2\,n_{p,\eta}$,

$$\bar{d}_{n,p}(X) = \inf_{(x_1,\ldots,x_n)\in\mathscr{I}_n} \|\bar{F}_{p,n}(X, x_1, \ldots, x_n)\|_{L^p} \le 2^{\frac{\eta}{p(p+\eta)}} 2C_{p,\eta}^{(0)} \frac{\|X\|_{L^{p+\eta}}}{n}.$$

Still calling upon (11.17), we note that, for every $n \in \{1, \ldots, 2n_{p,\eta}\}$, $\bar{d}_{n,p}(X) \le \bar{d}_{1,p}(X) = \inf_{x\in\mathbb{R}} \|X - x_1\|_{L^p} \le \|X\|_{L^p}$ so that

$$\bar{d}_{n,p}(X) \le 2n_{p,\eta}\frac{\|X\|_{L^{p+\eta}}}{n}.$$

Hence, setting $C_{p,\eta} = \max\left(2n_{p,\eta}, 2^{1+\frac{\eta}{p(p+\eta)}} C_{p,\eta}^{(0)}\right)$, yields

$$\forall\, n \ge 1, \quad \bar{d}_{n,p}(X) \le C_{p,\eta}\frac{\|X\|_{L^{p+\eta}}}{n}.$$

Now using that the function $\bar{F}$ is invariant by translation (property "Scaling (i)") combined with the fact that the real constant $C_{p,\eta}$ does not depend on the (distribution of the) random variable X implies that, for every $a \in \mathbb{R}$, $\bar{d}_{n,p}(X) = \bar{d}_{n,p}(X-a)$ so that

$$\bar{d}_{n,p}(X) \leq C_{p,\eta} \frac{\inf_{a\in\mathbb{R}} \|X-a\|_{L^{p+\eta}}}{n} = C_{p,\eta} \frac{\sigma_{p+\eta}(X)}{n}$$

which completes the proof. □

11.4.2 A d-Dimensional Non-asymptotic Upper-Bound for the Dual Quantization Error

Now, combining Theorem 11.6 and Proposition 11.3(b), we are in position to show Theorem 11.3 (the d-dimensional version of the extended Pierce Lemma) which provides a non-asymptotic upper-bound at the exact rate for dual quantization error moduli.

Proof of Theorem 11.3

(a) First note that $\bar{d}_{n,p}(X) = \bar{d}_{n,p}(X-a)$, $a \in \mathbb{R}^d$ (invariance by translation), so we may assume that X is $L^{p+\eta}$-centered i.e. $\sigma_{p+\eta,\|.\|}(X) = \|X\|_{L^{p+\eta}}$. When $d=1$, Theorem 11.6 solves the problem.

Let $d \geq 2$. Let $X = (X^1,\ldots,X^d)$ (X^i components of X). It follows form Proposition 11.3 that, if $\Gamma = \prod_{1\leq i\leq d} \Gamma_i$, with $\Gamma_i \subset \mathbb{R}$, $|\Gamma_i| = n_i$ with $n_1 \cdots n_d \leq n$. Then for every $\xi = (\xi^1,\ldots,\xi^d) \in \mathbb{R}^d$

$$\bar{F}^p_{\|.\|}(\xi;\Gamma) \leq C_{p,\|.\|} \bar{F}^p_{\ell^p}(\xi;\Gamma) = \sum_{j=1}^d \bar{F}^p(\xi^j;\Gamma_j)$$

where $C_{p,\|.\|} = \sup_{|\xi|_{\ell^p}=1} \|\xi\|^p$. Integrating with respect to the distribution of X yields $\bar{d}^p(X;\Gamma) \leq C_{p,\|.\|} \sum_{j=1}^d \bar{d}^p(X^j;\Gamma_j)$ which in turn easily implies

$$\bar{d}^p_n(X) \leq C_{p,\|.\|} \sum_{j=1}^d \bar{d}^p_{n_j}(X^j).$$

Now set $n_j = \lfloor n^{\frac{1}{d}} \rfloor$, $j = 1, \ldots, d$. It follows from Theorem 11.6 that

$$
\begin{aligned}
\bar{d}_n^p(X) &\le C_{p,\|.\|}^p C_{p,\eta} \sum_{j=1}^d \|X^j\|_{L^{p+\eta}}^p \lfloor n^{\frac{1}{d}} \rfloor^{-p} \\
&\le C_{p,\|.\|} C_{p,\eta} \sup_{k\ge 2} \Big(\frac{k^{\frac{1}{d}}}{k^{\frac{1}{d}} - 1} \Big)^p n^{-\frac{p}{d}} \sum_{j=1}^d \|X^j\|_{L^{p+\eta}}^p \\
&\le C_{p,\|.\|} C_{p,\eta} \left(\frac{2^{\frac{1}{d}}}{2^{\frac{1}{d}} - 1} \right)^p n^{-\frac{p}{d}} d^{\frac{\eta}{p+\eta}} \mathbb{E}\,|X|_{\ell^{p+\eta}}^{p+\eta} \\
&\le d^{\frac{\eta}{p+\eta}} C_{p,\|.\|} C_{p,\eta} \left(\frac{2^{\frac{1}{d}}}{2^{\frac{1}{d}} - 1} \right)^p \widetilde{C}_{\|.\|,p+\eta} \|X\|_{L^{p+\eta}}^{p+\eta} n^{-\frac{p}{d}}
\end{aligned}
$$

where $\widetilde{C}_{\|.\|,r} = \sup_{\|x\|=1} |x|_{\ell^r}^r$, $r > 0$.

(b) Let C be the smallest hypercube withe edges parallel to the coordinate axis containing conv(supp($\mathbb{P}_X$)). Up to a translation, which leaves $d_{n,p}(X)$ invariant, we may assume that $C = [0, L]^d$ where $0 \le L \le \mathrm{diam}_{\|.\|}(\mathrm{supp}(\mathbb{P}_X))$. The conclusion follows is obtained by following the lines of the proof of claim (a) once Inequality (11.14) is integrated with respect to $\mathbb{P}_X(d\xi)$ with $m = \lfloor n^{\frac{1}{d}} \rfloor$. □

11.5 Proof of the Sharp Rate Theorem

On the way to proving the sharp rate theorem, we have to establish few additional propositions.

Proposition 11.4 (Sub-linearity) *Let $P = \sum_{i=1}^m s_i P_i$ where $s_1, \ldots, s_m \in [0, 1]$, $\sum_{i=1}^m s_i = 1$ and let $n_1, \ldots, n_m \in \mathbb{N}$ such that $\sum_{i=1}^m n_i \le n$. Then,*

$$d_n^p(P) \le \sum_{i=1}^m s_i\, d_{n_i}^p(P_i).$$

Proof For $\varepsilon > 0$ and every $i = 1, \ldots, m$, choose $\Gamma_i \subset \mathbb{R}^d$, $|\Gamma_i| \le n_i$ such that

$$d^p(P_i; \Gamma_i) \le (1 + \varepsilon)\, d_{n_i}^p(P_i).$$

Set $\Gamma = \bigcup_{i=1}^{m} \Gamma_i$; from Proposition 11.2, we get

$$d_n^p(P) \le d_n^p(P;\Gamma) = \sum_{i=1}^{m} s_i \int F^p(\xi;\Gamma)\, P_i(d\xi)$$

$$\le \sum_{i=1}^{m} s_i \int F^p(\xi;\Gamma_i)\, P_i(d\xi) \le (1+\varepsilon)\sum_{i=1}^{m} s_i\, d_{n_i}^p(P_i).$$

Letting $\varepsilon \to 0$ completes the proof. □

Remark 11.1 Proposition 11.4 does not hold for $\bar{d}_n^p$ since $\bar{F}^p$ is not decreasing for the inclusion order on grids. This induces substantial difficulties in the proof of the sharp rate compared to the regular quantization setting.

Proposition 11.5 (Scaling Property) *Let $C = a + \rho[0,1]^d$ ($a \in \mathbb{R}^d$, $\rho > 0$) be a d-dimensional hypercube, with edges parallel to the coordinate axis and edge-length $\rho > 0$. Then, if $\mathscr{U}(C)$ denotes the uniform distribution over C, one has*

$$d_{n,p}(\mathscr{U}(C)) = \rho \cdot d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Proof Keeping in mind that $\lambda_d([0,\rho]^d) = \rho^d$, one derives that

$$d^p(\mathscr{U}(C);\{a+\rho x_1,\dots,a+\rho x_n\}) = \int_{[0,\rho]^d} \min_{\substack{\lambda\in\mathbb{R}^n \\ \text{s.t.} \left[\begin{smallmatrix}\rho x_1 & \dots & \rho x_n \\ 1 & \dots & 1\end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix}\xi \\ 1\end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^{n} \lambda_i \,\|\xi - \rho x_i\|^p \,\frac{\lambda_d(d\xi)}{\lambda_d\big([0,\rho]^d\big)}$$

$$= \int_{[0,1]^d} \min_{\substack{\lambda\in\mathbb{R}^n \\ \text{s.t.} \left[\begin{smallmatrix}\rho x_1 & \dots & \rho x_n \\ 1 & \dots & 1\end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix}\rho u \\ 1\end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^{n} \lambda_i \,\|\rho u - \rho x_i\|^p \,\lambda_d(du)$$

$$= \rho^p \int_{[0,1]^d} \min_{\substack{\lambda\in\mathbb{R}^n \\ \text{s.t.} \left[\begin{smallmatrix}x_1 & \dots & x_n \\ 1 & \dots & 1\end{smallmatrix}\right]\lambda = \left[\begin{smallmatrix}u \\ 1\end{smallmatrix}\right],\, \lambda\ge 0}} \sum_{i=1}^{n} \lambda_i \,\|u - x_i\|^p \,\lambda_d(du)$$

$$= \rho^p \cdot d^p\big(\mathscr{U}\big([0,1]^d\big);\{x_1,\dots,x_n\}\big).$$ □

Next lemma shows that it is also true for $\bar{d}_{n,p}$ that the convex hull spanned by a sequence of quantizers such that $\bar{d}_{n,p}(\mathbb{P}_X,\Gamma_n) \to 0$ asymptotically covers the interior of $\operatorname{supp}(\mathbb{P}_X)$. (This fact is trivial for $d_{n,p}$ if X has a compact support.)

Lemma 11.1 *Let $K = \operatorname{conv}\{a_1,\dots,a_k\} \subset \overbrace{\operatorname{supp}(\mathbb{P}_X)}^{\circ}$ be a set with $\mathring{K} \ne \emptyset$ and let Γ_n be a sequence of quantizers such that $\bar{d}_{n,p}(\mathbb{P}_X,\Gamma_n) \to 0$ as $n \to +\infty$. Then,* *there exists* $n_0 \in \mathbb{N}$ *such that, for all* $n \geq n_0$,

$$K \subset \mathrm{conv}(\Gamma_n).$$

Proof Set $a_0 = \frac{1}{k}\sum_{i=1}^{k} a_i$ and define for every $\rho > 0$

$$\widetilde{K}(\rho) = \mathrm{conv}\{\widetilde{a}_1(\rho), \ldots, \widetilde{a}_k(\rho)\} \quad \text{with} \quad \widetilde{a}_i(\rho) = a_0 + (1+\rho)(a_i - a_0).$$

Since $K \subset \overbrace{\mathrm{supp}(\mathbb{P}_X)}^{\circ}$, there exists $\rho_0 > 0$ such that $\widetilde{K} = \widetilde{K}(\rho_0) \subset \mathrm{supp}(\mathbb{P}_X)$. From now on, we denote $\widetilde{a}_i(\rho_0)$ by $\widetilde{a}_i$. Since moreover $\widetilde{a}_i \in \mathrm{supp}(\mathbb{P})$, there exists a sequence $(a_i^n)_{n\geq 1}$ having values in $\mathrm{conv}(\Gamma_n)$ and converging to $\widetilde{a}_i$. Otherwise, there would exist $\varepsilon_0 > 0$ and a subsequence (n') such that $B(\widetilde{a}_i, \varepsilon_0) \subset (\mathrm{conv}(\Gamma_{n'}))^c$. Then

$$\bar{d}_{n'}^{p}(X, \Gamma_{n'}) \geq \mathbb{E}\left[\mathrm{dist}(X, \Gamma_{n'})^p \mathbf{1}_{\{X \in B(\widetilde{a}_i, \varepsilon_0/2)\}}\right] \geq \left(\frac{\varepsilon_0}{2}\right)^p \mathbb{P}\big(X \in B(\widetilde{a}_i, \varepsilon_0/2)\big) > 0$$

since $\widetilde{a}_i \in \mathrm{supp}(\mathbb{P}_X)$. This contradicts the assumption on the sequence $(\Gamma_n)_{n\geq 1}$.

As K has a nonempty interior, $\mathrm{aff.\,dim}\{a_1, \ldots, a_k\} = \mathrm{aff.\,dim}\{\widetilde{a}_1, \ldots, \widetilde{a}_k\} = d$. Consequently, we may choose a subset $I^* \subset \{1, \ldots, k\}$, $|I^*| = d+1$, so that $\{\widetilde{a}_j : j \in I^*\}$ is an affinely independent system in $\mathbb{R}^d$ and, furthermore, there exists $n_0 \in \mathbb{N}$ such that the same holds for $\{a_j^n : j \in I^*\}$, $n \geq n_0$. Hence, we may write for $n \geq n_0$

$$\widetilde{a}_i = \sum_{j \in I^*} \mu_j^{n,i} a_j^n, \quad \sum_{j \in I^*} \mu_j^{n,i} = 1, \quad i = 1, \ldots, k. \tag{11.20}$$

This linear system has the unique asymptotic solution $\mu_j^{\infty,i} = \delta_{ij}$ (Kronecker symbol), which implies $\mu_j^{n,i} \to \delta_{ij}$ when $n \to +\infty$.

Now let $\xi \in K \subset \widetilde{K}$ and write

$$\xi = \sum_{i=1}^{k} \lambda_i a_i \quad \text{for some} \quad \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1.$$

One easily checks that it also holds

$$\xi = \sum_{i=1}^{k} \widetilde{\lambda}_i \widetilde{a}_i \ \text{ with } \widetilde{\lambda}_i = \frac{\rho_0}{k(1+\rho_0)} + \frac{\lambda_i}{1+\rho_0} \geq \frac{\rho_0}{k(1+\rho_0)} > 0 \quad \text{and} \quad \sum_{i=1}^{k} \widetilde{\lambda}_i = 1.$$

Furthermore, we may choose $n_1 \geq n_0$ such that, for every $n \geq n_1$,

$$\mu_i^{n,i} > \frac{1}{2} \quad \text{and} \quad \forall j \neq i,\ |\mu_j^{n,i}| \leq \frac{\rho_0}{4k(1+\rho_0)}.$$

Using (11.20), this leads to

$$\xi = \sum_{j \in I^*} \Big(\sum_{i=1}^{k} \widetilde{\lambda}_i \mu_j^{n,i} \Big) a_j^n$$

and

$$\sum_{i=1}^{k} \widetilde{\lambda}_i \mu_j^{n,i} > \widetilde{\lambda}_j \mu_j^{n,j} - \sum_{i=1, i \neq j}^{k} \widetilde{\lambda}_i |\mu_j^{n,i}| > \frac{\rho_0}{2k(1+\rho_0)} - \frac{\rho_0}{4k(1+\rho_0)} = \frac{\rho_0}{4k(1+\rho_0)} > 0$$

for every $j \in I^*$. Finally, one completes the proof by noting that

$$\sum_{j \in I^*} \sum_{i=1}^{k} \widetilde{\lambda}_i \mu_j^{n,i} = \sum_{i=1}^{k} \widetilde{\lambda}_i \sum_{j \in I^*} \mu_j^{n,i} = 1. \qquad \square$$

As already mentioned, Proposition 11.4 does not hold anymore for $\bar{d}_{n,p}$. As a consequence, we have to establish a "firewall Lemma", which will be a useful tool to overcome this problem in the non-compact setting.

Lemma 11.2 (Firewall) *Let $K \subset \mathbb{R}^d$ be compact and convex with $\mathring{K} \neq \varnothing$. Moreover, let $\varepsilon > 0$ be small enough so that*

$$K_\varepsilon = \Big\{ x \in K : \mathrm{dist}_{\ell^\infty}(x, K^c) \geq \varepsilon \Big\} \neq \emptyset.$$

Let $\Gamma_{\alpha,\varepsilon}$ be a subset of the lattice $\alpha\mathbb{Z}^d$ with edge-length $\alpha > 0$ satisfying

$$K \setminus K_\varepsilon \subset \mathrm{conv}(\Gamma_{\alpha,\varepsilon}) \text{ and } \forall\, x \in K \setminus K_\varepsilon,\ \mathrm{dist}_{\|\cdot\|}(x, \Gamma_{\alpha,\varepsilon}) \leq C_{\|\cdot\|}\alpha$$

where $C_{\|\cdot\|} > 0$ is a real constant only depending on the norm $\|\cdot\|$.

Then, for every grid $\Gamma \subset \mathbb{R}^d$ containing K and every $\eta \in (0, 1)$, it holds

$$\forall\, \xi \in K_\varepsilon,\ F^p(\xi; \Gamma) \geq \frac{1}{(1+\eta)^{p+d+1}} F^p(\xi; (\Gamma \cap \mathring{K}) \cup \Gamma_{\alpha,\varepsilon}) - (1+\eta)^{-d-1}\eta^{-p}(d+1)\, C_{\|\cdot\|}^p \alpha^p.$$

Remark 11.2 The lattice $\Gamma_{\alpha,\varepsilon}$ and its size will be carefully defined and estimated for the specified compact sets K when calling upon the firewall lemma in what follows.

Proof Let $\Gamma = \{x_1, \ldots, x_n\}$ and let $\xi \in K_\varepsilon$. Then we may choose $I = I(\xi) \subset \{1, \ldots, n\}$, $|I| \leq d+1$ such that

$$F^p(\xi; \Gamma) = \sum_{i \in I} \lambda_j \|\xi - x_i\|^p, \quad \sum_{i \in I} \lambda_i x_i = \xi,\ \lambda_i \geq 0,\ \sum_{i \in I} \lambda_i = 1.$$

If for every $x_i \in \Gamma \setminus \mathring{K}$ $\lambda_i = 0$, then $F^p(\xi;\Gamma) = F^p(\xi;\Gamma \cap \mathring{K})$ and our claim is trivial. Therefore, let $J(\xi) = \{i \,:\, x_i \in \Gamma \setminus \mathring{K},\, \lambda_i > 0\} \subset I(\xi)$ and choose one fixed $i_0 \in J(\xi)$. Let $\theta = \theta(i_0) \in (0,1)$ such that

$$\widetilde{x}_{i_0} = \xi + \theta(x_{i_0} - \xi) \in K \setminus K_\varepsilon \quad \text{and} \quad \frac{\theta^{p\wedge 1}}{\theta + \lambda_{i_0}(1-\theta)} \le 1 + \eta$$

(when $p \ge 1$, the right constraint is empty). Setting

$$\widetilde{\lambda}_i^0 = \frac{\lambda_i \theta}{\theta + \lambda_{i_0}(1-\theta)},\; i \in I \setminus \{i_0\}, \quad \widetilde{\lambda}_{i_0}^0 = \frac{\lambda_{i_0}}{\theta + \lambda_{i_0}(1-\theta)}$$

we obtain

$$\widetilde{\lambda}_{i_0}^0 \widetilde{x}_{i_0} + \sum_{i \in I\setminus\{i_0\}} \widetilde{\lambda}_i^0 x_i = \xi,\; \widetilde{\lambda}_i^0 \ge 0,\; \sum_{i\in I} \widetilde{\lambda}_i^0 = 1.$$

Consequently

$$\begin{aligned}
\widetilde{\lambda}_{i_0}^0 \|\xi - \widetilde{x}_{i_0}\|^p + \sum_{j \in I\setminus\{i_0\}} \widetilde{\lambda}_i^0 \|\xi - x_i\|^p &= \frac{\lambda_{i_0}\theta^p}{\theta + \lambda_{i_0}(1-\theta)} \|\xi - x_{i_0}\|^p \\
&\quad + \sum_{i \in I\setminus\{i_0\}} \frac{\lambda_i \theta}{\theta + \lambda_{i_0}(1-\theta)} \|\xi - x_i\|^p \\
&\le \frac{\theta^{p\wedge 1}}{\theta + \lambda_{i_0}(1-\theta)} \sum_{i\in I} \lambda_i \|\xi - x_i\|^p \\
&\le (1+\eta) \sum_{i \in I} \lambda_i \|\xi - x_i\|^p.
\end{aligned}$$

Repeating the procedure for every $i \in J(\xi)$ finally yields by induction the existence of $\widetilde{x}_i \in K \setminus K_\varepsilon$ and $\widetilde{\lambda}_i$, $i \in I$, such that

$$\sum_{i \in I : x_i \notin \mathring{K}} \widetilde{\lambda}_i \widetilde{x}_i + \sum_{i \in I : x_i \in \mathring{K}} \widetilde{\lambda}_i x_i = \xi,\; \widetilde{\lambda}_i \ge 0,\; \sum_{i \in I} \widetilde{\lambda}_i = 1$$

and

$$(1+\eta)^{|J(\xi)|} F^p(\xi;\Gamma) \ge \sum_{i \in I : x_i \notin \mathring{K}} \widetilde{\lambda}_i \|\xi - \widetilde{x}_i\|^p + \sum_{i \in I : x_i \in \mathring{K}} \widetilde{\lambda}_i \|\xi - x_i\|^p. \qquad (11.21)$$

Let us denote $\Gamma_{\alpha,\varepsilon} = \{a_1, \ldots, a_m\}$ and let $i_0 \in J(\xi)$ be such that $\widetilde{x}_{i_0}$ is a "modified" x_{i_0} (originally lying in $\Gamma \setminus \mathring{K}$). By construction $\widetilde{x}_{i_0} \in K \setminus K_\varepsilon \subset$

$\mathrm{conv}(\Gamma_{\alpha,\varepsilon})$ and there is $J_{i_0} \subset \{1,\ldots,m\}$ such that

$$F^p(\widetilde{x}_{i_0};\Gamma_{\alpha,\varepsilon}) = \sum_{j\in J_{i_0}} \mu_j^{i_0}\|\widetilde{x}_{i_0}-a_j\|^p,\ \sum_{j\in J_{i_0}} \mu_j^{i_0} x_j = \widetilde{x}_{i_0},\ \mu_j^{i_0}\geq 0,\ \sum_{j\in J_{i_0}} \mu_j^{i_0} = 1$$

and

$$\forall\, j\in J_{i_0},\quad \|\widetilde{x}_{i_0}-a_j\| \leq C_{\|\cdot\|}\,\alpha.$$

Using the elementary inequality

$$\forall\, p>0,\ \forall \eta>0,\ \forall\, u,v\geq 0,\quad (u+v)^p \leq (1+\eta)^p u^p + \Big(1+\frac{1}{\eta}\Big)^p v^p,$$

we derive that for every $j\in J_{i_0}$

$$\|\xi-a_j\|^p \leq \big(\|\xi-\widetilde{x}_{i_0}\|+\|\widetilde{x}_{i_0}-a_j\|\big)^p \leq (1+\eta)^p\|\xi-\widetilde{x}_{i_0}\|^p + \Big(1+\frac{1}{\eta}\Big)^p C^p_{\|\cdot\|}\,\alpha^p.$$

As a consequence,

$$\sum_{j\in J_{i_0}} \mu_j^{i_0}\|\xi-a_j\|^p \leq (1+\eta)^p\Big(\|\xi-\widetilde{x}_{i_0}\|^p + \eta^{-p}\,C^p_{\|\cdot\|}\,\alpha^p\Big)$$

which in turn implies

$$\|\xi-\widetilde{x}_{i_0}\|^p \geq \frac{1}{(1+\eta)^p}\sum_{j\in J_{i_0}} \mu_j^{i_0}\|\xi-a_j\|^p - \eta^{-p}\,C^p_{\|\cdot\|}\,\alpha^p.$$

Plugging this inequality in (11.21) yields and using that $|J(\xi)|\leq d+1$, we finally get

$$\begin{aligned}(1+\eta)^{|J(\xi)|}F^p(\xi;\Gamma) &\geq \sum_{i\in I: x_i\in \mathring{K}} \widetilde{\lambda}_i\|\xi-x_i\|^p + \frac{1}{(1+\eta)^p}\sum_{i\in I: x_i\notin \mathring{K}} \widetilde{\lambda}_i \sum_{j\in J_i}\mu_j^i\|\xi-a_j\|^p \\ &\quad - |J(\xi)|\eta^{-p} d\, C^p_{\|\cdot\|}\,\alpha^p \\ &\geq \frac{1}{(1+\eta)^p}F^p\big(\xi;(\Gamma\cap \mathring{K}\})\cup\Gamma_{\alpha,\varepsilon}\big) - \eta^{-p}\,(d+1)\,C^p_{\|\cdot\|}\,\alpha^p.\end{aligned}$$

□

Now we can establish the sharp rate for the uniform distribution $\mathscr{U}([0,1]^d)$.

Proposition 11.6 (Uniform Distribution) *For every* $p\geq 1$,

$$Q^{dq}_{\|\cdot\|,p,d} := \inf_{n\geq 1} n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) = \lim_n n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Proof Let $n, m \in \mathbb{N}$, $m < n$ and set $k = k(n,m) = \left\lfloor \left(\frac{n}{m}\right)^{1/d} \right\rfloor \geq 1$. Covering the unit hypercube $[0,1]^d$ by k^d translates $C_1, \ldots, C_{k^d}$ of the hypercube $\left[0, \frac{1}{k}\right]^d$, we arrive at $\mathscr{U}\big([0,1]^d\big) = k^{-d} \sum_{1 \leq i \leq k^d} \mathscr{U}(C_i)$. Hence, Proposition 11.4 yields

$$d_{n,p}^p\big(\mathscr{U}\big([0,1]^d\big)\big) \leq k^{-d} \sum_{i=1}^{k^d} d_m^p(\mathscr{U}(C_i)).$$

Furthermore, Proposition 11.5 implies

$$d_{m,p}(\mathscr{U}(C_i)) = k^{-1}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big),$$

so that we may conclude for all $n, m \in \mathbb{N}, m < n$,

$$d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) \leq k^{-1}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Thus, we get

$$\begin{aligned} n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) &\leq k^{-1}\, n^{1/d}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big) \\ &\leq \frac{k+1}{k}\, m^{1/d}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big), \end{aligned}$$

which yields, for every fixed integer $m \geq 1$,

$$\limsup_n n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) \leq m^{1/d}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big)$$

since $\lim_n k(n,m) = +\infty$. This finally implies

$$\lim_n n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) = \inf_{m \geq 1} m^{1/d}\, d_{m,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$ □

Proposition 11.7 *For every* $p \geq 1$,

$$Q^{dq}_{\|\cdot\|,p,d} = \lim_n n^{1/d}\, d_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) = \lim_n n^{1/d}\, \bar{d}_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Proof Since, for every compactly supported distribution P, we have $\bar{d}_{n,p}(P) \leq d_{n,p}(P)$, it remains to show

$$Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \leq \liminf_n n^{1/d}\, \bar{d}_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

For $\varepsilon \in (0, 1/2)$, let $C_\varepsilon = (1/2, \ldots, 1/2) + \frac{1-\varepsilon}{2}[-1,1]^d$ be the centered hypercube in $[0,1]^d$ with edge-length $1-\varepsilon$ and midpoint $(1/2, \ldots, 1/2)$. Moreover, let $(\Gamma_n)_{n\geq 1}$ be a sequence of quantizers such that, for every $n \geq 1$,

$$\bar{d}_p\big(\mathscr{U}\big([0,1]^d\big); \Gamma_n\big) \leq (1+\varepsilon)\bar{d}_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Owing to Lemma 11.1, as $C_\varepsilon \subset (0,1)^d$, there is an integer $n_\varepsilon \in \mathbb{N}$ such that

$$\forall n \geq n_\varepsilon, \quad C_\varepsilon \subset \operatorname{conv}(\Gamma_n).$$

We therefore get for any $n \geq n_\varepsilon$

$$\begin{aligned}(1+\varepsilon)^d \bar{d}_n^p\big(\mathscr{U}\big([0,1]^d\big)\big) &\geq \bar{d}^p\big(\mathscr{U}\big([0,1]^d\big); \Gamma_n\big)\\ &\geq \int_{C_\varepsilon} \bar{F}^p(\xi; \Gamma_n)^p d\xi\\ &= \int_{C_\varepsilon} F^p(\xi; \Gamma_n)^p d\xi = \lambda_d(C_\varepsilon) d^p\big(\mathscr{U}(C_\varepsilon), \Gamma_n\big)\\ &\geq (1-\varepsilon)^d d_n^p\big(\mathscr{U}(C_\varepsilon)\big) = (1-\varepsilon)^{d+p} d_n^p\big(\mathscr{U}\big([0,1]^d\big)\big)\end{aligned}$$

where we used the scaling property (Proposition 11.5) in the last line.

Hence, we obtain that, for every $0 < \varepsilon < 1/2$,

$$\liminf_n n^{1/d}\, \bar{d}_{n,p}\big(\mathscr{U}\big([0,1]^d\big)\big) \geq \frac{(1-\varepsilon)^{1+d/p}}{(1+\varepsilon)^{d/p}}\, Q^{\mathrm{dq}}_{\|\cdot\|,p,d}$$

so that letting $\varepsilon \to 0$ completes the proof. □

Proposition 11.8 *Let $P = \sum_{i=1}^m s_i\, \mathscr{U}(C_i)$, $\sum_{i=1}^m s_i = 1$, $s_i > 0$, $i = 1, \ldots, m$, where $C_i = a_i + [0,l]^d$, $i = 1, \ldots, m$, are pairwise disjoint hypercubes in $\mathbb{R}^d$ with common edge-length l. Set*

$$h := \frac{dP}{d\lambda_d} = \sum_{i=1}^m s_i l^{-d} \mathbf{1}_{C_i}.$$

Then

$$\lim_n n^{1/d}\, d_{n,p}(P) = \lim_n n^{1/d}\, \bar{d}_{n,p}(P) = Q^{dq}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)}.$$

Proof Since $d_{n,p}(P) \geq \bar{d}_{n,p}$, it suffices to show that

$$\limsup_n n^{1/d}\, d_{n,p}(P) \leq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)} \text{ and } \liminf_n n^{1/d}\, \bar{d}_{n,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)}.$$

For $n \in \mathbb{N}$, set

$$t_i = \frac{s_i^{d/(d+p)}}{\sum_{j=1}^m s_j^{d/(d+p)}} \qquad \text{and} \qquad n_i = \lfloor t_i n \rfloor,\ 1 \le i \le m.$$

Then, by Propositions 11.4 and 11.5, we get for every $n \ge \max_{1\le i\le m}(1/t_i)$

$$d_n^p(P) \le \sum_{i=1}^m s_i\, d_n^p\big(\mathscr{U}(C_i)\big) = l^p \sum_{i=1}^m s_i\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big).$$

Proposition 11.6 then yields

$$n^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big) = \left(\frac{n}{n_i}\right)^{\frac{p}{d}} n_i^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big) \longrightarrow t_i^{-\frac{p}{d}} Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \quad \text{as } n \to +\infty.$$

Noting that $\|h\|_{d/(d+p)} = l^p \Big(\sum s_i^{d/(d+p)}\Big)^{(d+p)/d}$, we get

$$\limsup_n n^{\frac{p}{d}}\, d_{n,p}^p(P) \le Q^{\mathrm{dq}}_{\|\cdot\|,p,d} l^p \sum_{i=1}^m s_i\, t_i^{-\frac{p}{d}} = Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|h\|_{d/(d+p)}.$$

Now, let us prove the reverse inequality. Let $\varepsilon \in (0, l/2)$ and let $C_{i,\varepsilon}$ denote the closed hypercube with the same center as C_i but with edge-length $l - \varepsilon$. For $\alpha \in (0, \varepsilon/2)$, we set $\widetilde{\alpha} = \frac{l}{\lceil l/\alpha \rceil}$ and we define the lattice

$$\Gamma_{\alpha,\varepsilon,i} = \big(a_i + \widetilde{\alpha}\mathbb{Z}^d\big) \cap \big(C_i \setminus C_{i,\varepsilon}\big) \bigcup \big\{\text{vertices of } C_i\big\}.$$

It is clear that $\mathrm{conv}(\Gamma_{\alpha,\varepsilon,i}) = C_i \subset C_i \setminus C_{i,\varepsilon}$ since it contains the vertices of C_i. Moreover, for every $\xi \in C_i \setminus C_{i,\varepsilon}$, $\mathrm{dist}_{\ell^\infty}(\xi, \Gamma_{\alpha,\varepsilon,i}) \le \alpha$ so that there exists a real constant $C_{\|\cdot\|} > 0$ only depending on the norm $\|.\|$ such that $\mathrm{dist}_{\|.\|}(\xi, \Gamma_{\alpha,\varepsilon,i}) \le C_{\|\cdot\|}\alpha$. Consequently, the lattice $\Gamma_{\alpha,\varepsilon,i}$ satisfies the assumption of the firewall lemma (Lemma 11.2).

On the other hand, easy combinatorial arguments show that the number of points m_i of $\Gamma_{\alpha,\varepsilon,i}$ falling in C_i satisfies $\lceil \frac{l}{\widetilde{\alpha}} \rceil^d \le m_i \le \big(\lceil \frac{l}{\widetilde{\alpha}} \rceil + 1\big)^d + 2^d$ whereas the number $m_{i,\varepsilon}$ of points falling in $C_{i,\varepsilon}$ satisfies $\big(\lceil \frac{l-\varepsilon}{\widetilde{\alpha}} \rceil - 1\big)^d \le m_{i,\varepsilon} \le \big(\lceil \frac{l-\varepsilon}{\widetilde{\alpha}} \rceil + 1\big)^d$ so that

$$\left\lceil \frac{l}{\widetilde{\alpha}} \right\rceil^d - \left(\left\lceil \frac{l-\varepsilon}{\widetilde{\alpha}} \right\rceil + 1\right)^d \le |\Gamma_{\alpha,\varepsilon,i}| \le \left(\left\lceil \frac{l}{\widetilde{\alpha}} \right\rceil + 1\right)^d + 2^d - \left(\left\lceil \frac{l-\varepsilon}{\widetilde{\alpha}} \right\rceil - 1\right)^d.$$

We define for every $\varepsilon \in (0, l/2)$ and every $\alpha \in (0, \varepsilon/2)$

$$g_{l,\varepsilon}(\alpha) = \alpha^d |\Gamma_{\alpha,\varepsilon,i}|.$$

Since $\frac{\alpha}{\tilde{\alpha}} \to 1$ and $2\alpha\left\lceil \frac{\varepsilon/2}{\tilde{\alpha}} \right\rceil \to \varepsilon$ as $\alpha \to 0$, we conclude from the above inequalities that

$$\forall \varepsilon \in (0, l/2), \quad \lim_{\alpha \to 0} g_{l,\varepsilon}(\alpha) = l^d - (l-\varepsilon)^d. \tag{11.22}$$

Let $\eta \in (0,1)$ and let $(\Gamma_n)_{n\geq 1}$ denote a sequence of n-quantizers such that $\bar{d}^p(P;\Gamma_n) \leq (1+\eta)d_n^p(P)$. It follows from Theorem 11.3 that $\bar{d}^p(P;\Gamma_n) \to 0$ for $n \to \infty$ so that Lemma 11.1 yields the existence of $n_\varepsilon \in \mathbb{N}$ such that for any $n \geq n_\varepsilon$

$$\bigcup_{1\leq i\leq m} C_{i,\varepsilon} \subset \text{conv}(\Gamma_n).$$

We then derive from Lemma 11.2 (firewall)

$$\begin{aligned}
\bar{d}^p(\mathscr{U}(C_i);\Gamma_n) &= l^{-d}\int_{C_i} \bar{F}^p(\xi;\Gamma_n)\,\lambda_d(d\xi)\\
&\geq l^{-d}\int_{C_{i,\varepsilon}} \bar{F}^p(\xi;\Gamma_n)\,\lambda_d(d\xi) = l^{-d}\int_{C_{i,\varepsilon}} F^p(\xi;\Gamma_n)\,\lambda_d(d\xi)\\
&\geq \frac{l^{-d}\,(l-\varepsilon)^d}{(1+\eta)^{p+d+1}}\, d^p\big(\mathscr{U}(C_{i,\varepsilon}); (\Gamma_n \cap \mathring{C}_i) \cup \Gamma_{\alpha,\varepsilon,i}\big)\\
&\quad - l^{-d}\,(l-\varepsilon)^d \frac{(1+\eta)^{-d-1}}{\eta^p}\,(d+1)C_{\|\cdot\|}\cdot \alpha^p.
\end{aligned}$$

At this stage, we set for every $\rho > 0$

$$\alpha_n = \alpha_n(\rho) = \left(\frac{m}{\rho n}\right)^{1/d} \tag{11.23}$$

and denote

$$n_i = |(\Gamma_n \cap \mathring{C}_i) \cup \Gamma_{\alpha_n,\varepsilon,i}|.$$

Proposition 11.5 yields $d_{n_i,p}(\mathscr{U}(C_{i,\varepsilon})) = (l-\varepsilon) d_{n_i,p}\big(\mathscr{U}\big([0,1]^d\big)\big)$, so that we get

$$
\begin{aligned}
n^{\frac{p}{d}} d_n^p(P) &\geq \frac{1}{1+\eta} \sum_{i=1}^m s_i\, n^{\frac{p}{d}}\, \bar{d}^p(\mathscr{U}(C_i); \Gamma_n) \\
&\geq \frac{l^{-d}(l-\varepsilon)^d}{(1+\eta)^{p+d+2}} \sum_{i=1}^m s_i\, n^{\frac{p}{d}}\, d^p\big(\mathscr{U}(C_{i,\varepsilon}); (\Gamma_n \cap \mathring{C}_i) \cup \Gamma_{\alpha_n,\varepsilon,i}\big) \\
&\quad - l^{-d}(l-\varepsilon)^d \frac{(1+\eta)^{-d-2}}{\eta^p} \sum_{i=1}^m s_i\,(d+1)\, C_{\|\cdot\|} \cdot \alpha^p \cdot n^{\frac{p}{d}} \\
&\geq \frac{l^{-d}(l-\varepsilon)^{d+p}}{(1+\eta)^{p+d+2}} \sum_{i=1}^m s_i\, n^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big) \\
&\quad - l^{-d}(l-\varepsilon)^d \frac{(1+\eta)^{-d-2}}{\eta^p}(d+1)\, C_{\|\cdot\|}\Big(\frac{m}{\rho}\Big)^{\frac{p}{d}}.
\end{aligned}
\tag{11.24}
$$

Since

$$
\frac{n_i}{n} \leq \frac{|\Gamma_n \cap \mathring{C}_i|}{n} + \frac{g_{l,\varepsilon}(\alpha_n)}{n\alpha_n^d} = \frac{|\Gamma_n \cap \mathring{C}_i|}{n} + \frac{\rho}{m} g_{l,\varepsilon}(\alpha_n),
$$

we conclude from (11.22) and (11.23) that

$$
\limsup_n \sum_{i=1}^m \frac{n_i}{n} \leq 1 + \rho\big(l^d - (l-\varepsilon)^d\big).
$$

