Multi-mean reverting processes : analytical and statistical approaches

Benoît Nieto

Institut Camille Jordan
Laboratoire de recherche en mathématiques Lyon/Saint-Étienne

Multi-Mean Reverting Processes: Analytical and Statistical Approaches

Benoît Nieto

Thèse de doctorat

UNIVERSITÉ DE LYON
ÉCOLE CENTRALE LYON

ÉCOLE CENTRALE LYON

Numéro d'ordre NNT : 2024ECDL0029

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON opérée au sein de l'École centrale de Lyon

École Doctorale 512
École Doctorale InfoMaths

Spécialité de doctorat : Mathématiques et applications
Discipline : Mathématiques

Soutenue publiquement le 23/09/2024 par

Benoît Nieto

Multi-Mean Reverting Processes: Analytical and Statistical Approaches

Devant le jury composé de :

M. Stefan Ankirchner	Professeur des universités, Jena University, Allemagne	Examinateur
M. Ahmed Kebaier	Professeur des universités, Université d'Évry	Rapporteur
Mme Adeline Leclerc Samson	Professeur des universités, Université Grenoble Alpes	Examinateur
M. Antoine Lejay	Directeur de recherche, Inria Nancy Grand-Est	Rapporteur
M. Paolo Pigato	Associate Professor, University of Rome Tor Vergata, Italie	Examinateur
M. Anthony Réveillac	Professeur des universités, Université de Toulouse III-INSA	Président
Mme Christophette Blanchet-Scalliet	Maître de conférences, École centrale de Lyon	Directrice
Mme Diana Dorobantu	Maître de conférences, ISFA Université de Lyon	Co-directrice

Remerciements

Je souhaite exprimer ma plus sincère reconnaissance à mes encadrantes, Christophette et Diana, pour leur accompagnement tout au long de ce projet de thèse. Leur expertise pointue et leurs conseils judicieux ont joué un rôle important dans l'orientation et l'avancement de mes recherches. Leur grande patience face aux défis rencontrés, ainsi que leur disponibilité, ont été pour moi un soutien précieux.

Au-delà de l'encadrement scientifique, je tiens à souligner la qualité des échanges que j'ai pu avoir avec elles, qui ont grandement enrichi mon parcours personnel et professionnel. Leurs retours constructifs, leur rigueur et leur bienveillance ont été des atouts majeurs tout au long de ce travail. Je leur suis reconnaissant de m'avoir guidé avec générosité en m'accordant leur confiance pour mener à bien cette thèse.

Je souhaite également exprimer ma gratitude à Sara Mazzonetto, pour la confiance qu'elle m'a accordée. Son expertise, sa disponibilité m'ont permis d'enrichir mes connaissances et de progresser dans mes travaux. Je la remercie pour nos échanges précieux, ainsi que pour son soutien constant. Je tiens aussi à la remercier chaleureusement pour ses accueils à Nancy, qui ont grandement facilité mon intégration lors de mes séjours. Je souhaite étendre ces remerciements à l'ensemble des personnes de Nancy qui m'ont réservé un accueil chaleureux et ont contribué à rendre ces moments particulièrement agréables et productifs.

J'exprime mes sincères remerciements à Antoine Lejay, rapporteur de cette thèse, pour les discussions enrichissantes que nous avons pu avoir. Je remercie Ahmed Kebaier, également rapporteur de cette thèse, pour le temps qu'il a consacré à l'évaluation de mes recherches.

Je remercie chaleureusement l'ensemble des membres du jury pour leur présence et leur participation à l'évaluation de cette thèse: Adeline Leclerc Samson, Paolo Pigato, Anthony Réveillac et Stefan Ankirchner. Je tiens à exprimer une reconnaissance particulière à Stefan Ankirchner pour l'accueil qu'il m'a réservé durant une semaine à Jena. Ses discussions enrichissantes ont été d'une grande aide pour approfondir certains aspects de mon travail. Je remercie également Paolo Pigato pour nos échanges qui ont contribué à éclairer et orienter certaines de mes réflexions.

J'exprime ma gratitude à l'ensemble des membres du laboratoire, permanents ou doctorants, pour leur soutien et nos échanges durant ces années. Je remercie le personnel administratif de Centrale Lyon pour leur disponibilité et leur efficacité. Leur aide a permis de faciliter de nombreux aspects pratiques de mon travail.

Je souhaite exprimer ma gratitude envers tous les enseignants qui ont marqué mon parcours, notamment durant mon master. Chaleureux remerciements à Cécile Mercadier et Clément Marteau pour leur engagement et leur dévouement dans la direction du master, contribuant ainsi largement à sa réussite. Un remerciement à Anne-Laure Fougères et

Thomas Lepoutre pour leur bienveillance et leur gentillesse lors de nos rencontres à l'ICJ.

J'exprime toute ma gratitude à l'ensemble des membres de ma famille. Un merci tout particulier à mes parents, dont le soutien indéfectible, l'encouragement constant et la confiance m'ont permis de mener à bien cette thèse. Un grand merci à mes frères et tous mes proches pour leur présence, leur patience et leur compréhension durant ce parcours.

Merci aux amis de longue date: Alex, Dorian, Germain, Hugo, Loïc, Nicolleti et Thed. Merci à ceux qui punissent les bas titubants avec des coups de pied sautés retardés: Edgar, Maxime, Nathan, Philippe. Puis merci aux ultras: Maxime, Régis, Loris, Manu (l'artiste aux multiples talents) pour ce qu'ils mettent.

Contents

List of Figures

List of Tables

Abstract

This thesis explores the theory and application of stochastic differential equations (SDEs), with a particular focus on parameter estimation and the behavior of processes with discontinuous coefficients.

The first part focuses on the Ornstein-Uhlenbeck (OU) process. We introduce an estimator for the OU parameters based on supremum observations of one trajectory. We derive an analytic expression for the supremum density and build an estimator using a pseudo-likelihood method. The statistical properties of this estimator, including consistency and asymptotic normality, are established using weak-dependence results. The effectiveness of our estimator is demonstrated through its application to both simulated and real data. Additionally, we explore the behavior of Parabolic Cylinder functions, which are involved in the law of the OU supremum. Specifically, we investigate the μ-zeros of the function $\mu \mapsto D_\mu(z)$ with respect to the real variable z, establishing a formula for the derivative of a zero and providing an asymptotic expansion for large positive z.

The second part investigates processes governed by SDEs with discontinuous coefficients. Introducing the threshold Ornstein-Uhlenbeck (T-OU) process, we provide explicit expressions for transition probability density and first hitting time density for the killed process. Then, we introduce the threshold CKLS (T-CKLS) process and focus on estimating its parameters using observations of a single trajectory. We study the asymptotic behavior of maximum likelihood and quasi-maximum likelihood estimators for drift parameters, as well as a volatility estimator. Some statistical properties under continuous and high-frequency observations over long time horizons are derived. Finally, we highlight the relevance of multiple thresholds through applications to simulated and real data.

Keywords: OU, CIR, CKLS, supremum law, parameters estimation, pseudo-likelihood, asymptotic normality, hitting time, self-exciting process , maximum likelihood, thresholds diffusion.

Resumé

Cette thèse traite de la théorie et des applications des équations différentielles stochastiques (EDS), en se concentrant particulièrement sur l'estimation des paramètres et le comportement des processus aux coefficients discontinus.

La première partie introduit un estimateur pour les paramètres du processus d'Ornstein-Uhlenbeck (OU), construit à partir d'observations du supremum d'une unique trajectoire. Une fois l'expression analytique pour la densité du supremum établie, nous procédons à la construction d'un estimateur en utilisant une méthode de pseudo-vraisemblance. Les propriétés statistiques de cet estimateur, à savoir la consistance et la normalité asymptotique, sont établies en utilisant des propriétés de faible dépendance de l'échantillon d'observations. L'efficacité de notre estimateur est démontrée à travers son application à des données simulées et réelles. De plus, nous étudions le comportement des fonctions Paraboliques Cylindrique, qui sont impliquées dans la loi du supremum de l'OU. Plus précisément, nous étudions les μ-zéros de la fonction $\mu \mapsto D_\mu(z)$ par rapport à la variable réelle z. Nous établissons une formule pour la dérivée d'un zéro et fournissons un développement asymptotique pour de grands z positifs.

La deuxième partie développe la théorie des processus solutions des EDS à coefficients discontinus. Après avoir introduit le processus d'Ornstein-Uhlenbeck à seuil (T-OU), nous établissons des expressions analytiques pour la densité de probabilité de transition et la densité de premier temps d'atteinte pour le processus tué. Ensuite, le processus CKLS avec seuil (T-CKLS) est introduit et nous nous concentrons sur l'estimation de ses paramètres de dérive et de volatilité en utilisant des observations d'une seule trajectoire. L'analyse du comportement asymptotique des estimateurs de maximum de vraisemblance et de quasi-maximum de vraisemblance pour les paramètres de dérive, ainsi qu'un estimateur de volatilité, est effectuée. Les propriétés statistiques sont obtenues à partir d'observations continues et à haute fréquence en temps long. Enfin, la pertinence d'une modélisation à plusieurs seuils est mise en évidence à travers des applications à des données simulées et réelles.

Mots clés: OU, CIR, CKLS, loi du supremum, estimation des paramètres, pseudo-vraisemblance, normalité asymptotique, premier temps d'atteinte, processus auto-excitant, maximum de vraisemblance, diffusion à seuils.

Introduction

In the field of stochastic processes, contributions have established the groundwork for understanding a wide array of phenomena. At the center of this framework lies the concept of Brownian motion, first observed by R. Brown in 1828 [23]. Brownian motion's random behavior paved the way for the development of stochastic differential equations (SDEs), offering a powerful tool for modeling systems subject to random fluctuations over time. One notable extension of SDEs is the Ornstein-Uhlenbeck process, introduced by L.S. Ornstein and G.E. Uhlenbeck in [104]. This process is known for its mean-reverting behavior and finds applications in various fields, including physics [110], finance [107], and biology [64]. In particular, it can be used to model the evolution of temperatures over time (see [37] and [4]). Moving on from our exploration of stochastic processes, the Cox-Ingersoll-Ross (CIR) process, introduced in [31], extends the understanding beyond the Ornstein-Uhlenbeck (OU) process. This process is used in fields such as finance for interest rate modeling and biology. Both of these processes are particular cases of the Chan–Karolyi–Longstaff–Sanders (CKLS) process, introduced in [27]. The CKLS process, which includes both the OU and CIR processes, is used in financial modeling, particularly in analyzing asset pricing dynamics and modeling interest rate fluctuations.