We may choose a subsequence (still denoted by (n)), such that

$$
n^{1/d}\, \bar{d}_{n,p}(P) \to \liminf_n n^{1/d}\, d_{n,p}(\mathbb{P}) \;\text{ and }\; \frac{n_i}{n} \to v_i \in \big[0, 1 + \rho(l^d - (l-\varepsilon)^d)\big].
$$

As a matter of fact, $v_i > 0$, for every $i = 1, \ldots m$: otherwise Proposition 11.6 would yield

$$
\begin{aligned}
n^{\frac{p}{d}}\, \bar{d}_{n,p}^p(P) &\geq \frac{l^{-d}(l-\varepsilon)^{d+p}}{(1+\eta)^{p+d+2}} \sum_{i=1}^m s_i \Big(\frac{n_i}{n}\Big)^{-\frac{p}{d}} n_i^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big) \\
&\quad - l^{-d}(l-\varepsilon)^d \frac{(1+\eta)^{p-d-2}}{\eta^p}(d+1) C_{\|\cdot\|} \cdot \Big(\frac{m}{\rho}\Big)^{\frac{p}{d}} \\
&\to +\infty \quad \text{as} \quad n \to +\infty
\end{aligned}
$$

which contradicts (a). Consequently, we may normalize the v_i's by setting

$$\widetilde{v}_i = \frac{v_i}{1+\rho(l^d-(l-\varepsilon)^d)},\ i=1,\dots,m,$$

so that $\sum_{i=1}^m \widetilde{v}_i \le 1$. We derive from Proposition 11.6 that

$$\begin{aligned}
&\liminf_n \sum_{i=1}^m s_i\, n^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big)\\
&\ge \sum_{i=1}^m s_i\, v_i^{-\frac{p}{d}} n_i^{\frac{p}{d}}\, d_{n_i}^p\big(\mathscr{U}\big([0,1]^d\big)\big)\\
&= Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\big(1+\rho\big(l^d-(l-\varepsilon)^d\big)\big)^{-\frac{p}{d}} \sum_{i=1}^m s_i\, \widetilde{v}_i^{-\frac{p}{d}}\\
&\ge Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\big(1+\rho\big(l^d-(l-\varepsilon)^d\big)\big)^{-\frac{p}{d}} \inf_{\sum_i y_i\le 1,\, y_i\ge 0} \sum_{i=1}^m s_i\, y_i^{-\frac{p}{d}}\\
&= Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\big(1+\rho\big(l^d-(l-\varepsilon)^d\big)\big)^{-\frac{p}{d}} \left(\sum_{i=1}^m s_i^{d/(d+p)}\right)^{(d+p)/d}.
\end{aligned}$$

Hence, we get from (11.24)

$$\begin{aligned}
&\liminf_n n^{\frac{p}{d}}\, \bar{d}^p_{n,p}(P)\\
&\ge \frac{l^{-d}\,(l-\varepsilon)^{d+p}}{(1+\eta)^{p+d+2}\big(1+\rho(l^d-(l-\varepsilon)^d)\big)^{\frac{p}{d}}}\, Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\left(\sum_{i=1}^m s_i^{d/(d+p)}\right)^{(d+p)/d}\\
&\quad - l^{-d}\,(l-\varepsilon)^d\,\frac{(1+\eta)^{-d-2}}{\eta^p}(d+1)\,C_{\|\cdot\|}\cdot\left(\frac{m}{\rho}\right)^{\frac{p}{d}}.
\end{aligned}$$

Letting $\varepsilon \to 0$ implies

$$\begin{aligned}
\liminf_n n^{\frac{p}{d}}\, \bar{d}^p_{n,p}(P) &\ge \frac{l^p}{(1+\eta)^{p+d+2}}\, Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\left(\sum_{i=1}^m s_i^{d/(d+p)}\right)^{(d+p)/d}\\
&\quad - \frac{(1+\eta)^{-d-2}}{\eta^p}(d+1)\,C_{\|\cdot\|}\left(\frac{m}{\rho}\right)^{\frac{p}{d}}\\
&= \frac{1}{(1+\eta)^{p+d+2}}\, Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\cdot \|h\|_{d/(d+p)}\\
&\quad - \frac{(1+\eta)^{-d-2}}{\eta^p} d\, C_{\|\cdot\|}\left(\frac{m}{\rho}\right)^{\frac{p}{d}}.
\end{aligned}$$

Letting successively ρ go to $+\infty$ and η go to 0 completes the proof. □

Proposition 11.9 *Assume that P is absolutely continuous w.r.t. λ_d with compact support and density h. Then*

$$\lim_n n^{\frac{p}{d}} d_{n,p}(P) = \liminf_n n^{\frac{p}{d}} \bar{d}_{n,p}(P) = Q^{dq}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)}.$$

Proof Since $d_{n,p}(P) \geq \bar{d}_{n,p}(P)$, it suffices to show that

$$\limsup_n n^{\frac{p}{d}} d_{n,p}(P) \leq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)} \text{ and } \liminf_n n^{\frac{p}{d}} \bar{d}_{n,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|h\|^{\frac{1}{p}}_{d/(d+p)}.$$

Preliminary Step Let $C = [-l/2, l/2]^d$ be a closed hypercube centered at the origin, parallel to the coordinate axis with edge-length l, such that $\mathrm{supp}(P) \subset C$. For $k \in \mathbb{N}$ consider the tessellation of C into k^d closed hypercubes with common edge-length l/k. To be precise, for every $\underline{i} = (i_1, \ldots, i_d) \in \mathbb{Z}^d$, we set

$$C_{\underline{i}} = \prod_{r=1}^{d} \Big[-\frac{l}{2} + \frac{i_r l}{k}, -\frac{l}{2} + \frac{(i_r+1)l}{k} \Big].$$

Then, set

$$h = \frac{dP}{d\lambda_d} \text{ and } P_k = \sum_{\substack{\underline{i} \in \mathbb{Z}^d \\ 0 \leq i_r < k}} P(C_{\underline{i}}) \,\mathscr{U}(C_{\underline{i}}),\ h_k = \frac{dP_k}{d\lambda_d} = \sum_{\substack{\underline{i} \in \mathbb{Z}^d \\ 0 \leq i_r < k}} \frac{P(C_{\underline{i}})}{\lambda_d(C_{\underline{i}})} \mathbf{1}_{C_{\underline{i}}},\ k \geq 1. \tag{11.25}$$

By differentiation of measures, we obtain $h_k \to h$, λ_d-$a.s.$ as $k \to +\infty$. Which in turn implies, owing to Scheffé's Lemma,

$$\lim_{k \to +\infty} \|h_k - h\|_1 = 0.$$

Furthermore,

$$\lim_{k \to +\infty} \|h_k\|_{d/(d+p)} = \|h\|_{d/(d+p)}$$

since $\|h_k - h\|_{d/(d+p)} \leq \Big(\lambda_d(C)\Big)^{\frac{p}{d}} \|h_k - h\|_1$ by Jensen's Inequality applied to the probability measure $\frac{\lambda_{d|C}}{\lambda_d(C)}$. Moreover, by Proposition 11.8 we have

$$\lim_n n^{1/d} d_{n,p}(P_k) = Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \|h_k\|^{\frac{1}{p}}_{d/(d+p)}. \tag{11.26}$$

Likewise, we define an inner approximation of P: first, define

$$C^k = \bigcup_{C_{\underline{i}} \subset \operatorname{supp}(P)} C_{\underline{i}}$$

the union of the hypercubes $C_{\underline{i}}$ lying in the interior of $\operatorname{supp}(P)$. Then set

$$\mathring{P}_k = \sum_{C_{\underline{i}} \subset \operatorname{supp}(P)} P(C_{\underline{i}})\, \mathscr{U}(C_{\underline{i}}) \quad \text{and} \quad \mathring{h}_k = \frac{d\mathring{P}_k}{d\lambda_d} = h_k \mathbf{1}_{C^k}.$$

We have as above that

$$\mathring{h}_k \to h, \quad \lambda_d\text{-a.s.} \quad \text{as} \quad k \to +\infty.$$

Consequently, we also have

$$\lim_{k\to\infty} \|\mathring{h}_k - h\|_1 = 0 \quad \text{and} \quad \lim_{k\to\infty} \|\mathring{h}_k\|_{d/(d+p)} = \|h\|_{d/(d+p)}.$$

We get likewise by Proposition 11.8 that, for every $k \in \mathbb{N}$,

$$\lim_n n^{1/d}\, d_{n,p}(\mathring{P}_k) = Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \|\mathring{h}_k\|^{\frac{1}{p}}_{d/(d+p)}. \tag{11.27}$$

(a) Let $0 < \varepsilon < 1$ and $n \ge 2^d/\varepsilon$. If we divide each edge of the hypercube C into

$$m = \left\lfloor (\varepsilon n)^{1/d} \right\rfloor - 1$$

intervals of equal length l/m, the interval endpoints define $m+1$ grid points on each edge. Denoting by $\Gamma_1 = \Gamma_1(\varepsilon, n)$ the product quantizer made up by this procedure, we clearly have

$$|\Gamma_1| = (m+1)^d = \left\lfloor (\varepsilon n)^{1/d} \right\rfloor^d =: n_1.$$

For this product quantizer, it follows from Proposition 11.3 that, for all $\xi \in C$,

$$F^p(\xi; \Gamma_1) \le C_{\|\cdot\|} \sum_{i=1}^{d} \left(\frac{l}{2m}\right)^p \le C_{\|\cdot\|,p,d}\, \frac{l^p}{(\varepsilon n)^{\frac{p}{d}}}.$$

For $n_2 = \lfloor (1-\varepsilon)n \rfloor$, let Γ_2 be an n_2-quantizer such that $d^p(P_k; \Gamma_2) \leq (1+\varepsilon) d^p_{n_2}(P_k)$. We clearly have $|\Gamma_1 \cup \Gamma_2| \leq n$ and

$$n^{\frac{p}{d}} \left| \int F^p(\xi; \Gamma_1 \cup \Gamma_2) dP_k(\xi) - \int F^p(\xi; \Gamma_1 \cup \Gamma_2) dP(\xi) \right|$$
$$\leq n^{\frac{p}{d}} \int F^p(\xi; \Gamma_1 \cup \Gamma_2) |h_k(\xi) - h(\xi)| d\lambda_d(\xi)$$
$$\leq C_{\|\cdot\|,p,d} \frac{l^p}{\varepsilon^{\frac{p}{d}}} \|h_k - h\|_1 = c_{1,\varepsilon} \|h_k - h\|_1$$

for $k \in \mathbb{N}$ and $n \geq \max\left\{\frac{2^d}{\epsilon}, \frac{1}{1-\varepsilon}\right\}$. This implies

$$n^{\frac{p}{d}} d^p_n(P) \leq n^{\frac{p}{d}} \int F^p(\xi; \Gamma_1 \cup \Gamma_2) dP(\xi)$$
$$\leq n^{\frac{p}{d}} \int F^p(\xi; \Gamma_1 \cup \Gamma_2) dP_k(\xi) + c_{1,\varepsilon} \|h_k - h\|_1$$
$$\leq n^{\frac{p}{d}} \int F^p(\xi; \Gamma_2) dP_k(\xi) + c_{1,\varepsilon} \|h_k - h\|_1$$
$$\leq (1+\varepsilon)\, n^{\frac{p}{d}} d^p_{n_2}(P_k) + c_{1,\varepsilon} \|h_k - h\|_1$$

so that we can conclude from (11.26) that

$$\limsup_n n^{\frac{p}{d}} d^p_n(P) \leq \frac{1+\varepsilon}{(1-\varepsilon)^{\frac{p}{d}}} \left(Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\right)^p \|h_k\|_{d/(d+p)} + c_{1,\varepsilon} \|h_k - h\|_1.$$

Letting first k go to infinity and then letting ε go to zero yields

$$\limsup_n n^{1/d} d^p_n(P) \leq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \|h_k\|^{\frac{1}{p}}_{d/(d+p)}.$$

(b) Assume now that Γ_3 is an n_2-quantizer such that $\bar{d}^p(P; \Gamma_3) \leq (1+\varepsilon)\, \bar{d}^p_{n_2}(P)$. Again it holds $|\Gamma_1 \cup \Gamma_3| \leq n$ and we derive as above

$$n^{\frac{p}{d}} \left| \int F^p(\xi; \Gamma_1 \cup \Gamma_3) d\mathring{P}_k(\xi) - \int F^p(\xi; \Gamma_1 \cup \Gamma_3) dP(\xi) \right| \leq c_{2,\varepsilon} \|\mathring{h}_k - h\|_1. \tag{11.28}$$

Moreover, Lemma 11.1 yields for every $k \in \mathbb{N}$ the existence of $n_{k,\varepsilon} \in \mathbb{N}$ such that, for all $n \geq n_{k,\varepsilon}$,

$$(1+\varepsilon)\, \bar{d}^{p}_{n_2}(P) \geq \bar{d}^{p}(P; \Gamma_3) \geq \int_{\mathrm{conv}(\Gamma_3)} F^p(\xi; \Gamma_3) dP(\xi)$$

$$\geq \int_{C^k} F^p(\xi; \Gamma_3) dP(\xi) \geq \int_{C^k} F^p(\xi; \Gamma_1 \cup \Gamma_3) dP(\xi).$$

Thus, we derive from (11.28) that, for every $n \geq \max\left(n_{k,\varepsilon}, \frac{2^d}{\varepsilon}, \frac{1}{1-\varepsilon}\right)$,

$$(1+\varepsilon)\, n^{\frac{p}{d}}\, \bar{d}^{p}_{n_2}(P) \geq n^{\frac{p}{d}} \int_{C^k} F^p(\xi; \Gamma_1 \cup \Gamma_3) dP(\xi)$$

$$\geq n^{\frac{p}{d}} \int_{C^k} F^p(\xi; \Gamma_1 \cup \Gamma_3) d\mathring{P}_k(\xi) - c_{2,\varepsilon} \|\mathring{h}_k - h\|_1$$

$$\geq n^{\frac{p}{d}} d^{p}_{n}(\mathring{P}_k) - c_{2,\varepsilon} \|\mathring{h}_k - h\|_1,$$

which yields, once combined with (11.27),

$$\frac{1+\varepsilon}{(1-\varepsilon)^{\frac{p}{d}}} \liminf_n n_2^{\frac{p}{d}}\, \bar{d}^{p}_{n_2,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \|\mathring{h}_k\|_{d/(d+p)} - c_{2,\varepsilon} \|\mathring{h}_k - h\|_1.$$

Letting first k go to $+\infty$ and then letting ε go to 0, we get

$$\liminf_n n^{\frac{1}{d}}\, \bar{d}_{n,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \|h\|^{\frac{1}{p}}_{d/(d+p)}. \qquad \square$$

Proposition 11.10 (Singular Distribution) *Assume that P is singular with respect to λ_d and has compact support. Then*

$$\limsup_n n^{\frac{p}{d}}\, \bar{d}_{n,p}(P) = 0.$$

Proof Let A be a Borel set such that $P(A) = 1$ and $\lambda_d(A) = 0$. Let $\varepsilon > 0$; by the outside regularity of λ_d, there exists an open set $O = O(\varepsilon) \supset A$ such that $\lambda_d(O) \leq \varepsilon$ (and $P(O) = 1$). Let C be an open hypercube with edges parallel to the coordinate axis, edge-length ℓ and containing the closure of A.

Let $C_k = \prod_{i=1}^d [c_{k,i}, c_{k,i} + \ell_i)$, $k \in \mathbb{N}$, be a countable partition of O consisting of nonempty half-open hypercubes, still with edges parallel to the coordinate axis (see, e.g. Lemma 1.4.2 in [4]).

Let $m = m(\varepsilon) \in \mathbb{N}$ such that $\sum_{k \geq m+1} P(C_k) \leq \varepsilon^{\frac{p}{d}} \ell^{-p}$.

Let $n \in \mathbb{N}$, $n \geq 2^{d+1}$ and let $n_1, \ldots, n_d \geq 2$ be integers such that the product $n_1^d + \cdots + n_m^d \leq n/2$. One designs a grid Γ as follows.

For every $k \in \{1, \ldots, m\}$, we consider the lattice of C_k of size n_i^d defined by

$$\prod_{i=1}^{d} \Big\{ c_{k,i} + \frac{r_i}{n_k - 1} \ell_i,\ r_i = 0, \ldots, n_k - 1,\ i = 1, \ldots, d \Big\}.$$

Then, one defines likewise the lattice of C of size $n_{m+1}^d \leq n/2$

$$\prod_{i=1}^{d} \Big\{ c_{k,i} + \frac{r_i}{n_{m+1} - 1} \ell_i,\ r_i = 0, \ldots, n_{m+1} - 1,\ i = 1, \ldots, d \Big\}.$$

The grid Γ is made up with all the points of the $m + 1$ above finite lattices.

Now let $\xi \in A$. It is clear from the definition of the function F_p that

$$F_p(\xi; \Gamma) \leq \begin{cases} C_{\|.\|} \big(\ell_k / n_k\big)^p & \text{if } \xi \in \bigcup_{k=1}^m C_k \\ C_{\|.\|} \big(\ell / n_{m+1}\big)^p & \text{if } \xi \in C \setminus \bigcup_{k=1}^m C_k \end{cases}$$

where $C_{\|.\|} > 0$ is a real constant only depending on the norm. As a consequence

$$\begin{aligned} d_n^p(P) &= \sum_{k=1}^{m} \int_{C_k} F^p(\xi; \Gamma) dP(\xi) + \int_{C \setminus \bigcup_{k=1}^m C_k} F^p(\xi; \Gamma) dP(\xi) \\ &\leq C_{\|.\|} \Big(\sum_{k=1}^{m} (\ell_k / n_k)^p P(C_k) + (\ell / n_{m+1})^p P\Big(C \setminus \bigcup_{k=1}^{m} C_k\Big)\Big). \end{aligned}$$

Set for every $k \in \{1, \ldots, m\}$, $n_k = \left\lfloor \dfrac{\ell_k (n/2)^{\frac{1}{d}}}{(\sum_{k'=1}^d \ell_{k'}^d)^{\frac{1}{d}}} \right\rfloor$ and $n_{m+1} = \lfloor (n/2)^{\frac{1}{d}} \rfloor$.

Note that

$$\sum_{k'=1}^{d} \ell_{k'}^d = \sum_{k=1}^{m} \lambda_d(C_k) \leq \lambda_d(O) \leq \varepsilon.$$

Elementary computations show that, for large enough n, all the integers n_k are greater than 1 and that

$$\sum_{k=1}^{m}(\ell_k/n_k)^p P(C_k) + (\ell/n_{m+1})^p P\Big(C \setminus \bigcup_{k=1}^{m} C_k\Big) \le (\sum_{k'=1}^{d} \ell_{k'}^d)^{\frac{p}{d}} (n/2)^{-\frac{p}{d}} P\big(\cup_{1\le k\le m} C_k\big)$$
$$+(n/2)^{-\frac{p}{d}} \ell^p P\Big(C \setminus \bigcup_{k=1}^{m} C_k\Big)$$

so that

$$\limsup_n n^{\frac{p}{d}} d_n^p(P) \le C_{\|.\|}(\varepsilon/2)^{\frac{p}{d}}$$

which in turn implies, by letting ε go to 0, that $\limsup_n n^{\frac{p}{d}} d_n^p(P) = 0$. □

Proof of Theorem 11.2 Claim (a) follows directly from Propositions 11.4, 11.9 and 11.10: assume $P = \rho P_a + (1-\rho)P_s$ where $P_a = \frac{h}{\rho}\lambda_d$ and P_s denote the absolutely continuous and singular parts of P respectively. The following inequalities hold true

$$\rho \bar{d}_{n,p}(P_a) \le \bar{d}_{n,p}(P) \le \rho \bar{d}_{n_1,p}(P_a) + (1-\rho)\bar{d}_{n_2,p}(P_s)$$

for every triplet of integers (n_1, n_2, n) with $n_1 + n_2 \le n$. Set $n_1 = n_1(n) = \lfloor (1-\varepsilon n)\rfloor$, $n_2 = n_2(n) = \lfloor \varepsilon n \rfloor$. Then we derive that

$$\rho Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \Big\| \frac{h}{\rho} \Big\|^{\frac{1}{p}}_{d/(d+p)} \liminf_n n^{\frac{p}{d}} \bar{d}_{n,p}(P_a) \le \liminf_n n^{\frac{p}{d}} \bar{d}_{n,p}(P)$$
$$\le \limsup_n n^{\frac{p}{d}} \bar{d}_{n,p}(P)$$
$$\le \rho(1-\varepsilon)^{-\frac{p}{d}} Q^{\mathrm{dq}}_{\|\cdot\|,p,d} \cdot \Big\| \frac{h}{\rho} \Big\|^{\frac{1}{p}}_{d/(d+p)}.$$

Letting ε go to 0 completes the proof.

Furthermore, part (c) was derived in [12, Sect. 5.1]. Hence, it remains to prove Claim (b).

Proof of Claim (b) Step 1. (Lower Bound) If X is compactly supported, the assertion follows from Proposition 11.9. Otherwise, set for every $R \in (0, \infty)$,

$$C_R = [-R, R]^d \text{ and } P(\cdot\,|C_k) = \frac{h\mathbf{1}_{C_k}}{P(C_k)} \cdot \lambda_d,\; k \in \mathbb{N}.$$

Proposition 11.9 yields again

$$\lim_n n^{\frac{1}{d}}\,\bar{d}_{n,p}(P(\cdot|C_k)) = Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\cdot \|h\mathbf{1}_{C_k}/P(C_k)\|^{\frac{1}{p}}_{d/(d+p)}, \tag{11.29}$$

so that $\bar{d}^p_{n,p}(P) \geq P(\cdot|C_k)\bar{d}^p_{n,p}(P(\cdot|C_k))$ implies for all $k \in \mathbb{N}$

$$\liminf_n n^{\frac{1}{d}}\,\bar{d}_{n,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\cdot \|h\mathbf{1}_{C_k}\|^{\frac{1}{p}}_{d/(d+p)}.$$

Sending k to infinity, we get at

$$\liminf_n n^{\frac{1}{d}}\,\bar{d}_{n,p}(P) \geq Q^{\mathrm{dq}}_{\|\cdot\|,p,d}\cdot \|h\|^{\frac{1}{p}}_{d/(d+p)}.$$

Step 2 (Upper Bound, $\mathrm{supp}(P) = \mathbb{R}^d$). Let $\rho \in (0,1)$. Set $K = C_{k+\rho}$ and $K_\rho = C_k$. Let $\Gamma_{k,\alpha,\rho}$ be the lattice grid associated to $K \setminus K_\rho$ with edge $\alpha > 0$ as defined in the proof of Proposition 11.8. It is straightforward that there exists a real constant $C > 0$ such that

$$\forall k \in \mathbb{N}, \forall \rho \in (0,1), \forall \alpha \in (0,\rho): \quad |\Gamma_{\alpha,\rho}| \leq C d\rho k^{d-1}\alpha^{-d}.$$

Let $\varepsilon \in (0,1)$. For every $n \geq 1$, set $\alpha_n = \widetilde{\alpha}_0 n^{-\frac{1}{d}}$ where $\widetilde{\alpha}_0 \in (0,1)$ is a real constant and

$$n_0 = |\Gamma_{k,\alpha_n,\rho}|, \quad n_1 = \lfloor (1-\varepsilon)(n-n_0)\rfloor, \quad n_2 = \lfloor \varepsilon(n-n_0)\rfloor$$

so that $\alpha_n \in (0.\rho)$, $n_0 + n_1 + n_2 \leq n$ and $n_i \geq 1$ for large enough n.

For every $\xi \in K_\rho = C_k$ and every grid $\Gamma \subset \mathbb{R}^d$ such that $\mathrm{conv}(\Gamma) \supset K_\rho$, we know by the "firewall" Lemma 11.2 that

$$F^p\big(\xi; (\Gamma \cap \mathring{K}) \cup \Gamma_{\alpha,\rho}\big) \leq (1+\eta)^p F^p(\xi;\Gamma) + (1+\eta)^p(1+1/\eta)^p C_{\|.\|}\alpha^p.$$

Let $\Gamma_1 = \Gamma_1(n_1,k)$ be an n_1 quantizer such that $d^p_{n_1}\big(P(.|C_k);\Gamma_1\big) \leq (1+\eta)d^p_{n_1}\big(P(\cdot\,|C_k)\big)$. Set $\Gamma_1' = \big((\Gamma_1 \cap \mathring{C}_{k+\rho}) \cup \Gamma_{k,\alpha_n,\rho}\big)$. One has $\Gamma_1' \subset C_{k+2\rho}$ for large enough n (so that $\alpha_n < \rho$).

Let moreover $\Gamma_2 = \Gamma_2(n_2,k)$ be an n_2 quantizer such that $\bar{d}^p_{n_2}\big(P(.|C^c_k);\Gamma_2\big) \leq (1+\eta)\bar{d}^p_{n_2}\big(P(.|C^c_k)\big)$. For $n \geq n_\rho$, we may assume that $C_{k+2\rho} \subset \mathrm{conv}(\Gamma_2)$ owing to Lemma 11.1 since $C_{k+2\rho} = \mathrm{conv}(C_{k+2\rho} \setminus C_{k+\frac{3}{2}\rho})$ and $C_{k+2\rho} \setminus C_{k+\frac{3}{2}\rho} \subset \overbrace{\mathrm{supp}\,P(.|C^c_k)}^{\circ}$. As a consequence $\Gamma_1' \subset \mathrm{conv}(\Gamma_2)$ so that $\mathrm{conv}(\Gamma_1') \subset \mathrm{conv}(\Gamma_2) = \mathrm{conv}(\Gamma)$ where $\Gamma = \Gamma_1' \cup \Gamma_2$ and

$$C_{k+\rho} \subset \mathrm{conv}(\Gamma) = \mathrm{conv}(\Gamma_2).$$

Now

$$\bar{d}_n^p(P) \le \int_{C_k} \Big(F^p(\xi;\Gamma)\mathbf{1}_{\{\xi\in \mathrm{conv}(\Gamma_2)\}} + \underbrace{d(\xi,\Gamma)^p\mathbf{1}_{\{\xi\notin \mathrm{conv}(\Gamma_2)\}}}_{=0}\Big)dP(\xi)$$
$$+\int_{C_k^c} \big(F^p(\xi;\Gamma)\mathbf{1}_{\{\xi\in \mathrm{conv}(\Gamma_2)\}} + d(\xi,\Gamma)^p\mathbf{1}_{\{\xi\notin \mathrm{conv}(\Gamma_2)\}}\big)\,dP(\xi).$$

Using that, for every $\xi \in C_k$,

$$F^p(\xi;\Gamma) \le F^p(\xi;\Gamma_1')$$
$$\le (1+\eta)^p\Big(F^p(\xi;\Gamma_1) + (1+1/\eta)^p\, C_{\|.\|}\,\alpha_n^p\Big)$$

implies

$$\bar{d}_n^p(P) \le P(C_k)(1+\eta)^p\Big((1+\eta)\,d_{n_1}^p(P(.|C_k)) + (1+1/\eta)^p\,C_{\|.\|}\,\widetilde{\alpha}_0\,n^{-\frac{1}{d}}\Big)$$
$$+P(C_k^c)\,(1+\eta)\,\bar{d}_{n_2}^p(P(.|C_k^c)).$$

Consequently

$$n^{\frac{p}{d}}\bar{d}_n^p(P) \le P(C_k)(1+\eta)^p\Big[(1+\eta)\left(\frac{n}{n_1}\right)^{\frac{p}{d}} n_1^{\frac{p}{d}}\,d_{n_1}^p(P(.|C_k)) + (1+1/\eta)^p C_{\|.\|}\widetilde{\alpha}_0\Big]$$
$$+(1+\eta)\left(\frac{n}{n_2}\right)^{\frac{p}{d}} P(C_k^c)\,n_2^{\frac{p}{d}}\,\bar{d}_{n_2}^p(P(.|C_k^c))$$

which in turn implies, using Proposition 11.9 for the modulus $d_{n,p}$ and the d-dimensional version of the extended Pierce Lemma (Theorem 11.3) for $\bar{d}_{n,p}$,

$$\limsup_n n^{\frac{p}{d}}\bar{d}_n^p(P) \le P(C_k)(1+\eta)^p\Big[\Big(\frac{(1+\eta)^{-p/d}}{(1-\varepsilon)(1-Cd\rho k^{d-1}\widetilde{\alpha}_0^{-d})}\Big)^{\frac{p}{d}} Q_{\|.\|}^{dq}\,\|h\mathbf{1}_{C_k}\|_{L^{\frac{d}{d+p}}}$$
$$+(1+1/\eta)^p C_{\|.\|}\widetilde{\alpha}_0\Big]$$
$$+P(C_k^c)\,(1+\eta)\;C_{p,d}\,\|X\mathbf{1}_{\{X\in C_k^c\}}\|_{L^{p+\delta}}^p\left(\frac{1}{\varepsilon(1-Cd\rho k^{d-1}\widetilde{\alpha}_0^{-d})}\right)^{\frac{p}{d}}.$$

One concludes by letting successively ρ, $\widetilde{\alpha}_0$, η go to 0, $k\to\infty$ and finally ε to 0.
Step 3. (Upper Bound: General Case). Let $\rho\in(0,1)$. Set $P_\rho = \rho P + (1-\rho)P_0$ where $P_0 = \mathcal{N}(0; I_d)$ (d-dimensional normal distribution). It is clear from the very definition of $\bar{d}_{n,p}$ that $\bar{d}_{n,p}(P) \le \frac{1}{\rho}\bar{d}_{n,p}(P_\rho)$ since $P \le \frac{1}{\rho}P_\rho$. The distribution P_ρ has $h_\rho = \rho h + (1-\rho)h_0$ as a density (with obvious notations) and one concludes

by noting that

$$\lim_{\rho \to 0} \|h_\rho\|_{d/(d+p)} = \|h\|_{d/(d+p)}$$

owing to the Lebesgue dominated convergence Theorem. □

Proof of Proposition 11.1 Using Hölder's inequality one easily checks that for $0 \le r \le p$ and $x \in \mathbb{R}^d$ it holds

$$|x|_{\ell^r} \le d^{\frac{1}{r}-\frac{1}{p}}\, |x|_{\ell^p}.$$

Moreover, for $m \in \mathbb{N}$ set $n = m^d$ and let Γ' be an optimal quantizer for $d_{m,p}\big(\mathscr{U}\big([0,1]\big)\big)$ (or at least $(1+\varepsilon)$-optimal for $\varepsilon > 0$). Denoting $\Gamma = \prod_{i=1}^d \Gamma'$, it then follows from Proposition 11.3(*b*) that

$$\begin{aligned} n^{\frac{p}{d}}\, d_n^p\big(\mathscr{U}\big([0,1]^d\big)\big) &\le n^{\frac{p}{d}}\, d^p\big(\mathscr{U}\big([0,1]^d\big);\Gamma\big) \\ &= m^p \sum_{i=1}^d d^p\big(\mathscr{U}\big([0,1]\big);\Gamma'\big) = d\, m^p\, d_m^p\big(\mathscr{U}\big([0,1]\big)\big). \end{aligned}$$

Combining both results and keeping in mind that $Q^{\mathrm{dq}}_{\|\cdot\|,p,d}$ holds as an infimum, we obtain for $r \in [0,p]$,

$$\big(Q^{\mathrm{dq}}_{|\cdot|_{\ell^r},p,d}\big)^p \le d^{\frac{p}{r}-1}\, n^{\frac{p}{d}}\, d^p_{n,|\cdot|_{\ell^p}}\big(\mathscr{U}\big([0,1]^d\big)\big) \le d^{\frac{p}{r}}\, m^p\, d_m^p\big(\mathscr{U}\big([0,1]\big)\big),$$

which finally proves the assertion by sending $m \to +\infty$. □

11.6 Concluding Remarks and Prospects

This result does not conclude the theoretical investigations about dual quantization (beyond the existence of optimal dual quantizers in the case $p = 1$, left open in [12]): the first one is to elucidate the asymptotic behaviour of the constant $Q^{dq}_{\|.\|,p,d}$ coming out in Theorem 11.2 as d goes to infinity, most likely by showing that $\lim_{d\to+\infty} \frac{Q^{dq}_{\|.\|,p,d}}{Q^{vq}_{\|.\|,p,d}} = 1$. From a practical point of view, is it possible to evaluate the mean dual quantization error induced by an optimal Voronoi quantization grid? An answer to that question would be very valuable for applications since many optimal quantization grids have been computed for various distributions (see e.g. [8] for Gaussian distributions).

Many natural questions solved in the optimal Voronoi quantization theory remain open. Among others "Is there a counterpart to the empirical measure theorem for (asymptotically) optimal quantizers?" (see Theorem 7.5, p. 96 in [5])? "How does dual quantization behave with respect to empirical distribution of i.i.d. n-samples of a given distribution?". Is it possible to develop an infinite dimensional "functional" dual quantization?

Appendix: Numerical Results for $\bar{d}_{n,2}(X)^2$

In order to support the heuristic argumentation on the intrinsic and rate optimal growth limitation of the truncation error $\mathbb{P}\big(X \notin C_n\big)$ induced by the extended dual quantization error modulus, we consider the two dimensional random variable

$$X = \Big(W_T, \sup_{0\leq t\leq T} W_t\Big),$$

where $(W_t)_{0\leq t\leq T}$ is a standard Brownian Motion. This example is motivated by the pricing of path-dependent (exotic) options, where this joint distribution plays an important role.

Using a variant of the CVLQ algorithm (see [12]) adapted for the dual quantization modulus inside C_n and the nearest neighbor mapping outside, we have computed a sequence of optimal grids together with the squared dual quantization error $\bar{d}_{n,2}(X)^2$ and the truncation error $\mathbb{P}\big(X \notin C_n\big)$.

These results are reported in Table 11.1 below.

Furthermore, we see in Fig. 11.1 a log-log plot for the convergence of the two rates $\bar{d}_{n,2}(X)^2$ and $\mathbb{P}\big(X \notin C_n\big)$.

The distortion rate $\bar{d}_{n,2}(X)^2$ shows here an absolute stable convergence rate (least-squares fit of the exponent yields -1.07192) which is consistent with the theoretical optimal rate of $n^{-\frac{2}{d}}$ since $d = 2$. Moreover, the truncation error $\mathbb{P}\big(X \notin C_n\big)$ outperforms also in this case the heuristically derived rate of n^{-1} and also outperforms the squared "inside" quantization error, which means that even for such an asymmetric and non-spherical distribution of the Brownian motion and its supremum, a second order rate can be achieved.

This confirms again the motivation of the extended dual quantization error as the correction penalization constraint on growth of the convex hull in order to preserve second order stationarity.

Table 11.1 Numerical results for the dual quantization X

n	$\bar{d}_{n,2}(X)^2$	$\mathbb{P}\big(X \notin C_n\big)$
50	0.04076	0.01784
100	0.01966	0.00795
150	0.01236	0.00412
200	0.00931	0.00141

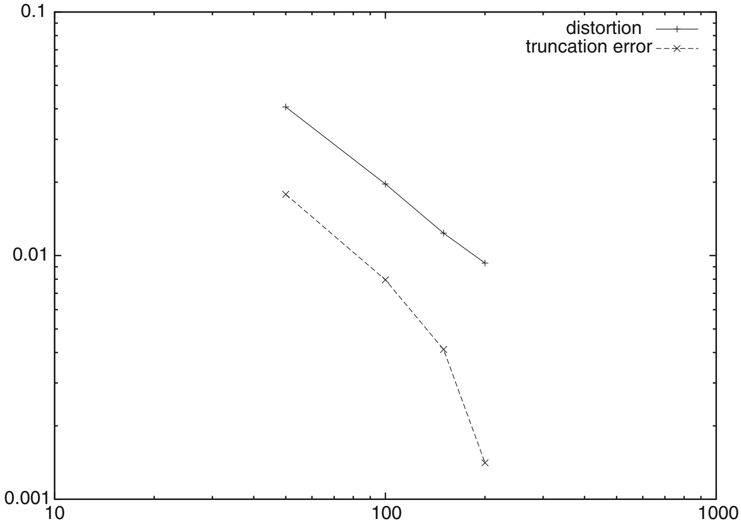


Fig. 11.1 Log-log plot of $\bar{d}_{n,2}(X)^2$ (distortion error) and $\mathbb{P}\big(X \notin C_n\big)$ (truncation) with respect to the grid size n

References

1. J.F. Bonnans, Z. Cen, T. Christel, Energy contracts management by stochastic programming techniques. Ann. Oper. Res. **200**, 199–222 (2012)
2. J.A. Bucklew, G.L. Wise, Multidimensional asymptotic quantization theory with r^{th} power distortion measures. IEEE Trans. Inf. Theory **28**(2), 239–247 (1982)
3. P. Capéraà, B. Van Cutsem, *Méthodes et modèles en statistique non paramétrique* (Les Presses de l'Université Laval, Sainte, 1988). Exposé fondamental. [Basic exposition], With a foreword by Capéraà, Van Cutsem and Alain Baille
4. P. Cohort, Limit theorems for random normalized distortion. Ann. Appl. Probab. **14**(1), 118–143 (2004)
5. S. Graf, H. Luschgy, *Foundations of Quantization for Probability Distributions*. Lecture Notes in Mathematics, vol. 1730 (Springer, Berlin, 2000)
6. R.M. Gray, D.L. Neuhoff, Quantization. IEEE Trans. Inf. **44**(6), 2325–2383 (1998)
7. H. Luschgy, G. Pagès, Functional quantization rate and mean regularity of processes with an application to Lévy processes. Ann. Appl. Probab. **18**(2), 427–469 (2008)
8. G. Pagès, J. Printems, www.quantize.maths-fi.com. Website devoted to quantization (2005)
9. G. Pagès, A. Sagna, Asymptotics of the maximal radius of an L^r-optimal sequence of quantizers. Bernoulli **18**(1), 360–389 (2012)
10. G. Pagès, B. Wilbertz, Optimal Delaunay and Voronoi quantization schemes for pricing American style options, in *Numerical Methods for Finance*, ed. by R. Carmona, P. Del Moral, P. Hu, N. Oudjane (Springer, New York, 2011)
11. G. Pagès, B. Wilbertz, Dual quantization for random walks with application to credit derivatives. J. Comput. Finance **16**(2), 33–60 (2012)

12. G. Pagès, B. Wilbertz, Intrinsic stationarity for vector quantization: foundation of dual quantization. SIAM J. Numer. Anal. **50**(2), 747–780 (2012)
13. J.N. Pierce, Asymptotic quantizing error for unbounded random variables. IEEE Trans. Inf. Theory **16**(1), 81–83 (1970)
14. V.T. Rajan, Optimality of the delaunay triangulation in $\mathbb{R}^d$, in *SCG '91: Proceedings of the Seventh Annual Symposium on Computational Geometry* (ACM, New York, 1991), pp. 357–363

Chapter 12
Cramér's Theorem in Banach Spaces Revisited

Pierre Petit

Abstract The text summarizes the general results of large deviations for empirical means of independent and identically distributed variables in a separable Banach space, without the hypothesis of exponential tightness. The large deviation upper bound for convex sets is proved in a nonasymptotic form; as a result, the closure of the domain of the entropy coincides with the closed convex hull of the support of the common law of the variables. Also a short original proof of the convex duality between negentropy and pressure is provided: it simply relies on the subadditive lemma and Fatou's lemma, and does not resort to the law of large numbers or any other limit theorem. Eventually a Varadhan-like version of the convex upper bound is established and embraces both results.

Keywords Cramér's theory · Large deviations · Subadditivity · Convexity · Fenchel-Legendre transformation

MSC 2010 Subject Classifications 60F10

12.1 Introduction

Cramér's original theorem (see [11]) about the large deviations of empirical means of independent and identically distributed real-valued random variables has led to an extensive literature. Several proofs of it were given by Chernoff, Bahadur, Ranga Rao, Petrov, Hammersley, and Kingman (see [2, 3, 10, 19, 22, 27]). The result was extended to higher dimensions by Sethuraman, Borovkov, Rogosin, Hoeffding, Sievers, Bartfai, and many others (see [5, 7, 21, 31–33]). At the same time, Sanov's

P. Petit (✉)
Institut de Mathématiques de Toulouse, UPS IMT, Toulouse Cedex, France
e-mail: pierre.petit@normalesup.org

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_12

theorem (see [30]) and its generalizations (see, e.g., [20]), and the study of large deviations of random processes (see, e.g., [34]) gave rise to Donsker and Varadhan's setting of large deviation principles in separable Banach spaces (see [16]). In this unifying setting, if we assume the exponential tightness of the sequence of empirical means, or equivalently the boundedness of the pressure in a neighborhood of the origin, then a full large deviation principle can be proved.

Independently, the physicist Lanford imported the subadditive argument, developed by him and Ruelle in statistical physics, into Cramér's theory (see [29] and [23]). Bahadur and Zabell (see [4]) took advantage of this new method to generalize Cramér's theory to locally convex spaces, to simplify some proofs, and to provide a good synthesis of the previous texts. By the way, they revealed that, if you replace the exponential tightness by the less restricting convex tightness, you still have the exponential decay for large deviation events associated with a convex set and the convex duality between negentropy and pressure. Among many others, the standard texts of Azencott, de Acosta, Deuschel, Stroock, Dembo, Zeitouni, and Cerf summarize the successive developments of the theory (see [1, 8, 12–14]).

Here, we prove the general results of Cramér's theory in separable Banach spaces without assuming extra hypotheses. Our arguments rely on geometrical and topological properties of Banach spaces, in the spirit of [4] and [8], and enable to complete some known partial conclusions. The main one is the large deviation upper bound for all convex sets, which is even valid in a nonasymptotic form. We deduce that the closure of the domain of the entropy coincides with the closed convex hull of the law of the variables. Another goal of the present text is to shed a new light on the theory, providing efficient and simple proofs. For instance, to prove the convex duality between the negentropy $-s$ and the pressure p, we prove the equality $p = (-s)^*$ using the convex tightness of the probability measures on a Banach space and Fatou's lemma (see [15] for a similar proof when the full large deviation principle is assumed), whereas usual proofs show the dual equality $s = -p^*$ by means of convex regularity and Cramér's theorem in $\mathbb{R}$, which in turn relies on an approximation by simpler variables (discrete in [10], bounded in [13]) and a limit theorem (Stirling's formula in [10], the law of large numbers in [13]). By the way, we intensively exploit the nice properties of convex sets to simplify proofs and establish the equivalence between convex regularity and convex tightness (which clarifies the appendix of [4]). It appears that our methods can be generalized to locally convex spaces, but technical points may have hidden the heart of our new proofs. We also show how Varadhan-like lemmas provide unifying results and, eventually, we prove a Varadhan-like lemma for concave functions which embraces both the nonasymptotic upper bound for convex sets and the equality $p = (-s)^*$.

After setting the stage and stating the results (Sect. 12.2), we first give a short proof of the weak large deviation principle (Sect. 12.3). Then we prove the large deviation upper bound for convex sets and deduce the clear identification of the closure of the domain of the entropy (Sect. 12.4). Section 12.5 is devoted to the proof of the convex duality between negentropy and pressure. Finally we prove the general convex upper bound à la Varadhan (Sect. 12.6). Except for the classic

Fenchel-Moreau theorem (see [25]), proofs of convex analysis are provided; complementary notions can be found in general texts like [25] and [28].

12.2 Setting and Results

Let $\mathscr{X}$ be a separable Banach space, $\mathscr{B}$ the Borel σ-algebra over $\mathscr{X}$, and μ a probability measure on $(\mathscr{X}, \mathscr{B})$. Let $(X_n)_{n\geqslant 1}$ be a sequence of independent and identically distributed random variables with law μ. For all $n \geqslant 1$, let $\overline{X}_n$ be the empirical mean $(X_1 + X_2 + \cdots + X_n)/n$.

Definition 12.1 The *entropy* of the sequence $(\overline{X}_n)_{n\geqslant 1}$ is the function $s : \mathscr{X} \to [-\infty, 0]$ defined by

$$\forall x \in \mathscr{X} \quad s(x) := \inf_{\varepsilon>0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(x,\varepsilon)\big)$$

where $B(x, \varepsilon)$ denotes the open ball of radius ε centered at x in $\mathscr{X}$.

By construction, the entropy s is the greatest function that satisfies the lower bound:

(LB) for all open subsets G, $\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in G\big) \geqslant \sup_{x\in G} s(x)$.

One says that the sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies a *large deviation principle* if, in addition, it satisfies the upper bound:

(UB) for all closed subsets F, $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in F\big) \leqslant \sup_{x\in F} s(x)$.

Conditions so that (UB) be satisfied, such as exponential tightness of the sequence $(\overline{X}_n)_{n\geqslant 1}$, are given in standard texts (see [1, 4, 8, 12–14, 16]). Here, as in [4] and [8], we are interested in weaker upper bounds that do not require additional hypotheses. For instance, the following result is well-known (see, e.g., [4, 16]).

Theorem 12.1 *The sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies a* weak large deviation principle*, i.e. it satisfies the compact upper bound:*

($\mathrm{UB_k}$) *for all compact subsets K,* $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in K\big) \leqslant \sup_{x\in K} s(x)$.

The upper bound is known also for open convex sets (see [4]), but the proof for closed convex sets is omitted. Here we prove the better nonasymptotic versions of them.

Theorem 12.2 *The sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies the nonasymptotic closed convex upper bound:*

($\mathrm{UB_{cc}}$) *for all closed convex subsets C and $n \geqslant 1$,* $\mathbb{P}\big(\overline{X}_n \in C\big) \leqslant \exp\Big(n \sup_{x\in C} s(x)\Big)$*;*

and the nonasymptotic open convex upper bound:

(UB$_{oc}$) *for all open convex subsets C and $n \geqslant 1$, $\mathbb{P}\big(\overline{X}_n \in C\big) \leqslant \exp\Big(n \sup_{x\in C} s(x)\Big)$.*

In particular, if C is an open convex subset, we get

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) = \sup_{n\geqslant 1} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) = \sup_{x\in C} s(x) \ .$$

The proof we give here does not rely on hypothesis ($\hat{\mathbf{C}}$) of [14, Sect. 3.1], or assumption 6.1.2 of [13], but simply on the convex tightness of μ introduced in [4] and it generalizes more easily.[1] Theorem 12.2 appears to be very convenient in the study of large deviations of means of independent and identically distributed random variables. For instance, consider the *domain* of the entropy $\mathrm{dom}(s) = \{s > -\infty\}$. Denote by co supp($\mu$) the convex hull of the support of the measure μ.

Theorem 12.3 *The closure of the domain of the entropy s is the closed convex hull of the support of the measure μ, i.e.*

$$\overline{\mathrm{dom}(s)} = \overline{\mathrm{co\ supp}(\mu)} \ .$$

The result is only partially proved in [4] and [8]. We give a complete proof. Another consequence of Theorem 12.2 is the link between entropy and pressure. Let $\mathscr{X}^*$ denote the topological dual of $\mathscr{X}$ and let $p : \mathscr{X}^* \to (-\infty, +\infty]$ be the *pressure*[2] of the sequence $(\overline{X}_n)_{n\geqslant 1}$ defined by

$$\forall \lambda \in \mathscr{X}^* \quad p(\lambda) := \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{n\lambda(\overline{X}_n)}\big) = \log \mathbb{E}\big(e^{\lambda(X_1)}\big)$$

which reduces to the log-Laplace transform of μ.

Theorem 12.4 *The pressure p and the negentropy $-s$ are convex-conjugate functions, i.e.*

$$\forall \lambda \in \mathscr{X}^* \quad p(\lambda) = \sup_{x\in\mathscr{X}} (\lambda(x) + s(x)) =: (-s)^*(\lambda) \tag{12.1}$$

and

$$\forall x \in \mathscr{X} \quad -s(x) = \sup_{\lambda\in\mathscr{X}^*} (\lambda(x) - p(\lambda)) =: p^*(x) \ . \tag{12.2}$$

[1]Hypothesis ($\hat{\mathbf{C}}$) of [14] and Assumption 6.1.2 (b) of [13] were introduced so as to complete the proofs of Appendix of [4], but they are not required to prove the first proposition of the appendix.

[2]Physically speaking, the function p should be interpreted as the opposite of a free energy, which is proportional to the pressure in the case of simple fluids.

Equation (12.2) is well-known (see, e.g., [4, 8, 13]) and standard proofs rely on three ingredients: Chebyshev's inequality, the open half-space upper bound,[3] which is a particular case of ($\mathsf{UB_{oc}}$), and Cramér's theorem in $\mathbb{R}$. Equation (12.1) follows from Eq. (12.2) by proving that p is convex and lower semi-continuous (see [8, Chapter 12]). Here we give a simple original proof of Eq. (12.1) from which we deduce Eq. (12.2). Even in $\mathscr{X} = \mathbb{R}$, it provides a new proof of Cramér's theorem (see [9]). Notice that Eq. (12.1) is similar to Varadhan's lemma (remember the first definition of the pressure p). The present proof relies on Varadhan-like versions of the lower bound and compact upper bound:

Lemma 12.1 *The sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies the lower bound à la Varadhan:*

(VLB) *for all lower semi-continuous functions $f : \mathscr{X} \to [-\infty, +\infty)$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \geqslant \sup_{x\in\mathscr{X}} (f(x) + s(x)) .$$

Lemma 12.2 *The sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies the compact upper bound à la Varadhan:*

($\mathsf{VUB_k}$) *for all upper semi-continuous functions $f : \mathscr{X} \to [-\infty, +\infty)$ such that $\{f > -\infty\}$ is relatively compact,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \leqslant \sup_{x\in\mathscr{X}} (f(x) + s(x)) .$$

Interestingly enough, Lemma 12.2 provides a Varadhan-like version of the convex upper bounds, which in turn implies Theorems 12.2 and 12.4:

Theorem 12.5 *The sequence $(\overline{X}_n)_{n\geqslant 1}$ satisfies the nonasymptotic convex upper bounds à la Varadhan:*

($\mathsf{VUB_{cc}}$) *for all upper semi-continuous concave functions $f : \mathscr{X} \to [-\infty, +\infty)$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) = \sup_{n\geqslant 1} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \leqslant \sup_{x\in\mathscr{X}} (f(x) + s(x)) ;$$

and:

($\mathsf{VUB_{oc}}$) *for all concave functions $f : \mathscr{X} \to [-\infty, +\infty)$ such that $C = \{f > -\infty\}$ is open and $f|_C$ is upper semi-continuous,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) = \sup_{n\geqslant 1} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \leqslant \sup_{x\in\mathscr{X}} (f(x) + s(x)) .$$

[3] The proof is even simpler using the closed half-space upper bound, which is a particular case of ($\mathsf{UB_{cc}}$).

12.3 Proof of Theorem 12.1

The proof of the weak large deviation principle relies on two key arguments: subadditivity and what may be called "the principle of the largest term" (see [24, Section 2]). The former is the purpose of Proposition 12.2 and the latter that of Proposition 12.3. Beforehand, we need two very handy properties of open convex sets.

Proposition 12.1 *Let C be an open convex subset of $\mathscr{X}$ containing* 0. *Then*

$$\bigcup_{t>0} tC = \mathscr{X} \ , \tag{12.3}$$

i.e. C is an absorbing *subset of $\mathscr{X}$, and*

$$\bigcup_{\delta\in(0,1)} (1-\delta)C = C \ . \tag{12.4}$$

Proof To show (12.3), let $x \in \mathscr{X}$. Since the mapping $a \in \mathbb{R} \mapsto ax \in \mathscr{X}$ is continuous and C is a neighborhood of 0, there is $\alpha > 0$ such that $\alpha x \in C$. Setting $t = 1/\alpha$, we get $x \in tC$. As for (12.4), let $x \in C$. Since the mapping $a \in \mathbb{R} \mapsto ax \in \mathscr{X}$ is continuous and C is a neighborhood of x, there is $\alpha > 0$ such that $(1+\alpha)x \in C$. Defining $\delta \in (0,1)$ by $1-\delta = 1/(1+\alpha)$, we get $x \in (1-\delta)C$, whence

$$C \subset \bigcup_{\delta\in(0,1)} (1-\delta)C$$

and the converse inclusion is trivial. □

Proposition 12.2 below is fundamental in Cramér's theory. Here is a short proof relying on the proposition above.

Proposition 12.2 *Let C be an open convex subset of $\mathscr{X}$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) = \sup_{n\geqslant 1} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) \ .$$

Proof The result is trivial if $C = \emptyset$. Now suppose $0 \in C$, otherwise consider $(X_n - x)_{n\geqslant 1}$ for some $x \in C$. Let $n, m \geqslant 1$ and write $n = qm + r$ the Euclidean division of n by m with $r \in \{1, 2, \dots, m\}$. Let $\delta \in (0,1)$. Using the convexity of C, the independence of $X_1, X_2, \dots, X_n$, and the fact that

$$\overline{X}_n = \frac{m}{n} \sum_{k=0}^{q-1} \left(\frac{1}{m} \sum_{i=km+1}^{(k+1)m} X_i \right) + \frac{1}{n} \sum_{i=mq+1}^{n} X_i \ ,$$

we get

$$\mathbb{P}\big(\overline{X}_n \in C\big) \geqslant \mathbb{P}\Big(\overline{X}_m \in \frac{n}{qm}(1-\delta)C\Big)^q \mathbb{P}\Big(X_1 \in \frac{n}{r}\delta C\Big)^r .$$

Since $r \leqslant m$ and C is an absorbing subset of $\mathscr{X}$ (see Proposition 12.1),

$$\mathbb{P}\Big(X_1 \in \frac{n\delta}{r}C\Big)^r \geqslant \mathbb{P}\Big(X_1 \in \frac{n\delta}{m}C\Big)^m \xrightarrow[n\to\infty]{} 1 .$$

Hence, remembering that $qm \leqslant n$,

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(\overline{X}_n \in C\big) &\geqslant \liminf_{n\to\infty}\frac{q}{n}\log\mathbb{P}\Big(\overline{X}_m \in \frac{n(1-\delta)}{qm}C\Big)\\ &\geqslant \frac{1}{m}\log\mathbb{P}\big(\overline{X}_m \in (1-\delta)C\big)\end{aligned}$$

and the proof is completed by taking the limit when $\delta \to 0$ (see Proposition 12.1), and then the supremum over $m \geqslant 1$. □

Notice that Proposition 12.2 is more generally valid for *algebraically open* convex sets that are measurable, i.e. measurable convex sets that satisfy properties (12.3) and (12.4) of Proposition 12.1 and their translates, i.e. measurable convex sets that are equal to their algebraic interior (see [35]).

The next simple but useful result is well-known and may be called the "principle of the largest term" (see, e.g., [24, Lemma 2.3]). We give its proof for the sake of completeness.

Proposition 12.3 *Let $u_1, u_2, \ldots, u_r$ be $[0, +\infty]$-valued sequences. Then*

$$\limsup_{n\to\infty}\frac{1}{n}\log\sum_{i=1}^{r}u_i(n) = \max_{1\leqslant i\leqslant r}\limsup_{n\to\infty}\frac{1}{n}\log u_i(n).$$

Proof From the double inequality

$$\max_{1\leqslant i\leqslant r}u_i(n) \leqslant \sum_{i=1}^{r}u_i(n) \leqslant r\max_{1\leqslant i\leqslant r}u_i(n),$$

we get

$$\limsup_{n\to\infty}\frac{1}{n}\log\sum_{i=1}^{r}u_i(n) = \limsup_{n\to\infty}\frac{1}{n}\log\max_{1\leqslant i\leqslant r}u_i(n).$$

Moreover

$$\limsup_{n\to\infty}\frac{1}{n}\log\max_{1\leqslant i\leqslant r}u_i(n)=\lim_{n\to\infty}\max_{1\leqslant i\leqslant r}\left(\sup_{k\geqslant n}\frac{1}{k}\log u_i(k)\right)$$
$$=\max_{1\leqslant i\leqslant r}\lim_{n\to\infty}\left(\sup_{k\geqslant n}\frac{1}{k}\log u_i(k)\right)\,,$$

since the function $\max : [-\infty, +\infty]^r \to [-\infty, +\infty]$ is continuous. □

Proof (Proof of Theorem 12.1) Let K be a compact subset of $\mathscr{X}$ and $\alpha > 0$. For all $x \in K$, apply Proposition 12.2 and choose $\varepsilon > 0$ such that

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(\overline{X}_n\in B(x,\varepsilon)\big)\leqslant\max(s(x)+\alpha,-1/\alpha)\;.$$

Since K is compact, there is a finite subcover $K \subset B_1 \cup B_2 \cup \cdots \cup B_r$ with $B_i = B(x_i, \varepsilon_i)$. Now apply Propositions 12.3 and 12.2 to get

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(\overline{X}_n\in K\big)\leqslant\limsup_{n\to\infty}\frac{1}{n}\log\sum_{i=1}^{r}\mathbb{P}\big(\overline{X}_n\in B_i\big)$$
$$=\max_{1\leqslant i\leqslant r}\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(\overline{X}_n\in B_i\big)$$
$$\leqslant\max_{1\leqslant i\leqslant r}\max(s(x_i)+\alpha,-1/\alpha)$$
$$\leqslant\max\left(\sup_{x\in K}s(x)+\alpha,-1/\alpha\right)$$

and finally let $\alpha \to 0$. □

12.4 Proofs of Theorems 12.2 and 12.3

To prove the convex upper bounds, we will simply extend the compact (convex) upper bound to convex sets using the convex tightness of the measures on $(\mathscr{X}, \mathscr{B})$. The idea can be traced back to [4] and the proof given here is shorter and complete.

Proposition 12.4 *Any probability measure ν on $(\mathscr{X}, \mathscr{B})$ is* convex tight*, i.e. for all $\alpha > 0$, there exists a compact convex subset K of $\mathscr{X}$ such that $\nu(K) > 1 - \alpha$.*

Proof Let ν be a probability measure on $(\mathscr{X}, \mathscr{B})$ and let $\alpha > 0$. Since $\mathscr{X}$ is metric, separable, and complete, ν is *tight*, i.e. there is a compact subset K_1 of $\mathscr{X}$ such that $\nu(K_1) > 1 - \alpha$ (see [6, Theorem 1.3]). Then $K = \overline{\mathrm{co}}(K_1)$ the closed convex hull of K_1 is compact (see [17, Theorem V.2.6]) and satisfies $\nu(K) > 1 - \alpha$. □

To prove ($\mathsf{UB_{cc}}$), we also need a fact similar to Proposition 12.2.

Proposition 12.5 *Let C be a measurable convex subset of $\mathscr{X}$. Then*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) = \sup_{n\geqslant 1} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) .$$

Proof Let $m, q \geqslant 1$. Since C is convex and $X_1, X_2, \ldots, X_{qm}$ are independent,

$$\mathbb{P}\big(\overline{X}_{qm} \in C\big) \geqslant \mathbb{P}\big(\overline{X}_m \in C\big)^q .$$

Hence

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C\big) \geqslant \limsup_{q\to\infty} \frac{1}{qm} \log \mathbb{P}\big(\overline{X}_{qm} \in C\big) \geqslant \frac{1}{m} \log \mathbb{P}\big(\overline{X}_m \in C\big) .$$

Take the supremum over $m \geqslant 1$ to conclude. □

Proof (Proof of ($\mathsf{UB_{cc}}$)*)* Let C be a closed convex subset of $\mathscr{X}$ and $N \geqslant 1$. By Proposition 12.4, the distribution of $\overline{X}_N$ is convex tight, whence, for all $\alpha > 0$, there exists a compact convex subset K of $\mathscr{X}$ such that

$$\frac{1}{N} \log \mathbb{P}\big(\overline{X}_N \in C\big) \leqslant \frac{1}{N} \log \mathbb{P}\big(\overline{X}_N \in C \cap K\big) + \alpha . \tag{12.5}$$

Applying Proposition 12.5 to the convex $C \cap K$ leads to

$$\frac{1}{N} \log \mathbb{P}\big(\overline{X}_N \in C \cap K\big) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C \cap K\big) .$$

Finally, the application of Theorem 12.1 to the compact $C \cap K$ yields

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in C \cap K\big) \leqslant \sup_{x\in C\cap K} s(x) \leqslant \sup_{x\in C} s(x) .$$

From (12.5), we get

$$\frac{1}{N} \log \mathbb{P}\big(\overline{X}_N \in C\big) \leqslant \sup_{x\in C} s(x) + \alpha .$$

Conclude by letting $\alpha \to 0$. □

A detailed observation of this last proof shows that it only requires the convex tightness of μ. Indeed, the convex tightness of μ implies the convex tightness of the distribution of $\overline{X}_N$, since, if K is convex, then

$$\mathbb{P}(\overline{X}_N \in K) \geqslant \mathbb{P}(X_1 \in K)^N .$$

This simple remark is fruitful: it permits to establish $(\mathsf{UB_{cc}})$ in a more general context and to avoid technical hypotheses. The proof of $(\mathsf{UB_{oc}})$ is in the same vein. We only need a nice property of open convex sets.

Proposition 12.6 *Let C be an open convex subset of $\mathscr{X}$ containing* 0. *Then,*

$$\bigcup_{\delta\in(0,1)} (1-\delta)\overline{C} = C \ .$$

Proof Given Proposition 12.1, it remains to show that, for all $\delta \in (0,1)$, $(1-\delta)\overline{C} \subset C$. Let $\delta \in (0,1)$ and let $x \in (1-\delta)\overline{C}$. Defining $\alpha > 0$ by $1+\alpha = 1/(1-\delta)$, we have $(1+\alpha)x \in \overline{C}$. Since $-C$ is a neighborhood of 0, $((1+\alpha)x - \alpha C) \cap C \neq \emptyset$, whence $x \in C$. □

Proposition 12.6 implies:

Proposition 12.7 *Any probability measure ν on $(\mathscr{X}, \mathscr{B})$ is* convex inner regular, *i.e. for all open convex subsets C of $\mathscr{X}$ and for all $\alpha > 0$, there exists a compact convex subset K of C such that $\nu(K) > \nu(C) - \alpha$.*

Proof Let ν be a probability measure on $\mathscr{X}$, let C be an open convex subset of $\mathscr{X}$, and let $\alpha > 0$. Using Proposition 12.4, there is a compact subset K_1 of $\mathscr{X}$ such that $\nu(K_1) > 1 - \alpha/2$. Using Proposition 12.6, we can choose $\delta \in (0,1)$ such that $\nu((1-\delta)\overline{C}) > \nu(C) - \alpha/2$. Finally, $K = K_1 \cap (1-\delta)\overline{C}$ is a compact convex subset of C such that $\nu(K) > \nu(C) - \alpha$. □

To sum up the previous proof, the convex inner regularity of a measure is equivalent to its convex tightness (in a general topological vector space). In a more general context, this argument completes the proof of [4, Appendix, Proposition 1] and gives a simpler condition than hypothesis $(\hat{\mathbf{C}})$ of [14, Sect. 3.1] or Assumption 6.1.2 of [13].

Proof (Proof of $(\mathsf{UB_{oc}})$*)* In inequality (12.5) of the proof of $(\mathsf{UB_{cc}})$, replace $C \cap K$ by a compact convex subset K of C given by Proposition 12.7 to obtain

$$\frac{1}{N} \log \mathbb{P}\big(\overline{X}_N \in C\big) \leqslant \sup_{x\in C} s(x) \ .$$

The last remark of Theorem 12.2 then follows from (LB). □

To prove Theorem 12.3, we show two intermediate and useful results. Remember that the *support* of the measure μ is the subset of $\mathscr{X}$ defined by

$$\mathrm{supp}(\mu) = \{x \in \mathscr{X} \ ; \ \forall \varepsilon > 0, \ \mu(B(x,\varepsilon)) > 0\} \ .$$

Proposition 12.8 *For any open ball B in $\mathscr{X}$,*

$$B \cap \mathrm{supp}(\mu) \neq \emptyset \iff \mu(B) > 0 \ .$$

Proof The direct implication is a mere consequence of the definition of $\text{supp}(\mu)$. And the converse one stems from the fact that $\mathscr{X}$ is second countable, so that we have $\mu(\text{supp}(\mu)) = 1$ (see [26, Theorem 2.1]). We provide another proof that relies on the convex inner regularity of μ. Consider an open ball B such that $B \cap \text{supp}(\mu) = \emptyset$. Let $\alpha > 0$. Use the convex inner regularity of μ to find a compact subset K of B such that $\mu(K) > \mu(B) - \alpha$. For all $x \in K$, there exists $\varepsilon > 0$ such that $\mu(B(x, \varepsilon)) = 0$. Extract a finite subcover $K \subset B_1 \cup B_2 \cup \cdots \cup B_r$ with $B_i = B(x_i, \varepsilon_i)$. Finally,

$$\mu(B) \leqslant \mu(K) + \alpha \leqslant \sum_{i=1}^{r} \mu(B_i) + \alpha = \alpha$$

and let $\alpha \to 0$. □

Proposition 12.9 *The entropy s is upper semi-continuous and concave.*

Proof To show that s is upper semi-continuous, take $t \in \mathbb{R}$ and $x \in \mathscr{X}$ such that $s(x) < t$. By the very definition of s, there is $\varepsilon > 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(x, \varepsilon)\big) < t \ .$$

For all $y \in B(x, \varepsilon)$, take δ such that $B(y, \delta) \subset B(x, \varepsilon)$ and write

$$s(y) \leqslant \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(y, \delta)\big) \leqslant \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(x, \varepsilon)\big) < t \ .$$

So s is upper semi-continuous. Now we prove that s is concave. Let $x, y \in \mathscr{X}$ and set $z = (x + y)/2$. Let $\varepsilon > 0$ and set $B_z = B(z, \varepsilon)$, $B_x = B(x, \varepsilon/2)$, and $B_y = B(y, \varepsilon/2)$. For all $n \geqslant 1$,

$$\mathbb{P}\big(\overline{X}_{2n} \in B_z\big) \geqslant \mathbb{P}\big(\overline{X}_n \in B_x\big)\mathbb{P}\big(\overline{X}_n \in B_y\big)$$

whence

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{2n} \log \mathbb{P}\big(\overline{X}_{2n} \in B_z\big) &\geqslant \lim_{n\to\infty} \frac{1}{2n} \log \big(\mathbb{P}\big(\overline{X}_n \in B_x\big)\mathbb{P}\big(\overline{X}_n \in B_y\big)\big) \\ &\geqslant \frac{s(x) + s(y)}{2} \ .\end{aligned}$$

Taking the infimum in ε, we get $s((x + y)/2) \geqslant (s(x) + s(y))/2$ and the concavity of s follows, since s is upper semi-continuous. □

Proof (Proof of Theorem 12.3) Since s is concave (see Proposition 12.9), $\text{dom}(s)$ is a convex subset of $\mathscr{X}$, so we only need to prove

$$\text{supp}(\mu) \subset \overline{\text{dom}(s)} \tag{12.6}$$

and

$$\operatorname{dom}(s) \subset \overline{\text{co}\,\operatorname{supp}(\mu)}\ . \tag{12.7}$$

Let $x \notin \overline{\operatorname{dom}(s)}$ and $\varepsilon > 0$ such that $B(x,\varepsilon) \cap \operatorname{dom}(s) = \emptyset$. The bound $(\mathsf{UB_{oc}})$ implies $\log \mu(B(x,\varepsilon)) = -\infty$. With Proposition 12.8, we get $B(x,\varepsilon) \cap \operatorname{supp}(\mu) = \emptyset$, so inclusion (12.6) is proved. Now, let $x \in \operatorname{dom}(s)$ and $\varepsilon > 0$. Showing that $B(x,\varepsilon) \cap \text{co}\,\operatorname{supp}(\mu) \neq \emptyset$ is enough to prove inclusion (12.7). There is $n \geqslant 1$ such that $\mathbb{P}(\overline{X}_n \in B(x,\varepsilon/2)) > 0$, i.e. $\mu^{\otimes n}(C) > 0$ where

$$C = \left\{(u_1, u_2, \ldots, u_n) \in \mathscr{X}^n\,;\ \frac{u_1 + u_2 + \cdots + u_n}{n} \in B(x, \varepsilon/2)\right\}.$$

Let Q be a countable dense subset of $\mathscr{X}$. Since C is an open subset of $\mathscr{X}^n$, $Q^n \cap C$ is a dense subset of C, whence

$$C \subset \bigcup_{(u_1,\ldots,u_n) \in Q^n \cap C} \prod_{i=1}^{n} B(u_i, \varepsilon/2)\ .$$

Since the union is countable and $\mu^{\otimes n}(C) > 0$, there is $(u_1, u_2, \ldots, u_n) \in C$ such that, for all integers $i \in \{1, 2, \ldots, n\}$, $\mu(B(u_i, \varepsilon/2)) > 0$. So, by Proposition 12.8, for all integers $i \in \{1, 2, \ldots, n\}$, there is $y_i \in B(u_i, \varepsilon/2) \cap \operatorname{supp}(\mu)$. Hence,

$$y := \frac{y_1 + y_2 + \cdots + y_n}{n} \in B\Big(\frac{u_1 + u_2 + \cdots + u_n}{n}, \varepsilon/2\Big) \subset B(x, \varepsilon)$$

and $y \in \text{co}\,\operatorname{supp}(\mu)$. □

Note that Theorem 12.3 implies Theorem 2.4(a), (b) of [4] and results 9.7 and 9.8 of [8].

12.5 Proof of Theorem 12.4

The *Fenchel-Legendre transform* of a function $g : \mathscr{X} \to [-\infty, +\infty]$ is the function on the dual space $g^* : \mathscr{X}^* \to [-\infty, +\infty]$ defined by

$$\forall \lambda \in \mathscr{X}^* \quad g^*(\lambda) = \sup_{x \in \mathscr{X}} \big(\lambda(x) - g(x)\big)\ .$$

Similarly, the Fenchel-Legendre transform of a function $h : \mathscr{X}^* \to [-\infty, +\infty]$ is the function $h^* : \mathscr{X} \to [-\infty, +\infty]$

$$\forall x \in \mathscr{X} \quad h^*(x) = \sup_{\lambda \in \mathscr{X}^*} \big(\lambda(x) - h(\lambda)\big)\ .$$

We say that the functions $g : \mathscr{X} \to [-\infty, +\infty]$ and $h : \mathscr{X}^* \to [-\infty, +\infty]$ are *convex conjugate functions* if $g^* = h$ and $h^* = g$.

Proposition 12.10 (Fenchel-Moreau Theorem) *A function* $g : \mathscr{X} \to (-\infty, +\infty]$ *satisfies* $g^{**} = g$ *if and only if* g *is lower semi-continuous and convex.*

Proof See, e.g., [25, Section 5]. □

Proof (Proof of Theorem 12.4) Knowing that s is upper semi-continuous and concave (see Proposition 12.9), and applying Proposition 12.10, we only need to prove $p = (-s)^*$. The classic proof of the inequality $p \geqslant (-s)^*$, or its equivalent $s \leqslant -p^*$, relies on Chebyshev's inequality (see, e.g., [4, Theorem 3.1]). Another proof consists in applying Lemma 12.1 (the proof of which is given below) to the continuous functions $f = \lambda \in \mathscr{X}^*$. The other inequality $p \leqslant (-s)^*$, or its equivalent $s \geqslant -p^*$, is usually proved via the open half-space upper bound and Cramér's theorem in $\mathbb{R}$ (see, e.g., [4, Section 3]). Let us see how we can get it via Lemma 12.2 (the proof of which is given below). Let $\lambda \in \mathscr{X}^*$ and let $\alpha > 0$. Since μ is convex tight (see Proposition 12.4) and using Fatou's lemma, there exists a compact convex subset K of $\mathscr{X}$ such that

$$\min\left(\log \mathbb{E}\big(e^{\lambda(X_1)}\big) - \alpha, 1/\alpha\right) \leqslant \log \mathbb{E}\big(e^{\lambda(X_1)}\mathbf{1}_K(X_1)\big) \ .$$

Since K is convex, for all $n \geqslant 1$, the conjunction of $X_1 \in K$, $X_2 \in K$, …, and $X_n \in K$ implies $\overline{X}_n \in K$. Hence, using the independence of the X_i's, we get

$$\begin{aligned}\log \mathbb{E}\big(e^{\lambda(X_1)}\mathbf{1}_K(X_1)\big) &\leqslant \inf_{n\geqslant 1} \frac{1}{n} \log \mathbb{E}\big(e^{n\lambda(\overline{X}_n)}\mathbf{1}_K(\overline{X}_n)\big)\\ &\leqslant \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{n(\lambda+\chi_K)(\overline{X}_n)}\big)\end{aligned}$$

where

$$\chi_K = \log \mathbf{1}_K$$

is the characteristic function of the convex set K. Finally, we apply Lemma 12.2 to the upper semi-continuous function $f = \lambda + \chi_K$ for which $\{f > -\infty\} = K$ is compact and we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{n(\lambda+\chi_K)(\overline{X}_n)}\big) \leqslant \sup_{x\in\mathscr{X}} (\lambda(x) + \chi_K(x) + s(x)) \leqslant (-s)^*(\lambda) \ .$$

Conclude the proof by letting $\alpha \to 0$. □

Proof (Proof of Lemma 12.1) Let $f : \mathscr{X} \to [-\infty, +\infty]$ be a lower semi-continuous function. Let $x \in \mathscr{X}$ and let $\alpha > 0$. There is $\varepsilon > 0$ such that, for all $y \in B(x, \varepsilon)$,

$$f(y) \geqslant \min(f(x) - \alpha, 1/\alpha) .$$

Hence,

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) &\geqslant \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)} \mathbf{1}_{B(x,\varepsilon)}(\overline{X}_n)\big) \\ &\geqslant \min(f(x) - \alpha, 1/\alpha) + \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(x, \varepsilon)\big) \\ &\geqslant \min(f(x) - \alpha, 1/\alpha) + s(x) .\end{aligned}$$

Taking the limit when $\alpha \to 0$ and the supremum over $x \in \mathscr{X}$, we get

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \geqslant \sup_{x\in\mathscr{X}} (f(x) \dot{+} s(x))$$

where $\dot{+}$ is the natural extension of the addition verifying $(-\infty) \dot{+} (+\infty) = -\infty$. The result reduces to (VLB) when $\{f = +\infty\} = \emptyset$. □

Proof (Proof of Lemma 12.2) Let $f : \mathscr{X} \to [-\infty, +\infty]$ be an upper semi-continuous function such that $K := \{f > -\infty\}$ is relatively compact. Let $\alpha > 0$. For all $x \in \mathscr{X}$, there is $\varepsilon > 0$ such that, for all $y \in B(x, \varepsilon)$,

$$f(y) \leqslant \max(f(x) + \alpha, -1/\alpha).$$

By the definition of $s(x)$ and Proposition 12.2, should we reduce ε, we may suppose that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\overline{X}_n \in B(x, \varepsilon)\big) \leqslant \max(s(x) + \alpha, -1/\alpha) .$$

Extract a finite subcover $\overline{K} \subset B_1 \cup B_2 \cup \cdots \cup B_r$ with $B_i = B(x_i, \varepsilon_i)$. For all $n \geqslant 1$,

$$\begin{aligned}\frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) &\leqslant \frac{1}{n} \log \sum_{i=1}^{r} \mathbb{E}\big(e^{nf(\overline{X}_n)} \mathbf{1}_{B_i}(\overline{X}_n)\big) \\ &\leqslant \frac{1}{n} \log \sum_{i=1}^{r} e^{n \max\left(f(x_i)+\alpha, -1/\alpha\right)} \mathbb{P}\big(\overline{X}_n \in B_i\big) .\end{aligned}$$

Taking the limit superior when $n \to \infty$ and applying the principle of the largest term (Proposition 12.3), we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big)$$
$$\leqslant \max_{1\leqslant i\leqslant r} \big(\max(f(x_i)+\alpha, -1/\alpha) + \max(s(x_i)+\alpha, -1/\alpha)\big)$$
$$\leqslant \sup_{x\in\mathscr{X}} \big(\max(f(x)+\alpha, -1/\alpha) + \max(s(x)+\alpha, -1/\alpha)\big) .$$

Letting $\alpha \to 0$, we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) \leqslant \sup_{x\in\mathscr{X}} (f(x) \dot{+} s(x))$$

where $\dot{+}$ is the natural extension of the addition such that $(-\infty) \dot{+} (+\infty) = +\infty$. The result reduces to $(\mathsf{VUB}_\mathsf{k})$ when $\{f = +\infty\} = \emptyset$. □

12.6 Proof of Theorem 12.5

The proof of Theorem 12.5 is a slight variant of that of Theorem 12.2. We need here a complete version of the subadditive lemma due to Fekete (see [18]). It is very well known when u is finite valued with a proof similar to that of Propositions 12.2 and 12.5.