This thesis focuses on studying processes that are solutions to stochastic differential equations (SDEs). Specifically, the focus is on parameter estimation of these processes, as well as on explicit computation of the laws associated with these processes. We mainly study SDEs with discontinuous coefficients. This study is divided into two parts: one part on the parameter estimation of a classical OU process, while the second part explores processes that are solutions to SDEs with discontinuous coefficients. In the first part, we focus on parameter estimation for an OU process based on observations of its supremum. As illustrated by [37] (also referenced in [4]), temperatures typically exhibit a tendency to cluster around a mean value, suggesting a natural fit for modeling their dynamics using a mean-reverting process. In [15], the authors propose parameter estimation of an OU process to perform risk assessments for phenomena like heatwaves. Because meteorological datasets commonly contain only daily temperature extremes, their estimation approach relies on a dataset composed of supremum values of the OU process. The authors introduced a least squares method based on the integral representation of the cumulative distribution function of the supremum. This representation is provided in [6]. Note that they lack statistical properties on their estimators, and moreover, their method is computationally

demanding. In this work, we consider a stationary OU process X solution of:

$$\begin{cases} dX_t = (a - bX_t)dt + \sqrt{\beta}dB_t, \\ X_0 \sim \mathcal{N}\left(\frac{a}{b}, \frac{\beta}{2b}\right), \end{cases}$$

with parameter $\theta = (a, b, \beta) \in \mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$, and B a Brownian motion independent of X_0.

We introduce an alternative estimation method for the parameter θ of the Ornstein-Uhlenbeck (OU) process. Our approach uses a pseudo-likelihood framework tailored to the supremum distribution of the OU process. As highlighted in [16], this distribution can be expressed as a series of Parabolic Cylinder functions, denoted $D_\mu(z)$, solution of the following differential equation:

$$\begin{cases} y''(z) + \left(\mu + \frac{1}{2} - \frac{1}{4}z^2\right) y(z) = 0, \\ y(z) \underset{z \to +\infty}{\sim} z^\mu e^{-z^2/4}. \end{cases}$$

To establish statistical properties for our estimator, such as consistency and asymptotic normality, it is essential to bound the supremum distribution. To this end, we focus on the asymptotic behavior of the zeros of the Parabolic Cylinder function. Indeed, the μ-zeros of the Parabolic Cylinder function play a crucial role in determining the law of the OU process running supremum ([6], [16]). Thus, the second aspect of this study involves exploring the properties of the μ-zeros of the Parabolic Cylinder functions. These functions belong to the set of special functions such as Bessel functions and Hankel functions. Various studies have investigated the zeros of special functions of order μ and argument z. For instance, Olver explores the z-zeros of Parabolic Cylinder functions for large values of μ [87]. However, fewer studies have focused on the μ-zeros. In works such as [78] and [56], the authors investigate the behavior of the μ-zeros of the Hankel function of the first kind. We provide a formula for the μ-zeros for large values of z and establish a formula for the derivative of a μ-zero, deducing some monotonicity results. Since the z-zeros of Hermite functions are related to those of Parabolic Cylinder functions, our analysis relies on the findings presented in [41]. Asymptotic expressions for the μ-zeros are obtained from the expansion provided by Olver [87]. This analysis contributes to a deeper understanding of the behavior of these special functions, which are essential for bounding the distribution of the OU supremum. One contribution of this work is to propose an explicit expression of this density based on the Parabolic Cylinder function, which is computationally inexpensive. By using the asymptotic expressions for the μ-zeros of the Parabolic Cylinder function, we obtain consistency and asymptotic normality of our estimator. Furthermore, our estimators outperform those presented in [15] in terms of relative root mean squared error (RMSE) and mean error (ME). We also provide an application to a meteorological dataset.

The second part of this thesis deals with stochastic processes governed by SDEs with discontinuous volatility and drift functions.

The analysis of discontinuous coefficients for the drift and volatility of SDEs represents a novel approach in the domain of stochastic modeling. These models, often categorized as

threshold diffusion models, exhibit dynamics that change abruptly when certain thresholds are crossed. In traditional SDEs, the coefficients of drift and volatility are assumed to be continuous functions. However, in many scenarios, discontinuities in these coefficients can significantly affect the behavior of the system. Threshold diffusion models capture such discontinuities by incorporating specific threshold values that trigger changes in the dynamics of the process. The analysis of these discontinuous coefficients presents unique challenges compared to traditional continuous coefficient models. Techniques from stochastic calculus and numerical analysis are often employed to study the properties and behavior of such models. Additionally, applications of threshold diffusion models are found in various fields including finance [34], biology, and ecology [24], where abrupt changes in dynamics are observed. Overall, the investigation of discontinuous coefficients in SDEs offers valuable insights into complex systems characterized by sudden shifts in behavior, contributing to advancements in stochastic modeling and its applications.

In the first stage, we focus on deriving explicit probability distributions associated with a stochastic differential equation featuring discontinuous coefficients. There is considerable interest in computing explicit laws for processes that are solutions of this this type of SDEs. For instance, in [34], the authors compute the transition probability of threshold Ornstein-Uhlenbeck (T-OU) and threshold Cox-Ingersoll-Ross (T-CIR) processes. Additionally, [35] determines the law of a drifted Skew Brownian motion, while [108] computes the law of the Skew Ornstein-Uhlenbeck (OU) process.

The computation of distribution densities associated with killed processes over constant boundaries has received significant attention in the literature. For example, in the works of V. Linetsky [76] and [75], the density of first hitting times for prominent processes like the Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes has been computed. Few studies have been conducted on computing the law of threshold killed processes.
In this work, we consider the T-OU process X solution of:

$$X_t = x_0 + \int_0^t a(X_s) - b(X_s)X_s \,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad t \geq 0,$$

where a, b and σ are discontinuous at a finite number of thresholds. We explicitly compute the transition probability density of the process X killed at a constant boundary, and the first hitting time density of this process.

Moving from computing distribution laws to estimating parameters involves understanding how estimation, simulation, and modeling are closely related. This connection is important because accurate calibration of models is necessary to represent the underlying dynamics of the system. The parametric estimation of threshold processes has recently attracted attention in the literature. For instance, in [72], the authors introduce a maximum likelihood estimator for a threshold drifted Brownian motion. They obtain the consistency and asymptotic normality results for both continuous and discrete observations, covering both ergodic and non-ergodic regimes of the process. Similarly, in [81], the authors focus on parameter estimation for a T-OU process, while in [70], the authors introduce an estimation approach for the diffusion parameter of a threshold Brownian motion.

In this part, we focused on parameter estimation for threshold processes using observations from a single trajectory, considering both continuous and discrete observation frameworks. We consider a multi-threshold CKLS (T-CKLS) process solution of the SDE:

$$X_t = x_0 + \int_0^t (a(X_s) - b(X_s)X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)|X_s|^{\gamma(X_s)}\,\mathrm{d}B_s, \quad t \geq 0,$$

where a, b, σ are piecewise constant coefficients and γ possibly discontinuous. We focus on the case $\gamma \in [0, 1]$. In this study, we examine both continuous-time observations and non-equally spaced discrete observations. The methods introduced include maximum likelihood estimation (MLE) and quasi-maximum likelihood estimation (QMLE) for the drift coefficients. Additionally, we provide an estimator for the diffusion coefficients σ, based on quadratic variation. We investigate the asymptotic properties of these estimators, deriving both a law of large numbers and a central limit theorem for the estimators of the drift coefficients, as well as establishing consistency for the estimator of the volatility.

Structure of the thesis: This thesis is organized as follows:

Chapter 1: In this chapter, we present essential tools in stochastic calculus and differential equations, beginning with foundational concepts like local time of semi-martingales and the existence and uniqueness of SDEs. Then, we explore diffusion process properties, including killed processes, Fokker-Planck equation, and boundary classification. Finally, we examine the Ornstein-Uhlenbeck and CKLS processes, investigating their properties to understand their dynamics.

Part 1:

Chapter 2: We propose an estimator for the Ornstein-Uhlenbeck parameters based on observations of its supremum. We derive an analytic expression for the supremum density. Making use of the pseudo-likelihood method based on the supremum density, our estimator is constructed as the maximal argument of this function. Using weak-dependency results, we prove some statistical properties on the estimator such as consistency and asymptotic normality. Finally, we apply our estimator to simulated and real data.

Chapter 3: We study the μ-zeros of the function $\mu \mapsto D_\mu(z)$ with respect to the real variable z. We establish a formula for the derivative of a zero and deduce some monotonicity results. Then we also give an asymptotic expansion for μ-zeros for large positive z.

Part 2:

Chapter 4: In this chapter, we propose an approach to studying diffusion processes with piecewise constant coefficients, focusing on the threshold Ornstein-Uhlenbeck (T-OU) process. We provide explicit solutions for the probability transition density and first hitting time density for the killed process. An explicit expression for the Laplace transform of the first hitting time is proposed for numerical testing.

Chapter 5: We consider a continuous-time process which is self-exciting and ergodic, called threshold Chan–Karolyi–Longstaff–Sanders (CKLS) process. We allow for the presence of several thresholds which determine changes in the dynamics. We study the asymptotic behavior of the maximum and quasi-maximum likelihood estimators of the drift parameters in the case of continuous time and discrete time observations. We show that for high frequency observations and infinite horizon the estimators satisfy the same asymptotic normality property as in the case of continuous time observations. We discuss diffusion coefficient estimation as well. Finally, we apply our estimators to simulated and real data to motivate considering (multiple) thresholds.

Chapter 1

Diffusion Processes: An introduction

In this chapter, we introduce the main tools in stochastic calculus and differential equations that will be used in this thesis. We begin by establishing foundational notions, exploring topics such as local time of semi-martingales, and existence and uniqueness of stochastic differential equations (SDEs).

Additionally, we focus on properties of diffusion processes. This includes discussions on killed processes, Fokker-Planck equation, boundary classification and process regimes.

Then, we introduce some well-known processes as solutions of SDEs, focusing on the Ornstein-Uhlenbeck process and the CKLS process.

Notation: In this thesis, we denote $\mathbb{N}^\star$ as the set of all strictly positive integers. Furthermore, $\mathbb{R}^\star$ is defined as the set of all non-zero real numbers, and $\mathbb{R}_+^\star$ as the set of all strictly positive real numbers.

1.1 Local time of semi-martingale

In the first part of this chapter, we introduce the local time and Itô-Tanaka formula for continuous semi-martingales. These results can be found in [91]. Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

Definition 1.1.1 (Semi-martingales)**.** *A continuous $(\mathcal{F}_t, \mathbb{P})$-semi-martingale is a continuous process X which can be written $X = M + A$ where M is a continuous local martingale and A a continuous adapted process of finite variation.*

Additional information on semi-martingales and local martingales can be found in [91]. Let us recall the Itô formula, which provides a powerful tool for expressing the differential of a continuous semi-martingale.

Theorem 1.1.2 (Itô formula)**.** *Let f belongs to $\mathcal{C}^2(\mathbb{R})$ and let $X = M + A$ be a continuous semi-martingale. Then, the process $f(X)$ is a continuous semi-martingale and*

$$t \geq 0, \quad f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X, X\rangle_s \quad \mathbb{P} - a.s.$$

Remark 1.1.3. *Note that the assumption $f \in \mathcal{C}^2(\mathbb{R})$ can be relaxed. In fact, if f admits an absolutely continuous derivative, then f'' exists Lebesgue-almost everywhere, and we have:*

$$t \geq 0, \quad f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X, X\rangle_s \quad \mathbb{P} - a.s.,$$

and $f(X)$ is a continuous semi-martingale. We refer to this formula as the generalized Itô formula.