Proposition 12.11 *Let u be a $[-\infty, +\infty]$-valued sequence. Suppose that u is* subadditive, *i.e. for all $m, n \geqslant 1$, $u(m+n) \leqslant u(m) \dot{+} u(n)$, where $\dot{+}$ is the natural extension of the addition such that $(-\infty) \dot{+} (+\infty) = +\infty$. Then*

$$\liminf_{n\to\infty} \frac{u(n)}{n} = \inf_{n\geqslant 1} \frac{u(n)}{n} . \tag{12.8}$$

If u is also controlled, *i.e. there is $N \geqslant 1$ such that, for all $n \geqslant N$, $u(n) < +\infty$, then*

$$\lim_{n\to\infty} \frac{u(n)}{n} = \inf_{n\geqslant 1} \frac{u(n)}{n} . \tag{12.9}$$

Proof Let u be a subadditive $[-\infty, +\infty]$-valued sequence. For $m \geqslant 1$, we have

$$\liminf_{n\to\infty} \frac{u(n)}{n} \leqslant \liminf_{q\to\infty} \frac{u(qm)}{qm} \leqslant \frac{u(m)}{m}$$

and Eq. (12.8) follows by taking the infimum over $m \geqslant 1$. Now suppose that u is also controlled. Let $m \geqslant N$. For all $n \geqslant m$, write $n = qm + r$ the Euclidean division of n by m with $r \in \{m, m+1, \dots, 2m-1\}$ and

$$u(n) \leqslant qu(m) \dotplus u(r) \leqslant \frac{n}{m} u(m) \dotplus \max_{m \leqslant i < 2m} u(i) .$$

Since, for all $i \geqslant m$, $u(i) < +\infty$, we get

$$\limsup_{n \to \infty} \frac{u(n)}{n} \leqslant \frac{u(m)}{m}$$

and Eq. (12.9) follows by taking the infimum over $m \geqslant 1$. □

We immediately deduce the useful property:

Proposition 12.12 *Let $f : \mathscr{X} \to [-\infty, +\infty]$ be a $\dotplus$-*concave *function, i.e. for all $x, y \in \mathscr{X}$ and $t \in (0, 1)$,*

$$f\big((1-t)x + ty\big) \geqslant (1-t) f(x) \dotplus t f(y) ,$$

where $\dotplus$ is the natural extension of the addition verifying $(-\infty) \dotplus (+\infty) = -\infty$. Then

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) = \sup_{n \geqslant 1} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) . \tag{12.10}$$

If, moreover, $C = \{f > -\infty\}$ is open, then

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) = \sup_{n \geqslant 1} \frac{1}{n} \log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big) . \tag{12.11}$$

Proof Let $f : \mathscr{X} \to [-\infty, +\infty]$ be a $\dotplus$-*concave* function. For all integers $m, n \geqslant 1$, since $(m+n) f(\overline{X}_{m+n}) \geqslant m f(\overline{X}_m) \dotplus n f((X_{m+1} + \cdots + X_{m+n})/n)$, we get

$$\mathbb{E}\big(e^{(m+n) f(\overline{X}_{m+n})}\big) \geqslant \mathbb{E}\big(e^{mf(\overline{X}_m)}\big)\mathbb{E}\big(e^{nf(\overline{X}_n)}\big) ,$$

so $u(n) := -\log \mathbb{E}\big(e^{nf(\overline{X}_n)}\big)$ is a subadditive sequence and Eq. (12.10) stems from Proposition 12.11. Suppose that $C = \{f > -\infty\}$ is open. Then, either, for all $n \geqslant 1$, $u(n) = +\infty$ and Eq. (12.11) is trivial; or there exists $m \geqslant 1$ such that $u(m) < +\infty$. Then $\mathbb{P}(\overline{X}_m \in C) > 0$. Using Proposition 12.2, we find that there exists $N \geqslant 1$ such that, for all $n \geqslant N$, $\mathbb{P}(\overline{X}_n \in C) > 0$, whence $u(n) < +\infty$. So u is controlled and Eq. (12.11) stems from Proposition 12.11. □

Proof (Proof of (VUB$_{\text{cc}}$)*)* Let $f : \mathscr{X} \to [-\infty, +\infty]$ be an upper semi-continuous $\dotplus$-concave function (if $\{f = +\infty\} = \emptyset$, f is simply upper semi-continuous and concave). The first equality stems from Proposition 12.12. Let us prove the inequality. Let $\alpha > 0$. Choose $N \geqslant 1$ such that

$$\min\Big(\sup_{n\geqslant 1}\frac{1}{n}\log\mathbb{E}\big(e^{nf(\overline{X}_n)}\big)-\alpha,\,1/\alpha\Big)\leqslant\frac{1}{N}\log\mathbb{E}\big(e^{Nf(\overline{X}_N)}\big)\,.$$

Let $\beta > 0$. By Proposition 12.4, the distribution of $\overline{X}_N$ is convex tight. Using Fatou's lemma, there exists a compact convex subset K of $\mathscr{X}$ such that

$$\min\Big(\frac{1}{N}\log\mathbb{E}\big(e^{Nf(\overline{X}_N)}\big)-\beta,\,1/\beta\Big)\leqslant\frac{1}{N}\log\mathbb{E}\big(e^{Nf(\overline{X}_N)}\mathbf{1}_K(\overline{X}_N)\big)\,. \qquad (12.12)$$

Applying Proposition 12.12 to the $\dot{+}$-concave function $f\dot{+}\chi_K$, we get

$$\frac{1}{N}\log\mathbb{E}\big(e^{Nf(\overline{X}_N)}\mathbf{1}_K(\overline{X}_N)\big)\leqslant\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big(e^{n(f\dot{+}\chi_K)(\overline{X}_n)}\big)\,.$$

Finally, we apply Lemma 12.2 (more precisely the slight generalization appearing in its proof) to the upper semi-continuous function $f\dot{+}\chi_K$ and get

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big(e^{n(f\dot{+}\chi_K)(\overline{X}_n)}\big)\leqslant\sup_{x\in\mathscr{X}}\big(f(x)\dot{+}s(x)\big)\,.$$

Conclude by letting $\alpha, \beta \to 0$. □

Proof (Proof of (VUB$_{\mathsf{oc}}$)*)* Let $f:\mathscr{X}\to[-\infty,+\infty]$ be a $\dot{+}$-concave function such that $C=\{f>-\infty\}$ is open and $f|_C$ is upper semi-continuous. The first equality stems from Proposition 12.12. To prove the inequality, suppose that, in inequality (12.12), K is a compact convex subset of C (see Proposition 12.7) and notice that $f\dot{+}\chi_K$ is upper semi-continuous. □

Acknowledgements I would like to thank Raphaël Cerf and Yann Fuchs for their careful reading, and the referee for his suggestions.

References

1. R. Azencott, Grandes déviations et applications, in *École d'Été de Probabilités de Saint-Flour VIII-1978*. Lecture Notes in Mathematics, vol. 774 (Springer, Berlin, 1980)
2. R.R. Bahadur, *Some Limit Theorems in Statistics* (SIAM, Philadelphia, 1971)
3. R.R. Bahadur, R. Ranga Rao, On deviations of the sample mean. Ann. Math. Stat. **31**(4), 1015–1027 (1960)
4. R.R. Bahadur, S.L. Zabell, Large deviations of the sample mean in general vector spaces. Ann. Probab. **7**(4), 587–621 (1979)
5. P. Bártfai, Large deviations of the sample mean in Euclidean spaces. Mimeograph Series No. 78-13, Statistics Department, Purdue University (1978)

6. P. Billingsley, *Convergence of Probability Measures*, 2nd edn. Wiley Series in Probability and Statistics: Probability and Statistics (Wiley, New York, 1999)
7. A.A. Borovkov, B.A. Rogozin, О центральной предельной теореме в многомерном случае. Teor. Verojatnost. i Primenen. **10**(1), 61–69 (1965). English translation: On the multidimensional central limit theorem. Theory Probab. Appl. **10**(1), 55–62 (1965)
8. R. Cerf, On Cramér's theory in infinite dimensions. Panoramas et Synthèses 23. Société Mathématique de France, Paris (2007)
9. R. Cerf, P. Petit, A short proof of Cramér's theorem in $\mathbb{R}$. Am. Math. Mon. **118**(10), 925–931 (2011)
10. H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Stat. **23**, 493–507 (1952)
11. H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités. Actual. Sci. Indust. **736**, 5–23 (1938)
12. A. De Acosta, On large deviations of sums of independent random vectors, in *Probability in Banach Spaces V*. Lecture Notes in Mathematics, vol. 1153 (Springer, Berlin, 1985)
13. A. Dembo, O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edn. (Springer, New York, 1998). First edition by Jones and Bartlett in 1993
14. J.D. Deuschel, D.W. Stroock, *Large Deviations* (Academic, New York, 1989)
15. I.H. Dinwoodie, Identifying a large deviation rate function. Ann. Probab. **21**(1), 216–231 (1993)
16. M.D. Donsker, S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time III. Commun. Pure Appl. Math. **29**, 389–461 (1976)
17. N. Dunford, J.T. Schwartz, *Linear Operators. Part I: General Theory* (Wiley, New York, 1958)
18. M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit. ganzzahligen Koeffizienten. Math. Z. **17**, 228–249 (1923)
19. J.M. Hammersley, Postulates for subadditive processes. Ann. Probab. **2**(4), 652–680 (1974)
20. A.B. Hoadley, On the probability of large deviations of functions of several empirical cdf's. Ann. Math. Stat. **38**, 360–381 (1967)
21. W. Hoeffding, On probabilities of large deviations, in *Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, CA, 1965/1966. Vol. I: Statistics (University of California Press, Berkeley, 1967), pp. 203–219
22. J.F.C. Kingman, Subadditive Processes, in *École d'Été de Probabilités de Saint-Flour V-1975*. Lecture Notes in Mathematics, vol. 539 (Springer, Berlin, 1976)
23. O.E. Lanford, Entropy and equilibrium states in classical statistical mechanics, in *Statistical Mechanics and Mathematical Problems*. Lecture Notes in Physics, vol. 20 (Springer, Berlin, 1973)
24. J.T Lewis, C.E. Pfister, W.G. Sullivan, Entropy, concentration of probability and conditional limit theorems. Markov Processes Relat. Fields **1**(3), 319–386 (1995)
25. J.J. Moreau, Fonctionnelles convexes. Séminaire sur les Équations aux Dérivées Partielles, Collège de France (1966–1967)
26. K.R. Parthasarathy, *Probability Measures on Metric Spaces* (AMS Chelsea Publishing, Providence, 2005). Reprint of the 1967 original
27. V.V. Petrov, О вероятностях больших уклонений сумм независимых случайных величин. Teor. Verojatnost. i Primenen **10**(2), 310–322 (1965). English translation: On the probabilities of large deviations for sums of independent random variables. Theory Probab. Appl. **10**(2), 287–298 (1965)
28. W. Rudin, *Functional Analysis* (McGraw-Hill, New York, 1973)
29. D. Ruelle, Correlation functionals. J. Math. Phys. **6**(2), 201–220 (1965)
30. I.N. Sanov, О вероятности больших отклонений случайных величин. Mat. Sb. **42**(1), 11–44 (1957). English translation: On the probability of large deviations of random variables. Sel. Transl. Math. Statist. Probab. I, 213–244 (1961)
31. J. Sethuraman, On the probability of large deviations of families of sample means. Ann. Math. Stat. **35**, 1304–1316 (1964)

32. J. Sethuraman, On the probability of large deviations of the mean for random variables in $D[0,\ 1]$. Ann. Math. Stat. **36**, 280–285 (1965)
33. G.L. Sievers, Multivariate probabilities of large deviations. Ann. Stat. **3**(4), 897–905 (1975)
34. S.R.S. Varadhan, Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261–286 (1966)
35. C. Zălinescu, *Convex Analysis in General Vector Spaces* (World Scientific, River Edge, 2002)

Chapter 13
On Martingale Chaoses

B. Rajeev

Abstract We extend Wiener's notion of 'homogeneous' chaos expansion of Brownian functionals to functionals of a class of continuous martingales via a notion of iterated stochastic integral for such martingales. We impose a condition of 'homogeneity' on the previsible sigma field of such martingales and show that under this condition the notions of purity, chaos representation property and the predictable representation property all coincide.

Keywords Martingale representation · Stochastic integral representation · Chaos expansion · CRP · PRP · Pure martingales

Mathematics Subject Classification (2000) Primary 60H10, 60H15; Secondary 60J60, 35K15

13.1 Introduction

The chaos expansion of Brownian functionals [10, 11, 13, 16, 28] has become the principal tool in what is called 'stochastic analysis'. Among other things an important application is to the notion of the stochastic derivative and its adjoint, the Skorokhod integral (see [11, 16, 18, 26]). Chaos expansions have been extended to Lévy processes (see [3, 4, 14, 19, 21, 22]), Azéma martingales [5–7]. The chaos expansion property is also related to Fock space expansions which are used in physics [8, 24] and in quantum probability [20].

In this paper we introduce the notion of chaos expansion for d-dimensional continuous local martingales $Y = (Y^1, \cdots, Y^d)$, which are pairwise orthogonal

B. Rajeev (✉)
Statistics and Mathematics Unit, Indian Statistical Institute, Bangalore Centre, Bangalore, India
e-mail: brajeev@isibang.ac.in

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_13

and have a common quadratic variation process $<Y>$ that increases to infinity with time, almost surely. These expansions are in terms of iterated, multiple stochastic integrals, denoted as $I^Y_{n,\alpha}(f)(t,\omega)$, $\alpha = (\alpha_1, \cdots, \alpha_n)$ where $1 \le \alpha_i \le d$, the vector α determines the order of integration, and $f \in L^2(\Delta_n)$, where Δ_n is the same as for Brownian motion (see Sect. 13.3 for definition). For example, an iterated integral of order two i.e. $n = 2$ and $\alpha_1 = 2, \alpha_2 = 3$ and $d > 3$ would be an integral of the form

$$I^Y_{2,\alpha}(f)(t,\omega) := \int_0^t \int_0^{t_2} f(<Y>_{t_1}, <Y>_{t_2})\, dY^2_{t_1}\, dY^3_{t_2}.$$

Such integrals have been defined in [17] for $d = 1$, when the martingale Y satisfies $<Y>_t = t$. It turns out in the present case that such integrals when evaluated at $t = \infty$, have all the properties as in the Brownian case viz. linearity, isometry and orthogonality. Indeed, such integrals turn out to be closely related (by time change) to the corresponding integral $I^B_{n,\alpha}(f)(t,\omega)$ for a Brownian motion $B = (B^1, \cdots, B^d)$, where we have the Dambis-Dubins-Schwartz (DDS) representation $Y = B \circ <Y>$ ([12], Chap. 2, Sect. 7). We use this relation as a definition of $I^Y_{n,\alpha}(f)$ and deduce all the properties from that of $I^B_{n,\alpha}(f)$ (Definition 13.2, and Theorem 13.2 in Sect. 3). In [23] we provide an alternate definition of $I^Y_{n,\alpha}(f)$ from first principles; in particular it does not use the definition for Brownian motion.

An immediate question that arises is whether the chaotic representation property (CRP) holds for the process Y i.e. whether every L^2 functional of the process Y can be expanded in terms of elements from its 'chaotic' subspaces. One answer to this question is an immediate consequence of the representation $Y = B \circ <Y>$ and time change viz. that the CRP is true iff Y is pure (see [25, 27], for the notion of purity and Proposition 13.11, Sect. 5 below).

A closely related notion to CRP is that of the predictable representation property (PRP) (see [25], Chap. V) and a natural question is its relationship with CRP and purity. It is known that there are martingales which have the PRP but are not pure. On the other hand pure martingales are extremal ([25], Chap. V, Proposition (4.11)) and hence have the PRP. We show that a certain measure theoretic condition on the previsible sigma field $\mathscr{P}(Y)$ of Y is sufficient for the equivalence of the PRP, CRP and the strong PRP, a notion that we introduce (see Theorem 13.4, Sect. 13.5, below).

To explain this further, we introduce a sigma field $\eta^{-1}\mathscr{P}(B)$ which is defined as the pull back of the previsible sigma field of B viz. $\mathscr{P}(B)$, by the map $\eta(s,\omega) := (<Y>_s, \omega)$ (see Definition 13.1(4), Sect. 13.2 below). Let $\phi(t,\omega) \equiv t$. Then our measure theoretic condition says that $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$ a.e. $\mu := d{<Y>}dP$ (Theorem 13.4) and a necessary condition is that the process $\phi(t,\omega) \equiv t$ should be measurable in the sigma field $\eta^{-1}\mathscr{P}(B)$. This latter condition may be viewed as a homogeneity condition on $\mathscr{P}(Y)$ (see Remark 13.1 following Corollary 13.2 to Theorem 13.4 in Sect. 13.5). In Theorem (5), following (Émery, 2015, personal communication), we show that when Y is pure this condition is also sufficient for the equality, $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$ to hold up to evanescent sets. We also obtain necessary and sufficient conditions for the law of the translates of Y, by processes of the form $h \circ \langle Y \rangle$ where h is a deterministic function on $[0, \infty)$, to be equivalent to the law of Y (Theorem 13.7) when Y is pure. In other words, we characterize the Cameron-Martin subspace for a pure martingale Y [1, 9].

The paper is organised as follows: in Sect. 13.2 we bring together some measure theoretic preliminaries. The definition of the iterated stochastic integrals and its properties are in Sect. 13.3. In Sect. 13.4 we introduce the space of kernels that arise in the chaos expansion. In Sect. 13.5 we discuss the relationship between $\mathscr{P}(Y)$ and $\eta^{-1}\mathscr{P}(B)$. We introduce a notion of 'strong PRP' and show that it is equivalent to that of CRP and purity. We discuss the equivalence between CRP and PRP (Theorem 13.4, Corollaries 13.1 and 13.2) using the measure theoretic condition $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$ and also discuss the connection between this condition and purity. In Sect. 13.6 we prove a version of the Cameron-Martin-Girsanov theorem for pure martingales.

13.2 Preliminaries

Let $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ be a filtered probability space with $(\mathscr{F}_t)$ taken to be right continuous and $\mathscr{F}_0$ containing all P null sets. Let $Y^1, \cdots, Y^d$ be continuous $\mathscr{F}_t$ local martingales with quadratic variation process $(< Y^i >_t)$, $i = 1, \cdots, d$, $Y_0^i \equiv 0$ and $< Y^i, Y^j >_t = 0, i \neq j, t \geq 0$, a.s. We will also assume throughout this paper that for $i \neq j$, $< Y^i >_t = < Y^j >_t =: \langle Y \rangle_t$. Let $Y_t := (Y_t^1, \cdots, Y_t^d)$. For $n \geq 1$, $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\alpha_i \in \{1, \cdots, d\}$, let (Y_t^α) be the $\mathbb{R}^n$ valued process defined by $Y_t^\alpha := (Y_t^{\alpha_1}, \cdots, Y_t^{\alpha_n})$. The right continuous inverse of $\langle Y \rangle$ will be denoted by τ. We will always assume that a.s., $\langle Y \rangle_\infty = \infty$. Then note that $Y = B \circ \langle Y \rangle$ where $B := (B^1, \cdots, B^d)$ is a standard d-dimensional Brownian motion (see [12, Thms. 7.2,7.3, Chapter 2]). For a filtration $\mathscr{F}_t$, $\mathscr{P}(\mathscr{F}_t)$ will denote the previsible sigma field associated with $\mathscr{F}_t$. The filtration generated by a process $X = (X_t)$ will be denoted by $\mathscr{F}^X := (\mathscr{F}_t^X)$. The corresponding previsible sigma field will be denoted by $\mathscr{P}(X) := \mathscr{P}(\mathscr{F}^X)$.

For a sigma field $\mathscr{G}$ on a set G and a function $T : (G, \mathscr{G}) \to (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, we use the notation $T \in \mathscr{G}$ to say that T is measurable with respect to $\mathscr{G}$ and the sigma field in the range viz. $\mathscr{B}(\mathbb{R}^d)$. Consider $\Omega' := [0, \infty) \times \Omega$ with the product sigma field $\mathscr{F}' = \mathscr{B}[0, \infty) \times \mathscr{F}_\infty^Y$, where for $0 \leq t < \infty$, $\mathscr{F}_t^Y := \sigma\{Y_s; 0 \leq s \leq t\}$ and $\mathscr{F}_\infty^Y = \sigma\{Y_s; 0 \leq s < \infty\}$. We make the following definitions, some of which are well known.

Definition 13.1 We make the following definitions, some of which are well known and others that are frequently used in the sequel.

1. We recall [27] that a continuous local martingale Y with the representation $Y = B \circ \langle Y \rangle$ as above is pure iff $\mathscr{F}_\infty^Y = \mathscr{F}_\infty^B$. In Sect. 13.5 below we will make a closer study of this notion.

2. For $A \in \mathscr{F}'$, we define $\mu(A) := E\int_0^\infty I_A(s,\omega)\,d{<}Y{>}_s,\ A \in \mathscr{F}'$. When ${<}Y{>}_t = t$ we shall denote the corresponding measure by μ_0.
3. Recall ([2], p. 86, Chapter IV, paras 6–8) that a set $E \in \mathscr{F}'$ is said to be evanescent iff $\exists\ A \in \mathscr{F},\ P(A) = 0$ such that the set $\{\omega : \exists\ t \geq 0, (t,\omega) \in E\} \subset A$. Note that a set E is evanescent iff the process $I_E(t,\omega)$ is indistinguishable from the zero process.
4. Let η be the map $\eta : \Omega' := [0,\infty)\times\Omega \to \Omega'$ defined as $\eta(s,\omega) := ({<}Y{>}_s, \omega)$. We define the sigma field $\eta^{-1}\mathscr{P}(B)$ on Ω' as follows:

$$\eta^{-1}\mathscr{P}(B) := \{A :\ A = \eta^{-1}A',\ A' \in \mathscr{P}(B)\}.$$

5. We define the map $\phi(t,\omega) := t$ for all $(t,\omega) \in [0,\infty)\times\Omega$. □

We will complete $\mathscr{F}'$ by all the μ null sets, and use the same notation for the completed sigma field. Define for $t \geq 0$, the stochastic process $Y^t : [0,\infty)\times\Omega \to \mathbb{R}^d$ by $Y^t_s(\omega) := Y_{t\wedge s}(\omega)$.

Before we state the next proposition we make some remarks on the map ϕ in Definition 13.1(5). For a fixed $t \geq 0$, $\phi(t,\omega)$ is the constant random variable t and in particular generates the trivial sigma field on Ω. However the situation is different when we let t vary and consider sigma fields on Ω' instead of Ω. The sigma field generated by ϕ on Ω' is easily seen to be $\mathscr{B}[0,\infty)\times\mathscr{E}$ where $\mathscr{E}$ denotes the trivial sigma field on Ω. The adjunction of ϕ to a collection of maps can crucially alter the structure of the generated sigma field on Ω'. In particular we note that in general, $\sigma\{Y^t : t \geq 0\} \neq \sigma\{\phi, Y^t : t \geq 0\}$ as can easily be seen by considering a stopped process (Y_s); for example, one that satisfies $Y_s = Y_{s\wedge 1}$ for every $s \geq 0$. We then have the following proposition.

Proposition 13.1 $\mathscr{P}(Y) = \sigma\{\phi, Y^t : t \geq 0\}$.

Proof It suffices to show that $\mathscr{P}(Y) \subseteq \sigma\{\phi, Y^t : t \geq 0\}$, the reverse inclusion being immediate from the continuity and adaptedness of the maps involved. Again it suffices to show that for every $t > 0$, and $A \in \mathscr{F}^Y_t$ the set $(t,\infty)\times A \in \sigma\{\phi, Y^t : t \geq 0\}$. By a monotone class argument it suffices to take A to be a finite dimensional set:

$$A = \{\omega : (Y_{t_1}, \cdots, Y_{t_n}) \in B\}$$

where $0 \leq t_1 < t_2 < \cdots < t_n \leq t$ and B a Borel set in $\mathbb{R}^{dn}$. But then

$$(t,\infty)\times A = \{(s,\omega) : \phi(s) > t,\ (Y^{t_1}_s(\omega)), \cdots Y^{t_n}_s(\omega)) \in B\}$$

and the RHS is clearly in $\sigma\{\phi, Y^u : u \geq 0\}$. □

Let $\Delta_1 = [0, \infty)$ and $L^2(\Delta_1, \lambda_1) =: L^2(\Delta_1)$ where λ_1 is the Lebesgue measure on $[0, \infty)$. For $f \in L^2(\Delta_1)$ we define the process

$$\{I^Y_{1,i}(f)(t);\ t \geq 0\} := \{\int_0^t f(<Y>_s)\, dY^i_s;\ t \geq 0\}$$

and the random variable

$$I_{1,i}(f) := I_{1,i}(f)(\infty) = \int_0^\infty f(<Y>_s)\, dY^i_s.$$

We will drop the Y and write $I_{1,i}(f)(t)$ when there is no risk of confusion. When $d = 1$, we drop the suffix i and just write $I_1(f)(t)$ or $I^Y_1(f)(t)$.

Define $\mathscr{M}_{1,t}(Y)$ to be the sigma field on Ω' generated by the first chaos viz.

$$\mathscr{M}_{1,t}(Y) := \sigma\{(s, \omega) \to I_1(f)(s \wedge t, \omega) : f \in L^2(\Delta_1)\}.$$

We note that by continuity of Y we have for all $t \geq 0$, $\mathscr{P}(Y) \supseteq \mathscr{M}_{1,t}(Y)$. We use the notation $Y^t_s(\omega) := Y_{t\wedge s}(\omega)$.

In the following two propositions we will take $d = 1$, although they are valid for higher dimensions. We will also use the fact, below and elsewhere, that the indistinguishability of two measurable processes implies that the subset of Ω on which they differ for some t is evanescent.

Proposition 13.2 *For every $t \geq 0$, we have up to evanescent sets,*

$$\mathscr{M}_{1,t}(Y) = \sigma\{I^{Y^t}_1(f)(s, \omega) : f \in L^2(\Delta_1), s \geq 0\}.$$

Proof We have for every $s \geq 0$, using $Y^t_s = \int_0^s I_{[0,t]}(u)\, dY_u$, a.s.

$$I^{Y^t}_1(f)(s) = \int_0^s f(< Y^t >_u) I_{[0,t]}(u)\, dY_u = \int_0^{t\wedge s} f(<Y>_u)\, dY_u = I_1(f)(s \wedge t, \omega)$$

for every $s \geq 0$. □

Combining the above two propositions with the fact that for each $t \geq 0$, $\{Y^t_s;\ s \geq 0\}$ is measurable with respect to the sigma field on $[0, \infty) \times \Omega$ generated by the maps $\{(s, \omega) \to I^{Y^t}_1(f)(s, \omega) : f \in L^2(\Delta_1)\}$ we get the following proposition. We use the notation $I^Y_1(f) := I^Y_1(f)(\infty) := \int_0^\infty f(<Y>_s)\, dY_s$.

Proposition 13.3 *We have the following equalities, up to evanescent sets:*

$$\begin{aligned}\mathscr{P}(Y) &= \sigma\{\phi, (s, \omega) \to I^{Y^t}_1(f)(s, \omega) : f \in L^2(\Delta_1), t \geq 0\}\\ &= \sigma\{\phi, \mathscr{M}_{1,t}(Y), t \geq 0\}.\end{aligned}$$

Proof By Proposition 13.2, it suffices to prove only the first equality. We have

$$I_1^{Y^t}(f) = \int_0^\infty f(<Y>_s) I_{[0,t]}(s)\, dY_s = \int_0^t f(<Y>_s)\, dY_s .$$

Hence the maps generating the sigma field in the RHS of the first equality in the statement of the proposition are $\mathscr{P}(Y)$ measurable. To get the reverse inclusion let $f_r(u) := I_{[0,r]}(u)$. Note that for each $s \geq 0$,

$$Y_s^t = \lim_{r \to \infty} I_1^{Y^t}(f_r)(s) ,$$

where the limit holds in probability, uniformly in $s \in [0, t]$; it follows that $(s, \omega) \to Y_s^t(\omega)$ is measurable w.r.t. $\sigma\{\phi, (s, \omega) \to I_1^{Y^t}(f)(s, \omega) : f \in L^2(\Delta_1), t \geq 0\}$, up to an evanescent set. Hence using Proposition 13.1,

$$\mathscr{P}(Y) = \sigma\{\phi, Y^t : t \geq 0\} \subseteq \sigma\{\phi, (s, \omega) \to I_1^{Y^t}(f)(s, \omega) : f \in L^2(\Delta_1), t \geq 0\} .$$

□

We state the following proposition whose proof is elementary.

Proposition 13.4 *Every $\eta^{-1}\mathscr{P}(B)$ measurable function $f(t, \omega)$ can be written as $f(t, \omega) = h(<Y>_t, \omega)$ for some h which is $\mathscr{P}(B)$ measurable.*

13.3 Martingale Chaoses

Let $\Delta_n := \{(t_1, \cdots, t_n) : 0 \leq t_1 < \cdots < t_n < \infty\}$. $L^2(\Delta_n, \lambda_n) =: L^2(\Delta_n)$ where λ_n is the Lebesgue measure on $\mathbb{R}^n$. To define the iterated multiple integral of $f \in L^2(\Delta_n)$ with respect to a one dimensional continuous martingale Y with quadratic variation $<Y>$, we first note that the intuitive definition viz.

$$I_n(f) := \int_0^\infty \int_0^{t_n} \cdots \int_0^{t_2} f(<Y>_{t_1}, \cdots, <Y>_{t_n})\, dY_{t_1} \cdots dY_{t_n} \tag{13.1}$$

where the RHS is an iterated multiple Ito integral of f w.r.t. Y, is problematic, since $I_{n-1}(f(., \cdots, t))$ is defined only up to null sets that depend on t. While it is possible to define the multiple iterated integrals $I_n(f)$, $f \in L^2(\Delta_n)$ by induction from first principles, a quicker way would be to exploit the relationship of Y with its DDS Brownian motion B and the fact that the quadratic variation processes are the same. We deal directly with the case of a vector valued continuous local martingale $Y = (Y^1, \cdots, Y^n)$, but in this case we need to also deal with the order of integration. For $f : \Delta_n \to \mathbb{R}$, for $1 < k < n$ and for $u = (u_1, \cdots, u_{n-k}) \in \Delta_{n-k}$, we use the notation f^u to denote the map $f^u : \Delta_k \to \mathbb{R}$, $f^u(t_1, \cdots, t_k) := f(t_1, \cdots, t_k, u_1, \cdots, u_{n-k}) I_{\Delta_n}(t_1, \cdots, t_k, u_1, \cdots, u_{n-k})$.

Note that if $f \in L^2(\Delta_n)$ we can choose a version $\tilde{f}$ such that for all $u = (u_1, \cdots, u_{n-k}) \in \Delta_{n-k}$, $\tilde{f}^u \in L^2(\Delta_k)$. We will not distinguish between f and $\tilde{f}$.

Let $(\mathcal{G}_t)$ be a filtration and $B = (B^1, \cdots, B^d)$ a standard $\mathcal{G}_t$-Brownian motion. The following theorem is well known and lists the main properties of the multiple integral with respect to B. We first recall that for $\alpha = (\alpha_1, \cdots, \alpha_n), \alpha_i \in \{1, \cdots, d\}$ the multiple integral $I^B_{n,\alpha}(f)$ is defined for $f \in L^2(\Delta_n)$ of the form $f(t_1, \cdots, t_n) = \prod_{i=1}^{n} f_i(t_i)$ with $f_i(t) := I_{(a_i,b_i]}(t)$, $a_1 < b_1 \le a_2 < \cdots < b_n$, as follows:

$$I^B_{n,\alpha}(f) := \prod_{i=1}^{n}(B^{\alpha_i}_{b_i} - B^{\alpha_i}_{a_i}) \; .$$

We extend the map $I^B_{n,\alpha}(f)$ linearly to the linear span of such functions, which we note is dense in $L^2(\Delta_n)$. We then have the following theorem (see for example [18]).

Theorem 13.1 *Let $n \ge 0$; $\alpha = (\alpha_1, \cdots, \alpha_n)$, $1 \le \alpha_i \le d$, $i = 1, \cdots, n$. Then the map $I^B_{n,\alpha}$ extends to the whole of $L^2(\Delta_n)$ and satisfies for $f, g \in \Delta_n$ and $a, b \in \mathbb{R}$*

$$I^B_{n,\alpha}(af + bg) = aI^B_{n,\alpha}(f) + bI^B_{n,\alpha}(g)$$

$$E(I^B_{n,\alpha}(f)I^B_{n,\alpha}(g)) = <f, g>_{L^2(\Delta_n)}. \tag{13.2}$$

Further,

$$E(I^B_{n,\alpha}(f)I^B_{m,\beta}(g)) = 0 \tag{13.3}$$

if either $n \ne m$ or $\alpha \ne \beta$. Finally if we define

$$I^B_{n,\alpha}(f)(t, \omega) := E[I^B_{n,\alpha}(f)|\mathcal{G}_t](\omega)$$

then $(I^B_{n,\alpha}(t), \mathcal{G}_t)$ is a continuous martingale which satisfies a.s. P, for every $t \ge 0$,

$$I^B_{n,\alpha}(f)(t) = \int_0^t I^B_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^s)(s)\, dB^{\alpha_n}_s \; .$$

Definition 13.2 Let $Y^1, \cdots, Y^d$ be continuous $\mathcal{F}_t$-local martingales with quadratic variation process $(< Y^i >_t)$, $i = 1, \cdots, d$. We assume that $Y^i_0 \equiv 0$, $< Y^i, Y^j > \equiv 0$, $i \ne j$, $< Y^i > \equiv < Y^j > =: <Y>_t$ and $<Y>_\infty = \infty, a.s.$. Then $Y_t := (Y^1_t, \cdots, Y^d_t) = B \circ <Y>_t$ where $B := (B^1, \cdots, B^d)$ is a standard $\mathcal{G}_t := \mathcal{F}^Y_{\tau_t}$-Brownian motion with $\tau_t := \inf\{s > 0 : <Y>_s > t\}$.

For $n = 0$, $\Delta_0 := \{0\}$, $f : \Delta_0 \to \mathbb{R}$ define $I_0(f)(t, \omega) := f(0) \in \mathbb{R}$ for every $t \ge 0$, $\omega \in \Omega$. We define for $n \ge 1$, $f \in L^2(\Delta_n)$ and $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\alpha_i \in \{1, \cdots, d\}$, the process

$$\{I^Y_{n,\alpha}(f)(t); t \ge 0\} := \{I^B_{n,\alpha}(f)(<Y>_t);\; t \ge 0\}$$

where $I^B_{n,\alpha}(f)(t,\omega)$ is the iterated multiple Wiener-Ito integral of $f \in L^2(\Delta_n)$. We define the random variable $I^Y_{n,\alpha}(f)(\omega) := I^Y_{n,\alpha}(f)(\infty,\omega) = I^B_{n,\alpha}(f)(\omega)$. □

Let $\mathscr{H}_t := \mathscr{G}_{<Y>_t}$. Then since $<Y>_t$ is a $\mathscr{G}_t$-stopping time, $(\mathscr{H}_t)$ is a filtration. Note that since $Y_t = B \circ <Y>_t$, we have $\mathscr{F}^Y_t \subseteq \mathscr{H}_t$. We note that $I^Y_{n,\alpha}(f)(t,\omega) := I^B_{n,\alpha}(f)(<Y>_t)$ is a square integrable $\mathscr{H}_t$-martingale; this follows from Doob's optional sampling theorem and the fact that $I^B_{n,\alpha}(f)(\omega) \in L^2(\Omega)$.

In the following theorem we list the main properties of the multiple integral $I^Y_{n,\alpha}(f)(t,\omega)$.

Theorem 13.2 *Let $n \geq 0$; $\alpha = (\alpha_1, \cdots, \alpha_n)$, $1 \leq \alpha_i \leq d$, $i = 1, \cdots, n$. Then the map $I^Y_{n,\alpha}$ satisfies for $f, g \in \Delta_n$ and $a, b \in \mathbb{R}$*

$$I^Y_{n,\alpha}(af + bg) = aI^Y_{n,\alpha}(f) + bI^Y_{n,\alpha}(g)$$
$$E(I^Y_{n,\alpha}(f)I^Y_{n,\alpha}(g)) = <f, g>_{L^2(\Delta_n)}. \tag{13.4}$$

Further,

$$E(I^Y_{n,\alpha}(f)I^Y_{m,\beta}(g)) = 0 \tag{13.5}$$

if either $n \neq m$ or $\alpha \neq \beta$. Finally $(I^Y_{n,\alpha}(f)(t), \mathscr{F}^Y_t)$ is a continuous martingale which satisfies a.s. P, for every $t \geq 0$,

$$I^Y_{n,\alpha}(f)(t) := \int_0^t I^Y_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^u)(s)|_{u=<Y>_s}\, dY^{\alpha_n}_s. \tag{13.6}$$

Proof The proofs of linearity, isometry and orthogonality are immediate consequences of the definition of $I^Y_{n,\alpha}(f)$, the fact that $I^Y_{n,\alpha}(f) = I^B_{n,\alpha}(f)$ a.s. P and the fact that $<Y>_\infty = \infty$ a.s. The last statement in the theorem follows from the corresponding statement for Brownian motion in Theorem 13.1 and time change:

$$\begin{aligned} I^Y_{n,\alpha}(f)(t) &= \int_0^{<Y>_t} I^B_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^s)(s)\, dB^{\alpha_n}_s \\ &= \int_0^t I^Y_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^u)(s)|_{u=<Y>_s}\, dY^{\alpha_n}_s \end{aligned}$$

where for every $u \geq 0$ we have by definition,

$$I^Y_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^u)(s) = I^B_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^u)(<Y>_s)\ .$$

Here we note that the fact that the map

$$(s,\omega) \longrightarrow I^Y_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^u)(s)|_{u=<Y>_s}(\omega)$$

is $\mathscr{P}(Y)$ measurable and in particular that $I^Y_{n,\alpha}(f)(t)$ is $\mathscr{F}^Y_t$ adapted follows by a monotone class argument; that it is square integrable $d{<}Y{>}\,dP$ follows by time change and the fact that $I^Y_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^s)(\tau_s) = I^B_{n-1,(\alpha_1,\cdots,\alpha_{n-1})}(f^s)(s)$. In particular it follows that $(I^Y_{n,\alpha}(f)(t), \mathscr{F}^Y_t)$ is a continuous martingale. This completes the proof. □

13.4 Kernels Associated with Chaos Expansions

In this section we define the kernels associated with the multiple stochastic integrals with respect to the martingale Y and the kernels associated with the chaos expansions of functionals of Y. As in the previous section we consider continuous vector orthogonal martingales $Y = (Y^1, \cdots, Y^d)$ with a common quadratic variation process ${<}Y{>}$. Our chaos expansions will be for elements of $L^2(\Omega', \mathscr{P}(Y), \mu)$ rather than for those of $L^2(\Omega, \mathscr{F}^Y_\infty, P)$ where we recall that $d\mu = d{<}Y{>}\,dP$ on $\Omega' := [0, \infty) \times \Omega$. We abbreviate $L^2(\Omega', \mathscr{P}(Y), \mu)$ as $L^2(\mu)$. In the following, for $f \in L^2(\Delta_{n+1})$ we set $h_{n,\alpha}(f)(s, t, \omega) := I_{n,\alpha}(f^s)(t, \omega)$. To be more precise it is a $\mathscr{B}[0, \infty) \times \mathscr{P}(Y^\alpha)$-measurable 'process' and for each $s \geq 0$, is indistinguishable from the process $(I_{n,\alpha}(f^s)(t, \omega))_{t\geq 0}$. We sometimes use the notation $I_{n,\alpha}(f^s)(t, \omega)|_{s={<}Y{>}_t}$ for $h_{n,\alpha}(f)({<}Y{>}_t, t, \omega)$ where $f \in \Delta_{n+1}$.

For $n \geq 1$, $\alpha = (\alpha_1, \cdots, \alpha_n)$ we define the set of kernels of type (n, α) as follows:

$$H_{n,\alpha}(Y) := \{F : F(t, \omega) = h_{n,\alpha}(f)({<}Y{>}_t, t, \omega),\ \text{a.s.}(P) \text{ for all } t \geq 0, \\ \text{for some } f \in L^2(\Delta_{n+1})\}$$

For $n = 0$, we define $H_0(Y) := \{h({<}Y{>}_t) : h \in L^2(\Delta_1)\}$.

We have the following

Proposition 13.5 *For every $n \geq 0$ and $\alpha = (\alpha_1, \cdots, \alpha_n)$, $H_{n,\alpha}(Y)$ is a closed subspace of $L^2(d\mu)$.*

Proof The result is immediate from the fact that the map $f \to h_{n,\alpha}(f)({<}Y{>}_t, t, \omega) : L^2(\Delta_{n+1}) \to L^2(\Omega', \mathscr{P}(Y), \mu)$ is an isometry and hence its range $H_{n,\alpha}(Y)$ is closed. □

Proposition 13.6 *$H_{n,\alpha}(Y)$ is orthogonal to $H_{m,\beta}(Y)$ if either $m \neq n$ or if $\alpha \neq \beta$.*

Proof Let $F \in H_{n,\alpha}(Y)$ and $G \in H_{m,\beta}(Y)$. Suppose $F(t, \omega) = I_{n,\alpha}(f^s)(t, \omega)|_{s={<}Y{>}_t}$, $G(t, \omega) = I_{m,\beta}(g^s)(t, \omega)|_{s={<}Y{>}_t}$ where $f \in L^2(\Delta_{n+1})$, $g \in L^2(\Delta_{m+1})$. Then, using Eqs. (13.5) and (13.6), we have

$$< F, G>_{L^2(\mu)} = E(I_{n+1,\alpha}(f) I_{m+1,\beta}(g)) = 0$$

if either $m \neq n$ or if $\alpha \neq \beta$. □

We define the closed subspace of kernels of type n viz. $H_n(Y)$ of $L^2(d\mu)$ for $n \geq 1$ as

$$H_n(Y) := \bigoplus_{\alpha=(\alpha_1,\cdots,\alpha_n)} H_{n,\alpha}(Y)$$

where $\alpha_i \in \{1, 2, \cdots, d\}$. Let $H(Y)$ be the closed subspace of $L^2(d\mu)$ which is the direct sum of the $H_n(Y), n \geq 0$:

$$H(Y) := \bigoplus_{n=0}^{\infty} H_n(Y).$$

As the calculations in Proposition 13.8 below show for $d = 1$, the usual chaos expansion of elements in $L^2(\mathscr{F}^Y)$ are obtained by integrating elements of the d-fold direct sum of $H(Y)$ with respect to Y.

Recall that $Y = B \circ <Y>$. Define an operator T acting on $\mathscr{P}(B)$ measurable functions $f(t, \omega)$ with range $\mathscr{P}(Y)$ measurable functions, as follows: $T(f)(t, \omega) := f(<Y>_t, \omega)$. Then T extends to a linear operator, $T : L^2(d\mu_0) \to L^2(d\mu)$ where $d\mu_0 := dt\, dP$.

Proposition 13.7 *T is an isometric isomorphism from $L^2(d\mu_0)$ into $L^2(d\mu)$ that maps $H_{n,\alpha}(B)$ onto $H_{n,\alpha}(Y)$. Further we have*

$$H(Y) = T(H(B)) \ .$$

Proof The proof of the first part of the proposition is immediate from the fact that $T(f) = f \circ \eta$ and that $\mu \circ \eta^{-1} = \mu_0$.

To prove that $T : H_{n,\alpha}(B) \to H_{n,\alpha}(Y)$ is onto, let

$$F \in H_{n,\alpha}(Y),\ F(t, \omega) = h_{n,\alpha}(f)(<Y>_t, t, \omega) = I^Y_{n,\alpha}(f^s)(t, \omega)|_{s=<Y>_t},\ f \in \Delta_{n+1} \ .$$

If one defines $G(t, \omega) := g_{n,\alpha}(f)(t, t, \omega)$ where $g_{n,\alpha}(f)(s, t, \omega) := I^B_{n,\alpha}(f^s)(t, \omega)$ then $G \in H_{n,\alpha}(B)$ and it is easy to verify using the definition of the iterated integrals of Y, that

$$TG(t, \omega) = g_{n,\alpha}(f)(<Y>_t, <Y>_t, \omega) = h_{n,\alpha}(f)(<Y>_t, t, \omega) = F(t, \omega).$$

Thus $T : H_{n,\alpha}(B) \to H_{n,\alpha}(Y)$ is onto.

To show that the $H(Y) = T(H(B))$, we first note that, by what has just been proved, T is an isometric isomorphism, and hence preserves direct sums. Consequently,

$$H_n(Y) = \bigoplus_{\alpha=(\alpha_1,\cdots,\alpha_n)} H_{n,\alpha}(Y) = \bigoplus_{\alpha=(\alpha_1,\cdots,\alpha_n)} T(H_{n,\alpha}(B)) = T(H_n(B)).$$

Therefore,

$$H(Y) := \bigoplus_{n=0}^{\infty} H_n(Y) = \bigoplus_{n=0}^{\infty} T(H_n(B)) = T(\bigoplus_{n=0}^{\infty} H_n(B)) = T(H(B)) \,. \qquad \square$$

In the following two propositions we use the following notation for the Ito integral: For a martingale Y and previsible process F we denote by $I^Y(F) := \int_0^\infty F(s)\, dY_s$. The following Proposition is easily seen to be true—with appropriate notation—even for $d > 1$.

Proposition 13.8 *Let $d = 1$ and $F \in H(Y)$. Then $I^Y(F)$ has a chaos expansion as follows: for each $n \geq 1$, there exists unique $f_n \in L^2(\Delta_n)$ such that almost surely P,*

$$I^Y(F) = \sum_{n \geq 1} I_n^Y(f_n).$$

Proof From the definition of $H(Y)$ we can write the previsible process $F := (F(t, \omega))$ as the sum

$$F(t, \omega) = \sum_n I_n^Y(f_{n+1}^s)(t, \omega)|_{s=<Y>_t}$$

for a sequence $\{f_n\}$, $f_n \in L^2(\Delta_{n+1})$, $n \geq 0$. Integrating this expression from 0 to ∞ with respect to Y, we get the required expansion for $I^Y(F)$. $\square$

The next result is a consequence of the chaotic representation property (CRP) for Brownian motion. Recall that $d\mu_0 := dt\, dP$.

Proposition 13.9 $L^2(\Omega', \mathscr{P}(B), \mu_0) = H(B)$.

Proof It is clear from the definitions that $H(Y) \subseteq L^2(\mu)$. Taking $Y = B$, the RHS in the statement is contained in the LHS. To show the reverse inclusion, let $f \in L^2(\Omega', \mathscr{P}(B), \mu_0) =: L^2(\mu_0)$ and $F := \int_0^\infty f(s)\, dB_s^1 =: I^{B^1}(f)$. Then by CRP for Brownian motion there exist $f_{n,\alpha} \in L^2(\Delta_n)$ such that

$$F = \sum_{n \geq 1, \alpha} I_{n,\alpha}^B(f_{n,\alpha}) = \sum_{i=1}^{d} \sum_{n \geq 1, \alpha : \alpha_n = i} I^{B^i}(I_{n-1,(\alpha_1, \cdots, \alpha_{n-1})}^B(f_{n,\alpha}^s))$$

$$= \sum_{i=1}^{d} I^{B^i}(\sum_{n \geq 1, \alpha : \alpha_n = i} I_{n-1,(\alpha_1, \cdots, \alpha_{n-1})}^B(f_{n,\alpha}^s)) = \sum_{i=1}^{d} I^{B^i}(g_i)$$

where

$$g_i(s, \omega) := \sum_{n \geq 1, \alpha : \alpha_n = i} I_{n-1,(\alpha_1, \cdots, \alpha_{n-1})}^B(f_{n,\alpha}^s)(s, \omega).$$

Note that $g_i \in \bigoplus_{n=0}^{\infty} H_n(B) =: H(B)$. From the definition of F it follows that $f = g_1$. In particular $f = g_1 \in H(B)$. □

13.5 Integral Representations

Let $n \geq 1$ and $\alpha = (\alpha_1, \cdots, \alpha_n)$ be a multi index. We define sigma fields on $[0, \infty) \times \Omega$ as follows $\mathscr{M}_{n,\alpha,t} := \sigma\{(s, \omega) \to I_{n,\alpha}(f)(s \wedge t, \omega) : f \in L^2(\Delta_n)\}$. We then have the following proposition. Recall the notation $\phi(u, \omega) := u$. In the following proposition all multiple integrals are with respect to $Y = (Y^1, \cdots, Y^d)$.

Proposition 13.10 *We have* $\mathscr{M}_{n,\alpha,t} \subseteq \sigma\{\phi,\ \mathscr{M}_{1,i,s};\ 0 \leq s \leq t,\ 1 \leq i \leq d\}$.

Proof We first note the following: let $\mathscr{G}^t := (\mathscr{G}^t_s)$ be the filtration

$$\mathscr{G}^t_s := \sigma\{I_{1,i}(g)(u \wedge t);\ u \leq s,\ g \in L^2(\Delta_1),\ 1 \leq i \leq d\}\ .$$

Then

$$\mathscr{P}(\mathscr{G}^t) = \sigma\{\phi,\ \mathscr{M}_{1,i,s};\ 0 \leq s \leq t,\ 1 \leq i \leq d\}.$$

This can be seen as in Proposition 13.1, Sect. 13.2, using the fact that the previsible sigma field $\mathscr{P}(\mathscr{G}^t)$ is generated by sets of the form $[0, \infty) \times A$, $A \in \mathscr{G}^t_0$ and $(s, \infty) \times A$, $A \in \mathscr{G}^t_s$, $0 \leq s \leq t$. Since $\mathscr{G}^t_s$ is generated by finite dimensional cylinder sets of the form

$$A = \{\omega : (I_{1,i}(g)(s_1 \wedge t), \cdots, I_{1,i}(g)(s_n \wedge t)) \in B\}$$

where $g \in L^2(\Delta_1)$, $1 \leq i \leq d$, and $0 \leq s_1 \leq \cdots \leq s_n \leq s$, $n \geq 1$ we can proceed as in Proposition 13.1 to show that $\mathscr{P}(\mathscr{G}^t) \subseteq \sigma\{\phi,\ \mathscr{M}_{1,i,s};\ 0 \leq s \leq t,\ 1 \leq i \leq d\}$. The reverse inequality follows from the continuity and adaptedness of the processes involved.

We will use an inductive argument to show that if $f \in L^2(\Delta_n)$ and t is fixed, then $\sigma\{I_{n,\alpha}(f)(s \wedge t, \omega)\} \subseteq \sigma\{\phi,\ \mathscr{M}_{1,i,s};\ 0 \leq s \leq t,\ 1 \leq i \leq d\}$. Suppose the above holds for $0 \leq n \leq k$. It suffices then to prove the claim for $n = k + 1$ and $f = \bigotimes_{l=1}^{k+1} f_l$. Then from Eq. (13.6) we have

$$I_{k+1,\alpha}(f)(s \wedge t) = \int_0^{s \wedge t} I_{k,(\alpha_1,\cdots,\alpha_k)}(\bigotimes_{l=1}^{k} f_l)(u)\ dI_{1,\alpha_{k+1}}(f_{k+1})(u \wedge t).$$

By the induction hypothesis and the observation made at the beginning of the proof, $I_{k,(\alpha_1,\cdots,\alpha_k)}(\bigotimes_{l=1}^{k} f_l)(u \wedge t)$ is measurable w.r.t. $\mathscr{P}(\mathscr{G}^t)$. By the continuity of the stochastic integral and the fact that it is adapted to the filtration $(\mathscr{G}_s^t)$ it follows that $I_{k+1,\alpha}(f)(s \wedge t, \omega)$ is measurable $\mathscr{P}(\mathscr{G}^t)$ and we are done. □

We may, by virtue of the above proposition, restrict our analysis to the sigma fields generated by the first chaos. We use the notation

$$\mathscr{M}_1(Y) := \sigma\{I_1^{Y^i}(f)(s,\omega) : \ f \in L^2(\Delta_1), \ i = 1, \cdots, d\}$$

for the sigma field generated on $[0,\infty) \times \Omega$ by the first chaos. Recall the map η defined in Sect. 13.2, Definition 13.1(4).

Lemma 13.1 *We have the following inclusions up to evanescent sets:*

$$\mathscr{M}_1(Y) \subseteq \eta^{-1}\mathscr{P}(B) \subseteq \mathscr{P}(Y).$$

Proof (Émery, 2015, personal communication). The first inclusion is an immediate consequence of the definition of $I_1^Y(f)(t)$. To show the second inclusion, recall that τ is the right continuous inverse of $<Y>$ and $\mathscr{G}_t := \mathscr{F}_{\tau_t}^Y$. Since $\mathscr{F}_t^B \subseteq \mathscr{G}_t$ for every $t \geq 0$, we have $\mathscr{P}(B) \subseteq \mathscr{P}(\mathscr{G})$. Hence $\eta^{-1}\mathscr{P}(B) \subseteq \eta^{-1}\mathscr{P}(\mathscr{G})$. It suffices to show that $\eta^{-1}\mathscr{P}(\mathscr{G}) \subseteq \mathscr{P}(Y)$. This follows from the fact that for $t > 0$, $A \in \mathscr{G}_t$ we have,

$$\begin{aligned} \eta^{-1}((t,\infty) \times A) &= \{(s,\omega) : \eta(s,\omega) \in (t,\infty) \times A\} \\ &= \{(s,\omega) : <Y>_s(\omega) > t, \omega \in A\} \\ &= \{(s,\omega) : \tau_t(\omega) < s, \omega \in A\} \\ &= \{(s,\omega) : T(\omega) < s < \infty\} \\ &=:]\!]T,\infty[\![\end{aligned}$$

where $T := \tau_t \mathbf{I}_A + \infty \mathbf{I}_{A^c}$ is an $(\mathscr{F}_t^Y)$ stopping time. Consequently the RHS in the above equality is in $\mathscr{P}(Y)$ and this completes the proof of the second inclusion. □

In the following three definitions Y is a continuous d-dimensional local martingale $Y = B \circ <Y>$ with B a d-dimensional standard Brownian motion and $<Y>_t \uparrow \infty$, a.s.

Definition 13.3 We say that $F \in L^2(\mathscr{F}_\infty^Y)$ has the strong previsible representation (SPR) property with respect to Y if there exists $f_i \in L^2(\mathscr{P}(B), \mu_0)$, $i = 1, \cdots, d$ such that

$$F = EF + \sum_{i=1}^{d} \int_0^\infty f_i(<Y>_s, \omega)\, dY_s^i \ . \tag{13.7}$$

Definition 13.4 We will say that $F \in L^2(\mathscr{F}^Y_\infty)$ has the chaotic representation (CR) property with respect to Y if there exist $f_{n,\alpha} \in L^2(\Delta_n)$ for $n \geq 0, \alpha = (\alpha_1, \cdots, \alpha_n)$ such that

$$F = \sum_{n,\alpha} I^Y_{n,\alpha}(f_{n,\alpha})$$

where the equality holds in $L^2(\mathscr{F}^Y_\infty)$.

Let P be any one of the properties CR, SPR or PR (the previsible representation property). We define for any property P pertaining to the elements of $L^2(\mathscr{F}^Y_\infty)$, the subsets

$$H_P := \{F \in L^2(\mathscr{F}^Y_\infty) : F \text{ has property} P\}.$$

We thus have the closed subspaces of $L^2(\mathscr{F}^Y_\infty)$ viz. H_{CR}, H_{PR}, H_{SPR}, the maximal subspaces with the chaotic representation property, the previsible representation property and the strong previsible representation property, respectively. Note that H_{SPR} is closed follows by time change.

Recall the definition of a pure local martingale (Definition 13.1, Sect. 13.2). It is clear from the above definitions, and the definition of $I^Y_{n,\alpha}(f)$ and time change that purity is equivalent to the chaotic representation property (CRP) and the strong previsible representation property. We state this as our next proposition.

Proposition 13.11 *The following are equivalent:*

a) Y has strong PRP, i.e. $L^2(\mathscr{F}^Y_\infty) = H_{SPR}$.
b) Y has the CRP, i.e. $L^2(\mathscr{F}^Y_\infty) = H_{CR}$.
c) Y is pure.

Even when CRP fails for Y we have the following result.

Theorem 13.3 *Let Y be as above. Then $H_{CR} = H_{SPR}$*

Proof The proof follows by observing that both sides of the equality in the statement equal $L^2(\mathscr{F}^B_\infty)$: For the RHS this follows by observing that by time change, the stochastic integral in Eq. (13.7) is just $EF + \sum_{i=1}^{d} \int_0^\infty f_i(s,\omega) dB^i_s$. This shows $H_{SPR} \subseteq L^2(\mathscr{F}^B_\infty)$. The reverse inclusion follows from the PRP for Brownian motion.

For the LHS we note that in the chaos expansion for F in Definition 13.4, $I^Y_{n,\alpha}(f) = I^B_{n,\alpha}(f)$ and consequently $H_{CR} \subseteq L^2(\mathscr{F}^B_\infty)$. The reverse inclusion follows by using the CRP for B. □

It is well known that the PRP for Y is not equivalent to its purity (see [25], Proposition (4.11), Chap V.). The following result provides a sufficient condition viz. μ a.e. $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$ for this equivalence.

Theorem 13.4 *The following are equivalent:*

1) a.e. μ, $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$*;*
2) $H_{PR} = H_{CR} = H_{SPR}$.

Consequently when $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$*, a.e.* μ*, the CRP, the PRP and the strong PRP are all equivalent.*

Proof 1) implies 2). Suppose $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B)$, a.e. ($\mu$). As we have noted above $H_{CR} = H_{SPR}$ and by Lemma 13.1 above, $H_{SPR} \subseteq H_{PR}$. The reverse inclusion follows since by our assumption any $\mathscr{F}_t^Y$-previsible process $f(t,\omega)$ must be of the form $g(<Y>_t, \omega)$ with $g(t,\omega)$ a $\mathscr{F}_t^B$ previsible process. Thus 2) holds.

2) implies 1). It suffices to show $L^2(\eta^{-1}\mathscr{P}(B), \mu) = L^2(\mathscr{P}(Y), \mu)$. By Lemma 13.1, $\eta^{-1}\mathscr{P}(B) \subseteq \mathscr{P}(Y)$ and so we have $L^2(\eta^{-1}\mathscr{P}(B), \mu) \subseteq L^2(\mathscr{P}(Y), \mu)$. Conversely let $f \in L^2(\mathscr{P}(Y), \mu)$. Let $1 \le i \le d$. Then $\int_0^\infty f(s,\omega)\, dY_s^i \in H_{PR} = H_{SPR}$ where the last equality holds by assumption. But this implies $f \in L^2(\eta^{-1}\mathscr{P}(B), \mu)$ and consequently $L^2(\mathscr{P}(Y), \mu) \subseteq L^2(\eta^{-1}\mathscr{P}(B), \mu)$. □

Corollary 13.1 *Suppose that Y has CRP. Then it has PRP; Consequently* $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B),\ a.e.\ \mu$.

Proof Since $\eta^{-1}\mathscr{P}(B) \subseteq \mathscr{P}(Y)$, a.e. ($\mu$) we always have $H_{CR} = H_{SPR} \subseteq H_{PR}$. In particular if $H_{CR} = L^2(\mathscr{F}^Y)$ then 2) and consequently 1) holds. □

Corollary 13.2 *Suppose that Y has PRP and in addition* $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B),\ a.e.\ \mu$*. Then Y has CRP.*

Proof This is immediate from '1) implies 2)' part of the theorem. □

Remark 13.1 A natural question here is to get sufficient conditions for $\mathscr{P}(Y) = \eta^{-1}\mathscr{P}(B),\ a.e.\ \mu$ to hold. Note that $\phi \in \eta^{-1}\mathscr{P}(B)$ a.e. μ is a necessary condition for the equality to hold. We do not know if it is sufficient. Consider the case $d = 1$, and $<Y>_t$ is strictly increasing so $t = \tau_{<Y>_t}$, where τ is the inverse of $<Y>_t$. If $\tau \in \mathscr{P}(B)$, then it follows that $\phi \in \eta^{-1}\mathscr{P}(B)$. The following result (Émery, 2015, personal communication) is in this direction.

Theorem 13.5 *Let Y be a continuous* $\mathbb{R}^d$ *valued local martingale such that* $< Y^i, Y^j > = \delta_{ij} <Y>$*, with* $<Y>_\infty = \infty\ a.s.$ *The following are equivalent:*

(i) $\eta^{-1}\mathscr{P}(B) = \mathscr{P}(Y)$*, where the equality holds up to evanescent sets;*
(ii) Y is pure and ϕ *is measurable with respect to* $\eta^{-1}\mathscr{P}(B)$*;*
(iii) Y is pure and $<Y>$ *is almost surely strictly increasing.*

Proof We first observe that

$$\eta^{-1}\mathscr{P}(B) = \mathscr{P}(Y) \implies \phi \in \eta^{-1}\mathscr{P}(B) \implies <Y> \text{ is strictly increasing.} \tag{13.8}$$

The first implication is because ϕ is continuous and deterministic and consequently $\mathscr{P}(Y)$ measurable. Secondly, if $\phi \in \eta^{-1}\mathscr{P}(B)$, then $\phi = H \circ \eta$ for some $\mathscr{P}(B)$-measurable H, so for all (t, ω), one has $t = H(<Y>_t(\omega), \omega)$ and hence for each ω the map $t \to <Y>_t(\omega)$ is injective. This shows (13.8).

The next step is to show that

$$\eta^{-1}\mathscr{P}(B) = \mathscr{P}(Y) \text{ and } <Y> \text{ is strictly increasing } \implies Y \text{ is pure.} \quad (13.9)$$

Indeed $\eta^{-1}\mathscr{P}(B) = \mathscr{P}(Y)$ entails $]\!]t, \infty[\![\ \in \eta^{-1}\mathscr{P}(B)$. When $<Y>$ is strictly increasing, this becomes $\eta^{-1}]\!]<Y>_t, \infty[\![\ \in \eta^{-1}\mathscr{P}(B)$, which in turn is equivalent to $]\!]<Y>_t, \infty[\![\in \mathscr{P}(B)$ because η is then a bijection. Consequently each $<Y>_t$ is an $\mathscr{F}^B$-stopping time. Hence $Y_t = B_{<Y>_t} \in \mathscr{F}^B_{<Y>_t} \subseteq \mathscr{F}^B_\infty$, so $\mathscr{F}^Y_t \subseteq \mathscr{F}^B_\infty$ and Eq. (13.9) is established.

The implications (i) $\implies$ (ii) $\implies$ (iii) are immediate consequences of (13.8) and (13.9). We now show (iii) $\implies$ (i). Assuming (iii), take T any $\mathscr{F}^Y$ stopping time. As Y is pure $<Y>_T$ is an $\mathscr{F}^B$ stopping time, from which follows $]\!]<Y>_T, \infty[\![\ \in \mathscr{P}(B)$, and consequently $\eta^{-1}]\!]<Y>_T, \infty[\![\ \in \eta^{-1}\mathscr{P}(B)$. But since $<Y>$ is strictly increasing $\eta^{-1}]\!]<Y>_T, \infty[\![\ = \ ]\!]T, \infty[\![$. Hence $]\!]T, \infty[\![\ \in \eta^{-1}\mathscr{P}(B)$ for every $\mathscr{F}^Y$-stopping time T, giving $\mathscr{P}(Y) \subseteq \eta^{-1}\mathscr{P}(B)$. □

Call τ^+ (instead of our usual τ) the right-continuous inverse of $<Y>$ and τ^- the left-continuous inverse of $<Y>$. Note that $\tau_t^- := \inf\{s > 0 : <Y>_s \geq t\}$. Define the subsets Ω^+, Ω^- of Ω' as follows: $\Omega^+ :=$ Range τ^+ and $\Omega^- :=$ Range τ^-; the restricted maps $\eta|_{\Omega^+} : \Omega^+ \to \Omega'$ and $\eta|_{\Omega^-} : \Omega^- \to \Omega'$ are two bijections. Recall that $\mathscr{G}_t = \mathscr{F}^Y_{\tau_t}$. With this notation, purity is easily characterized in terms of $\eta^{-1}\mathscr{P}(B)$ (Émery, 2015, personal communication).

Proposition 13.12 *The following are equivalent:*

(i) Y *is pure;*
(ii) $\mathscr{F}^B = \mathscr{G}$;
(iii) $\mathscr{P}(B) = \mathscr{P}(\mathscr{G})$;
(iv) $\eta^{-1}\mathscr{P}(B) = \eta^{-1}\mathscr{P}(\mathscr{G})$;
(v) $\eta^{-1}\mathscr{P}(B)$ *and* $\eta^{-1}\mathscr{P}(\mathscr{G})$ *have the same restriction to the subset* Ω^+;
(vi) $\eta^{-1}\mathscr{P}(B)$ *and* $\eta^{-1}\mathscr{P}(\mathscr{G})$ *have the same restriction to the subset* Ω^-.

Proof (i) $\Leftrightarrow$ (ii) is classical. (ii) $\Leftrightarrow$ (iii) is due to the fact that a filtration, say $(\mathscr{H}_t)$, is always characterized by its previsible σ-field, because $\mathscr{H}_t = \{A \subseteq \Omega : (t, \infty) \times A$ is $\mathscr{H}$previsible$\}$. (iii) $\Leftrightarrow$ (iv), (iii) $\Leftrightarrow$ (v) and (iii) $\Leftrightarrow$ (vi) hold because the three maps $\eta : \Omega' \to \Omega'$, $\eta|_{\Omega^-} : \Omega^- \to \Omega'$, $\eta|_{\Omega^+} : \Omega^+ \to \Omega'$ are onto (and if a map $f : E \to F$ is onto, then f^{-1} acting on subsets of F is into; so, when f is onto, a σ-field $\mathscr{A}$ on F is characterized by $f^{-1}\mathscr{A}$). □

This leads to the question of clarifying the links between $\eta^{-1}\mathscr{P}(B)$ and $\mathscr{P}(Y)$. We have the following result (Émery, 2015, personal communication).

Theorem 13.6 *Let Y be a continuous $\mathbb{R}^d$ valued local martingale such that $<Y^i, Y^j> = \delta_{ij}<Y>$, with $<Y>_\infty = \infty$ a.s. Then Y is pure iff $\mathscr{P}(Y)$ and $\eta^{-1}\mathscr{P}(B)$ have the same restriction to Ω^-.*

Proof We show that $\mathscr{P}(Y)$ and $\eta^{-1}\mathscr{P}(\mathscr{G})$ have the same restriction to Ω^-. Then the result follows from the equivalence between (i) and (vi) in Proposition 13.12.

To prove our claim we shall first show that if T is an $\mathscr{F}^Y$-stopping time, $<Y>_T$ is a $\mathscr{G}_t$-stopping time, which satisfies

$$]\!]T, \infty[\![\,\cap\, \Omega^- = \eta^{-1}(]\!]<Y>_T, \infty[\![) \cap \Omega^- . \tag{13.10}$$

Indeed if T is an $\mathscr{F}^Y$-stopping time, $\{<Y>_T \leqslant s\} = \{\tau_s \geqslant T\} \in \mathscr{F}^Y_{\tau_s} = \mathscr{G}_s$, so $<Y>_T$ is a $\mathscr{G}_t$-stopping time. Moreover, putting $S = <Y>_T$, one has $\tau_S^- = \tau^-_{<Y>_T} \leqslant T \leqslant \tau_{<Y>_T} = \tau_S$. Hence the symmetric difference $]\!]T, \infty[\![\,\triangle\,]\!]\tau_S, \infty[\![$ equals $]\!]T, \tau_S]\!]$, and is included in $]\!]\tau_S^-, \tau_S]\!]$, which does not meet Ω^-. Consequently, $]\!]T, \infty[\![\cap \Omega^- =]\!]\tau_S, \infty[\![\cap \Omega^-$. Equation (13.10) now follows from $]\!]\tau_S, \infty[\![= \eta^{-1}]\!]S, \infty[\![= \eta^{-1}]\!]<Y>_T, \infty[\![$.

According to Eq. (13.10), for each $\mathscr{F}^Y_t$-stopping time T there exists a set $\Gamma \in \eta^{-1}\mathscr{P}(\mathscr{G})$ such that $]\!]T, \infty[\![\,\cap\, \Omega^- = \Gamma \cap \Omega^-$. Now, when T ranges over all $\mathscr{F}^Y_t$-stopping times, the sets $]\!]T, \infty[\![$ generate the sigma field $\mathscr{P}(Y)$; consequently $\mathscr{P}(Y)|_{\Omega^-} \subseteq (\eta^{-1}\mathscr{P}(\mathscr{G}))|_{\Omega^-}$. The reverse inclusion $\eta^{-1}\mathscr{P}(\mathscr{G}) \subseteq \mathscr{P}(Y)$ is already known; so $\eta^{-1}\mathscr{P}(\mathscr{G})$ and $\mathscr{P}(Y)$ always have the same restriction to the subset Ω^- of Ω'. □

13.6 Change of Measure

On $C([0,\infty), \mathbb{R}^d)$, define for each $t \geq 0$, $Y_t(\omega) := \omega(t) = (Y_1(t), \cdots, Y_d(t))$, the coordinate random variables. The sigma fields $\mathscr{F}_t := \sigma\{Y_s; s \leq t\}$ and $\mathscr{F}_\infty := \sigma\{Y_s; s \geq 0\}$. Let P be a probability on $C([0,\infty), \mathbb{R}^d)$ such that $(Y_t, \mathscr{F}_t)$ is a martingale with $Y_0 = 0$ a.s. We will assume that $<Y^i, Y^j>_t = \delta_{ij}<Y>_t$ and that $<Y> \uparrow \infty$ a.s. Let $h(t)$ denote a fixed, deterministic, $\mathbb{R}^d$-valued continuous function. Define the probability measure P_h on $C([0,\infty), \mathbb{R}^d)$ as follows:

$$P_h(A) := P(\omega : Y_.(\omega) + h \circ <Y>_.(\omega) \in A).$$

For a continuous martingale (M_t) with quadratic variation $<M>$ we define $exp(M)_t := exp(M_t - \frac{1}{2}<M>_t)$.

Let $\tau_t := \inf\{s > 0 : <Y>_s > t\}$. The increasing process (τ_t) induces a map $\tau : C([0,\infty), \mathbb{R}^d) \to C([0,\infty), \mathbb{R}^d)$ given by $\tau(\omega) := \omega'$ where $\omega'(t) := \omega(\tau_t)$. Define $\tau^{-1}\mathscr{F} := \sigma\{\tau^{-1}(A) : A \in \mathscr{F}\}$. Then note that $\tau^{-1}\mathscr{F} = \sigma\{Y_{\tau_u} : u \geq 0\} \subseteq \mathscr{F}$ where the equality can be seen by considering finite dimensional sets and the inclusion follows by the measurability of each Y_{τ_u}, $u \geq 0$. Define $\tau^{-1}(\mathscr{F})_t := \sigma\{\tau^{-1}(A) : A \in \mathscr{F}_t\}$. As before, $\tau^{-1}(\mathscr{F})_t = \sigma\{Y_{\tau_u} : 0 \leq u \leq t\} \subseteq \mathscr{F}_{\tau_t}$.

Note that $\tau : (C([0,\infty), \mathbb{R}^d), \tau^{-1}\mathscr{F}) \to (C([0,\infty), \mathbb{R}^d), \mathscr{F})$ given by the map $\tau(\omega) := \omega \circ \tau = Y_{\tau_\cdot}(\omega)$ is measurable and $B_t := Y_{\tau_t}$ is a standard d-dimensional Brownian motion. Note that $\mathscr{F}^B := \sigma\{B_u;\ u \geq 0\} = \sigma\{Y_{\tau_u};\ u \geq 0\} = \tau^{-1}\mathscr{F}$. We write $P_1 \sim P_2$ if P_1 and P_2 are mutually absolutely continuous measures on $C([0,\infty), \mathbb{R}^d)$.

Theorem 13.7 *Let $(Y_t, \mathscr{F}_t)$ be a pure martingale. Then $P_h \sim P$ on $\mathscr{F}$ if and only if $h(t) := \int_0^t \dot{h}(s)\, ds$, $\dot{h} = (\dot{h}_1, \cdots, \dot{h}_d)$, $\dot{h}_i \in L^2[0,\infty)$. Moreover, in this case, we have $P_h(A) = \int_A exp(M)_t\, dP$, $A \in \mathscr{F}_t$ where $M_t := \int_0^t \dot{h} \circ <Y>_s\, dY_s$.*

Proof We follow the proof in [15] given for Brownian motion. Suppose that $P_h \sim P$ on $\mathscr{F}$. Then $P_h(A) = \int_A Z_t\, dP$ where (Z_t) is a uniformly integrable $(\mathscr{F}_t, P)$ martingale. Since Y is pure it has the CRP and hence the strong PRP. Hence there exists $V \in \mathscr{P}(B) = \mathscr{P}((\tau^{-1}\mathscr{F}_t))$, $\int_0^\infty V^2 \circ <Y>_s\, d<Y>_s < \infty$, a.s. (P) such that $Z_t = exp(\int_0^t V \circ <Y>_s\, dY_s)$.

Hence under P_h,

$$\tilde{Y}_t := Y_t - \int_0^t V \circ <Y>_s\, d<Y>_s$$

is an $\mathscr{F}_t$-martingale. Since $< \tilde{Y} >_t = <Y>_t$ a.s. P_h, we have by Lévy's characterization $(\tilde{Y}_{\tau_t})$ is a Brownian motion under P_h. In particular,

$$(B_t - h_t) + h_t = B_t = \tilde{Y}_{\tau_t} + \int_0^t V(s)\, ds.$$

By a time change argument it follows that under P_h, $B - h$ is a Brownian motion. It follows by the uniqueness of the semi-martingale decomposition for (B_t) that $h(t) := \int_0^t \dot{h}(s)\ ds$ where, almost surely, $\dot{h}(s) = V(s)$, a.e. (s). In particular, it follows that $\dot{h} \in L^2[0,\infty)$.

Conversely, we assume $h(t) := \int_0^t \dot{h}(s)\, ds$, $\dot{h} = (\dot{h}_1, \cdots, \dot{h}_d)$, $\dot{h}_i \in L^2[0,\infty)$. Define

$$Q(A) := \int_A exp(M)_t\, dP, \quad A \in \mathscr{F}$$

where (M_t) is as in the statement of the Theorem. Then Q is a probability on $\mathscr{F}$. We note that under Q, $\{Y_t - \int_0^t \dot{h} \circ <Y>_s\, d<Y>_s, \mathscr{F}_t\}$ is a martingale. We then have

$$Q \circ \tau^{-1}(A) = Q\{\omega : Y \circ \tau \in A\} = P\{Y \circ \tau \in A - \int_0^\cdot \dot{h}(s)\, ds\}$$

$$= P\{Y \circ \tau + (\int_0^\cdot \dot{h}(<Y>_s)\, d<Y>_s) \circ \tau \in A\}$$

$$= P\{Y + \int_0^{\cdot} \dot{h}(<Y>_s)\, d<Y>_s \in \tau^{-1}(A)\}$$
$$= P_h\{\tau^{-1}(A)\}$$

where the second equality follows because under Q, $Y \circ \tau - \int_0^{\cdot} \dot{h}(s)\, ds$ is a Brownian motion. Thus P_h is equivalent to P on the sigma field $\tau^{-1}(\mathscr{F}) = \mathscr{F}^B = \mathscr{F}$, where the last equality follows from purity. □

Acknowledgements The author would like to thank Michel Émery for extensive discussions over e-mail.

References

1. R.H. Cameron, W.T. Martin, Transformation of Wiener integrals under translations. Ann. Math. **45**, 386–396 (1944)
2. C. Dellacherie, P.A. Meyer, *Probabilities and Potential* (North Holland, Amsterdam, 1978)
3. G. Di Nunno, B. Øksendal, F. Proske, White noise analysis for Lévy processes. J. Funct. Anal. **206**, 109–148 (2004)
4. G. Di Nunno, B. Øksendal, F. Proske, *Malliavin Calculus for Lévy Processes with Applications to Finance* (Springer, Berlin, 2009)
5. M. Émery, On the Azéma martingales, in *Séminaire de Probabilités XXIV*. Lecture Notes in Mathematics, vol. 1426 (Springer, Berlin, 1990), pp. 442–447
6. M. Émery, Quelques cas de représentation chaotique, in *Séminaire de Probabilités XXV*. Lecture Notes in Mathematics, vol. 1485 (Springer, Berlin, 1991), pp. 10–23
7. M. Émery, Chaotic representation property of certain Azema martingales. Ill. J. Math. **50**(2), 395–411 (2006)
8. V. Fock, Konfigurationsraum und zweite Quantelung. Z. Phys. **75**, 622–647 (1932)
9. I.V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. Theor. Prob. Appl. **5**, 285–301 (1960)
10. A. Goswami, B.V. Rao, Conditional expectation of odd chaos given even. Stoch. Stoch. Rep. **35**, 213–214 (1991)
11. C. Houdré, V. Pérez-Abreu, *Chaos Expansions, Multiple Wiener-Itô Integrals and Their Applications* (CRC, Boca Raton, 1994)
12. N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North Holland, Amsterdam, 1981)
13. K. Itô, Multiple Wiener integral. J. Math. Soc. Jpn. **3**, 157–169 (1951)
14. K. Itô, Spectral type of the shift transformation of differential processes with stationary increments. Trans. Am. Math. Soc. **81**, 252–263 (1956)
15. O. Kallenberg, *Foundations of Modern Probability* (Springer, Berlin, 2010)
16. V. Mandrekar, P.R. Masani, *Proceedings of the Nobert Wiener Centenary Congress, 1994*. Proceedings of Symposia in Applied Mathematics, vol. 52 (American Mathematical Society, Providence, 1997)
17. P.A. Meyer, Un cours sur les intégrales stochastiques, in *Séminaire de Probabilités X*. Lecture Notes in Mathematics, vol. 511 (Springer, Berlin, 1976), pp. 245–400
18. D. Nualart, *The Malliavin Calculus and Related Topics* (Springer, Berlin, 1995)
19. D. Nualart, W. Schoutens, Chaotic and predictable representation for Lévy processes. Stoch. Process. Appl. **90**, 109–122 (2000)

20. K.R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*. Monographs in Mathematics (Birkhäuser, Boston, 1992)
21. N. Privault, *Stochastic Analysis in Discrete and Continuous Settings* (Springer, Berlin, 2009)
22. B. Rajeev, Martingale representations for functionals of Lévy processes. Sankhya Ser. A, **77**(Pt 2), 277–299 (2015)
23. B. Rajeev, Iterated stochastic integrals and purity. Pre-print (2016)
24. M.C. Reed, B. Simon, *Methods of Modern Mathematical Physics*, vol. II (Academic, New York, 1975)
25. D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 3rd edn. (Springer, Berlin, 1999)
26. D.W. Stroock, Homogenous chaos revisited, in *Séminaire de Probabilités XXI*. Lecture Notes in Mathematics, vol. 1247 (Springer, Berlin, 1987), pp. 1–7
27. D.W. Stroock, M. Yor, On extremal solutions of Martingale problems. Ann. Sci. École Norm. Sup. **13**, 95–164 (1980)
28. N. Wiener, The homogeneous chaos. Am. J. Math. **60**, 897–936 (1938)

Chapter 14
Explicit Laws for the Records of the Perturbed Random Walk on $\mathbb{Z}$

Laurent Serlet

Abstract We consider a nearest neighbor random walk on $\mathbb{Z}$ which is perturbed when it reaches its extrema, as considered before by several authors. We give invariance principles for the signs of the records, the values of the records, the times of the records, the number of visited points, with explicit asymptotic Laplace transforms and/or densities.

Keywords Perturbed random walk · Once-reinforced random walk · Perturbed Brownian motion · Records · Invariance principle · Recurrence

14.1 Introduction

We consider a process $(X_n)_{n\geq 0}$ with values in $\mathbb{Z}$, started at 0 ($X_0 = 0$), which is a nearest neighbor random walk on $\mathbb{Z}$ that is, for every $n \geq 0$, we have $X_{n+1} \in \{X_n - 1, X_n + 1\}$. We denote its maximum and minimum up to time n by $\overline{X}_n = \max\{X_0, X_1, \ldots, X_n\}$ and $\underline{X}_n = \min\{X_0, X_1, \ldots, X_n\}$. We say that $(X_n)_{n\geq 0}$ is a perturbed random walk (**PRW**) with reinforcement parameters $\beta, \gamma \in (0, +\infty)$ if the transition probability

$$\mathbb{P}\left(X_{n+1} = X_n + 1 \,\middle|\, X_0, X_1, \ldots, X_n\right)$$

is equal to

- $1/2$ if $\underline{X}_n < X_n < \overline{X}_n$ or $n = 0$
- $1/(1+\beta)$ if $X_n = \overline{X}_n$ and $n \geq 1$
- $\gamma/(1+\gamma)$ if $X_n = \underline{X}_n$ and $n \geq 1$.