We now seek to introduce a generalization of the Itô formula. We respectively denote $(\ell_t^x(X))_{t\geq 0}$ as the right local time. The local time of a semi-martingale can be constructed in several different ways. We mainly focus on the construction based on Tanaka formula. In the next results, the positive and negative parts of $x \in \mathbb{R}$ are denoted $x^+ = x \vee 0$, $x^- = (-x) \vee 0$. We also denote f_l' and f_r' as the left and right derivatives of the function f.

Theorem 1.1.4. *(Tanaka formula) For $x \in \mathbb{R}$ and X a continuous semi-martingale, there exists an increasing continuous process $(\ell_t^x(X))_{t\geq 0}$ such that:*

$$\frac{1}{2}\ell_t^x(X) = (X_t - x)^+ - (X_0 - x)^+ - \int_0^t \mathbb{1}_{(X_s > x)}dX_s, \tag{1.1.1}$$

$$\frac{1}{2}\ell_t^x(X) = (X_t - x)^- - (X_0 - x)^- + \int_0^t \mathbb{1}_{(X_s \leq x)}dX_s.$$

In particular, $(X - x)^+$ and $(X - x)^-$ are semi-martingales.

Theorem 1.1.5. *(Ito-Tanaka formula, Theorem 1.5, Chapter VI in [91]) Let X be a continuous semi-martingale. If f is the difference of two convex functions, we have:*

$$f(X_t) = f(X_0) + \int_0^t f_l'(X_s)dX_s + \frac{1}{2}\int_{\mathbb{R}} \ell_t^x(X)f''(dx).$$

In particular, $f(X)$ is a semi-martingale.

Note that Itô-Tanaka formula extends Itô formula for functions that are differences of two convex functions.

Theorem 1.1.6 (Theorem 1.7, Chapter VI in [91])**.** *Let X a continuous semi-martingale, there exists a modification of the process $\{\ell_t^x : x \in \mathbb{R}, t \in \mathbb{R}_+\}$ such that $(x, t) \mapsto \ell_t^x$ is a.s. continuous in t and cadlag in x. Additionally, if $X = M + A$, we have:*

$$\ell_t^x(X) - \ell_t^{x^-}(X) = 2\int_0^t \mathbb{1}_{(X_s = x)}dX_s = 2\int_0^t \mathbb{1}_{(X_s = x)}dA_s, \tag{1.1.2}$$

where $\ell_t^{x^-}(X) = \lim_{h\to 0, h<0} \ell_t^{x+h}$.

Remark 1.1.7. *Note that, from* (1.1.2)*, we have:*

$$\frac{1}{2}\ell_t^{x^-}(X) = (X_t - x)^- - (X_0 - x)^- + \int_0^t \mathbb{1}_{(X_s<x)} dX_s. \tag{1.1.3}$$

In particular, we can prove that $\ell_t^x(-X) = \ell_t^{(-x)^-}(X)$*. Then there exists a modification of* $\{\ell_t^{x^-} : x \in \mathbb{R}, t \in \mathbb{R}_+\}$ *such that* $(x,t) \mapsto \ell_t^{x^-}$ *is continuous in time and caglad in* x*. Furthermore, if* X *is a local martingale, Theorem 1.1.6 implies that there exists a bicontinuous modification of the family* ℓ_t^x *of local times.*

Building upon the preceding theorem, we define an analogous Itô-Tanaka formula, which enables us to introduce the concept of symmetric local time.

Corollary 1.1.8. *(Symmetric Itô-Tanaka formula) If* f *is the difference of two convex functions and if* X *is a continuous semi-martingale,*

$$f(X_t) = f(X_0) + \int_0^t \frac{f_l'(X_s) + f_r'(X_s)}{2} dX_s + \frac{1}{2}\int_{\mathbb{R}} L_t^x(X) f''(dx),$$

where $L_t^x(X)$ *is the symmetric local time at* x*, and defined such that:*

$$L_t^x(X) = \frac{\ell_t^x(X) + \ell_t^{x^-}(X)}{2}.$$

Furthermore, $f(X)$ *is a semi-martingale.*

Proof. Summing (1.1.1) and (1.1.3), leads to:

$$L_t^x(X) = |X_t - x| - |X_0 - x| - \int_0^t \operatorname{sgn}(X_s - x) dX_s,$$

where $\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases}$

Then, applying a similar reasoning as Theorem 1.5, Chapter VI in [91], we conclude. □

Remark 1.1.9. *Note that, from Theorem 1.1.6 and Remark 1.1.7, there exists a modification* $\{L_t^x : x \in \mathbb{R}, t \in \mathbb{R}_+\}$ *such that* $(x,t) \mapsto L_t^x$ *is continuous in time and laglad in* x*. In the case where* X *is a continuous local martingale, there exists an a.s. continuous modification in* t *and* x *of* $\{L_t^x : x \in \mathbb{R}, t \in \mathbb{R}_+\}$*.*

Proposition 1.1.10. *The measure* $dL_t^x(X)$ *is a.s. carried by the set* $\{t : X_t = x\}$*.*

Proof. Applying Itô formula to $(|X_t - x|^2)_{t\geq 0}$, we have:

$$|X_t - x|^2 = |X_0 - x|^2 + 2\int_0^t |X_s - x|\, d(|X - x|)_s + \langle |X - x|, |X - x| \rangle_t. \tag{1.1.4}$$

Then, using Corollary 1.1.8 in (1.1.4), we obtain:

$$|X_t - x|^2 = |X_0 - x|^2 + 2\int_0^t |X_s - x|\mathrm{sgn}(X_s - x)\, dX_s + 2\int_0^t |X_s - x|\, dL_s^x(X) + \langle X, X\rangle_t. \tag{1.1.5}$$

Comparing equality (1.1.5) with the equality given by applying Itô-formula:

$$(X_t - x)^2 = (X_0 - x)^2 + 2\int_0^t (X_s - x) dX_s + \langle X, X\rangle_t,$$

we have $\int_0^t |X_s - x|\, dL_s^x(X) = 0$ a.s.. Then the conclusion holds. □

Comparing Itô-Tanaka formula and Itô formula, we get the following result.

Proposition 1.1.11. *(Occupation times formula, Corollary 1.6, Chapter VI in [91]) For all $t \geq 0$, and for every measurable positive function f, we have:*

$$\int_0^t f(X_s) d\langle X, X\rangle_s = \int_{-\infty}^{\infty} f(x) L_t^x(X) dx.$$

Remark 1.1.12. *We have defined the right and symmetric local time of a continuous semi-martingale X at x, using Itô-Tanaka formula. From the occupation times formula in Proposition 1.1.11, the symmetric and the right local time can be constructed as follows:*

$$L_t^x(X) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{(x-\varepsilon \leq X_s \leq x+\varepsilon)} d\langle X, X\rangle_s \quad \textit{and} \quad \ell_t^x(X) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{(x \leq X_s \leq x+\varepsilon)} d\langle X, X\rangle_s.$$

1.2 Stochastic Differential Equations

1.2.1 Definitions and uniqueness

We refer to stochastic differential equations (SDEs), the equations of the form:

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s, \quad t \geq 0, \tag{1.2.1}$$

where b is the drift function and σ is the volatility function. In this section, we remind the classical results on the existence and uniqueness of the solution of equation (1.2.1). Discussions on the extension of these results are covered later in this thesis. We first introduce the concepts of solutions, as well as the notions of existence and uniqueness for SDEs. These definitions are given in [54, 67, 91].

Definition 1.2.1. *A solution of SDE* (1.2.1) *is a pair (X, B) of adapted processes defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}), \mathbb{P})$ and such that:*

- *B is a standard $(\mathcal{F}_t)$- Brownian motion.*

- *X is continuous and, for all $t \geq 0$:*

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad \mathbb{P} - a.s.$$

Definition 1.2.2. *(Strong and weak solution) A solution (X, B) is said to be a strong solution if X is adapted to the filtration $\bar{\mathcal{F}}^B : (\bar{\mathcal{F}}_t^B)_{t\geq 0}$, i.e., the filtration of B completed with respect to $\mathbb{P}$. By contrast, a solution that is not strong is called a weak solution.*

Definition 1.2.3. *(Uniqueness in law) There is uniqueness in law for* (1.2.1) *if for every (X, B) and $(\tilde{X}, \tilde{B})$ solutions to* (1.2.1) *with $X_0 = \tilde{X}_0$ in law, then the laws of X and $\tilde{X}$ are equal.*

Remark 1.2.4. *Note that, uniqueness in law can hold even for solutions defined on two different filtered probability space.*

Definition 1.2.5. *There is pathwise uniqueness for* (1.2.1) *if for every (X, B) and $(\tilde{X}, B)$ solutions to* (1.2.1) *on the same filtered probability space with $X_0 = \tilde{X}_0$ a.s., then $X = \tilde{X}$ a.s..*

Strong solution to (1.2.1) can also be defined up to a hitting time τ. In the following definition, the hitting time τ_∞ is referred to as the explosion time.

Definition 1.2.6. *(Strong solution up to an explosion time) On the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, we say that (X, τ_∞, B) is a strong solution of* (1.2.1)*, if:*

1. *τ_∞ is a stopping time and $X_{[0,\tau_\infty)}$ is $\bar{\mathcal{F}}^B$-adapted.*

2. *$X_{[0,\tau_\infty)}$ is continuous and, on $\{t < \tau_\infty\}$, we have:*

$$X_t = X_0 + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dB_s, \quad \mathbb{P} - a.s.$$

3. *On the event $\{\tau_\infty < +\infty\}$, $\lim_{t\to\tau_\infty} |X_t| = +\infty$ $\mathbb{P}$-a.s..*

The following results state the link between the previous notions of uniqueness and existence.

Theorem 1.2.7. *(Yamada Watanabe [111]) Consider the equation* (1.2.1)*.*

1. *Pathwise uniqueness property implies uniqueness in law.*

2. *Weak existence and pathwise uniqueness imply strong existence.*

Let b and σ be two Lipschitz functions. In this case, strong existence and uniqueness results for (1.2.1) are provided by the following theorem.

Theorem 1.2.8 (Theorem 7.1 in [67])**.** *Assume that $X_0 \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then* (1.2.1) *admits a unique strong solution X.*

Assume now that b and σ are locally Lipschitz functions, *i.e.*, Lipschitz on every compact set. The following theorem states the existence and uniqueness of a strong solution to (1.2.1) up to an explosion time τ_∞.

Theorem 1.2.9 (Theorem 3.1 in [48])**.** *Assume that $X_0 \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a unique strong solution (X, τ_∞, B) of* (1.2.1)*.*

1.2.2 Weak existence and uniqueness in law

We now outline the key results that establish the properties of weak existence and uniqueness in law, which are used throughout this thesis.

SDE without drift: Let us first recall some basic results on the weak existence and uniqueness in law of the solution to (1.2.1) without drift. We have the following equation:

$$X_t = X_0 + \int_0^t \sigma(X_t) dB_t, \quad t \geq 0. \tag{1.2.2}$$

Proposition 1.2.10 (Proposition 1 in [43] and Problem 5.3, Chapter 5 in [54]). *Suppose that weak existence holds for* (1.2.2) *up to an explosion time* τ_∞*. Then, we have* $\tau_\infty = +\infty$ *a.s..*

In [44], the authors provide a necessary and sufficient result for weak existence and uniqueness in law for the equation (1.2.2). Let us introduce the following sets:

$$S_\sigma = \left\{ x \in \mathbb{R}, \forall \epsilon > 0 \int_{x-\epsilon}^{x+\epsilon} \frac{1}{\sigma^2(y)} dy = \infty \right\} \text{ and } N_\sigma = \{x \in \mathbb{R}, \sigma(x) = 0\}.$$

Proposition 1.2.11 (Weak existence and uniqueness in law to (1.2.2)). *Weak existence holds (without explosion) for equation* (1.2.2) *with arbitrary initial distribution* X_0 *if and only if* $S_\sigma \subset N_\sigma$*. In that case, uniqueness in law holds for every initial distribution* X_0 *if and only if* $S_\sigma = N_\sigma$*.*

SDE with drift: We return to the general framework for the case where the drift term is nonzero. Let us introduce the following hypotheses:

(ND): $\forall x \in \mathbb{R}$, $\sigma^2(x) > 0$,

(LI): $\forall x \in \mathbb{R}, \exists \varepsilon > 0$ such that $\int_{x-\varepsilon}^{x+\varepsilon} \frac{|b(y)|}{\sigma^2(y)} \, dy < \infty$.

Under these assumptions, we fix a number $c \in \mathbb{R}$ and define the scale function (see Section 1.3 for more details on this function):

$$\forall x \in \mathbb{R}, \quad s(x) = \int_c^x \exp\left(-2 \int_c^u \frac{b(v)}{\sigma^2(v)} \, dv \right) du.$$

The function s has a continuous, strictly positive derivative, and s'' exists almost everywhere. Furthermore, s has a continuously differentiable inverse $\tilde{s} : (s(-\infty), s(+\infty)) \to \mathbb{R}$ and such that $\tilde{s}'(y) = \frac{1}{s'(\tilde{s}(y))}$. We extend s to $[-\infty, +\infty]$ and $\tilde{s}$ to $[s(-\infty), s(\infty)]$.

Proposition 1.2.12 (Proposition 5.13, Chapter 5 in [54]). *Assume (ND) and (LI) hold, then the process* X *is a weak (or strong) solution of equation* (1.2.1) *if and only if the process* Y *is a weak (or strong) solution of:*

$$Y_t = Y_0 + \int_0^t \tilde{\sigma}(Y_s) dB_s, \quad t \geq 0,$$

where $s(-\infty) < Y_0 < s(+\infty)$ *a.s., and*

$$\tilde{\sigma}(y) = \begin{cases} s'(\tilde{s}(y))\sigma(\tilde{s}(y)) & \text{if } s(-\infty) < y < s(+\infty), \\ 0 & \text{otherwise.} \end{cases}$$

The process X *may explode in finite time, but the process* Y *does not.*

The following theorem is a direct application of Proposition 1.2.11 and Proposition 1.2.12.

Theorem 1.2.13 (Theorem 5.15, Chapter 5 in [54])**.** *Assume that* b *is a locally integrable function at every point in* $\mathbb{R}$*, and conditions (ND) and (LI) hold. Then, for every initial distribution* X_0*, the equation* (1.2.1) *has a weak solution up to an explosion time, and this solution is unique in law.*

We also recall the following two classical results that will be used in the upcoming section.

Theorem 1.2.14 (Theorem 2.3 in [48])**.** *Assume that* b *and* σ *are continuous, if* $X_0 \in L^2(\Omega, \mathcal{F}, \mathbb{P})$*, then there exists a weak solution* (X, τ_∞, B) *of* (1.2.1)*.*

Proposition 1.2.15 ([111], Corollary 1.14, Chapter IX in [91])**.** *Assume* σ *and* b *are bounded measurable functions and for some* $\epsilon > 0$*, we have* $|\sigma| \geq \epsilon$*. Then, weak existence and uniqueness in law hold for* (1.2.1)*.*

Remark 1.2.16. *In the case where* b *and* σ *are at most linear growth, we have* $\tau_\infty = +\infty$ *a.s. (see Theorem 2.4 in [48]). We can also use the result in Section 1.3.3 to conclude whether* $\tau_\infty = +\infty$ *a.s. or not.*

1.2.3 Pathwise uniqueness

We are now focusing on the pathwise uniqueness property of the SDE given by (1.2.1). The following results are taken from [66] and [91].
We denote $\rho : [0, \infty) \to [0, \infty)$, an increasing function such that $\int_{\mathcal{V}(0^+)} \frac{dx}{\rho(x)} = +\infty$, where $\mathcal{V}(0^+)$ is a neighbourhood of 0^+. Let us state the key results that lead to pathwise uniqueness for (1.2.1).

Proposition 1.2.17. *If* X^1*,* X^2 *are two solutions of* (1.2.1) *on the same filtered probability space such that* $X_0^1 = X_0^2$ *a.s., then* $X^1 \vee X^2$ *is a solution to* (1.2.1) *if and only if* $L^0(X^2 - X^1) = 0$ *a.s..*

Proof. Applying Itô-Tanaka formula to $X^1 \vee X^2 = X^1 + (X^2 - X^1)^+$, we have:

$$X_t^1 \vee X_t^2 = (X_0^1 \vee X_0^2) + \int_0^t b(X_s^1 \vee X_s^2)ds + \int_0^t \sigma(X_s^1 \vee X_s^2)dB_s + \frac{1}{2}L_t^0(X^2 - X^1).$$

Then the conclusion holds. □

Lemma 1.2.18. *Let* X^1*,* X^2 *be two solutions of* (1.2.1) *on the same filtered probability space such that* $X_0^1 = X_0^2$ *a.s.. If uniqueness in law holds for* (1.2.1) *and* $L^0(X^2 - X^1) = 0$ *a.s., then pathwise uniqueness holds for* (1.2.1)*.*

Proof. Using Proposition 1.2.17, X^1 and $X^1 \vee X^2$ are solutions of (1.2.1). Uniqueness in law implies $X^1 = X^1 \vee X^2$, then $X^1 = X^2$ a.s.. □

Lemma 1.2.19. *Assume X is a continuous semi-martingale and for all $t \in \mathbb{R}_+^\star$:*

$$A_t = \int_0^t \mathbb{1}_{(0<X_s<\epsilon)} \frac{d\langle X, X\rangle_s}{\rho(X_s)} < \infty \text{ a.s.,} \tag{1.2.3}$$

then $L_t^0(X) = 0$ a.s..

Proof. For all $t \in \mathbb{R}_+^\star$, suppose that (1.2.3) holds. Then, using the occupation times formula in Proposition 1.1.11, we have:

$$A_t = \int_0^\epsilon \rho(x)^{-1} L_t^x(X) dx.$$

If there exists a Lebesgue-non-negligible set B such that $L_t^0(X) > 0$ $\mathbb{P}$-a.s., then $A_t = \infty$ with strictly positive probability. Then, by contradiction, $L_t^0(X) = 0$ a.s.. □

Corollary 1.2.20. *Suppose σ and b are two bounded measurable functions such that σ satisfies:*

$$\forall (x, y) \in \mathbb{R}, \quad (\sigma(x) - \sigma(y))^2 \leq \rho(|x - y|). \tag{1.2.4}$$

If X^1 and X^2 are two solutions of (1.2.1) on the same filtered probability space, then we have:

$$L_t^0(X^2 - X^1) = 0 \text{ a.s.}$$

Proof. Using (1.2.4), we have:

$$\int_0^t \mathbb{1}_{(X_s^2 - X_s^1 > 0)} \frac{d\langle X^2 - X^1\rangle_s}{\rho(X_s^2 - X_s^1)} = \int_0^t \frac{(\sigma(X_s^2) - \sigma(X_s^1))^2}{\rho(X_s^2 - X_s^1)} \mathbb{1}_{(X_s^2 - X_s^1 > 0)} ds \leq t.$$

Applying Lemma 1.2.19 on $X = X^2 - X^1$, the conclusion holds. □

Corollary 1.2.21. *Suppose σ and b are two bounded measurable functions such that for all $x \in \mathbb{R}$, $\sigma(x) > \epsilon$. Furthermore, assume that there exists an increasing bounded function f, such that*

$$\forall x, y \in \mathbb{R}, \quad (\sigma(x) - \sigma(y))^2 \leq |f(x) - f(y)|.$$

If X^1 and X^2 are two solutions of (1.2.1) on the same filtered probability space, then we have:

$$L_t^0(X^1 - X^2) = 0 \text{ a.s.}$$

Proof. Let $X = X^1 - X^2$, the idea is to apply Lemma 1.2.19. Without loss of generality, we focus on the case $\rho(x) = x$. So, we need to prove that:

$$\forall t \geq 0, \quad \int_0^t \frac{\mathbb{1}_{(X_s>0)}}{X_s} d\langle X, X\rangle_s < \infty \text{ a.s.} \tag{1.2.5}$$

Let $\delta > 0$ be fixed, we have:

$$\mathbb{E}\left[\int_0^t \frac{\mathbb{1}_{(X_s>\delta)}}{X_s} d\langle X, X\rangle_s\right] = \mathbb{E}\left[\int_0^t \frac{(\sigma(X_s^1)-\sigma(X_s^2))^2}{X_s^1 - X_s^2}\mathbb{1}_{(X_s^1-X_s^2>\delta)} ds\right],$$
$$\leq \mathbb{E}\left[\int_0^t \frac{f(X_s^1)-f(X_s^2)}{X_s^1 - X_s^2}\mathbb{1}_{(X_s^1-X_s^2>\delta)} ds\right].$$