L. Serlet (✉)
Université Clermont Auvergne, Laboratoire de Mathématiques Blaise Pascal (CNRS UMR 6620), Aubière, France
e-mail: Laurent.Serlet@math.univ-bpclermont.fr

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_14

When $\beta = \gamma = 1$, we obtain a standard random walk (**SRW**). When $\beta = \gamma$, we obtain a symmetric perturbed random walk (**SPRW**) with parameter $\beta \in (0, +\infty)$. We interpret the case $\beta > 1$ as a self-attractive walk whereas for $\beta \in (0, 1)$ the walk is self-repulsive.

This process belongs to the broad class of processes with reinforcement which has generated an important amount of literature. Pemantle gives in [7] a nice survey with lots of references. More precisely, our PRW is sometimes called the once-reinforced random walk. This once-reinforced random walk can also be defined in $\mathbb{Z}^d$ for $d > 1$ and some fundamental questions are still open in these dimensions but, in the present paper, we stay in dimension 1 which enables a much easier treatment and in particular, explicit computations of laws.

In [4] and [5], Davis introduces a diffusive rescaling by setting $X_t^n = \frac{1}{\sqrt{n}} X_{nt}$ (after linear interpolation of X between integer times) and he proves that the process $(X_t^n)_{t \geq 0}$ converges in law to a process $(W_t)_{t \geq 0}$ which is the unique solution of the equation

$$W_t = B_t - (\beta - 1) \sup_{s \leq t} W_s + (\gamma - 1) \inf_{s \leq t} W_s \tag{14.1}$$

where $(B_t)_{t \geq 0}$ is a Brownian motion.

The solutions of (14.1) have been studied by several authors under the name "Brownian motion perturbed at its extrema", see for instance [2–5, 8, 10] and the references therein.

The present paper uses a different approach since it is based on explicit computations for the random walk and we get results on the limiting continuous objects as by-products. The methods are similar to those in [9] where we treated the case of the reflecting random walk perturbed at the maximum. We will of course refer to this paper for several proofs which are identical to that case. However since our approach is based on explicit computations, the non-reflecting case generates different formulas than the reflecting case and the non-reflecting case also adds new questions related to the signs, that we address in the present paper.

Since the PRW behaves as a SRW when it stays away from the extrema, we concentrate on the study of the process when it reaches an extremum, in particular for the first time and in that case we call it a *record time*. More precisely, we set

$$V_k = \overline{X}_k - \underline{X}_k + 1$$

for the number of visited points up to time k i.e. the number of distinct values in the set $\{X_0, X_1, \ldots, X_k\}$. Then we define $T_0 = 0$ and for $n \geq 1$, we call

$$T_n = \inf\{k \geq 1;\ V_k = n + 1\}$$

the time of the n-th record. Then $R_n = X_{T_n}$ is the value of the n-th record and the sign of R_n is denoted by $\chi_n \in \{-1, 1\}$. As we will see the sequence $(\chi_n)_{n \geq 1}$ is a time-inhomogeneous Markov chain for which the transition matrix is easily computed. As a consequence we will derive an invariance principle. We will also note that the record values $(R_n)_{n\geq 0}$ can be reconstructed from the sequence of signs $(\chi_n)_{n\geq 1}$ and thus derive an invariance principle for the record values.

Then we introduce the rescaled record time process $(\tau_t^n)_{t\geq 0}$ by

$$\tau_t^n = \frac{1}{n^2}\, T_{[nt]} \tag{14.2}$$

where $[\cdot]$ denotes the integer part. We want an invariance principle for this process. First we work conditionally on the record signs and then without conditioning. Unfortunately in this latter case, we are unable to obtain a result in the general case and we restrict to symmetric perturbation ($\beta = \gamma$).

This invariance principle has consequences on the process of the number of visited points $(V_k)_{k\geq 1}$. As in the standard case, V_k is of order $\sqrt{k}$ and we obtain in particular the asymptotic law of $V_k/\sqrt{k}$. Finally we will examine the possibility of "positive recurrence" for the PRW.

The paper is organized as follows. The next section is a precise statement of our main results, ending with two open questions. The following sections are devoted to proofs, beginning with a section of technical preliminary lemmas.

14.2 Statement of the Results

Most of the processes that we consider in this section have their trajectories in the space $\mathrm{d}([0,+\infty),\mathbb{R})$ of càdlàg functions that we endow with the usual Skorohod topology. Weak convergence of probability laws on this space is simply called in the sequel "convergence in law". However special care will be needed in Proposition 14.2 where we have to restrict to compact intervals of $(0,+\infty)$. In the sequel the notations cosh, sinh and tanh refer to the usual functions of hyperbolic trigonometry.

Let $(X_n)_{n\geq 0}$ denote a PRW with parameters $\beta,\gamma \in (0,+\infty)$. First note that the sequence of record values $(R_n)_{n\geq 0}$ is easily reconstructed from the sequence of record signs $(\chi_n)_{n\geq 1}$ because, for any $n \geq 1$,

$$R_n = \chi_n \sum_{k=1}^{n} \mathbf{1}_{\{\chi_k=\chi_n\}} \tag{14.3}$$

and it justifies that we first focus on $(\chi_n)_{n\geq 1}$. We start with an easy fact.

Proposition 14.1 *The sequence of the signs of records $(\chi_n)_{n\geq 1}$ of the PRW is a time-inhomogeneous Markov chain with transition matrix*

$$Q_n = \begin{pmatrix} \frac{\beta+n}{\gamma+\beta+n} & \frac{\gamma}{\gamma+\beta+n} \\ \frac{\beta}{\gamma+\beta+n} & \frac{\gamma+n}{\gamma+\beta+n} \end{pmatrix}. \tag{14.4}$$

We notice that the off-diagonal terms of this transition matrix are of order $1/n$ so we speed up time by factor n to get a limit in law.

14.2.1 Asymptotic Results for the Signs of Records

Proposition 14.2 *Let the rescaled sequence of the record signs of the PRW be defined by*

$$\forall t > 0, \quad \chi_t^n = \chi_{[nt]}. \tag{14.5}$$

There is a process $(x_t)_{t\in(0,+\infty)}$ *with values in* $\{-1,+1\}$ *uniquely defined in law such that, for any* $a > 0$*,*

$$\mathbb{P}(x_a = -1) = \frac{\beta}{\beta+\gamma}; \quad \mathbb{P}(x_a = 1) = \frac{\gamma}{\beta+\gamma} \tag{14.6}$$

and $(x_t)_{t\in(a,+\infty)}$ *is a time-inhomogeneous Markov jump process with generator*

$$\begin{pmatrix} -\frac{\gamma}{t} & \frac{\gamma}{t} \\ \frac{\beta}{t} & -\frac{\beta}{t} \end{pmatrix} \tag{14.7}$$

and transition probability matrix

$$T(s,t) = \frac{1}{\beta+\gamma}\begin{pmatrix} \beta+\gamma\left(\frac{s}{t}\right)^{\beta+\gamma} & \gamma\left(1-\left(\frac{s}{t}\right)^{\beta+\gamma}\right) \\ \beta\left(1-\left(\frac{s}{t}\right)^{\beta+\gamma}\right) & \gamma+\beta\left(\frac{s}{t}\right)^{\beta+\gamma} \end{pmatrix}. \tag{14.8}$$

For all $0 < a < b$*, the sequence of processes* $(\chi_t^n)_{t\in[a,b]}$ *converges in law to* $(x_t)_{t\in[a,b]}$*, as laws on the space* $\mathbb{D}([a,b],\{-1,1\})$ *of càdlàg functions from* $[a,b]$ *to* $\{-1,1\}$ *endowed with the Skorohod topology.*

Let us remark that setting $\tilde{x}_t = x_{e^t}$ for $t \in \mathbb{R}$ gives a new process $(\tilde{x}_t)_{t\in\mathbb{R}}$ which is a time-homogeneous Markov jump process on $\{-1,1\}$ or, in other words, an alternating renewal process.

14.2.2 Consequences on the Sequence of Records

We derive a corollary which is the continuous-time counterpart of (14.3).

Corollary 14.3 *Let* $(y_t)_{t>0}$ *be the process defined by*

$$\forall t > 0, \quad y_t = x_t \int_0^t \mathbf{1}_{\{x_u = x_t\}}\, du$$

where $(x_t)_{t>0}$ *is the process introduced in Proposition 14.2.*

Then, the process $(R_{[nt]}/n)_{t>0}$ *converges in law to the process* $(y_t)_{t>0}$.

In particular R_n/n converges in law to y_1. One must not be deceived by the formula

$$\frac{R_n}{n} = \begin{cases} -\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{\chi_k=-1\}} & \text{if } \chi_n = -1 \\ \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{\chi_k=1\}} & \text{if } \chi_n = 1 \end{cases} \tag{14.9}$$

(which is a reformulation of (14.3)) and the fact that $(\chi_n)_{n\geq 1}$ is a Markov chain converging in law to the probability $\beta/(\beta+\gamma)\ \delta_{-1} + \gamma/(\beta+\gamma)\ \delta_{+1}$. In the time-homogeneous case, (14.9) would imply by the ergodic Theorem that R_n/n accumulates almost surely on the two limit points $-\frac{\beta}{\beta+\gamma}$ and $\frac{\gamma}{\beta+\gamma}$. But here the Markov chain $(\chi_n)_{n\geq 1}$ is time-inhomogeneous and the almost sure behaviour of R_n/n is completely different as one can see in the following proposition which holds whatever the values of $\beta, \gamma \in (0, +\infty)$.

Proposition 14.4 *Almost surely, the sequence* $\{R_n/n;\ R_n > 0\}$ *has limsup 1 and liminf 0 and similarly* $\{R_n/n;\ R_n < 0\}$ *has limsup 0 and liminf -1 that is, for all* $\varepsilon \in (0, 1)$ *and* $N \geq 1$ *there exist* $n_1, n_2, n_3, n_4 \geq N$ *such that*

$$\frac{R_{n_1}}{n_1} > 1-\varepsilon,\ 0 < \frac{R_{n_2}}{n_2} < \varepsilon,\ -\varepsilon < \frac{R_{n_3}}{n_3} < 0,\ \frac{R_{n_4}}{n_4} < -1+\varepsilon.$$

14.2.3 Invariance Principle for the Record Times, Conditionally on the Signs

We pass to the study of $(\tau_t^n)_{t\geq 0}$ the rescaled process of the record times of the PRW as defined by (14.2).

We have seen in the previous subsection the convergence in law of the rescaled record signs (on compact sets away from 0). By the Skorohod representation Theorem, we could suppose—concerning the properties that involve the law—that this convergence holds almost surely. In that case, we want to show the convergence of the conditional law of (rescaled) record times knowing these (rescaled) record signs. Let us introduce some notation. Let $(x^n(t))_{t>0}$ be a sequence of càdlàg functions taking their values in $\{-1, +1\}$ which converges to a function $(x(t))_{t>0}$ with respect to the Skorohod topology when t varies in any compact sets of $(0, +\infty)$. We set, for $0 \leq s < t$, $\mathcal{D}(x; s, t) = \{r \in (s, t);\ x(r-) \neq x(r)\}$ for the set of discontinuities of $x(\cdot)$ and similarly $\mathcal{D}(x^n; \cdot, \cdot)$ for $x^n(\cdot)$. We suppose that $\mathcal{D}(x^n; 0, +\infty) \subset \frac{1}{n}\mathbb{N}$ for all n and that $\mathcal{D}(x; s, t)$ is finite for all $0 < s < t < +\infty$.

Proposition 14.5 *As* $n \to +\infty$, *the conditional law of* $(\tau_t^n)_{t\geq 0}$ *knowing* $(\chi_t^n)_{t>0} = (x^n(t))_{t>0}$ *converges weakly to the law of a process* $(\tau_t^{(x)})_{t\geq 0}$*—defined conditionally on* $(x(t))_{t>0}$*—such that it has independent non-negative increments with* *distribution given, for* $0 < s < t$, *by the Laplace transform*

$$\mathbb{E}\left[e^{-\frac{\mu^2}{2}(\tau_t^{(x)}-\tau_s^{(x)})}\right] = \left(\prod_{r\in\mathcal{D}(x;s,t)} \frac{\mu\, r}{\sinh(\mu\, r)}\right) \times \exp\left[\int_s^t \delta(x(u))\left(\frac{1}{u} - \mu \coth \mu u\right) du\right] \qquad (14.10)$$

where

$$\delta(y) = \beta\,\mathbf{1}_{\{y=1\}} + \gamma\,\mathbf{1}_{\{y=-1\}}.$$

In the case of the SPRW (i.e. $\beta = \gamma$*), this formula simplifies into*

$$\mathbb{E}\left[e^{-\frac{\mu^2}{2}(\tau_t^{(x)}-\tau_s^{(x)})}\right] = \left(\frac{\sinh(\mu\, s)}{\mu\, s}\right)^{\beta} \left(\prod_{r\in\mathcal{D}(x;s,t)} \frac{\mu\, r}{\sinh(\mu\, r)}\right) \left(\frac{\mu\, t}{\sinh(\mu\, t)}\right)^{\beta}. \qquad (14.11)$$

14.2.4 Invariance Principle for the Record Times of the SPRW

Our goal is to state an invariance principle for $(\tau_t^n)_{t\geq 0}$ without any conditioning. But the same approach as the one leading to Proposition 14.5 stumbles over a computational difficulty that we will explain later and we are compelled to restrict to the case of symmetric perturbation.

Theorem 14.6 *Let* $(\tau_t^n)_{t\geq 0}$ *be the rescaled record process of the SPRW (*$\beta = \gamma$*).*

Then, as $n \to +\infty$*, the process* $(\tau_t^n)_{t\geq 0}$ *converges in law to a process* $(\tau_t)_{t\geq 0}$ *with independent non-negative increments whose law is given, for* $0 < s < t$*, by the Laplace transform*

$$\mathbb{E}\left[e^{-\frac{\mu^2}{2}(\tau_t-\tau_s)}\right] = \left(\frac{\cosh(\frac{\mu}{2}\, s)}{\cosh(\frac{\mu}{2}\, t)}\right)^{2\beta}.$$

This process has strictly increasing trajectories and is self-similar:

$$\forall a > 0, \quad (\tau_{a\,t})_{t\geq 0} \overset{(d)}{=} (a^2\, \tau_t)_{t\geq 0}. \qquad (14.12)$$

For any $t > 0$*, the density of* τ_t *on* $(0, +\infty)$ *is a (signed) mixture of* $1/2$*-stable laws:*

$$\phi_{\tau_t} : x \to \frac{2^{2\beta}}{\sqrt{2\pi}} \sum_{k=0}^{+\infty} \binom{-2\,\beta}{k} \frac{(\beta+k)\, t}{x^{3/2}}\; e^{-\frac{(\beta+k)^2\, t^2}{2x}} \qquad (14.13)$$

(using the usual notation (14.18)). Moreover we have the representation

$$\tau_t = \int_0^t \int_{\mathbb{R}_+} x \, \mathcal{N}(ds \; dx)$$

where $\mathcal{N}(ds \; dx)$ is a Poisson point measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $f_s(x)\, ds\, dx$ where

$$f_s(x) = \frac{2\,\pi^2}{s^3}\,\beta \sum_{n=1}^{+\infty} (2n-1)^2\, e^{-\frac{(2n-1)^2\pi^2}{2}\frac{x}{s^2}}.$$

14.2.5 Criterion of Positive Recurrence for the PRW

From the value of the transition matrix (14.4), it is clear that the signs of records cannot be asymptotically constant; indeed the infinite product of the diagonal terms of that transition matrix tends to 0. The PRW, whatever its parameters $\beta, \gamma \in (0, +\infty)$, is thus recurrent in the sense that every level is visited infinitely often, as it is the case for the SRW. But compared to the SRW which is null recurrent, the PRW can also become "positive recurrent" if the reinforcement parameters are high enough that is if the process is sufficiently self-attractive. We define the notion of positive recurrence in the non-Markov setting of PRW as follows but note that this property clearly implies that the return time to any level has finite mean. For $q \geq 1$, we introduce

$$C_q = \inf\{j > q;\;\; \chi_j = -\chi_q\}$$

as the index of the first record after record q of opposite sign. We say that the PRW is *positive recurrent* if and only if, for every integer $q \geq 1$,

$$\mathbb{E}\left(T_{C_q} - T_q\right) < +\infty.$$

Theorem 14.7 *The PRW is positive recurrent if and only if $\beta, \gamma \in (2, +\infty)$.*

14.2.6 Number of Visited Points for the SPRW

The number of visited points of the PRW is the inverse of the record time process: $V_k = \inf\{n \geq 1;\;\; T_n > k\}$. In the case of the SPRW ($\beta = \gamma$) the invariance principle stated in Theorem 14.6, with limit process $(\tau_t)_{t \geq 0}$, implies an invariance principle for $(V_k)_{k \geq 1}$.

Proposition 14.8 *Let $(Y_s)_{s\geq 0}$ be the non-decreasing process defined by $Y_s = \inf\{t;\ \tau_t > s\}$. This process has continuous trajectories and is self-similar: $\forall a > 0, \quad (Y_{a\,s})_{s\geq 0} \overset{(d)}{=} (\sqrt{a}\ Y_s)_{s\geq 0}$. Its marginal laws are*

$$\forall s > 0, \quad Y_s \overset{(d)}{=} \sqrt{\frac{1}{\tau_{1/\sqrt{s}}}} \overset{(d)}{=} \sqrt{\frac{s}{\tau_1}} \tag{14.14}$$

and, for any $s > 0$, the variable Y_s admits the density on $\mathbb{R}_+$ given by:

$$\phi_{Y_s}(x) = \frac{2\,s}{x^3}\ \phi_{\tau_1}\left(\frac{s}{x^2}\right) \tag{14.15}$$

$$= \frac{2^{2\beta+1}}{\sqrt{2\pi\ s}} \sum_{k=0}^{+\infty} \binom{-2\beta}{k} (\beta+k)\ e^{-\frac{(\beta+k)^2\,x^2}{2\,s}}. \tag{14.16}$$

Moreover, as $k \to +\infty$, the process $\left(\frac{V_{[ks]}}{\sqrt{k}}\right)_{s\geq 0}$ converges in law to the process $(Y_s)_{s\geq 0}$.

As a corollary of the invariance principle for the PRW proved by Davis, we have the representation

$$Y_s = \max_{u\in[0,s]} W_u - \min_{u\in[0,s]} W_u \tag{14.17}$$

where (W_u) is the perturbed Brownian motion as introduced in (14.1).

Also the process $(\tau_t)_{t\geq 0}$ obtained in Theorem 14.6 is the inverse of $(Y_s)_{s\geq 0}$. As a consequence it can be interpreted in terms of the perturbed Brownian motion via the representation (14.17) of $(Y_s)_{s\geq 0}$ given above.

14.2.7 Open Questions

Here are two questions we were unable to solve:

- find the law of y_1; this will describe how “non-centered” the range can be, asymptotically;
- obtain a generalization of Theorem 14.6 to the case $\beta \neq \gamma$.

As we will see later, both questions amount to find the value of an infinite matrix product. Alternatively the question about the law of y_1 can also be stated in terms of alternating renewal process.

14.3 Preliminary Lemmas

In the sequel we will use the classical notation of matrix exponential:

$$\exp[M] = \sum_{k=0}^{+\infty} \frac{1}{k!} M^k$$

for any real or complex square matrix M. We also recall the usual notation for generalized binomial coefficients

$$\binom{-\beta}{k} = \prod_{j=1}^{k} \frac{-\beta - j + 1}{j} = \frac{(-1)^k}{k!} \frac{\Gamma(k+\beta)}{\Gamma(\beta)} \tag{14.18}$$

where $\Gamma(\cdot)$ is the classical Gamma function. Let $0 < s < t$ be fixed. We introduce a notion of approximate equality of two quantities $a(k,n)$ and $b(k,n)$ up to terms of order $1/n^2$ by

$$a(k,n) \approx b(k,n) \Leftrightarrow \sup_{n \geq 1;\ [ns] \leq k \leq [nt]} n^2 \ |a(k,n) - b(k,n)| \ < +\infty. \tag{14.19}$$

Denote by $\mathcal{M}_{d\times d}$ the set of real $d \times d$ matrices. If $A(k,n) = \big(A_{x,y}(k,n)\big)_{1\leq x,y\leq d}$ and $B(k,n) = \big(B_{x,y}(k,n)\big)_{1\leq x,y\leq d}$ belong to $\mathcal{M}_{d\times d}$, we extend the previous notion by setting $A(k,n) \approx B(k,n)$ if and only if $A_{x,y}(k,n) \approx B_{x,y}(k,n)$ for all x, y.

Lemma 14.9 *Let us suppose that $g(\cdot)$ is a $\mathcal{M}_{d\times d}$-valued function which is continuous on $[s,t]$ and $f(k,n)$ is a $\mathcal{M}_{d\times d}$-valued function such that*

$$f(k,n) \approx I + \frac{1}{n}\, g\left(\frac{k}{n}\right).$$

Then

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} f(k,n) = \lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} \exp\left[\frac{1}{n}\, g\left(\frac{k}{n}\right)\right] \tag{14.20}$$

provided the limit on the right-hand side exists. Moreover, when $g(x)\ g(y) = g(y)\ g(x)$ for all $x, y \in [s,t]$, then

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} f(k,n) = \exp\left[\int_s^t g(x)\ dx\right]. \tag{14.21}$$

Proof As all matrix norms induce the same topology we may choose a matrix norm $||\cdot||$ which has the supplementary property that $||A\,B|| \le ||A||\,||B||$ for all matrices A, B and, as a consequence, $||\exp[A]|| \le e^{||A||}$. Also, using this property of the norm, the assumption easily implies that

$$f(k,n) = \exp\left[\frac{1}{n}\, g\left(\frac{k}{n}\right)\right] + \frac{1}{n^2}\, R(k,n) \text{ with } \sup_{n\ge 1}\ \sup_{[ns]\le k\le [nt]} ||R(k,n)|| < +\infty.$$

Then we have

$$\prod_{k=[ns]}^{[nt]} f(k,n) = \prod_{k=[ns]}^{[nt]} \exp\left[\frac{1}{n}\, g\left(\frac{k}{n}\right)\right] + \text{Rem}$$

where the remainder term is

$$\text{Rem} = \sum_{j=1}^{[nt]-[ns]+1}\ \sum_{J:\#(J)=j}\ \prod_{k=[ns]}^{[nt]} \left(\exp\left[\frac{1}{n}\, g\left(\frac{k}{n}\right)\right] \mathbf{1}_{\{k\notin J\}} + \frac{1}{n^2}\, R(k,n)\, \mathbf{1}_{\{k\in J\}}\right).$$

But the term in the product on the right-hand side has a norm bounded by $e^{H_\infty/n}$ if $k \notin J$ and by R_∞/n^2 if $k \in J$ where H_∞ and R_∞ are two constants. It follows that

$$\begin{aligned}
||\text{Rem}|| &\le \sum_{j=1}^{[nt]-[ns]+1} \binom{[nt]-[ns]+1}{j} \left(e^{\frac{H_\infty}{n}}\right)^{[nt]-[ns]+1-j} \left(\frac{R_\infty}{n^2}\right)^j \\
&= \left(e^{\frac{H_\infty}{n}} + \frac{R_\infty}{n^2}\right)^{[nt]-[ns]+1} - \left(e^{\frac{H_\infty}{n}}\right)^{[nt]-[ns]+1} \\
&= \left(e^{\frac{H_\infty}{n}}\right)^{[nt]-[ns]+1} \left[\left(1 + \frac{R_\infty\, e^{-\frac{H_\infty}{n}}}{n^2}\right)^{[nt]-[ns]+1} - 1\right].
\end{aligned}$$

The first term is bounded and the second one tends to zero; hence the remainder term converges to zero and we get (14.20). When the commutation property is satisfied by g, the classical property of the matrix exponential entails

$$\prod_{k=[ns]}^{[nt]} \exp\left[\frac{1}{n}\, g\left(\frac{k}{n}\right)\right] = \exp\left[\sum_{k=[ns]}^{[nt]} \frac{1}{n}\, g\left(\frac{k}{n}\right)\right]$$

and (14.21) follows immediately by a Riemann sum argument.

Lemma 14.10 *Let $(Y_n)_{n\geq 1}$ be a sequence of nonnegative random variables, ν be a probability on $\mathbb{R}_+$ and denote its Laplace transform by*

$$L_\nu(\mu) = \int_0^{+\infty} e^{-\frac{\mu^2}{2}x}\ \nu(dx).$$

Assume that

$$\lim_{n\to+\infty} \mathbb{E}\left[\left(\frac{1}{\cosh(\mu/n)}\right)^{Y_n}\right] = L_\nu(\mu)$$

uniformly with respect to μ belonging to a compact neighborhood of any positive value.

Then $\frac{Y_n}{n^2}$ converges in law to ν.

Proof Easy and omitted.

The following elementary lemma is for instance a part of Lemma 12 of [9] but we recall it for the convenience of the reader.

Lemma 14.11 (Time Spent in a Strip by a SRW) *Let $(X_n)_{n\geq 0}$ be a SRW started at 1 and ξ be the hitting time of $\{0, k\}$. Then*

$$J_{+1}(k,\lambda) = \mathbb{E}\left[(\cosh\lambda)^{-(1+\xi)}\ \mathbf{1}_{\{X_\xi=0\}}\right] = 1 - \frac{\tanh\lambda}{\tanh(k\lambda)} \tag{14.22}$$

and

$$J_{-1}(k,\lambda) = \mathbb{E}\left[(\cosh\lambda)^{-(1+\xi)}\ \mathbf{1}_{\{X_\xi=k\}}\right] = \frac{\tanh\lambda}{\sinh(k\lambda)}. \tag{14.23}$$

Lemma 14.12 (Time Needed for the PRW to Exit the Strip of Visited Points) *Let $k \in \{1, 2, 3, \dots\}$ and $(X_n)_{n\geq 0}$ be the Markov chain on $\mathbb{Z}$ whose transition probabilities $(p(x, y);\ x, y \in \mathbb{Z})$ are given by*

- $p(x, x+1) = p(x, x-1) = 1/2$ *if* $x \in \{1, \dots, k-1\}$
- $p(0, 1) = \frac{\gamma}{1+\gamma}$, $p(0, -1) = \frac{1}{1+\gamma}$
- $p(k, k+1) = \frac{1}{1+\beta}$, $p(k, k-1) = \frac{\beta}{1+\beta}$,

other transition probabilities being irrelevant for what follows. Let ζ be the hitting time of $\{-1, k+1\}$. We use the notation $\mathbb{E}_a[\,\cdot\,] = \mathbb{E}[\,\cdot \mid X_0 = a]$ and define,

$$G_k(\lambda, 1, 1) = \mathbb{E}_k\left[(\cosh\lambda)^{-\zeta}\ \mathbf{1}_{\{X_\zeta=k+1\}}\right] \tag{14.24}$$

$$G_k(\lambda, -1, -1) = \mathbb{E}_0\left[(\cosh\lambda)^{-\zeta}\ \mathbf{1}_{\{X_\zeta=-1\}}\right] \tag{14.25}$$

$$G_k(\lambda, 1, -1) = \mathbb{E}_k\left[(\cosh\lambda)^{-\zeta}\ \mathbf{1}_{\{X_\zeta=-1\}}\right] \tag{14.26}$$

$$G_k(\lambda, -1, 1) = \mathbb{E}_0\left[(\cosh\lambda)^{-\zeta}\ \mathbf{1}_{\{X_\zeta=k+1\}}\right]. \tag{14.27}$$

Then we have

$$G_k(\lambda, 1, 1) = \frac{\cosh\lambda\ \sinh(k\lambda) + \gamma \sinh\lambda\ \cosh(k\lambda)}{D_k(\lambda)} \tag{14.28}$$

$$G_k(\lambda, -1, -1) = \frac{\cosh\lambda\ \sinh(k\lambda) + \beta \sinh\lambda\ \cosh(k\lambda)}{D_k(\lambda)} \tag{14.29}$$

$$G_k(\lambda, 1, -1) = \frac{\beta\ \sinh\lambda}{D_k(\lambda)} \tag{14.30}$$

$$G_k(\lambda, -1, 1) = \frac{\gamma\ \sinh\lambda}{D_k(\lambda)} \tag{14.31}$$

where

$$D_k(\lambda) = \sinh(k\lambda)\ \left(1 + (1 + \beta\gamma) \sinh^2\lambda\right) + \frac{\beta + \gamma}{2}\ \sinh(2\lambda)\ \cosh(k\lambda). \tag{14.32}$$

Proof Let us start with $G_k(\lambda, -1, -1)$. To simplify notation we set $z = 1/(\cosh\lambda)$. We condition on the value of X_1 to get

$$G_k(\lambda, -1, -1) = \frac{1}{1+\gamma} z + \frac{\gamma}{1+\gamma} \mathbb{E}_1 \left[z^{1+\xi+\zeta\circ\theta_\xi}\ \mathbf{1}_{\{X_{\xi+\zeta\circ\theta_\xi} = -1\}} \right]$$

where ξ is the duration needed to reach $\{0, k\}$ and $\zeta \circ \theta_\xi$ is the duration needed after that time to hit -1 or $k+1$. We now use the strong Markov property at the stopping time ξ to get

$$\begin{aligned}
\mathbb{E}_1 \left[z^{1+\xi+\zeta\circ\theta_\xi} \mathbf{1}_{\{X_{\xi+\zeta\circ\theta_\xi} = -1\}} \right] &= \mathbb{E}_1 \left[z^{1+\xi}\ \mathbf{1}_{\{X_\xi = 0\}}\ \mathbb{E}_0 \left(z^\zeta\ \mathbf{1}_{\{X_\zeta = -1\}} \right) \right] \\
&\quad + \ \mathbb{E}_1 \left[z^{1+\xi}\ \mathbf{1}_{\{X_\xi = k\}}\ \mathbb{E}_k \left(z^\zeta\ \mathbf{1}_{\{X_\zeta = -1\}} \right) \right] \\
&= J_{+1}(k, \lambda)\, G_k(\lambda, -1, -1) + J_{-1}(k, \lambda)\, G_k(\lambda, 1, -1)
\end{aligned}$$

where $J_{-1}(k, \lambda)$ and $J_1(k, \lambda)$ are the functions introduced in Lemma 14.11. So we get a first equation on the G_k's as displayed on the first line below and we add three more equations by similar reasoning:

$$\begin{aligned}
G_k(\lambda, -1, -1) &= \frac{1}{1+\gamma} z + \frac{\gamma}{1+\gamma} J_{+1}(k, \lambda)\, G_k(\lambda, -1, -1) \\
&\quad + \frac{\gamma}{1+\gamma} J_{-1}(k, \lambda)\, G_k(\lambda, 1, -1) \\
G_k(\lambda, 1, 1) &= \frac{1}{1+\beta} z + \frac{\beta}{1+\beta} J_{+1}(k, \lambda)\, G_k(\lambda, 1, 1) \\
&\quad + \frac{\beta}{1+\beta} J_{-1}(k, \lambda)\, G_k(\lambda, -1, 1)
\end{aligned}$$

$$G_k(\lambda,-1,1) = \frac{\gamma}{1+\gamma} J_{+1}(k,\lambda)\, G_k(\lambda,-1,1) + \frac{\gamma}{1+\gamma} J_{-1}(k,\lambda)\, G_k(\lambda,1,1)$$

$$G_k(\lambda,1,-1) = \frac{\beta}{1+\beta} J_{+1}(k,\lambda)\, G_k(\lambda,1,-1) + \frac{\beta}{1+\beta} J_{-1}(k,\lambda)\, G_k(\lambda,-1,-1).$$

From the second and third equations we derive

$$G_k(\lambda,-1,1) = \frac{\frac{1}{1+\beta}\frac{\gamma}{1+\gamma}\, z\, J_{-1}(k,\lambda)}{\left(1-\frac{\beta}{1+\beta}J_1(k,\lambda)\right)\left(1-\frac{\gamma}{1+\gamma}J_1(k,\lambda)\right)-\frac{\beta}{1+\beta}\frac{\gamma}{1+\gamma}J_{-1}^2(k,\lambda)}.$$

We replace $J_{-1}(k,\lambda)$ and $J_1(k,\lambda)$ by their explicit values in terms of hyperbolic trigonometric functions. After a few lines of computation, we get (14.31). Then, by the third equation of the system above we obtain (14.28). Finally, (14.30) and (14.29) can be obtained by the substitution $\gamma \leftrightarrow \beta$.

14.4 Signs of Records of the PRW

We start with the **proof of Proposition 14.1**. The fact that $(\chi_n)_{n\geq 1}$ is a (time-inhomogeneous) Markov chain is clear. For $n \geq 1$ and $x, y \in \{-1, 1\}$, the transition probabilities are given by

$$p_n(x,y) = \mathbb{P}(\chi_{n+1} = y \mid \chi_n = x) = \lim_{\lambda\to 0} G_n(\lambda,x,y) \tag{14.33}$$

where the G_n's were introduced in (14.24)–(14.27) (for the second equality above apply Lebesgue's dominated convergence Theorem). Using the explicit values given by (14.28)–(14.31), it is straightforward to compute these limits and this completes the proof of (14.4).

We pass to the **proof of Proposition 14.2**. The requirements on $(x_t)_{t\in(0,+\infty)}$ impose the finite-dimensional marginal laws thus the uniqueness in law of $(x_t)_{t\in(0,+\infty)}$ is clear. The existence of this law is a consequence of the standard Kolmogorov extension Theorem, the compatibility condition following from the invariance of the probability defined by (14.6) for the transition matrices $T(s,t)$.

In order to prove the convergence in law of $(\chi_t^n)_{t\in[r,A]}$, for any $0 < r < A$, we first show the tightness. Recalling for instance Corollary 7.4 of [6], it suffices, in the present context, to show that, for any $\eta > 0$, we may find $\delta > 0$ such that for all n large enough, the probability that $(\chi_t^n)_{t\in[r,A]}$ has two jumps separated by less than

δ is lower than η. But this probability is lower than

$$\sum_{k=[nr]}^{[nA]} \sum_{j=1}^{[n\delta]} \mathbb{P}\left(\chi_{k+1} = -\chi_k\,,\ \chi_{k+j+1} = -\chi_{k+j}\right)$$
$$\leq \sum_{k=[nr]}^{[nA]} \sum_{j=1}^{[n\delta]} \frac{\beta \vee \gamma}{\beta+\gamma+k} \quad \frac{\beta \vee \gamma}{\beta+\gamma+k+j}. \tag{14.34}$$

Using the usual expansion of the partial sums of the harmonic series, we have

$$\sum_{j=1}^{[n\delta]} \frac{\beta \vee \gamma}{\beta+\gamma+k+j} \leq c\ \log\left(1+\frac{n\delta}{k}\right)$$

so that we can bound the expression (14.34) above by

$$c \sum_{k=[nr]}^{[nA]} \frac{1}{k} \log\left(1+\frac{n\delta}{k}\right) \leq c \int_{([nr]-1)/n}^{[nA]/n} \frac{1}{x} \log\left(1+\frac{\delta}{x}\right)\, dx$$

and this quantity tends to 0 as $\delta \downarrow 0$, uniformly in n large enough. This ends the proof of tightness.

To complete the proof of the proposition, it suffices to show firstly that, for $r > 0$, the law of χ_r^n converges to the law given by (14.6) which will follow from

$$\lim_{n\to+\infty} \prod_{k=1}^{[nr]} Q_k = \frac{1}{\beta+\gamma} \begin{pmatrix} \beta & \gamma \\ \beta & \gamma \end{pmatrix} \tag{14.35}$$

and secondly that the transition kernels also converge that is, for all $t > s > 0$,

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} Q_k = T(s,t). \tag{14.36}$$

We will only prove (14.36) since (14.35) is similar. Note that $Q_k = I + \frac{1}{\beta+\gamma+k} A$ where

$$A = \begin{pmatrix} -\gamma & \gamma \\ \beta & -\beta \end{pmatrix} = \Omega \begin{pmatrix} 0 & 0 \\ 0 & -\beta-\gamma \end{pmatrix} \Omega^{-1}$$

with

$$\Omega = \begin{pmatrix} 1 & \gamma \\ 1 & -\beta \end{pmatrix} \text{ and } \Omega^{-1} = \frac{1}{\beta+\gamma} \begin{pmatrix} \beta & \gamma \\ 1 & -1 \end{pmatrix}.$$

We deduce that

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} Q_k = \Omega \begin{pmatrix} 1 & 0 \\ 0 & L \end{pmatrix} \Omega^{-1} \tag{14.37}$$

where

$$L = \lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} \left(1 - \frac{\beta+\gamma}{\beta+\gamma+k}\right) \tag{14.38}$$

but this limit L is easily shown to be $(s/t)^{\beta+\gamma}$, for instance by expressing the product in values of the Gamma function and using the fact that, for $a > 0$, as $x \to +\infty$,

$$\frac{\Gamma(x+a)}{\Gamma(x)} \sim x^a. \tag{14.39}$$

Finally we check that the matrix product on the right-hand side of (14.37) is equal to $T(s,t)$ and it concludes the proof.

14.5 Record Values of the PRW

Let us start with the **proof of Corollary 14.3**. We rewrite (14.3) as

$$R_k = \chi_k \left(1 + \int_1^k \mathbf{1}_{\{\chi_{[u]}=\chi_k\}}\, du\right)$$

Changing the variable in the integral and substituting $[nt]$ for k we get

$$\frac{1}{n} R_{[nt]} = \chi_t^n \left(\frac{1}{n} + \int_{1/n}^{[nt]/n} \mathbf{1}_{\{\chi_y^n=\chi_t^n\}}\, dy\right).$$

For our purpose of convergence in law, by Skorohod representation Theorem and Proposition 14.2, we can suppose that, almost surely, (χ_t^n) converges to (x_t) in the Skorohod topology over every compact of $(0,+\infty)$. It follows easily that the right-hand side above converges—again with respect to the Skorohod topology over all compacts of $(0,+\infty)$—to the process (y_t) defined in the statement of the proposition.

We now give a **proof of Proposition 14.4**. As was noticed by Formula (14.3), the sequence $(R_n)_{n\geq 1}$ can be reconstructed from the sequence of record signs $(\chi_n)_{n\geq 1}$. The idea is to show that there are long sequences of records with the same sign.

Let q be a large integer. For $k \geq 1$ we introduce the events

$$A_k^1 = \{\chi_{q^{4k}} = 1\},$$

$$A_k^2 = \left\{\forall j \in [q^{4k}, q^{4k+1}], \quad \chi_j = \chi_{q^{4k}} = 1\right\},$$

$$A_k^3 = \left\{\exists! j_0 \in (q^{4k+1}, q^{4k+2}), \quad \chi_{j_0+1} = -\chi_{j_0}\right\},$$

$$A_k^4 = \left\{\forall j \in [q^{4k+2}, q^{4k+3}], \quad \chi_j = \chi_{q^{4k+2}} = -1\right\},$$

and $A_k = A_k^1 \cap A_k^2 \cap A_k^3 \cap A_k^4$. The probability of A_k^1 converges to $\gamma/(\beta+\gamma)$ as $k \to +\infty$. The probability of $A_k^1 \cap A_k^2$ is equal to $\mathbb{P}(A_k^1)$ multiplied by

$$\prod_{j=q^{4k}}^{q^{4k+1}-1} \left(1 - \frac{\beta}{\beta+\gamma+j}\right) = \frac{\Gamma(\beta+\gamma+q^{4k})\ \Gamma(\gamma+q^{4k+1})}{\Gamma(\gamma+q^{4k})\ \Gamma(\beta+\gamma+q^{4k+1})}.$$

This term converges to $q^{-\beta} > 0$ using (14.39).