The function f can be approximated by a sequence of uniformly bounded increasing functions $(f_n)_{n\geq 0}$ in $\mathcal{C}^1(\mathbb{R})$ such that $f_n(x) \xrightarrow[n\to\infty]{} f(x)$ and $\sup_{n\in\mathbb{N},x\in\mathbb{R}} |f_n(x)|) \leq M = \sup_{x\in\mathbb{R}} |f(x)|$. Then we have:

$$\mathbb{E}\left[\int_0^t \frac{f(X_s^1)-f(X_s^2)}{X_s^1 - X_s^2}\mathbb{1}_{(X_s^1-X_s^2>\delta)} ds\right] = \lim_{n\to\infty} \mathbb{E}\left[\int_0^t \frac{f_n(X_s^1)-f_n(X_s^2)}{X_s^1 - X_s^2}\mathbb{1}_{(X_s^1-X_s^2>\delta)} ds\right],$$
$$= \lim_{n\to\infty} \mathbb{E}\left[\int_0^t \int_0^1 f_n'(Z_s^u)\mathbb{1}_{(X_s^1-X_s^2>\delta)} du ds\right],$$
$$\leq \lim_{n\to\infty} \int_0^1 \mathbb{E}\left[\int_0^t f_n'(Z_s^u) ds\right] du.$$

where for all $u \in [0,1]$: $Z_s^u = X_s^2 + u(X_s^1 - X_s^2)$. Note that Z^u is solution to :

$$Z_t^u = Z_0^u + \int_0^t \tilde{b}(u, X_s^1, X_s^2) ds + \int_0^t \tilde{\sigma}(u, X_s^1, X_s^2) dB_s,$$

where $\tilde{b}(u, X_s^1, X_s^2) = b(X_s^2) + u(b(X_s^1) - b(X_s^2))$ and $\tilde{\sigma}(u, X_s^1, X_s^2) = \sigma(X_s^2) + u(\sigma(X_s^1) - \sigma(X_s^2))$. As the function σ and b are bounded, we have $\tilde{\sigma} \geq \epsilon$ and $|\tilde{b}| + |\tilde{\sigma}| \leq M$. Applying the symmetric Itô-Tanaka formula (see Corollary 1.1.8), we can easily prove that $\sup_{u\in[0,1],x\in\mathbb{R}} \mathbb{E}[L_t^x(Z^u)] \leq C < \infty$, where C is independent of δ. We deduce that,

$$\mathbb{E}\left[\int_0^t \frac{f(X_s^1)-f(X_s^2)}{X_s^1 - X_s^2}\mathbb{1}_{(X_s^1-X_s^2>\delta)} ds\right] \leq \lim_{n\to\infty} \int_0^1 \mathbb{E}\left[\int_0^t f_n'(Z_s^u) ds\right] du,$$
$$\leq \frac{1}{\epsilon^2} \lim_{n\to\infty} \int_0^1 \mathbb{E}\left[\int_{\mathbb{R}} f_n'(x) L_t^x(Z^u) dx\right] du,$$
$$\leq \frac{2MC}{\epsilon^2}.$$

Finally, we have:

$$\mathbb{E}\left[\int_0^t \frac{\mathbb{1}_{(X_s>\delta)}}{X_s} d\langle X, X\rangle_s\right] \leq \frac{2MC}{\epsilon^2}.$$

Letting δ tend to 0 leads to (1.2.5). Then, from Lemma 1.2.19, the conclusion holds. □

Making use of the previous results, pathwise uniqueness can be deduced for (1.2.1) under several hypotheses on b and σ.

Theorem 1.2.22 (Theorem 3.5, Chapter IX in [91])**.** *Pathwise uniqueness holds for* (1.2.1) *in each of the following cases:*

1. $|\sigma(x) - \sigma(y)|^2 \leq \rho(|x-y|)$, $|\sigma| \geq \epsilon > 0$ *and* b *and* σ *bounded.*

2. $|\sigma(x) - \sigma(y)|^2 \leq |f(x) - f(y)|$ *where* f *is increasing and bounded,* $\sigma \geq \epsilon > 0$ *and* b *is bounded.*

3. $|\sigma(x) - \sigma(y)|^2 \leq \rho(|x - y|)$ *and* b *is bounded and* K*-Lipschitz continuous.*

Proof. Let X^1 and X^2 be two solutions of equation (1.2.1) on the same filtered probability space and such that $X_0^1 = X_0^2$ a.s..

In the case 1. and 2., since $|\sigma| > \epsilon$, from Proposition 1.2.15, weak existence and uniqueness in law hold for (1.2.1). Then from Corollary 1.2.20 and Corollary 1.2.21, we have, for all $t \geq 0$, $L_t^0(X^1 - X^2) = 0$ a.s.. We conclude by using Lemma 1.2.18.

Let us now focus on 3., suppose that b is K-Lipschitz continuous. From Itô-Tanaka formula and Lemma 1.2.19, we have:

$$|X_t^1 - X_t^2| = \int_0^t \operatorname{sgn}(X_s^1 - X_s^2) d(X_s^1 - X_s^2).$$

As σ is bounded, we can prove that:

$$|X_t^1 - X_t^2| - \int_0^t \operatorname{sgn}(X_s^1 - X_s^2)(b(X_s^1) - b(X_s^2))ds = \int_0^t \operatorname{sgn}(X_s^1 - X_s^2)(\sigma(X_s^1) - \sigma(X_s^2))dB_s,$$

is martingale. Then, $\mathbb{E}\left[\int_0^t \sigma(X_s^1) - \sigma(X_s^2)dB_s\right] = 0$ and we have:

$$\begin{aligned}\mathbb{E}\left[|X_t^1 - X_t^2|\right] &= \mathbb{E}\left[\int_0^t \operatorname{sgn}(X_s^1 - X_s^2)(b(X_s^1) - b(X_s^2))ds\right], \\ &\leq K\mathbb{E}\left[\int_0^t |X_s^1 - X_s^2|ds\right].\end{aligned}$$

Grönwall Lemma leads to $X^1 = X^2$ a.s. and the conclusion holds for 3..

□

Under the hypotheses of Theorem 1.2.22, we deduce a comparison theorem for processes solution of (1.2.1).

Theorem 1.2.23 (Comparison Theorem)**.** *For* $i = 1, 2$*, let* X^i *be the solution of the following SDE:*

$$t \geq 0, \quad X_t^i = X_0^i + \int_0^t b_i(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s.$$

Assume further that for all $x \in \mathbb{R}$*,* $b_1(x) \geq b_2(x)$*, one of the two functions* b_1 *or* b_2 *is* K*-Lipschitz, and* $X_0^1 \geq X_0^2$ *a.s.. Then, under the hypotheses of Theorem 1.2.22,* $X_t^1 \geq X_t^2$ *for all* $t \geq 0$ *a.s..*

Proof. Without loss of generality, suppose that b_1 is Lipschitz. Then, from Tanaka formula:

$$\begin{aligned}\mathbb{E}\left[(X_t^2 - X_t^1)^+\right] &= \mathbb{E}\left[\int_0^t \mathbb{1}_{(X_s^2 > X_s^1)}(b_2(X_s^2) - b_1(X_s^1))ds\right] \\ &\leq \mathbb{E}\left[\int_0^t \mathbb{1}_{(X_s^2 > X_s^1)}(b_1(X_s^2) - b_1(X_s^1))ds\right] \\ &\leq K\int_0^t \mathbb{E}\left[(X_t^2 - X_t^1)^+\right] ds.\end{aligned}$$

Using Grönwall Lemma, we have $\mathbb{E}\left[(X_t^2 - X_t^1)^+\right] = 0$ for all $t \geq 0$. □

We refer to [48] and [91] for further discussions on the existence and uniqueness of solutions, as well as for examples and remarks related to Theorem 1.2.22.

1.3 Diffusion processes

1.3.1 Infinitesimal generator, scale function and speed measure

Let I be an interval with left endpoint $l_1 \geq -\infty$ and right endpoint $l_2 \leq +\infty$. Define X as the process that takes values in I and such that X is the unique strong solution of:

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad t \geq 0, \tag{1.3.1}$$

where $X_0 = x_0 \in I$, b and σ are two measurable functions that satisfy:

(ND)': $\forall x \in I$, $\sigma^2(x) > 0$,

(LI)': $\forall x \in I, \exists \varepsilon > 0$ such that $\int_{x-\varepsilon}^{x+\varepsilon} \frac{|b(y)|}{\sigma^2(y)}\, dy < \infty$.

Here I is called the state space of the process. The solution of (1.3.1) is defined up to the first exit time $\tau_I := \tau_{l_1} \wedge \tau_{l_2}$, where for $i = 1, 2$, $\tau_{l_i} = \inf\{t > 0 : X_t = l_i\}$. Next, we denote $\mathbb{P}(.|X_0 = x_0) = \mathbb{P}_{x_0}(.)$.

Proposition 1.3.1 ([53],[91]). *The process X is a strong Markov process with continuous paths on I, such that $\mathbb{P}_x(\tau_y < \infty) > 0$ for all $x, y \in I$.*

We denote $(P_t)_{t\geq 0}$ as the semigroup of the process X.

Definition 1.3.2 (Definition 1.1, Chapter VII in [91]). *The infinitesimal generator $\mathcal{L}$ of X is defined by*

$$\mathcal{L}f(x) = \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}, \quad \forall x \in I, \tag{1.3.2}$$

with $f \in Dom(\mathcal{L}) = \{f \in C_0(I) | \mathcal{L}f \in C_0(I)\}$.

Proposition 1.3.3. *The infinitesimal generator $\mathcal{L}$ of X is given by:*

$$\mathcal{L} : Dom(\mathcal{L}) \to Dom(\mathcal{L})$$
$$f \mapsto \frac{1}{2}\sigma^2 f'' + bf',$$

with $f \in Dom(\mathcal{L}) = \{f \in C_0(I) | \mathcal{L}f \in C_0(I)\}$.

Proof. Let $f \in C^2(I)$ with compact support, using Itô formula, we have:

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)b(X_s) + \frac{1}{2}f''(X_s)\sigma^2(X_s)ds + \int_0^t f'(X_s)\sigma(X_s)dB_s.$$

Then, from Definition 1.3.2, for $x \in I$:

$$\begin{aligned}\mathcal{L}f(x) &= \lim_{t\to 0}\frac{1}{t}\mathbb{E}_x\left[\int_0^t f'(X_s)b(X_s) + \frac{1}{2}f''(X_s)\sigma^2(X_s)ds\right],\\ &= \mathbb{E}_x\left[\lim_{t\to 0}\frac{1}{t}\int_0^t f'(X_s)b(X_s) + \frac{1}{2}f''(X_s)\sigma^2(X_s)ds\right],\\ &= b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x).\end{aligned}$$

□

Definition 1.3.4. *Let X be a solution of (1.3.1). A scale function is a continuous, increasing function from I to $\mathbb{R}$ such that, for $x \in [a, b] \subset I$,*

$$\mathbb{P}_x(\tau_a < \tau_b) = \frac{s(x) - s(b)}{s(a) - s(b)}.$$

Note that from Definition 1.3.4, the scale function is unique up to a multiplicative and an additive constant. Furthermore, the continuity and increasing property of the scale function can be deduced from the strong Markov property of X.

Proposition 1.3.5. *The process $s(X_{\cdot\wedge\tau_I})$ is a local martingale. Furthermore, the scale function satisfies*

$$\frac{1}{2}\sigma^2(x)s''(x) + b(x)s'(x) = 0, \tag{1.3.3}$$

and for $x, c \in I$, we have:

$$s(x) = \int_c^x \exp\left(-2\int_c^u \frac{b(v)}{\sigma^2(v)}\,dv\right)\,du. \tag{1.3.4}$$

Proof. Let $\tau < \tau_I$ be a stopping time for the process X. Applying the strong Markov property, we have:

$$\begin{aligned}\mathbb{E}_x\left[\frac{s(X_\tau) - s(b)}{s(a) - s(b)}\right] &= \mathbb{E}_x\left[\mathbb{E}\left[\mathbb{1}_{(\tau_a<\tau_b)}|\mathcal{F}_\tau\right]\right],\\ &= \frac{s(x) - s(b)}{s(a) - s(b)}.\end{aligned}$$

Then $s(X_{\cdot\wedge\tau_I})$ is a local martingale. The differential equation (1.3.3) is obtained by applying the general Itô formula (see Remark 1.1.3). □

Let $\tau_{[a,b]} = \{t > 0 : X_t \notin [a,b]\}$, we now introduce the speed measure of the process X.