Thus we claim that $\mathbb{P}(A_k^1 \cap A_k^2)$ is bounded from below by a positive constant. Now $\mathbb{P}(A_k^1 \cap A_k^2 \cap A_k^3)$ is equal to $\mathbb{P}(A_k^1 \cap A_k^2)$ multiplied by

$$\sum_{j_0=q^{4k+1}+1}^{q^{4k+2}-1} \prod_{j=q^{4k+1}}^{j_0-1} \left(1 - \frac{\beta}{\beta+\gamma+j}\right) \frac{\beta}{\beta+\gamma+j_0} \prod_{j=j_0+1}^{q^{4k+2}-1} \left(1 - \frac{\gamma}{\beta+\gamma+j}\right)$$

which is bounded from below by a positive constant, by the same arguments as above. Now $\mathbb{P}(A_k) = \mathbb{P}(A_k^1 \cap A_k^2 \cap A_k^3 \cap A_k^4)$ is equal to $\mathbb{P}(A_k^1 \cap A_k^2 \cap A_k^3)$ multiplied by

$$\prod_{j=q^{4k+2}}^{q^{4k+3}-1} \left(1 - \frac{\gamma}{\beta+\gamma+j}\right)$$

and repeating once more the same arguments we obtain finally that $\mathbb{P}(A_k)$ is bounded from below by a positive constant. We deduce that with positive probability the events A_k holds infinitely often. Note that on A_k, we have

$$\frac{R_{q^{4k+1}}}{q^{4k+1}} \geq \frac{q^{4k+1} - q^{4k}}{q^{4k+1}} = 1 - \frac{1}{q},$$

$$\frac{R_{q^{4k+3}}}{q^{4k+3}} \leq -\frac{q^{4k+3} - q^{4k+2}}{q^{4k+3}} = -\left(1 - \frac{1}{q}\right),$$

$$0 > \frac{R_{j_0+1}}{j_0+1} \geq \frac{-q^{4k}}{q^{4k+1}} = -\frac{1}{q},$$

$$0 < \frac{R_{j_1}}{j_1} \leq \frac{q^{4k+2}}{q^{4k+3}} = \frac{1}{q},$$

where j_1 denotes the time of the first change of sign of χ_n after time q^{4k+3}. From these remarks we deduce that the statements of Proposition 14.4 hold with positive probability. To conclude with probability 1 and thus complete the proof, we use the following zero-one law.

Proposition 14.13 (Zero-One Law) *Every event in the asymptotic σ-algebra*

$$\mathcal{A}(\chi) = \bigcap_n \sigma(\chi_k;\ k \geq n)$$

has probability zero or one.

Proof Let A belong to $\mathcal{A}(\chi)$ and take $B \in \sigma(\chi_k;\ k \leq m)$ of the form

$$B = \{\chi_1 = x_1, \dots, \chi_m = x_m\}$$

where $x_1, \dots, x_m \in \{-1, 1\}$. Take any $n > m$. Since $A \in \sigma(\chi_k;\ k \geq n)$ we may write

$$\begin{aligned}\mathbb{P}(A \mid B) - \mathbb{P}(A) &= \mathbb{P}(A|\chi_m = x_m) - \mathbb{P}(A)\\ &= \mathbb{P}(A|\chi_n = 1)\,[\mathbb{P}(\chi_n = 1|\chi_m = x_m) - \mathbb{P}(\chi_n = 1)]\\ &\quad + \mathbb{P}(A|\chi_n = -1)\,[\mathbb{P}(\chi_n = -1|\chi_m = x_m) - \mathbb{P}(\chi_n = -1)].\end{aligned} \tag{14.40}$$

By the same computation as the one leading to (14.35), we obtain

$$\begin{aligned}&\lim_{n\to+\infty}\begin{pmatrix}\mathbb{P}(\chi_n = -1|\chi_m = -1) & \mathbb{P}(\chi_n = 1|\chi_m = -1)\\ \mathbb{P}(\chi_n = -1|\chi_m = 1) & \mathbb{P}(\chi_n = 1|\chi_m = 1)\end{pmatrix}\\ &= \lim_{n\to+\infty}\prod_{k=m}^{n-1} Q_k = \frac{1}{\beta+\gamma}\begin{pmatrix}\beta & \gamma\\ \beta & \gamma\end{pmatrix}\end{aligned}$$

so that both quantities in square brackets in (14.40) tend to zero as $n \to +\infty$. As a consequence $\mathbb{P}(A|B) = \mathbb{P}(A)$ i.e. A is independent of B. Since this holds for all B of the specified form, in particular for all m, we deduce that $\mathcal{A}(\chi)$ is independent of the σ-algebra generated by all the variables χ_i, $i \geq 1$. But this σ-algebra contains $\mathcal{A}(\chi)$. Hence $\mathcal{A}(\chi)$ is independent of itself which ends the proof.

14.6 Record Times of the PRW: Conditional Case

Let us **prove Proposition 14.5**. We first concentrate on the convergence of finite-dimensional marginal laws. We will give the justification of tightness at the end of Sect. 14.7.

For $x, y \in \{-1, 1\}$ and $k \geq 1$, the quantity

$$\tilde{G}_k(\lambda, x, y) = \mathbb{E}\left[\left(\frac{1}{\cosh\lambda}\right)^{T_{k+1}-T_k} \;\middle|\; \chi_k = x,\ \chi_{k+1} = y\right] \tag{14.41}$$

is equal to $G_k(\lambda, x, y)/p_k(x, y)$ because of (14.24)–(14.27) and (14.33). For further reference we gather the explicit values in a matrix:

$$\begin{pmatrix} \tilde{G}_k(\lambda,-1,-1) & \tilde{G}_k(\lambda,-1,1) \\ \tilde{G}_k(\lambda,1,-1) & \tilde{G}_k(\lambda,1,1) \end{pmatrix} = \frac{k+\beta+\gamma}{D_k(\lambda)} \cdots$$
$$\cdots \times \begin{pmatrix} \frac{\cosh\lambda\ \sinh k\lambda + \beta \sinh\lambda \cosh k\lambda}{k+\beta} & \sinh\lambda \\ \sinh\lambda & \frac{\cosh\lambda\ \sinh k\lambda + \gamma \sinh\lambda \cosh k\lambda}{k+\gamma} \end{pmatrix}, \tag{14.42}$$

recalling the $D_k(\lambda)$ is given by (14.32). The increments of the record times $T_{k+1} - T_k,\ k \geq 1$ are independent random variables, even conditionally on the record signs $(\chi_k)_{k\geq 1}$. It follows that

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{\cosh\lambda}\right)^{T_{[nt]}-T_{[ns]}} \;\middle|\; (\chi_k)_{k\geq 1} = (x^n(k/n))_{k\geq 1}\right] \\
&= \prod_{k=[ns]}^{[nt]-1} \tilde{G}_k\left(\lambda, x^n\left(\frac{k}{n}\right), x^n\left(\frac{k+1}{n}\right)\right) \\
&= \prod_{\substack{[ns] \leq k < [nt] \\ x^n(k/n) \neq x^n((k+1)/n)}} (\sinh\lambda)\ \frac{k+\beta+\gamma}{D_k(\lambda)} \prod_{\substack{[ns] \leq k < [nt] \\ x^n(k/n) = x^n((k+1)/n)}} \\
&\quad \times \frac{(k+\beta+\gamma)\ \left(\cosh\lambda \sinh(k\lambda) + \tilde{\delta}(x^n(k/n)) \sinh\lambda \cosh(k\lambda)\right)}{D_k(\lambda)\ \left(k + \tilde{\delta}(x^n(k/n))\right)}
\end{aligned}$$

where $\tilde{\delta}(x) = \gamma$ if $x = 1$ and $\tilde{\delta}(x) = \beta$ if $x = -1$. We denote the set of discontinuities of the càdlàg function $x^n(\cdot)$ over $[s, t]$ by $\mathcal{D}(x^n; s, t) = \{r \in (s, t);\quad x^n(r-) \neq x^n(r)\}$. In order to get the asymptotic behaviour, we regroup the terms and set $\lambda = \frac{\mu}{n}$:

$$\mathbb{E}\left[\left(\frac{1}{\cosh\frac{\mu}{n}}\right)^{T_{[nt]}-T_{[ns]}} \,\Big|\, (\chi_k)_{k\geq 1} = \left(x^n\left(\frac{k}{n}\right)\right)_{k\geq 1}\right]$$
$$= \prod_{k=[ns]}^{[nt]-1} \frac{(k+\beta+\gamma)\quad \left(\cosh\frac{\mu}{n}\sinh\left(\frac{k\mu}{n}\right) + \tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\sinh\frac{\mu}{n}\cosh\left(\frac{k\mu}{n}\right)\right)}{D_k\left(\frac{\mu}{n}\right)\quad \left(k+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)}$$
$$\times \prod_{\frac{k+1}{n}\in\mathcal{D}(x^n;s,t)} \frac{\sinh\left(\frac{k\mu}{n}\right)\left(k+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)}{\left(\cosh\frac{\mu}{n}\sinh\left(\frac{k\mu}{n}\right) + \tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\sinh\frac{\mu}{n}\cosh\left(\frac{k\mu}{n}\right)\right)}. \tag{14.43}$$

Let us now perform asymptotic expansions up to the order $1/n^2$ in the sense of (14.19). We obtain that

$$D_k\left(\frac{\mu}{n}\right) \approx \sinh\left(\frac{k\mu}{n}\right) + (\beta+\gamma)\frac{\mu}{n}\cosh\left(\frac{k\mu}{n}\right).$$

Also we get

$$\frac{(k+\beta+\gamma)}{\left(k+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)} \approx 1 + \frac{1}{n}\left(\beta+\gamma-\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)\frac{1}{k/n}$$
$$\approx 1 + \frac{1}{n}\,\delta\left(x^n\left(\frac{k}{n}\right)\right)\frac{1}{k/n},$$

recalling that $\delta(y) = \beta+\gamma-\tilde{\delta}(y)$. Then we do similarly for the other terms in (14.43). Now we use Lemma 14.9 (1-dimensional case) to deduce that

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]-1} \frac{(k+\beta+\gamma)\quad \left(\cosh\frac{\mu}{n}\sinh\left(\frac{k\mu}{n}\right) + \tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\sinh\frac{\mu}{n}\cosh\left(\frac{k\mu}{n}\right)\right)}{D_k\left(\frac{\mu}{n}\right)\quad \left(k+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)}$$
$$= \exp\left(\int_s^t \delta(x(u))\left(\frac{1}{u} - \mu\coth\mu u\right)\, du\right). \tag{14.44}$$

To be precise we apply this Lemma a finite number of time, on each interval where $\delta(x^n(\cdot))$ is constant and at the limit, reunite all the integrals over these intervals into a single one.

The set $\mathcal{D}(x;s,t)$ is finite and the convergence of $(x^n(r))_{s\le r\le t}$ toward $(x(r))_{s\le r\le t}$ with respect to the Skorohod topology implies that, for any continuous function φ,

$$\lim_{n\to+\infty} \prod_{\frac{k+1}{n}\in\mathcal{D}(x^n;s,t)} \varphi(k/n) = \prod_{r\in\mathcal{D}(x;s,t)} \varphi(r).$$

As a consequence we deduce that,

$$\lim_{n\to+\infty} \prod_{k\in\mathcal{D}(x^n;s,t)} \frac{\sinh\left(\frac{k\mu}{n}\right)\left(k+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\right)}{\left(\cosh\frac{\mu}{n}\sinh\left(\frac{k\mu}{n}\right)+\tilde{\delta}\left(x^n\left(\frac{k}{n}\right)\right)\sinh\frac{\mu}{n}\cosh\left(\frac{k\mu}{n}\right)\right)}$$
$$= \prod_{r\in\mathcal{D}(x;s,t)} \frac{\mu r}{\sinh\mu r}\,.$$

Moreover by inspecting all the proof we see that the limits above are uniform for μ varying in any compact neighborhood of a fixed positive value. By Lemma 14.10, the proof of the convergence of finite-dimensional marginal laws is complete.

14.7 Record Times of the PRW: Unconditional Case

We start with the general case $\beta,\gamma\in(0,+\infty)$ to see how far we can go before being compelled to restrict to $\beta=\gamma$. The main step is to compute the limit in law of a rescaled increment $(T_{[nt]}-T_{[ns]})/n^2$ and, as before, this is done by computing the limit of

$$\mathbb{E}\left[\left(\frac{1}{\cosh\frac{\mu}{n}}\right)^{T_{[nt]}-T_{[ns]}}\right].$$

We set $\lambda=\mu/n$ and $z=1/\cosh\lambda$, as before. By the repeated use of the Markov property and the definition of $G_k(\cdot,\cdot,\cdot)$, we get, for fixed $x_{[ns]}\in\{-1,1\}$,

$$\mathbb{E}\left[z^{T_{[nt]}-T_{[ns]}}\;\middle|\;\chi_{[ns]}=x_{[ns]}\right]$$
$$=\sum_{\substack{x_k\in\{-1,1\}\\ [ns]<k\le[nt]}} \mathbb{E}\left[\left(\prod_{k=[ns]}^{[nt]-1} z^{T_{k+1}-T_k}\right)\prod_{k=[ns]+1}^{[nt]}\mathbf{1}_{\{\chi_k=x_k\}}\;\middle|\;\chi_{[ns]}=x_{[ns]}\right]$$

$$= \sum_{\substack{x_k \in \{-1,1\} \\ [ns] < k \le [nt]}} \prod_{k=[ns]}^{[nt]-1} G_k(\lambda, x_k, x_{k+1})$$

$$= \sum_{x_{[nt]}\in\{-1,1\}} \left(\prod_{k=[ns]}^{[nt]-1} G_k(\lambda) \right) (x_{[ns]}, x_{[nt]}) \qquad (14.45)$$

where $G_k(\lambda) = (G_k(\lambda, x, y))_{x,y\in\{-1,1\}}$ is a 2×2 matrix. We recall that

$$G_k(\lambda) = \frac{\cosh\lambda \ \sinh(k\lambda)}{D_k(\lambda)} \begin{pmatrix} 1 + \beta \tanh\lambda \operatorname{cotanh}(k\,\lambda) & \frac{\gamma \ \tanh\lambda}{\sinh(k\lambda)} \\ \frac{\beta \ \tanh\lambda}{\sinh(k\lambda)} & 1 + \gamma \tanh\lambda \operatorname{cotanh}(k\,\lambda) \end{pmatrix}.$$

Changing λ into μ/n and without conditioning, Formula (14.45) writes as

$$\mathbb{E}\left[\left(\frac{1}{\cosh \mu/n} \right)^{T_{[nt]}-T_{[ns]}} \right]$$

$$= \sum_{x\in\{-1,1\}} \mathbb{P}(\chi_{[ns]} = x) \sum_{y\in\{-1,1\}} \left(\prod_{k=[ns]}^{[nt]-1} G_k(\mu/n) \right) (x, y)\,. \qquad (14.46)$$

We want to pass to the limit $n \to +\infty$. It is easy to see, using again the notation (14.19), that

$$G_k\left(\frac{\mu}{n}\right) \approx I + \frac{1}{n}\tilde{H}\left(\frac{k}{n}\right)$$

where

$$\tilde{H}(x) = \begin{pmatrix} -\gamma\,\mu \operatorname{cotanh}(\mu\,x) & \gamma\,\mu \ \sinh^{-1}(\mu\,x) \\ \beta\,\mu \ \sinh^{-1}(\mu\,x) & -\beta\,\mu \operatorname{cotanh}(\mu\,x) \end{pmatrix}.$$

Then we would like to apply Lemma 14.9 and conclude that

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]-1} G_k(\mu/n) = \lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]-1} \exp\left[\frac{1}{n}\,\tilde{H}\left(\frac{k}{n}\right) \right]$$

provided the limit on the right-hand side exists. Unfortunately we are not able to prove the existence of the limit in the general case. The problem is that the matrices $\tilde{H}(\cdot)$ do not commute in the general case. We listed this limit as one of the open problems of Sect. 14.2.7. But in the particular case of symmetric perturbation

$\beta = \gamma$, the matrices $\tilde{H}(\cdot)$ do commute and we deduce by Lemma 14.9 that

$$\lim_{n\to+\infty} \prod_{k=[ns]}^{[nt]} \exp\left[\frac{1}{n}\,\tilde{H}\left(\frac{k}{n}\right)\right] = \exp\left[\int_s^t \tilde{H}(x)\,dx\right].$$

Moreover in this case we check that

$$\int_s^t \tilde{H}(x)\,dx = H(t) - H(s)$$

with

$$H(t) = \begin{pmatrix} -\beta\log\sinh(\mu\,t) & \beta\log\tanh(\mu\,t/2) \\ \beta\log\tanh(\mu\,t/2) & -\beta\log\sinh(\mu\,t) \end{pmatrix}.$$

Moreover we know that the probabilities $\mathbb{P}(\chi_{[ns]} = x),\ \ x \in \{-1, 1\}$ appearing in (14.46) simply converge in this case to $1/2$. Combining all these facts, the passage to the limit in (14.46) gives

$$\lim_{n\to+\infty} \mathbb{E}\left[\left(\frac{1}{\cosh\frac{\mu}{n}}\right)^{T_{[nt]}-T_{[ns]}}\right] = \frac{1}{2}\,(1\ \ 1)\ \ \exp\left[H(t) - H(s)\right]\ \ \begin{pmatrix}1\\1\end{pmatrix}$$

where, of course, the right-hand side should be read as a product of three matrices. But the matrix $H(t) - H(s)$ has the special form

$$H(t) - H(s) = \beta\,\begin{pmatrix} -a & b \\ b & -a \end{pmatrix}$$

where

$$a = \log\left(\frac{\sinh(\mu\,t)}{\sinh(\mu\,s)}\right) \text{ and } b = \log\left(\frac{\tanh\frac{\mu\,t}{2}}{\tanh\frac{\mu\,s}{2}}\right).$$

So, computing the exponential is easily done via the diagonalization

$$\begin{pmatrix} -a & b \\ b & a \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} -a+b & 0 \\ 0 & -a-b \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and we obtain finally

$$\lim_{n\to+\infty} \mathbb{E}\left[\left(\frac{1}{\cosh\frac{\mu}{n}}\right)^{T_{[nt]}-T_{[ns]}}\right] = \left(\frac{\cosh(\frac{\mu}{2}\,s)}{\cosh(\frac{\mu}{2}\,t)}\right)^{2\beta}$$

which leads to the desired convergence of finite-dimensional marginal laws, via the usual argument.

Now we want to address the problem of tightness of the laws of the processes $\big((\tau_t^n)_{t\geq 0},\ n \geq 1\big)$. We can use for instance the criterion stated in [1, Theorem 15.6] which consists, for any $T > 0$, in finding a nondecreasing continuous function F such that, for all $0 \leq t_1 \leq t \leq t_2 \leq T$ and all n large enough,

$$\mathbb{E}\left[(\tau_t^n - \tau_{t_1}^n)\,(\tau_{t_2}^n - \tau_t^n)\right] \leq [F(t_2) - F(t_1)]^2. \tag{14.47}$$

Note that only the case $t_2 - t_1 \geq 1/n$ has to be considered, otherwise the left-hand side vanishes. Let us recall that trivially

$$\mathbb{E}(\tau_t^n - \tau_s^n) = \frac{1}{n^2} \sum_{k=[ns]}^{[nt]-1} \mathbb{E}(T_{k+1} - T_k)\,. \tag{14.48}$$

By the Definition (14.41) we derive easily

$$\mathbb{E}\left(T_{k+1} - T_k \mid \chi_k = x,\ \chi_{k+1} = y\ \right) = -\lim_{\lambda\to 0} \frac{1}{\lambda}\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{G}_k(\lambda, x, y)\,. \tag{14.49}$$

The expressions of $\tilde{G}_k(\lambda, x, y)$ for $x, y \in \{-1, 1\}$ are explicitly given by (14.42). So it suffices to differentiate $\tilde{G}_k(\lambda, x, y)$ and replace every hyperbolic trigonometric function by its Taylor expansion around 0 (up to order 3) to get the value of the limit in (14.49). The computations are a bit tedious and left to the reader but the important fact is that there exists a constant c such that, for all k,

$$\mathbb{E}(T_{k+1} - T_k \mid \chi_{k+1} = \chi_k) \leq c\ k \tag{14.50}$$

and

$$\mathbb{E}(T_{k+1} - T_k \mid \chi_{k+1} = -\chi_k) \leq c\ k^2\,. \tag{14.51}$$

Moreover we have seen that, for a certain (other) constant c, $\mathbb{P}(\chi_{k+1} = -\chi_k) \leq \frac{c}{k}$ so that we conclude that $\mathbb{E}(T_{k+1} - T_k) \leq c\ k$ and as a consequence,

$$\mathbb{E}(\tau_t^n - \tau_s^n) \leq c\ ([nt] - [ns])/n.$$

Using the independence of the increments, (14.47) easily follows, with a linear function $F(\cdot)$ and tightness is assured.

Let us come back shortly to the conditional case where we claim that a similar proof of tightness can be constructed on every compact interval of $(0, +\infty)$. Indeed the same argument works on a time interval where the signs of the corresponding records are constant. Any time interval $[\varepsilon, T]$ with $T > \varepsilon > 0$ can be decomposed for every n into a (finite) partition such that on each interval of this partition the signs of the corresponding records, given by $x^n(\cdot)$ are constant. Because of the convergence in Skorohod topology of $x^n(\cdot)$ toward $x(\cdot)$, these partitions converge to the partition ruling the sign of $x(\cdot)$ over $[\varepsilon, T]$. So the relative compactness of the conditional laws over $[\varepsilon, T]$ can be deduced from the tightness guaranteed on each sub-interval of the partition.

The **end of the proof of Theorem 14.6** is similar to the proof of Theorem 1 in [9], except the multiplication by a power of 2 from places to places. Also we omit the **proof of Proposition 14.8** which is identical to the proof of Proposition 2 in [9].

14.8 Positive Recurrence

We now want to **prove Theorem 14.7**. By symmetry, it suffices to prove that $\beta > 2$ is equivalent to

$$\mathbb{E}\left(T_{\inf\{j>q;\ \chi_j=-1\}} - T_q \,\middle|\, \chi_q = 1\right) < +\infty.$$

This (conditional) expectation equals

$$\sum_{j=q}^{+\infty} \left\{ \left(\sum_{k=q}^{j-1} \mathbb{E}(T_{k+1} - T_k \mid \chi_{k+1} = \chi_k = 1) \right) + \mathbb{E}(T_{j+1} - T_j \mid \chi_j = 1 = -\chi_{j+1}) \right\}$$
$$\times \quad \mathbb{P}\left(\chi_{q+1} = \cdots = \chi_j = 1 = -\chi_{j+1} \mid \chi_q = 1\right). \tag{14.52}$$

But, by Proposition 14.1,

$$\mathbb{P}\left(\chi_{q+1} = \cdots = \chi_j = 1 = -\chi_{j+1} \mid \chi_q = 1\right)$$
$$= \left(\prod_{k=q}^{j-1} \left(1 - \frac{\beta}{\beta + \gamma + k} \right) \right) \frac{\beta}{\beta + \gamma + j} \tag{14.53}$$

and it is easy to see that this quantity is equivalent to $C\ j^{-\beta-1}$ where C is a constant.

We discussed at the end of the previous section the procedure to get the conditional means of $T_{k+1} - T_k$, see (14.50) and (14.51). This procedure shows also that, as $k, j \to +\infty$,

$$\mathbb{E}(T_{k+1} - T_k \mid \chi_{k+1} = \chi_k = 1) \sim \frac{2\beta}{3}\, k$$

and

$$\mathbb{E}(T_{j+1} - T_j \mid \chi_j = 1 = -\chi_{j+1}) \sim \frac{j^2}{3}.$$

Combining this with the estimate for (14.53) already obtained, we get that the expression of the (conditional) expectation given by (14.52) behaves like $\sum_j j^{1-\beta}$ hence is finite for $\beta > 2$ as announced.

References

1. P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968)
2. P. Carmona, F. Petit, M. Yor, Beta variables as times spent in $[0, +\infty)$ by certain perturbed Brownian motions. J. Lond. Math. Soc. **58**, 239–256 (1998)
3. L. Chaumont, R.A. Doney, Y. Hu, Upper and lower limits of doubly perturbed Brownian motion. Ann. Inst. H. Poincaré Probab. Stat. **36**(2), 219–249 (2000)
4. B. Davis, Weak limits of perturbed random walks and the equation $Y_t = B_t + \alpha \sup\{Y_s : s \leq t\} + \beta \inf\{Y_s ; s \leq t\}$. Ann. Probab. **24**(4), 2007–2023 (1996)
5. B. Davis, Brownian motion and random walk perturbed at extrema. Probab. Theory Relat. Fields **113**, 501–518 (1999)
6. S.N. Ethier, T.G. Kurtz, *Markov Processes; Characterization and Convergence* (Wiley, New York, 2009)
7. R. Pemantle, A survey of random processes with reinforcement. Probab. Surv. **4**, 1–79 (2007)
8. M. Perman, W. Werner, Perturbed Brownian motions. Probab. Theory Relat. Fields **108**, 357–383 (1997)
9. L. Serlet, Hitting times for the perturbed reflecting random walk. Stoch. Process. Appl. **123**, 110–130 (2013)
10. W. Werner, Some remarks on perturbed reflecting Brownian motion, in *Séminaire de Probabilités XXIX*. Lecture Notes in Mathematics, vol. 1613 (Springer, Berlin, 1995), pp. 37–43

Chapter 15
A Potential Theoretic Approach to Tanaka Formula for Asymmetric Lévy Processes

Hiroshi Tsukada

Abstract In this paper, we shall introduce the Tanaka formula from viewpoint of the Doob–Meyer decomposition. For symmetric Lévy processes, if the local time exists, Salminen and Yor (Tanaka formula for symmetric Lévy processes. In: Séminaire de Probabilités XL. Lecture notes in mathematics, vol. 1899, Springer, Berlin, pp. 265–285, 2007) obtained the Tanaka formula by using the potential theoretic techniques. On the other hand, for strictly stable processes with index $\alpha \in (1, 2)$, we studied the Tanaka formula by using Itô's stochastic calculus and the Fourier analysis. In this paper, we study the Tanaka formula for asymmetric Lévy processes via the potential theoretic approach. We give several examples for important processes. Our approach also gives the invariant excessive function with respect to the killed process in the case of asymmetric Lévy processes, and it generalized the result in Yano (J Math Ind 5(A):17–24, 2013).

Keywords Local time · Lévy process · Resolvent · Excessive function

MSC 2010 60J55, 60G51, 60J45

15.1 Introduction

In this paper, we shall focus on local times for Lévy processes. It is known that there are several definitions of local times for different stochastic processes, see Geman and Horowitz [6]. Thus, we define a local time $L = \{L_t^x : x \in \mathbb{R}, t \geq 0\}$ for a one-dimensional Lévy process X by the occupation density which means

H. Tsukada (✉)
Graduate School of Science, Osaka City University, Osaka, Japan

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIX*,
Lecture Notes in Mathematics 2215, https://doi.org/10.1007/978-3-319-92420-5_15

random variables $L = \{L_t^x : x \in \mathbb{R}, t \geq 0\}$ satisfying for each non-negative Borel measurable function f and $t \geq 0$,

$$\int_0^t f(X_s)ds = \int_{\mathbb{R}} f(x)L_t^x dx, \quad \text{a.s.,}$$

and is chosen as

$$L_t^x = \limsup_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t 1_{\{|X_s - x| < \varepsilon\}} ds \left(:= \lim_{\varepsilon \downarrow 0} \sup_{0<\delta<\varepsilon} \frac{1}{2\delta} \int_0^t 1_{\{|X_s - x| < \delta\}} ds \right),$$

by Lebesgue's differentiation theorem, see Bertoin [1, Chap.V].

For a real-valued Brownian motion $B = (B)_{t \geq 0}$, it is well known that the Tanaka formula holds:

$$|B_t - x| - |B_0 - x| = \int_0^t \operatorname{sgn}(B_s - x) dB_s + L_t^x,$$

where L_t^x denotes the local time of the Brownian motion at level x. It is well known as an important expression to understand the reflection problem (see, e.g. Chung and Williams [5]) and the Ray–Knight theorem (see, Jeulin [8]). It also represents that the local time L^x can be understood as a bounded variation process in the Doob–Meyer decomposition on the positive submartingale $|B - x|$. Our goal in this paper is to construct the Tanaka formula from the viewpoint of the Doob–Meyer decomposition.

The Tanaka formula has already studied for symmetric stable processes with index $\alpha \in (1, 2)$ by Yamada [13], for symmetric Lévy processes by Salminen and Yor [10]. In this paper, we are interested in asymmetric Lévy processes, while the formula has been obtained for strictly stable processes in [12]. We shall construct the Tanaka formula for asymmetric Lévy processes based upon the potential theoretic approach as stated in [10]. Moreover, it will clearly extend the original Tanaka formula for Brownian motions to our formula for Lévy processes.

In [12], we have already obtained the Tanaka formula for strictly stable processes with index $\alpha \in (1, 2)$ via Itô's stochastic calculus. By using the Fourier transform, we could obtain the fundamental solution F of the infinitesimal generator $\mathscr{L}$ for strictly stable processes:

$$\mathscr{L}(F * \phi)(x) = \phi(x), \quad \phi \in \mathscr{S}(\mathbb{R}),$$

where $F * \phi$ is the convolution of F and ϕ. By using Itô's stochastic calculus and the scaling property, we could construct the Tanaka formula for a strictly stable process $S = (S_t)_{t \geq 0}$ with index $\alpha \in (1, 2)$:

$$F(S_t - x) - F(S_0 - x) = M_t^x + L_t^x,$$

where the process $(M_t^x)_{t\geq 0}$ given by

$$M_t^x := \int_0^t \int_{\mathbb{R}\setminus\{0\}} \{F(S_{s-} - x + h) - F(S_{s-} - x)\}\tilde{N}(ds, dh),$$

is a square integrable martingale and L_t^x is the local time at level x. Here, $\tilde{N}(ds, dh)$ is the compensated Poisson random measure. But it is not clear whether a similar representation can be obtained for general Lévy processes, or not, because it is very difficult to find the fundamental solution of the infinitesimal generator for Lévy processes.

In [10], Salminen and Yor have constructed the Tanaka formula for a symmetric Lévy process $X = (X_t)_{t\geq 0}$ via the potential theoretic approach, if the local time exists. By using the continuous resolvent density r_q, they could construct the Tanaka formula:

$$h(X_t - x) - h(x) = M_t^x + L_t^x$$

where $h(x) := \lim_{q\downarrow 0}(r_q(0) - r_q(x))$ which is called a renormalized zero resolvent, $M_t^x := -\lim_{q\downarrow 0} M_t^{q,x}$ is a martingale and L_t^x is the local time at level x. But the expression of the martingale part M_t^x was not given.

In [14] and [15], Yano obtained an invariant excessive function h with respect to the killed process:

$$\mathbb{E}_x^0[h(X_t)] = h(x)$$

where $\mathbb{E}_x^0$ is the expectation with respect to the law of a Lévy process X starting at x killed upon hitting zero. The function h is associated with the Tanaka formula for the local time at level zero because the local time for such a process at level zero becomes zero. In the symmetric case, Yano [14] needed a necessary and sufficient condition for the existence of local times, and Salminen and Yor [10] also needed the same condition, but in the asymmetric case, Yano [15] needed sufficient conditions for the existence of the function and its expression. Our result also gives the existence and its expression under weaker conditions than the ones in [15].

In Sect. 15.2, we shall give the preliminaries about resolvent operators of Lévy processes and a connection between the local time and the resolvent density. In Sect. 15.3, the convergence and its expression of the renormalized zero resolvent are mentioned. In Sect. 15.4, the Doob–Meyer decomposition can be constructed in the case of asymmetric Lévy processes. And then, we obtain the Tanaka formula for asymmetric Lévy processes and the invariant excessive function with respect to the killed process. In Sect. 15.5, we give several examples that satisfy the conditions introduced in Sects. 15.2 and 15.3.

15.2 Preliminaries

Let $\mathscr{S}(\mathbb{R})$ be the Schwartz space of rapidly decreasing functions on $\mathbb{R}$. We denote the law of processes starting at x and the corresponding expectation by $\mathbb{P}_x$ and $\mathbb{E}_x$ respectively.

Consider a Lévy process $X = (X_t)_{t\geq 0}$ on $\mathbb{R}$ with the Lévy–Khintchine representation given by

$$\mathbb{E}_0[e^{iuX_t}] = e^{t\eta(u)},$$

where the Lévy symbol η of X can be represented as

$$\eta(u) = ibu - \frac{1}{2}au^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{iuy} - 1 - iuy1_{|y|\leq 1}\right) \nu(dy)$$

for a drift parameter $b \in \mathbb{R}$, a Gaussian coefficient $a \geq 0$ and a Lévy measure ν on $\mathbb{R} \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}\setminus\{0\}} (y^2 \wedge 1)\nu(dy) < \infty.$$

We note that the Lévy symbol η is continuous. Let θ and ω be the real and imaginary parts of η respectively:

$$\theta(u) := \Re\eta(u) = -\frac{1}{2}au^2 + \int_{\mathbb{R}\setminus\{0\}} (\cos(uy) - 1)\, \nu(dy),$$

$$\omega(u) := \Im\eta(u) = bu + \int_{\mathbb{R}\setminus\{0\}} \left(\sin(uy) - uy1_{|y|\leq 1}\right) \nu(dy),$$

for $u \in \mathbb{R}$. Remark that $\theta \leq 0$, θ is even and ω is odd.

The resolvent operator of a Lévy process X is defined by

$$R_q f(x) := \mathbb{E}_x \left[\int_0^\infty e^{-qt} f(X_t)dt\right], \quad q > 0, x \in \mathbb{R}$$

for all bounded Borel measurable function f.

Using the Fourier transform of $f \in L^1(\mathbb{R})$ defined by

$$\mathscr{F}[f](u) := \int_{\mathbb{R}} e^{-iux} f(x)dx, \quad u \in \mathbb{R},$$

and the inverse Fourier transform defined by

$$\mathscr{F}^{-1}[f](x) := \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} f(u)du, \quad x \in \mathbb{R},$$

the resolvent operator is also represented as follows:

Proposition 15.1 ([1, Proposition I.9]) *For any* $f \in \mathscr{S}(\mathbb{R})$ *and* $x \in \mathbb{R}$,

$$R_q f(x) = \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\mathscr{F}[f](u)\right](x), \quad q > 0.$$

For $q > 0$, the resolvent kernel $R_q(x, dy)$ is defined by

$$R_q f(x) = \int_{\mathbb{R}} f(y) R_q(x, dy)$$

for all bounded Borel measurable function f and $x \in \mathbb{R}$. If there exists its density with respect to the Lebesgue measure, then we define the resolvent density by $r_q(x)$ such that

$$R_q f(x) = \int_{\mathbb{R}} f(y) r_q(y - x) dy.$$

Let T_0 be the first hitting time to 0 of X:

$$T_0 := \inf\{t > 0 : X_t = 0\}.$$

We say that 0 is regular for itself if $\mathbb{P}_0(T_0 = 0) = 1$, and irregular for itself otherwise. From the Blumenthal zero-one law, 0 is irregular if $\mathbb{P}_0(T_0 = 0) = 0$.

We introduce the following conditions:

(A1) The Lévy symbol η satisfies that

$$\int_{\mathbb{R}} \Re\left(\frac{1}{q-\eta(u)}\right) du < \infty, \quad \text{for all } q > 0,$$

(A2) 0 is regular for itself.

It is known that the condition **(A1)** is equivalent to the existence of the resolvent density. See [1, 3, 9].

Lemma 15.1 ([1, Theorem II.16]) *The condition* ***(A1)*** *holds if and only if the resolvent kernel* $R_q(0, dy)$ *is absolutely continuous with respect to the Lebesgue measure and has the bounded resolvent density.*

Using the Fourier transform for $L^2(\mathbb{R})$-functions, the resolvent density can be represented as follows:

Proposition 15.2 *Suppose that the condition* ***(A1)*** *hold. The bounded continuous resolvent density can be expressed as:*

$$r_q(x) = \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\right](-x) \quad a.e.$$

for all $q > 0$.

Proof Since we have for $q > 0$

$$\left|\frac{1}{q-\eta(u)}\right|^2 \le \frac{\Re(q-\eta(u))}{q|q-\eta(u)|^2} = \frac{1}{q}\Re\left(\frac{1}{q-\eta(u)}\right),$$

it follows by the condition **(A1)** that

$$\frac{1}{q-\eta(u)} \in L^2(\mathbb{R}).$$

Hence, by Proposition 15.1 and Parseval's theorem, we have for all $\phi \in \mathscr{S}(\mathbb{R})$,

$$\begin{aligned} R_q\phi(x) &= \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\mathscr{F}[\phi](u)\right](x) \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{iux}}{q-\eta(u)}\mathscr{F}[\phi](u)du \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathscr{F}\left[\frac{e^{iux}}{q-\eta(u)}\right](y)\phi(y)dy \\ &= \int_{\mathbb{R}} \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\right](x-y)\phi(y)dy. \end{aligned}$$

By Lemma 15.1, we then have for all $\phi \in \mathscr{S}(\mathbb{R})$,

$$\int_{\mathbb{R}} \left(r_q(y) - \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\right](-y)\right)\phi(y)dy = 0.$$

Since r_q is integrable, we have

$$r_q(x) = \mathscr{F}^{-1}\left[\frac{1}{q-\eta(u)}\right](-x) \quad \text{a.e.}$$

for all $q > 0$. □

It is known that the condition **(A2)** is equivalent to the continuity of the resolvent density. See [1–3, 7, 9].

Lemma 15.2 ([1, Theorem II.19]) *Suppose that the condition **(A1)** holds. Then, the followings hold for all $q > 0$:*

*(i) The condition **(A2)** holds if and only if there exist a bounded continuous resolvent density r_q such that*

$$R_q f(x) = \int_{\mathbb{R}} f(y)r_q(y-x)dy,$$

for all bounded Borel measurable function f and that

$$\mathbb{E}_x[e^{-qT_0}] = \frac{r_q(-x)}{r_q(0)}, \quad x \in \mathbb{R}.$$

(ii) If r_q is continuous, then

$$r_q(0) = \frac{1}{\pi}\int_0^\infty \Re\left(\frac{1}{q-\eta(u)}\right) du,$$

and for all $x \in \mathbb{R}$

$$2r_q(0) - \{r_q(x) + r_q(-x)\} = \frac{2}{\pi}\int_0^\infty \Re\left(\frac{1-\cos(ux)}{q-\eta(u)}\right) du.$$

We introduce the following condition which is stronger than the condition **(A1)**:

(A) The Lévy symbol η satisfies that

$$\frac{1}{q-\eta(u)} \in L^1(\mathbb{R}), \quad \text{for all } q > 0.$$

Corollary 15.1 *Suppose that the condition **(A)** holds. The bounded continuous resolvent density r_q can be expressed as:*

$$r_q(x) = \frac{1}{\pi}\int_0^\infty \Re\left(\frac{e^{-iux}}{q-\eta(u)}\right) du$$

for all $q > 0$ and $x \in \mathbb{R}$.