Theorem 1.3.6 (Theorem 3.6, Chapter VII in [91]). *There exists a unique strictly positive locally finite measure m defined on $int(I)$ such that, for $x, a, b \in I$ with $a < x < b$ and $f : \mathbb{R} \to \mathbb{R}$, the function $x \mapsto \mathbb{E}_x[f(\tau_{[a,b]})]$ is continuous on the interval $[a,b]$ and we have:*

$$\mathbb{E}_x[\tau_{[a,b]}] = \int_a^b G_{[a,b]}(x,y)m(dy),$$

where,

$$G_{[a,b]}(x,y) = \begin{cases} \dfrac{(s(x)-s(a))(s(b)-s(y))}{s(b)-s(a)} & \text{if } a \le x \le y \le b, \\ \dfrac{(s(b)-s(x))(s(y)-s(a))}{s(b)-s(a)} & \text{if } a \le y \le x \le b, \\ 0 & \text{otherwise.} \end{cases}$$

An alternative definition for the speed measure can be given by the following results.

Theorem 1.3.7 (Theorem 3.12, Chapter VII in [91]). *The speed measure m is defined as the measure such that $\mathcal{L}$ can be written as:*

$$\mathcal{L}f = \frac{d}{dm}\frac{d}{ds}f, \tag{1.3.5}$$

where:

$$\frac{d}{ds}f(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{s(x+h)-s(x)} \text{ and } \frac{d}{dm}g(x) = \lim_{h\to 0}\frac{g(x+h)-g(x)}{\int_x^{x+h} m(dy)}.$$

In this thesis, we mainly consider processes that are solutions of (1.3.1). In this case, the scale function is always in $\mathcal{C}^1(I)$ and m is always absolutely continuous with respect to the Lebesgue measure.

Proposition 1.3.8. *Let X be the solution of (1.3.1) then the speed measure m associated to this process is given by:*

$$m(dx) = \frac{2}{\sigma^2(x)s'(x)}dx. \tag{1.3.6}$$

Proof. Using (1.3.5), we have:

$$\mathcal{L}f(x) = \frac{d}{dm}\frac{d}{ds}f(x) = \frac{1}{m(x)s'(x)}f''(x) + \frac{2b(x)}{m(x)s'(x)\sigma^2(x)}f'(x).$$

Then we conclude by identification, using Proposition 1.3.3. □

Proposition 1.3.9. *Let X be the solution of (1.3.1). Then, for all $x \in int(I)$, we have:*

$$\int_0^t \mathbb{1}_{(X_s=x)}ds = 0, \quad \mathbb{P}\text{-a.s.}$$

In particular, we have $L_t^x(X) = \ell_t^x$ a.s..

Proof. This result follows from a direct application of Proposition 1.1.11 and Theorem 1.1.6. □

1.3.2 Killed process and Fokker-Planck equation

Let X, the process solution of:

$$X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dB_s, \quad t \geq 0, \tag{1.3.7}$$

where $x_0 \in (-\infty, c)$, $c \in \mathbb{R}$. Here, b and σ are two Lipschitz functions such that $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+$. From the previous section, these conditions imply that X is the unique non-exploding strong solution of (1.3.7). Here, the conditions on b and σ are sufficient for the results introduced in this section.

We introduce the process $\bar{X}$, which is the process X killed at the constant boundary c, *i.e.*,

$$\bar{X}_t = \begin{cases} X_t \text{ on } \{t < \tau_c\}, \\ \Delta \quad \text{otherwise,} \end{cases}$$

where Δ is the cemetery.

Proposition 1.3.10. *The infinitesimal generator $\mathcal{L}$ of $\bar{X}$ is given by:*

$$\mathcal{L} : Dom(\mathcal{L}) \to Dom(\mathcal{L})$$

$$f \mapsto \frac{1}{2}\sigma^2 f'' + bf',$$

where $Dom(\mathcal{L}) = \{f \in C_0((-\infty, c)) | f(c) = 0, \mathcal{L}f \in C_0((-\infty, c))\}$.

We denote $p(x_0, t) = \mathbb{P}_{x_0}(t < \tau_c)$. The following proposition states the Fokker-Planck equation associated to p.

Proposition 1.3.11. *The function p satisfies the following Fokker-Planck equation:*

$$\begin{cases} \frac{\partial}{\partial t}p(x_0, t) = \frac{\sigma^2(x_0)}{2}\partial^2_{x_0^2}p(x_0, t) + b(x_0)\partial_{x_0}p(x_0, t), & (t, x_0) \in \mathbb{R}^\star_+ \times (-\infty, c), \\ p(c, t) = 0, & t \in \mathbb{R}^\star_+, \\ p(x_0, 0) = 1, & x_0 \in (-\infty, c]. \end{cases}$$

Proof. This result is a straight application of Theorem 4.4.5 in [46]. □

Remark 1.3.12. *Let $c_1, c_2 \in \mathbb{R}$ such that $c_1 < c_2$. Similar results can be obtained for the process X killed when it leaves the interval $[c_1, c_2]$. In this case, $x_0 \in [c_1, c_2]$ and for all $t \in \mathbb{R}^\star_+$, the boundary conditions are $p(c_1, t) = p(c_2, t) = 0$.*

1.3.3 Boundary classification

In this section, we provide the classification of boundary conditions for the solutions of stochastic differential equations given by (1.3.1).

Definition 1.3.13 ([49] and [19]). *The boundary conditions are classified as follows: for all $l_1 < x < l_2$,*

1. *The left-hand endpoint l_1 is called exit if,*
$$\int_{l_1}^{x} \int_{y}^{x} m(dz)s'(y)dy < \infty.$$
2. *and entrance if,*
$$\int_{l_1}^{x} \int_{l_1}^{y} m(dz)s'(y)dy < \infty.$$

For all $l_1 < x < l_2$,

1. *The left-hand endpoint l_2 is called exit if,*
$$\int_{x}^{l_2} \int_{x}^{y} m(dz)s'(y)dy < \infty.$$
2. *and entrance if,*
$$\int_{x}^{l_2} \int_{y}^{l_2} m(dz)s'(y)dy < \infty.$$

- *A boundary point, which is both entrance and exit, is called non-singular, i.e., the process can reach this point with positive probability.*
- *A boundary point, which is neither entrance nor exit, is called natural, i.e., the boundary point cannot be reached in finite time. It is not possible to start the process from this boundary.*
- *A boundary point, which is entrance but not exit, cannot be reached from an interior point of I. It is possible to start the process from this boundary.*
- *A boundary point, which is exit but not entrance, can be reached from an interior point of I with positive probability. It is not possible to start the process from this boundary.*

In this thesis, we study the behavior of processes with both singular and non-singular boundaries. We define the notion of instantaneous reflecting boundary in the following definition.

Definition 1.3.14. *Let X be the solution of (1.2.1) and let $i = 1, 2$ such that $\{l_i\}$ is a non-singular boundary, then $\{l_i\}$ is said to be instantaneously reflecting if and only if $\mathbb{P}_x\left(\int_0^t \mathbb{1}_{(X_s=l_i)}ds = 0\right) = 1$.*

Remark 1.3.15. *For $i = 1, 2$, if $\{l_i\}$ is an instantaneously reflecting boundary, then according to Proposition 1.3.1, for all $y \in (l_1, l_2)$, we have $\mathbb{P}_{l_i}(\tau_y < \infty) > 0$.*

1.3.4 Regime of the process

In this section, we define the conditions for a non-exploding strong solution X of (1.2.1) to be ergodic. The regime of the process is characterized by properties on the scale function s and the speed measure m defined above. We mainly focus on the case where $\{l_1\}$ and $\{l_2\}$ are unattainable boundaries. Some remarks will be done in the case where one boundary is instantaneously reflecting.

Definition 1.3.16 (Invariant measure)**.** *A measure μ on $(I, \mathcal{B}(I))$ is called invariant if for all $t \geq 0$, we have:*

$$\int_I \mu(dx) P_t(dy) = \mu(dy).$$

Definition 1.3.17 (Recurrent process)**.** *The process X with values in I is said to be recurrent if:*

$$\mathbb{P}_x(\tau_y < \infty) = 1, \text{ for all } x, y \in I,$$

otherwise the process is transient. Furthermore, if:

$$\mathbb{E}_x[\tau_y] < \infty \text{ for all } x, y \in I,$$

then X is positive recurrent, otherwise the process is null-recurrent.

The recurrence property for the process X can be deduced from the scale function. The following proposition states the relation between the scale function behavior and the recurrence property.

Proposition 1.3.18. *The process X is recurrent if and only if $s(l_1) = -\infty$ and $s(l_2) = +\infty$.*

Proof. Let $l_1 < a < x < b < l_2$, from Definition 1.3.4,

$$\mathbb{P}_x(\tau_a < \tau_b) = \frac{s(x) - s(b)}{s(a) - s(b)},$$

Letting $b \to l_2$, since X is non-exploding, we have:

$$\mathbb{P}_x(\tau_a < \infty) = \frac{1 - \frac{s(x)}{s(l_2)}}{1 - \frac{s(a)}{s(l_2)}}.$$

Then, for all $x > a$ $\mathbb{P}_x(\tau_a < \infty) = 1$ if and only if $s(l_2) = +\infty$. We also have,

$$\mathbb{P}_x(\tau_b < \tau_a) = \frac{s(x) - s(a)}{s(b) - s(a)}.$$

The same argument can be applied by letting $a \to l_1$. As the scale function is strictly increasing, we conclude that $s(l_1) = -\infty$ and $s(l_2) = +\infty$ if and only if X is recurrent. □

Proposition 1.3.19. *Let X be a recurrent process with speed measure m, and fix two measurable functions $f, g : I \to \mathbb{R}^+$ with $f \in L^1(m)$ and $g > 0$ m-almost-surely. Then*

$$\lim_{t\to\infty} \frac{\int_0^t f(X_s)\,ds}{\int_0^t g(X_s)\,ds} = \frac{\int_I f(x)m(dx)}{\int_I g(x)m(dx)}, \quad \mathbb{P}\text{-}a.s.$$

Proof. The proof of this result is omitted. For further details, see Lemma 20.14 [53] and [77]. □

By abuse of language, positive recurrent processes are often called "ergodic". In the next, we refer to an ergodic process as a positive recurrent process. We also denote $\mathbb{E}_\mu[.]$, when we are referring to the expectation of a random variable under the measure μ.

Proposition 1.3.20. *(Existence invariant measure) Any ergodic process X solution of* (1.2.1) *admits an invariant measure.*

Proof. The proof of this result is given in Lemma 20.18 in [53]. □

Proposition 1.3.21. *The process X is ergodic if and only if $\int_I m(dx) < \infty$. In this case, we have $\mu(dx) = \frac{m(dx)}{\int_I m(dx)}$.*

Proof. Letting $a \to l_1$ and $b \to l_2$ in Theorem 1.3.6, one can easily verify that:

1. For $x < b$, we have:

$$\mathbb{E}_x[\tau_b] = (s(b) - s(x)) \int_{l_1}^{x} m(dy) + \int_x^b (s(b) - s(y))m(dy).$$

2. For $x > a$, we have:

$$\mathbb{E}_x[\tau_a] = (s(x) - s(a)) \int_a^{l_2} m(dy) + \int_a^x (s(y) - s(a))m(dy).$$

Then for all $x, y \in I$, $\mathbb{E}_x[\tau_y] < \infty$ if and only if $\int_I m(dx) < \infty$. So X is ergodic and from Proposition 1.3.20, it admits an invariant measure μ. Let $f : I \mapsto \mathbb{R}_+$ with bounded support, using Proposition 1.3.19, Fubini Theorem and dominated convergence, we have:

$$\frac{1}{t}\int_0^t \mathbb{E}_\mu[f(X_s)]ds \xrightarrow[t\to\infty]{} \frac{\int_I f(x)\ m(dx)}{\int_I m(dx)},$$

then we get $\int_I f(x)\mu(dx) = \frac{\int_I f(x)\ m(dx)}{\int_I m(dx)}$ and the conclusion holds. □

Remark 1.3.22. *The previous results can be adapted in the case where $\{l_1\}$ is an instantaneously reflecting barrier. In fact, for the recurrence property, it suffices to have $s(l_2) = +\infty$ and $s(l_1) < +\infty$. For further details on the ergodic behaviour of the process X with differents boundary conditions, we refer to Theorem 20.15 in [53].*

1.4 Some examples of stochastic differential equations

In this section, we introduce some well-known processes solution of SDEs used in this thesis.

1.4.1 Ornstein-Uhlenbeck process

In the following, we present the Ornstein-Uhlenbeck (OU) process, originally introduced in [104].

Definition 1.4.1 (Ornstein-Uhlenbeck (OU) process)**.** *Let $(a, b, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+^\star$, $x_0 \in \mathbb{R}$ and B a Brownian motion. The OU process is solution of:*

$$X_t = x_0 + \int_0^t a - bX_s ds + \sigma B_t, \quad t \geq 0. \tag{1.4.1}$$

Proposition 1.4.2. *There exists a unique strong solution to* (1.4.1)*. Additionally, the solution is given by:*

$$X_t = x_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) + \sigma \int_0^t e^{b(s-t)} dB_s, \quad t \geq 0, \tag{1.4.2}$$

Furthermore, the process X is a Gaussian process with mean and covariance given by:

$$\mathbb{E}[X_t] = x_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) \quad \textit{and} \quad \mathbb{C}ov(X_t, X_s) = \sigma^2 e^{-b(t+s)} \left(\frac{e^{2b(t \wedge s)} - 1}{2b} \right).$$

Proof. Existence of a unique strong solution to (1.4.1) follows from Theorem 1.2.8. The solution (1.4.2) can be obtained by applying Itô formula on $(e^{bt} X_t)_{t \geq 0}$. Proving that X is a gaussian process follows directly from (1.4.2). □

In finance, the solution of equation (1.4.1) is commonly referred as a Vasicek model. Typically, the parameter b is selected to be positive, resulting in $\mathbb{E}[X_t] \xrightarrow[t \to +\infty]{} \frac{a}{b}$, which is why this process is described as having a mean-reverting property.

Proposition 1.4.3. *The process X is ergodic if and only if $b \in \mathbb{R}_+^\star$.*

Proof. The speed measure and scale function of the process X are given by:

$$s(x) = \int^x e^{\frac{1}{\sigma^2}(by^2 - 2ay)} dy \quad \text{and} \quad m(dx) = \frac{2}{\sigma^2} e^{-\frac{1}{\sigma^2}(bx^2 - 2ax)} dx.$$

Then, using Proposition 1.3.18 and Proposition 1.3.21, we have $s(x) \xrightarrow[x \to \pm\infty]{} \pm\infty$ and $\int_{\mathbb{R}} m(dy) < \infty$ if and only if $b \in \mathbb{R}_+^\star$. □

1.4.2 Chan–Karolyi–Longstaff–Sanders process

The Chan–Karolyi–Longstaff–Sanders (CKLS) process is a stochastic process used in finance. It was first introduced in [27] for modeling the term structure of interest rates.

Definition 1.4.4 (CKLS process)**.** *Let $(a, b, \sigma, \gamma) \in \mathbb{R}_+^\star \times \mathbb{R}_+ \times \mathbb{R}_+^\star \times [1/2, 1]$ and B a Brownian motion, the CKLS process X is solution of:*

$$X_t = x_0 + \int_0^t a - bX_s ds + \int_0^t \sigma |X_s|^\gamma dB_s, \quad t \geq 0, \tag{1.4.3}$$

where $x_0 > 0$.

This process includes several different models such as the Cox-Ingersoll-Ross (CIR) process ($\gamma = \frac{1}{2}$) and the OU process ($\gamma = 0$) introduced in Definition 1.4.1.

Proposition 1.4.5. *There exists a unique strong solution to* (1.4.3)*.*

Proof. Existence of a unique non-exploding strong solution follows from Theorem 1.2.14, Theorem 1.2.22 and Remark 1.2.16. In fact, we have:

$$|\sigma x^\gamma - \sigma y^\gamma|^2 \leq \rho(|x - y|) = \sigma^2 |x - y|^{2\gamma}.$$

□

Lemma 1.4.6. *For the process X solution of the SDE* (1.4.3)*, the following holds:*

1. *The state space is $I = [0, +\infty)$ and the point $\{0\}$ is instantaneously reflecting, if $\gamma = \frac{1}{2}$ and $a \leq \frac{\sigma^2}{2}$.*
2. *Otherwise, the state space is $I = (0, +\infty)$ and $\{0\}$ is an unattainable boundary.*

Proof. 1. The state space can be directly deduce from Theorem 1.2.23 and Definition 1.3.13. In fact, taking $a = 0$ and $x_0 = 0$, one can verify that 0 is solution of (1.4.3). According to Theorem 1.2.23, we have $X_t \geq 0$ a.s.. Then, using the scale function, the speed measure, and Definition 1.3.13, we can establish that $\{0\}$ is a non-singular boundary. It remains now to prove that $\{0\}$ is an instantaneously reflecting boundary. Using Theorem 1.1.6, we have:

$$\ell_t^0(X) - \ell_t^{0^-}(X) = 2a \int_0^t \mathbb{1}_{(X_s = 0)} ds.$$

As the process is positive, we have $\ell_t^{0^-}(X) = 0$ a.s.. Making use of the occupation times formula, we obtain:

$$t \geq \int_0^t \mathbb{1}_{(0 < X_s)} ds = \int_0^t \frac{\mathbb{1}_{(0<X_s)}}{\sigma^2 X_s} d\langle X, X \rangle_s = \int_0^\infty \frac{L_t^y(X)}{\sigma^2 y} dy.$$

Then $L_t^0(X) = 0$ a.s., and the process X spends no time in zero (see Definition 1.3.14).

2. Using the boundary classification, we can prove that $\{0\}$ cannot be reached from an interior point of $(0, +\infty)$ (see Definition 1.3.13).

□

From the previous lemma, we can remove the absolute values on the SDE (1.4.3). The following lemma states an alternative construction of the CIR process ($\gamma = 1/2$ in (1.4.3)).

Lemma 1.4.7 (Construction of a CIR process)**.** *Let* $(a, b, \sigma) \in \mathbb{R}_+^\star \times \mathbb{R} \times \mathbb{R}_+^\star$ *and* $\|Y_t\|^2 = \sum_{j=1}^n (Y_t^{(j)})^2$ *where* $(Y_t^{(j)})_{j=1}^n$ *satisfy for all* $j = 1, \dots, n$ *the SDE:*

$$Y_t^{(j)} = y_0 - \frac{b}{2}\int_0^t Y_s^{(j)} ds + \frac{\sigma}{2}\tilde{B}_t^{(j)}, \quad t \geq 0,$$

with $y_0 \in \mathbb{R}$ *and* $\tilde{B} = (\tilde{B}^{(j)})_{j=1}^n$ *a* n*-dimensional Brownian motion. Then,* $X = \|Y\|^2$ *is solution of:*

$$X_t = x_0 + \int_0^t \left(n\sigma^2/4 - bX_t\right) dt + \int_0^t \sigma\sqrt{X_t} dB_t, \quad t \geq 0, \tag{1.4.4}$$

where $x_0 = ny_0^2$ *and* B *is a Brownian motion.*

Proof. The proof follows by using a similar reasoning to Theorem 1.2 in Chapter XI of [91]. By Itô-formula, it holds:

$$\begin{aligned} dX_t &= \sum_{j=1}^n 2Y_t^{(j)} dY_t^{(j)} + \frac{1}{4}\sum_{j=1}^n \sigma^2 dt = \left(n\sigma^2/4 - bX_t\right) dt + \sum_{j=1}^n \sigma Y_t^{(j)} d\tilde{B}_t^{(j)}, \\ &= \left(n\sigma^2/4 - bX_t\right) dt + dM_t. \end{aligned}$$

where $M_t := \sum_{j=1}^n \int_0^t \sigma Y_s^{(j)} d\tilde{B}_s^{(j)}$. The process M is a local martingale and, by Levy-characterization theorem, there exists a Brownian motion B such that $M_t = \int_0^t \sigma\sqrt{X_s} dB_s$. Therefore X satisfies (1.4.4). The proof is thus completed. □

Remark 1.4.8. *In the case where* $b = 0$ *in the previous lemma, a Square Bessel process (see e.g. [52]) can be constructed as the norm of a sequence of Brownian motions.*

Part I

Parameters inference of the Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck (OU) process was introduced by G.E. Uhlenbeck and L.S Ornstein in [104]. This model provides a stochastic framework to describe systems that revert to an average or equilibrium state.

In finance, the OU process has found extensive use, particularly in interest rate modeling. One prominent application is the Vasicek model, introduced by O. Vasicek in [107]. This model uses the OU process to capture the dynamics of interest rates over time, assuming they follow a mean-reverting pattern with fluctuations around a long-term mean level.

In addition to finance, the OU process is also widely used in fields such as neuroscience [63], and other natural sciences to analyze and model various types of dynamic phenomena. In particular, the OU process can be used to model the dynamics of temperatures. For instance, as demonstrated by [37] (see also [4]), temperatures tend to distribute around a mean axis, making it natural to model their evolution through a mean-reverting process. Given the effectiveness of mean-reverting processes for temperature modeling, estimating their parameters becomes crucial for accurate risk assessment. In [15], the authors estimate the parameters of an OU process to conduct risk measures for phenomena such as heatwaves. Since meteorological datasets typically only include daily temperature extremes, their estimation method is built from a sample of suprema from a single trajectory of the OU process. The authors proposed a least squares method to estimate the OU parameters. They used the cumulative distribution function of the supremum in integral form obtained from the findings in [6] on the hitting times of the OU process. Note that they do not possess statistical properties on their estimators, and furthermore, their method is computationally expensive.