Proof By Proposition 15.2 and the condition **(A)**, we have

$$r_q(x) = \frac{1}{\pi}\int_0^\infty \Re\left(\frac{e^{-iux}}{q-\eta(u)}\right) du,$$

and it follows from the dominated convergence theorem that $r_q(\cdot)$ is continuous. □

From Corollary 15.1 and Lemma 15.2(i), we have the following:

Corollary 15.2 *If the condition **(A)** holds, then the conditions **(A1)** and **(A2)** hold.*

Remark 15.1 In the case of a 1-stable, not strictly stable, process which is called an asymmetric Cauchy process, the conditions **(A1)** and **(A2)** hold but the condition **(A)** does not hold.

We introduce the following conditions that are associated with the conditions **(A1)** and **(A2)**:

(A3) The process X is the type C, i.e.,

$$\text{either} \quad a > 0 \quad \text{or} \quad \int_{|y|\leq 1} |y|\nu(dy) = \infty,$$

(A4) The process X is not a compound Poisson process.

The following was proved by Kesten [9], and another proof was given by Bretagnolle [3].

Lemma 15.3 ([9] and [3]) *The conditions* ***(A1)*** *and* ***(A3)*** *hold if and only if the conditions* ***(A2)*** *and* ***(A4)*** *hold. Furthermore, under the condition* ***(A1)****, the condition* ***(A2)*** *holds if and only if the condition* ***(A3)*** *holds.*

Remark 15.2 From Corollary 15.2 and Lemma 15.3, we have that if the condition **(A)** holds, then the conditions **(A1)–(A4)** hold.

In order to construct the Tanaka formula via the techniques in the potential theory, we use a connection between the local time and the resolvent density.

Lemma 15.4 ([1, Lemma V.3]) *Suppose that the conditions* ***(A1)*** *and* ***(A2)*** *hold. For any $x \in \mathbb{R}$, denote by dL_t^x the Stieltjes measure of the increasing function $L_\cdot^x$. Then, it holds that*

$$\mathbb{E}_y \left[\int_0^\infty e^{-qt} dL_t^x \right] = r_q(x - y), \quad q > 0, y \in \mathbb{R}.$$

Remark 15.3 In [1, Chap.V], the condition **(A1)** holds if and only if the occupation measure μ_t satisfying for each non-negative Borel measurable function f and $t \geq 0$,

$$\int_0^t f(X_s)ds = \int_{\mathbb{R}} f(x)\mu_t(dx),$$

has the density in $L^2(dx \otimes d\mathbb{P}_0)$ as the Radon–Nikodym derivative. Therefore, if the condition **(A1)** holds, then local times for Lévy processes exist. Moreover, under the condition **(A1)**, the local time L_t^x is continuous almost surely with respect to t if the condition **(A2)** holds. In the symmetric case, if the condition **(A1)** holds, then the condition **(A2)** holds.

Remark 15.4 By Blumenthal and Getoor [2], it can be considered as the potential theoretic definition of local times, i.e. the local time can be defined as a positive additive functional L_t^x such that

$$\mathbb{E}_0 \left[\int_0^\infty e^{-qt} dL_t^x \right] = r_q(x).$$

15.3 Renormalized Zero Resolvent

Now, we set

$$h_q(x) := r_q(0) - r_q(-x), \quad q > 0, x \in \mathbb{R}.$$

Since $0 \le r_q(y) \le r_q(0)$ for all $y \in \mathbb{R}$ by Lemma 15.2(i), we have $h_q \ge 0$. In [15], the limit $h := \lim_{q\downarrow 0} h_q$ is called the renormalized zero resolvent if the limit exists, which is known as a harmonic function for the killed process under some conditions.

But its convergence of h_q is not clear for the asymmetric case, and Yano [15] needed the following conditions:

(L1) The Lévy symbol η satisfies that

$$\int_0^\infty \frac{1}{q - \theta(u)} du < \infty, \quad \text{for all } q > 0,$$

(L2) The process X is the type C, that is the same condition as **(A3)**,
(L3) The real and imaginary parts of the Lévy symbol η have measurable derivatives on $(0, \infty)$ which satisfy

$$\int_0^\infty (u^2 \wedge 1) \frac{|\theta'(u)| + |\omega'(u)|}{\theta(u)^2 + \omega(u)^2} du < \infty.$$

However, our condition **(A)** is weaker than the condition **(L1)**, and if the condition **(A)** holds, then the condition **(L2)** holds. Moreover, we shall introduce the condition **(B)** which is weaker than the condition **(L3)** under the condition **(A)**:

(B) The Lévy symbol η satisfies that

$$\int_0^1 \left| \Im \left(\frac{u}{\eta(u)} \right) \right| du < \infty.$$

Proposition 15.3 *Suppose that the condition **(A)** holds. If the condition **(L3)** holds, then the condition **(B)** holds.*

Proof By the condition **(A)**, we have $\theta(u) \neq 0$ if $u \neq 0$. Using the inequality: $1 - \cos(x) \ge x^2/4$ for $|x| \le 1$, we have for all $|u| \le 1$

$$\begin{aligned} |\eta(u)| &\ge -\theta(u) \\ &\ge \left(\frac{a}{2} + \int_{|y| \le |u|^{-1}} \frac{1 - \cos(uy)}{(uy)^2} y^2 \nu(dy) \right) u^2 \end{aligned}$$

$$\geq \left(\frac{a}{2} + \frac{1}{4}\int_{|y|\leq|u|^{-1}} y^2\nu(dy)\right)u^2$$
$$\geq \left(\frac{a}{2} + \frac{1}{4}\int_{|y|\leq 1} y^2\nu(dy)\right)u^2 \geq 0.$$

Hence, by using integration by parts, we have

$$\begin{aligned}\int_0^1 \left|\Im\left(\frac{u}{\eta(u)}\right)\right| du \leq \int_0^1 \frac{u}{|\eta(u)|}du &= \left[\frac{u^2}{2|\eta(u)|}\right]_0^1 - \int_0^1 \frac{u^2}{2}\left(\frac{1}{|\eta(u)|}\right)' du \\ &= \frac{1}{|\eta(1)|} + \int_0^1 \frac{u^2(\theta(u)\theta'(u) + \omega(u)\omega'(u))}{(\theta(u)^2 + \omega(u)^2)\sqrt{\theta(u)^2 + \omega(u)^2}}du \\ &\leq \frac{1}{|\eta(1)|} + \int_0^1 \frac{u^2(|\theta'(u)| + |\omega'(u)|)}{\theta(u)^2 + \omega(u)^2}du < \infty.\end{aligned}$$

The proof is now complete. □

Under the conditions **(A)** and **(B)**, we obtain the following renormalized zero resolvent for general Levy processes:

Theorem 15.1 *Suppose that the conditions* **(A)** *and* **(B)** *hold. For all* $x \in \mathbb{R}$,

$$\lim_{q\downarrow 0} h_q(x) = \frac{1}{\pi}\int_0^\infty \Re\left(\frac{e^{iux} - 1}{\eta(u)}\right) du =: h(x).$$

In order to show Theorem 15.1 and establish the Tanaka formula, we need the following lemma:

Lemma 15.5 *Suppose that the condition* **(A)** *holds. Then, the followings hold:*

(i) $\quad |\eta(u)| \to \infty \quad as\ |u| \to \infty.$

(ii) $\quad \displaystyle\int_c^\infty \left|\frac{1}{\eta(u)}\right| du < \infty \quad for\ all\ c > 0.$

(iii) $\quad \displaystyle\int_0^c \left|\frac{u^2}{\eta(u)}\right| du < \infty \quad for\ all\ c > 0.$

(iv) $\quad \displaystyle\lim_{q\downarrow 0}\int_{\mathbb{R}} \left|\frac{q}{q - \eta(u)}\right| du = 0.$

Proof

(i) Since r_1 is integrable, it follows that

$$|\mathscr{F}[r_1](u)| = \left|\frac{1}{1-\eta(-u)}\right| \geq \frac{1}{1+|\eta(-u)|},$$

by Proposition 15.2. Hence, by the Riemann–Lebesgue theorem, we have $|\eta(u)| \to \infty$ as $|u| \to \infty$.

(ii) We know $\theta(u) \neq 0$ if $u \neq 0$. By the condition **(A)** and the assertion (i), we have

$$\left|\frac{\eta(u)}{1-\eta(u)}\right| \to 1 \quad \text{as } |u| \to \infty.$$

Hence, the required result follows.

(iii) By the argument in the proof of Proposition 15.3, it follows that for $0 < u \leq 1$,

$$\left|\frac{\eta(u)}{u^2}\right| \geq \frac{a}{2} + \frac{1}{4}\int_{|y|\leq 1} y^2 \nu(dy) > 0.$$

Hence, the required result follows from the dominated convergence theorem.

(iv) Since we have for each $q < 1$

$$\left|\frac{q}{q-\eta(u)}\right| \leq 1 \wedge \left|\frac{1}{\eta(u)}\right|,$$

it follows from the dominated convergence theorem that

$$\lim_{q\downarrow 0}\int_{\mathbb{R}}\left|\frac{q}{q-\eta(u)}\right| du = \int_{\mathbb{R}}\lim_{q\downarrow 0}\left|\frac{q}{q-\eta(u)}\right| du = 0.$$

The proof is now complete. □

Now, we shall prove Theorem 15.1.

Proof (Proof of Theorem 15.1) By Corollary 15.1, we have

$$\begin{aligned} h_q(x) &= \frac{1}{\pi}\int_0^\infty \Re\left(\frac{1-e^{iux}}{q-\eta(u)}\right) du \\ &= \frac{1}{\pi}\int_0^\infty \Re\left(\frac{1-\cos(ux)}{q-\eta(u)}\right) du + \frac{1}{\pi}\int_0^\infty \Im\left(\frac{\sin(ux)}{q-\eta(u)}\right) du. \end{aligned}$$

Using the inequality: $1-\cos(y) \leq y^2 \wedge 2$ for $y \in \mathbb{R}$, we have

$$\left|\Re\left(\frac{1-\cos(u)}{q-\eta(u)}\right)\right| \leq \frac{u^2 \wedge 2}{|\eta(u)|} \in L^1(\mathbb{R}),$$

by Lemma 15.5(ii) and (iii). Hence, it follows from the dominated convergence theorem that

$$\int_0^\infty \Re\left(\frac{1-\cos(u)}{q-\eta(u)}\right)du \to \int_0^\infty \Re\left(\frac{\cos(u)-1}{\eta(u)}\right)du,$$

as $q \downarrow 0$.

By the condition **(B)** and Lemma 15.5(ii), we have

$$\left|\Im\left(\frac{\sin(u)}{q-\eta(u)}\right)\right| \leq \left|\Im\left(\frac{u\wedge 1}{\eta(u)}\right)\right| \leq \left|\Im\left(\frac{u}{\eta(u)}\right)\right| \wedge \left|\frac{1}{\eta(u)}\right| \in L^1(\mathbb{R}).$$

Hence, it follows from the dominated convergence theorem that

$$\int_0^\infty \Im\left(\frac{\sin(ux)}{q-\eta(u)}\right)du \to -\int_0^\infty \Im\left(\frac{\sin(ux)}{\eta(u)}\right)du,$$

as $q \downarrow 0$. □

15.4 Tanaka Formula

Using Lemma 15.4, we can construct the Doob–Meyer decomposition as stated in [10, Proposition 1].

Proposition 15.4 *Suppose that the conditions **(A1)** and **(A2)** hold. For each $q > 0$, $t \geqslant 0$ and $x \in \mathbb{R}$, it holds that*

$$r_q(-X_t + x) = r_q(-X_0 + x) + M_t^{q,x} + q\int_0^t r_q(-X_s + x)ds - L_t^x,$$

where $M_t^{q,x}$ is a martingale with respect to the natural filtration $\{\mathcal{G}_t\}_{t\geq 0}$ of X.

Proof By Lemma 15.4 and by the Markov property, we have

$$\begin{aligned}\mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathcal{G}_s\right] &= \int_0^s e^{-qu}dL_u^x + \mathbb{E}_{X_s}\left[\int_0^\infty e^{-q(s+u)}dL_u^x\right]\\ &= \int_0^s e^{-qu}dL_u^x + e^{-qs}r_q(-X_s + x).\end{aligned}$$

By using integration by parts, we obtain

$$\begin{aligned}
&q\int_0^t e^{qs}\int_0^s e^{-qu}dL_u^x ds\\
&= e^{qt}\int_0^t e^{-qu}dL_u^x - L_t^x\\
&= e^{qt}\left(\mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathscr{G}_t\right] - e^{-qt}r_q(-X_t+x)\right) - L_t^x\\
&= e^{qt}\mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathscr{G}_t\right] - r_q(-X_t+x) - L_t^x.
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
&r_q(-X_t+x) - q\int_0^t r_q(-X_s+x)ds + L_t^x\\
&= -q\int_0^t e^{qs}\mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathscr{G}_s\right]ds + e^{qt}\mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathscr{G}_t\right]
\end{aligned}$$

For the sake of simplicity of notations, we shall write

$$\begin{aligned}
Y_t &:= \mathbb{E}_{X_0}\left[\int_0^\infty e^{-qu}dL_u^x|\mathscr{G}_t\right],\\
Z_t &:= -q\int_0^t e^{qs}Y_s ds + e^{qt}Y_t.
\end{aligned}$$

Since $Z_0 = r_q(-X_0+x)$, we will show that Z_t is a martingale with respect to the natural filtration $\{\mathscr{G}_t\}_{t\geq 0}$. By Fubini's theorem, we have for all $0 \leq v < t$,

$$\begin{aligned}
\mathbb{E}_{X_0}[Z_t|\mathscr{G}_v] &= -q\int_0^t e^{qs}\mathbb{E}_{X_0}[Y_s|\mathscr{G}_v]ds + e^{qt}\mathbb{E}_{X_0}[Y_t|\mathscr{G}_v]\\
&= -q\int_0^v e^{qs}Y_s ds - q\int_v^t e^{qs}Y_v ds + e^{qt}Y_v\\
&= -q\int_0^v e^{qs}Y_s ds + e^{qv}Y_v = Z_v,
\end{aligned}$$

and the required result follows. □

Now we will establish the Tanaka formula for general Lévy processes.

Theorem 15.2 *Suppose that the conditions* ***(A)*** *and* ***(B)*** *hold. Let h and $M^{q,x}$ be the same as in Theorem 15.1 and Proposition 15.4 respectively. Then, for each $t \geq 0$*

and $x \in \mathbb{R}$, *it holds that*

$$h(X_t - x) = h(X_0 - x) + M_t^x + L_t^x,$$

where $M_t^x := -\lim_{q\downarrow 0} M_t^{q,x}$ *is a martingale.*

Proof From the Doob–Meyer decomposition (Proposition 15.4), let $q \downarrow 0$, we have

$$h(X_t - x) = h(X_0 - x) - \lim_{q\downarrow 0}\left(M_t^{q,x} + q\int_0^t r_q(-X_s + x)ds\right) + L_t^x.$$

by Theorem 15.1. Recall that $0 \le r_q(y) \le r_q(0)$ for all $y \in \mathbb{R}$, and then it follows that

$$0 \le q\int_0^t r_q(-X_s + x)ds \le q r_q(0)t.$$

Hence, by Lemma 15.5(iv), we have

$$q\int_0^t r_q(-X_s + x)ds \to 0 \quad \text{as } q \downarrow 0.$$

It remains to show that $M_t^x := -\lim_{q\downarrow 0} M_t^{q,x}$ is a martingale. Thus, it is enough to prove that

$$\mathbb{E}_0|M_t^x + M_t^{q,x}| \to 0 \quad \text{as } q \downarrow 0,$$

because then $M_t^{q,x}$ is a uniformly integrable martingale. We know that

$$\begin{aligned}|M_t^x + M_t^{q,x}| \le |h(X_t - x) - h_q(X_t - x)| + |h(X_0 - x) - h_q(X_0 - x)| \\ + q\int_0^t r_q(-X_s + x)ds.\end{aligned}$$

By Theorem 15.1, the second term on the above right-hand side converges to 0 as $q \downarrow 0$. By Lemma 15.5(iv), the third term convergences to 0 as $q \downarrow 0$.

It remains to prove the convergence of the first term as $q \downarrow 0$. By Lemma 15.2(ii), we have

$$\begin{aligned}h_q(x) &\le h_q(x) + h_q(-x) \\ &= \frac{2}{\pi}\int_0^\infty \Re\left(\frac{1-\cos(ux)}{q-\eta(u)}\right)du \\ &\le \frac{2}{\pi}\int_0^\infty \frac{1-\cos(ux)}{|\eta(u)|}du.\end{aligned}$$

By Fubini's theorem and by Lemma 15.5(ii) and (iii), we have

$$\begin{aligned}
&\mathbb{E}_0\left[\int_0^\infty \frac{1-\cos(u(X_t-x))}{|\eta(u)|}du\right] \\
&= \int_0^\infty \frac{1-\Re\exp\{t\eta(u)-iux\}}{|\eta(u)|}du \\
&= \int_0^\infty \frac{1-\cos(t\omega(u)-ux)\exp\{t\theta(u)\}}{|\eta(u)|}du \\
&\le \int_0^1 \frac{1-\cos(t\omega(u)-ux)-t\theta(u)}{|\eta(u)|}du + \int_1^\infty \left|\frac{2}{\eta(u)}\right|du \\
&\le \int_0^1 \frac{(t\omega(u)-ux)^2}{|\eta(u)|}du + \int_1^\infty \left|\frac{2}{\eta(u)}\right|du + t \\
&\le 2\int_0^1 \frac{(t\omega(u))^2+(ux)^2}{|\eta(u)|}du + \int_1^\infty \left|\frac{2}{\eta(u)}\right|du + t < \infty.
\end{aligned}$$

Hence, it follows from the dominated convergence theorem that

$$\mathbb{E}_0|M_t^x + M_t^{q,x}| \to 0 \quad \text{as } q \downarrow 0.$$

The proof is now complete. □

Remark 15.5 From Theorem 15.2, we obtain the invariant excessive function with respect to the killed process. Indeed, when we denote the law of the process starting at x killed upon hitting zero and the corresponding expectation by $\mathbb{P}_x^0$ and $\mathbb{E}_x^0$ respectively, under the conditions **(A)** and **(B)** we have

$$\mathbb{E}_x^0[h(X_t)] = h(x), \quad t \ge 0, x \in \mathbb{R},$$

because $\mathbb{E}_x^0[L_t^0] = 0$.

15.5 Examples

We shall introduce examples satisfying the conditions **(A)** and **(B)**. Because the condition **(A)** is a sufficient condition to have local times and explicit resolvent densities, we give examples with a focus on satisfying the condition **(B)**.

Example 15.1 (Strictly Stable Processes) Let X be a strictly stable process of index $\alpha \in (1, 2)$ with the Lévy measure ν on $\mathbb{R} \setminus \{0\}$ given by

$$\nu(dy) = \begin{cases} c_+|y|^{-\alpha-1}dy & \text{on } (0,\infty), \\ c_-|y|^{-\alpha-1}dy & \text{on } (-\infty,0), \end{cases}$$

where $\alpha \in (1,2)$, and c_+ and c_- are non-negative constants such that $c_+ + c_- > 0$, and with the drift parameter b given by

$$b = -\int_{|y|>1} y\nu(dy).$$

The Lévy symbol η of X is represented as

$$\eta(u) = -d|u|^\alpha \left(1 - i\beta \mathrm{sgn}(u) \tan \frac{\pi\alpha}{2}\right),$$

where $d > 0$ and $\beta \in [-1,1]$ are given by

$$d = \frac{c_+ + c_-}{2c(\alpha)}, \quad \beta = \frac{c_+ - c_-}{c_+ + c_-}$$

with

$$c(\alpha) = \frac{1}{\pi}\Gamma(\alpha+1)\sin\frac{\pi\alpha}{2}.$$

See Sato [11] on details.

By $\alpha \in (1,2)$, we have for $q > 0$

$$\left|\frac{1}{q-\eta(u)}\right| \le \frac{1}{q-\theta(u)} = \frac{1}{q+d|u|^\alpha} \in L^1(\mathbb{R}).$$

Hence, the condition **(A)** holds.

By $-\alpha+1 \in (-1,0)$, we have

$$\int_0^1 \left|\Im\left(\frac{u}{\eta(u)}\right)\right| du \le \int_0^1 \frac{u}{|\theta(u)|} du = \int_0^1 \frac{1}{d}|u|^{-\alpha+1} du < \infty.$$

Hence, the condition **(B)** holds.

In this case, it can be represented by

$$h(x) = c(-\alpha)\frac{1-\beta \mathrm{sgn}(x)}{d\left(1+\beta^2\tan^2(\pi\alpha/2)\right)}|x|^{\alpha-1}.$$

The result is consistent with [12].

Remark 15.6 This process also satisfies Yano's conditions **(L1)–(L3)**.

Remark 15.7 In [12], by using the Fourier transform, we could find the fundamental solution F of the infinitesimal generator for a strictly stable process $S = (S_t)_{t\geq 0}$ with index $\alpha \in (1, 2)$. Moreover, we have $F = h$. By using Itô's stochastic calculus, we have the martingale part M_t^x of the Tanaka formula can be represented as the explicit form:

$$M_t^x := \int_0^t \int_{\mathbb{R}\setminus\{0\}} \{F(S_{s-} - x + h) - F(S_{s-} - x)\}\tilde{N}(ds, dh).$$

Thus, we could study the property of local times from the Tanaka formula. On the other hand, for general Lévy processes, even if the renormalized zero resolvent and the local time exist, we could not use Itô's stochastic calculus, because we do not know the explicit form of the renormalized zero resolvent.

Example 15.2 (Truncated Stable Processes) A truncated stable process is a Lévy process with the Lévy measure ν on $\mathbb{R} \setminus \{0\}$ given by

$$\nu(dy) = \begin{cases} c_+|y|^{-\alpha-1}1_{\{y\leq 1\}}dy & \text{on } (0, \infty), \\ c_-|y|^{-\alpha-1}1_{\{y\geq -1\}}dy & \text{on } (-\infty, 0), \end{cases}$$

where $\alpha \in (1, 2)$, and c_+ and c_- are non-negative constants such that $c_+ + c_- > 0$, and without a drift parameter b.

Using the inequality: $1 - \cos(x) \geq x^2/4$ for $|x| \leq 1$, we have for all $u \geq 1$,

$$\begin{aligned} |\theta(u)| &\geq \frac{1}{4} \int_{|y|\leq u^{-1}} (uy)^2 \nu(dy) \\ &= \frac{c_+ + c_-}{4} \int_0^{u^{-1}} u^2 y^{-\alpha+1} dy = \frac{c_+ + c_-}{4(2-\alpha)} u^\alpha. \end{aligned}$$

Hence, by $\alpha \in (1, 2)$, the condition **(A)** holds.

By the argument in the proof of Proposition 15.3, we have

$$\int_0^1 \left| \Im\left(\frac{u}{\eta(u)}\right) \right| du = \int_0^1 \frac{|u\omega(u)|}{|\eta(u)|^2} du \leq \frac{1}{c_1^2} \int_0^1 \left| \frac{\omega(u)}{u^3} \right| du$$

where the positive constant c_1 is given by

$$c_1 = \frac{1}{4} \int_{|y|\leq 1} y^2 \nu(dy).$$

Using the inequality: $|\sin(x) - x| \le |x|^3$ for all $x \in \mathbb{R}$, we have

$$\left|\frac{\omega(u)}{u^3}\right| = \left|\int_{|y|\le 1} \frac{\sin(uy) - uy}{u^3}\nu(dy)\right|$$
$$\le \int_{|y|\le 1} \left|\frac{\sin(uy) - uy}{u^3}\right| \nu(dy)$$
$$\le \int_{|y|\le 1} |y|^3 \nu(dy) < \infty.$$

Hence, the condition **(B)** holds.

Remark 15.8 Using the inequalities: $1 - \cos(x) \le x^2$ and $|x - \sin(x)| \le |x|^3$ for $x \in \mathbb{R}$, we have for $u \in \mathbb{R}$

$$|\theta(u)| \le c_2 u^2, \quad |\omega(u)| \le c_3 |u|^3,$$

where the positive constants c_2 and c_3 are given by

$$c_2 = \int_{|y|\le 1} y^2 \nu(dy), \quad c_3 = \int_{|y|\le 1} |y|^3 \nu(dy).$$

We have for $u \in \mathbb{R}$

$$\theta'(u) = -\int_{|y|\le 1} y \sin(uy)\nu(dy),$$
$$\omega'(u) = \int_{|y|\le 1} y(\cos(uy) - 1)\nu(dy).$$

Using the inequalities: $1 - \cos(x) \ge x^2/2$ and $|\sin(x)| \ge |x|/2$ for all $|x| \le 1$, we have for $0 \le u \le 1$

$$|\theta'(u)| \ge c_4 u, \quad |\omega'(u)| \ge c_5 u^2,$$

where the positive constant c_4 and the non-negative constant are given by

$$c_4 = \frac{1}{2}\int_{|y|\le 1} y^2 \nu(dy), \quad c_5 = \frac{1}{4}\left|\int_{|y|\le 1} y^3 \nu(dy)\right|.$$

Thus, we have for $0 < u \le 1$

$$\frac{u^2(|\theta'(u)| + |\omega'(u)|)}{\theta(u)^2 + \omega(u)^2} \ge \frac{c_4 u^3 + c_5 u^4}{c_2^2 u^4 + c_3^2 u^6} \ge \frac{c_4}{(c_2^2 + c_3^2)u} > 0$$

Hence, this process does not satisfy Yano's condition **(L3)**.

Remark 15.9 If a Lévy process with a Lévy measure having a bounded support and with the drift parameter b given by

$$b = -\int_{|y|>1} y\nu(dy),$$

then the condition **(B)** holds by the same argument as stated in Example 15.2, but Yano's condition **(L3)** does not hold by the same argument as stated in Remark 15.8.

Example 15.3 (Tempered Stable Processes) A tempered stable process is a Lévy process with the Lévy measure ν on $\mathbb{R} \setminus \{0\}$ given by

$$\nu(dy) = \begin{cases} c_+|y|^{-\alpha_+-1}e^{-\lambda_+|y|}dy & \text{on } (0, \infty), \\ c_-|y|^{-\alpha_--1}e^{-\lambda_-|y|}dy & \text{on } (-\infty, 0), \end{cases}$$

where $\alpha_+, \alpha_- \in (1, 2)$, and c_+, c_-, λ_+ and λ_- are non-negative constants such that $c_+ + c_- > 0$, and with the drift parameter b given by

$$b = -\int_{|y|>1} y\nu(dy).$$

The processes have studied as models for stock price behavior in finance. See Carr et al. [4] on details.

We have for all $u \geq 1$,

$$\begin{aligned} |\theta(u)| &\geq \frac{1}{4}\int_{|y|\leq u^{-1}} (uy)^2\nu(dy) \\ &\geq \frac{u^2}{4}\left(c_+e^{-\lambda_+}\int_0^{u^{-1}} y^{-\alpha_++1}dy + c_-e^{-\lambda_-}\int_0^{u^{-1}} y^{-\alpha_-+1}dy\right) \\ &= \frac{c_+e^{-\lambda_+}}{4(2-\alpha_+)}u^{\alpha_+} + \frac{c_-e^{-\lambda_-}}{4(2-\alpha_-)}u^{\alpha_-}. \end{aligned}$$

Hence, by $\alpha_+, \alpha_- \in (1, 2)$, the condition **(A)** holds.

In the case of $[\lambda_+, \lambda_- > 0]$, $[c_+, \lambda_+ = 0 \;\&\; \lambda_- > 0]$ or $[c_-, \lambda_- = 0 \;\&\; \lambda_+ > 0]$, we have for all $0 < u \leq 1$,

$$\left|\frac{\omega(u)}{u^3}\right| \leq \int_{\mathbb{R}\setminus\{0\}} \left|\frac{\sin(uy) - uy}{u^3}\right| \nu(dy) \leq \int_{\mathbb{R}\setminus\{0\}} |y|^3\nu(dy) < \infty.$$

Hence, the condition **(B)** holds.

In the case of $[c_+ > 0 \;\&\; \lambda_+ = 0]$ or $[c_- > 0 \;\&\; \lambda_- = 0]$, we have for all $0 \le u \le 1$,

$$|\theta(u)| \ge \left(\int_0^\infty (1-\cos(uy))\,\nu(dy)\right) \vee \left(\int_{-\infty}^0 (1-\cos(uy))\,\nu(dy)\right) = \frac{c_+ \vee c_-}{2c(\alpha)}|u|^{\alpha_+ \vee \alpha_-}$$

by the Lévy symbol of Example 15.1. We then have for $0 < u \le 1$,

$$\left|\Im\left(\frac{u}{\eta(u)}\right)\right| = \left|\frac{u\omega(u)}{\theta(u)^2 + \omega(u)^2}\right| \le \frac{u}{2|\theta(u)|} \le \frac{c(\alpha)}{c_+ \vee c_-} u^{-(\alpha_+ \vee \alpha_-)+1},$$

Hence, by $-(\alpha_+ \vee \alpha_-) + 1 \in (-1, 0)$, the condition **(B)** holds.

Remark 15.10 In the case of $[\lambda_+, \lambda_- > 0]$, $[c_+, \lambda_+ = 0 \;\&\; \lambda_- > 0]$ or $[c_-, \lambda_- = 0 \;\&\; \lambda_+ > 0]$, we have

$$\int_{|y|>1} |y|^3 \nu(dy) < \infty.$$

Hence, by the similar argument as stated in Remark 15.8, Yano's condition **(L3)** does not hold.

But, in the case of $[c_+ > 0 \;\&\; \lambda_+ = 0]$ or $[c_- > 0 \;\&\; \lambda_- = 0]$, the condition **(L3)** holds.

Example 15.4 (Integrable Processes That Are Not Martingales) Suppose that the condition **(A)** holds, and that a Lévy measure ν satisfies

$$\int_{|y|>1} |y|\nu(dy) < \infty,$$

and that a drift parameter b satisfies

$$b \ne -\int_{|y|>1} y\nu(dy).$$

Using the inequality: $|\sin(x) - x1_{|x|\le 1}| \le |x|^3 \wedge |x|$ for all $x \in \mathbb{R}$, we have

$$\left|\Im\left(\frac{u}{\eta(u)}\right)^{-1}\right| \ge \left|\frac{\omega(u)}{u}\right| = \left|b + \int_{|y|\le 1} \frac{\sin(uy) - uy}{u}\nu(dy) + \int_{|y|>1} \frac{\sin(uy)}{u}\nu(dy)\right| \to \left|b + \int_{|y|>1} y\nu(dy)\right| > 0,$$

as $u \downarrow 0$. By the dominated convergence theorem, the condition **(B)** follows.

Example 15.5 (Spectrally Negative or Positive Processes) A Lévy process with no positive (negative) jumps is called a spectrally negative (positive) process. The processes have studied as models for insurance risk and dam theory.

Suppose that the condition **(A)** holds, and that a Lévy measure ν has a support in $(-\infty, 0)$ and satisfies

$$\int_{|y|>1} |y|\nu(dy) < \infty.$$

In the case of a drift parameter b such that

$$b \neq -\int_{|y|>1} y\nu(dy),$$

these processes are in Example 15.4.

We consider the case of the drift parameter b given by

$$b = -\int_{|y|>1} y\nu(dy).$$

We have for all $x \in \mathbb{R}$,

$$\begin{aligned} 0 \leq h_q(x) &= \frac{1}{\pi}\int_0^\infty \Re\left(\frac{1-\cos(ux)}{q-\eta(u)}\right)du + \frac{1}{\pi}\int_0^\infty \Im\left(\frac{\sin(ux)}{q-\eta(u)}\right)du. \\ &\leq h_q(x) + h_q(-x) = \frac{2}{\pi}\int_0^\infty \Re\left(\frac{1-\cos(ux)}{q-\eta(u)}\right)du. \end{aligned}$$

Hence, by Lemma 15.5(ii) and (iii), we have

$$\begin{aligned} &\left|\int_0^1 \Im\left(\frac{\sin(u)}{q-\eta(u)}\right)du\right| \\ &\quad\leq \left|\int_0^\infty \Re\left(\frac{1-\cos(u)}{q-\eta(u)}\right)du\right| + \left|\int_1^\infty \Im\left(\frac{\sin(u)}{q-\eta(u)}\right)du\right| \\ &\quad\leq \int_0^\infty \frac{|u|^2\wedge 1}{|\eta(u)|}du + \int_1^\infty \left|\frac{1}{\eta(u)}\right|du < \infty. \end{aligned}$$

Since we have $\omega(u) \geq 0$ for all $u \geq 0$, we have the following function:

$$\Im\left(\frac{\sin(u)}{q-\eta(u)}\right)\left(= \frac{\omega(u)\sin(u)}{(q-\theta(u))^2+\omega(u)^2}\right) \quad \text{on } (0,1]$$

is increasing as $q \downarrow 0$. Hence, by the monotone convergence theorem, the condition **(B)** follows.

Suppose that the condition **(A)** holds, and that a Lévy measure ν has a support in $(0, \infty)$ and satisfies

$$\int_{|y|>1} |y|\nu(dy) < \infty.$$

In this case, the condition **(B)** holds by the same argument as the spectrally negative case.

Acknowledgements I would like to thank Professor Atsushi Takeuchi of Osaka City University and Professor Kouji Yano of Kyoto University for their valuable advice.

The author was partially supported by JSPS-MAEDI Sakura program.

References

1. J. Bertoin, *Lévy Processes* (Cambridge University Press, Cambridge, 1996)
2. R.M. Blumenthal, R.K. Getoor, Local times for Markov processes. Probab. Theory Relat. Fields **3**(1), 50–74 (1964)
3. J. Bretagnolle, Résultats de Kesten sur les processus à accroissements indépendants, in *Séminaire de Probabilités V*. Lecture Notes in Mathematics, vol. 191 (Springer, Berlin, 1971), pp. 21–36
4. P. Carr, H. Geman, D. Madan, M. Yor, The fine structure of assert returns: an empirical investigation. J. Bus. **75**, 305–332 (2002)
5. K.L. Chung, R.J. Williams, *Introduction to Stochastic Integration*, 2nd edn. (Birkhäuser, Boston, 1990)
6. D. Geman, J. Horowitz, Occupation densities. Ann. Probab. **8**, 1–67 (1980)
7. R.K. Getoor, H. Kesten, Continuity of local times of Markov processes. Compos. Math. **24**, 277–303 (1972)
8. T. Jeulin, Ray–Knight's theorem on Brownian local times and Tanaka's formula, in *Seminar on Stochastic Processes, 1983* (Birkhäuser, Boston, 1984), pp. 131–142
9. H. Kesten, Hitting probabilities of single points for processes with stationary independent increments. Mem. Am. Math. Soc. **93**, 129 pp. (1969)
10. P. Salminen, M. Yor, Tanaka formula for symmetric Lévy processes, in *Séminaire de Probabilités XL*. Lecture Notes in Mathematics, vol. 1899 (Springer, Berlin, 2007), pp. 265–285
11. K. Sato, *Lévy Processes and Infinitely Divisible Distributions* (Cambridge University Press, Cambridge, 1999)
12. H. Tsukada, Tanaka formula for strictly stable processes, to appear in Probability and Mathematical Statistics
13. K. Yamada, Fractional derivatives of local times of α-stable Levy processes as the limits of occupation time problems, in *Limit Theorems in Probability and Statistics II*. (János Bolyai Mathematical Society, Budapest, 2002), pp. 553–573
14. K. Yano, Excursions away from a regular point for one-dimensional symmetric Lévy processes without Gaussian part. Potential Anal. **32**(4), 305–341 (2010)
15. K. Yano, On harmonic function for the killed process upon hitting zero of asymmetric Lévy processes. J. Math. Ind. **5**(A), 17–24 (2013)

LECTURE NOTES IN MATHEMATICS

Springer

Editorial Policy

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications – quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome.

   Manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They may be based on specialised lecture courses. Furthermore, the manuscripts should provide sufficient motivation, examples and applications. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this "lecture notes" character. For similar reasons it is unusual for doctoral theses to be accepted for the Lecture Notes series, though habilitation theses may be appropriate.

2. Besides monographs, multi-author manuscripts resulting from SUMMER SCHOOLS or similar INTENSIVE COURSES are welcome, provided their objective was held to present an active mathematical topic to an audience at the beginning or intermediate graduate level (a list of participants should be provided).

   The resulting manuscript should not be just a collection of course notes, but should require advance planning and coordination among the main lecturers. The subject matter should dictate the structure of the book. This structure should be motivated and explained in a scientific introduction, and the notation, references, index and formulation of results should be, if possible, unified by the editors. Each contribution should have an abstract and an introduction referring to the other contributions. In other words, more preparatory work must go into a multi-authored volume than simply assembling a disparate collection of papers, communicated at the event.

3. Manuscripts should be submitted either online at www.editorialmanager.com/lnm to Springer's mathematics editorial in Heidelberg, or electronically to one of the series editors. Authors should be aware that incomplete or insufficiently close-to-final manuscripts almost always result in longer refereeing times and nevertheless unclear referees' recommendations, making further refereeing of a final draft necessary. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter, a bibliography and several sample chapters. Parallel submission of a manuscript to another publisher while under consideration for LNM is not acceptable and can lead to rejection.

4. In general, **monographs** will be sent out to at least 2 external referees for evaluation.

   A final decision to publish can be made only on the basis of the complete manuscript, however a refereeing process leading to a preliminary decision can be based on a pre-final or incomplete manuscript.

   Volume Editors of **multi-author works** are expected to arrange for the refereeing, to the usual scientific standards, of the individual contributions. If the resulting reports can be forwarded to the LNM Editorial Board, this is very helpful. If no reports are forwarded or if other questions remain unclear in respect of homogeneity etc, the series editors may wish to consult external referees for an overall evaluation of the volume.

5. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include

   – a table of contents;
   – an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   – a subject index: as a rule this is genuinely helpful for the reader.
   – For evaluation purposes, manuscripts should be submitted as pdf files.

6. Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online. After acceptance of the manuscript authors will be asked to prepare the final LaTeX source files (see LaTeX templates online: https://www.springer.com/gb/authors-editors/book-authors-editors/manuscriptpreparation/5636) plus the corresponding pdf- or zipped ps-file. The LaTeX source files are essential for producing the full-text online version of the book, see http://link.springer.com/bookseries/304 for the existing online volumes of LNM). The technical production of a Lecture Notes volume takes approximately 12 weeks. Additional instructions, if necessary, are available on request from lnm@springer.com.

7. Authors receive a total of 30 free copies of their volume and free access to their book on SpringerLink, but no royalties. They are entitled to a discount of 33.3 % on the price of Springer books purchased for their personal use, if ordering directly from Springer.

8. Commitment to publish is made by a *Publishing Agreement*; contributing authors of multiauthor books are requested to sign a *Consent to Publish form*. Springer-Verlag registers the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: a brief written (or e-mail) request for formal permission is sufficient.

Addresses:
Professor Jean-Michel Morel, CMLA, École Normale Supérieure de Cachan, France
E-mail: moreljeanmichel@gmail.com

Professor Bernard Teissier, Equipe Géométrie et Dynamique,
Institut de Mathématiques de Jussieu – Paris Rive Gauche, Paris, France
E-mail: bernard.teissier@imj-prg.fr

Springer: Ute McCrory, Mathematics, Heidelberg, Germany,
E-mail: lnm@springer.com