In this part, our goal is to propose an alternative estimation method for this problem. Additionally, this method should be less computationally expensive, and we aim to have statistical properties on our estimators. We focus on parameter estimation for an OU process based on observations of its supremum. The distribution associated with the OU supremum, according to [16], is represented as a series of Parabolic Cylinder functions. Studying the asymptotic behavior of the μ-zeros of the Parabolic Cylinder functions is a crucial step in deriving the statistical properties of estimators. Therefore, the second part of this work addresses the properties of the μ-zeros of the Parabolic Cylinder functions. More precisely, in Chapter 2, we derive an analytical expression for the supremum density. Making use of the pseudo-likelihood method based on the supremum density, our estimator is constructed as the maximum argument of this function. Using weak-dependence results, we establish some statistical properties of the estimator such as consistency and asymptotic normality. Finally, we apply our estimator to simulated and real data. In Chapter 3, we study some properties of μ-zeros of Parabolic Cylinder functions. This analysis contributes to a deeper understanding of the behavior of these special functionss, which are essential for bounding the distribution of the OU supremum. Let $D_\mu(z)$ be the Parabolic Cylinder function. The zeros of the function $\mu \mapsto D_\mu(z)$ are studied with respect to the real variable z. A formula for the derivative of a zero is established, leading to some monotonicity results. Additionally, an asymptotic expansion for μ-zeros is provided when z is a large positive real number. We conclude this part by suggesting

potential directions for further research and highlighting areas for improvement.

In this part only, we introduce a slight change in notation. Instead of considering an OU process with parameters (a, b, σ) as in (1.4.1), we consider (a, b, β) with $\beta = \sigma^2$. This change will allow us to be consistent with the parameterization used in [15].

The works presented in this part were published in *Statistical Inference for Stochastic Processes* [18] and in *Le Mathematiche* [17].

Chapter 2

A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations

2.1 Introduction

Ornstein-Uhlenbeck processes are often used in various fields such as finance [92, 106], temperature modeling [4, 22, 37, 38], medicine [103], physics [110], neuroscience [64]. Estimation comes together with simulation and modeling, as models need to be calibrated. Estimators of discretely observed diffusions are presented in [3]. In [64], the authors propose an estimation based on observations of first hitting times.

We consider a stationary OU process X solution of:

$$\begin{cases} dX_t = (a - bX_t)dt + \sqrt{\beta}dB_t, \\ X_0 \sim \mathcal{N}\left(\frac{a}{b}, \frac{\beta}{2b}\right), \end{cases} \tag{2.1.1}$$

with parameter $\theta = (a, b, \beta) \in \mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$, and B a Brownian motion independent of X_0. In this chapter, we focus on the parameter estimation of X when we observe $(S^1, \cdots, S^N)$, $N \in \mathbb{N}^\star$ a set of suprema observations taken over a single trajectory.

This approach has already been proposed in [15], using an estimator constructed by the least squares method. Few statistical results have been proved on this estimator and it is computationally expensive to deal with. Our goal is to provide a new estimator with good statistical properties and less computational cost.

Even if the observations $(S^1, \cdots, S^N)$ are dependent, the sequence has some mixing properties that imply asymptotic independence. So we estimate the parameter θ using a

pseudo-likelihood method introduced in the early 70's in [13]. Our estimator is then the maximal argument of the pseudo-likelihood function:

$$\mathcal{L}(\theta) = \prod_{i=1}^{N} f_{S^i}(S^i, \theta),$$

where f_{S^i} is the probability density function associated to the random variable S^i.

The estimator requires an expression of the probability density f_{S^i}. One contribution of this work is to propose an explicit expression of this density based on the Parabolic Cylinder function which is numerically inexpensive. Moreover, with this expression we obtain the consistency and asymptotic normality properties of the estimator.

Outline: In Section 2.2, we present some properties of the OU supremum sequence, then we derive the probability density function of the supremum. The estimation method and the estimator statistical properties are presented in Section 2.3. In Section 2.4, some numerical experiments on simulated data are performed. We also present an application of our procedure to a dataset of daily temperature extreme values from Paris [60]. Proofs are collected in Section 2.5 and some auxiliary results in Section 2.6.

2.2 Some results related to the supremum of an OU process

Let X be an OU process defined by (2.1.1), $(t_n)_{n\geq 0}$ a sequence of time such that $t_0 = 0$ and for $i \geq 1$, $t_i - t_{i-1} = \Delta$, where $\Delta > 0$ is fixed. We denote $(S^{i,0})_{i\geq 1}$ the following sequence of suprema observations on time windows of size Δ:

$$S^{i,0} = \sup_{s\in[t_{i-1},t_i]} X_s.$$

2.2.1 Properties of the suprema sequence $(S^{i,0})_{i\geq 1}$

In order to present the estimation method, we give some properties of the sequence $(S^{i,0})_{i\geq 1}$. As the sequence of observations are dependent, some weak dependence notions as mixing properties (see [20, 39] e.g.) are required to get statistical results on the estimator.

The following result is induced by the properties of the stationary OU process.

Proposition 2.2.1. *The sequence $(S^{i,0})_{i\geq 1}$ is stationary, ergodic and exponentially ρ-mixing.*

Proof. Since the process X is stationary and ergodic, then by Theorem 3.5.8 in [95], the sequence $(S^{i,0})_{i\geq 1}$ is also stationary and ergodic.

Using the definition in [20], for $p \in \mathbb{N}^\star$, the sequence $(S^{i,0})_{i\geq 1}$ is ρ-mixing if it verifies:

$$\rho(p) = \sup_{n\geq 1} \sup_{f\in\mathcal{L}^2\left(\mathcal{F}_0^n\right), g\in\mathcal{L}^2\left(\mathcal{F}_{n+p}^{+\infty}\right)} |\mathrm{Corr}(f,g)| \xrightarrow[p\to+\infty]{} 0, \tag{2.2.1}$$

with $\mathcal{F}_0^n = \sigma(S^{i,0}, 1 \leq i \leq n)$ and $\mathcal{F}_{n+p}^{+\infty} = \sigma(S^{i,0}, i \geq n+p)$.

From Theorem 2.1 in [47], for all $0 \leq s \leq t$, for all functions $\tilde{f} : \mathcal{C}^0([0,s],\mathbb{R}) \to \mathbb{R}$, $\tilde{g} : \mathcal{C}^0([t,+\infty],\mathbb{R}) \to \mathbb{R}$ such that $\tilde{f}$ and $\tilde{g}$ are square integrable with respect to the law of X_0, we have:

$$\left|\mathrm{Cov}\left[\tilde{f}\left((X_u)_{u\leq s}\right), \tilde{g}\left((X_v)_{v\geq t}\right)\right]\right| \leq e^{-b(t-s)}\sqrt{\mathrm{Var}\left[\tilde{f}\left((X_u)_{u\leq s}\right)\right]\mathrm{Var}\left[\tilde{g}\left((X_v)_{v\geq t}\right)\right]}. \tag{2.2.2}$$

Therefore, process X is exponentially ρ-mixing.

However, the following inclusions are verified:

$$\sigma(S^{i,0}, 1 \leq i \leq n) \subset \sigma(X_t, 0 \leq t \leq t_n) \quad \text{and} \quad \sigma(S^{i,0}, i \geq n+p) \subset \sigma(X_t, t \geq t_{n+p-1}).$$

Then, the exponentially ρ-mixing property of $(S^{i,0})_{i\geq 1}$ is induced by (2.2.2). □

2.2.2 Supremum law

In this section, we give some results on the law of $S^{1,0}$. Since $S^{1,0} = \sup_{s\in[0,\Delta]} X_s$, we can use the existing results on supremum of an OU process. The cumulative distribution function has already been introduced in [15]. The authors used the Bessel formulation (see [6] e.g.) which is numerically expensive. In [16], the authors give the supremum cumulative distribution of the non-stationary OU process with parameter $\theta = (0, b, 1)$ in terms of Parabolic Cylinder function. Using this result, we easily obtain the cdf and the density of the supremum of a stationary OU process with parameter $\theta = (a, b, \beta)$.

Proposition 2.2.2 (Probability density of the supremum)**.** *Let X be an OU process solution of (2.1.1), with parameter $\theta = (a,b,\beta) \in \mathbb{R}\times\mathbb{R}_+^\star\times\mathbb{R}_+^\star$. For $\Delta > 0$ and $c \in \mathbb{R}$, the probability density of the supremum $S^{1,0}$ is given by:*

$$\begin{aligned} f_\Delta(c,\theta) = -\sqrt{\frac{b}{\beta\pi}} e^{-\frac{b}{\beta}\left(c-\frac{a}{\beta}\right)^2} \sum_{n\geq 1} e^{-b\mu_{n,c,\theta}\Delta} \Bigg[& -\Delta b \frac{D^2_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\partial_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)^2} \\ & + 2\frac{D_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)\partial_\mu D_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\partial_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)^2} \\ & - \frac{D^2_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)\partial^2_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\partial_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)^3} \Bigg], \end{aligned} \tag{2.2.3}$$

with $\mu_{n,c,\theta}$ the positive (ordered) zeros of the function $\mu \mapsto D_\mu\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)$ and $D_\mu(.)$ the Parabolic Cylinder function (see Section 2.6).

Proof. For $x_0, c \in \mathbb{R}$ such that $c > x_0$ and $\Delta > 0$, by replacing c, x_0, Δ and b respectively by $c - \frac{a}{b}$, $x_0 - \frac{a}{b}$, $\beta\Delta$ and $\frac{b}{\beta}$ in Proposition 3 in [16], we get the cumulative distribution of the supremum for the non-stationary OU process with parameter $\theta = (a, b, \beta)$.

Integration with respect to the invariant measure $X_0 \sim \mathcal{N}\left(\frac{a}{b}, \frac{\beta}{2b}\right)$ and Formula (2.6.3) give the cumulative distribution function of $S^{1,0}$ for the stationary OU process:

$$\mathbb{P}(S^{1,0} < c) = -\frac{e^{-(c-\frac{a}{b})^2\frac{b}{\beta}}}{\sqrt{2\pi}} \sum_{n\geq 1} e^{-b\mu_{n,c,\theta}\Delta} \frac{D_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\mu_{n,c,\theta}\partial_\mu D_{\mu_{n,c,\theta}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}. \tag{2.2.4}$$

Then, making use of Proposition 2.6.1, the series in Equation (2.2.4) is differentiable. We easily get the probability density of the random variable $S^{1,0}$. □

The new cdf expression (2.2.4) is less expensive than the one in [15].

Remark 2.2.3. *For $\Delta > 0$, the support of the random variable $S^{1,0}$ is $\mathbb{R}$.*

2.3 Estimation problem

In this section, we introduce the estimation method of a stationary OU process parameters. As the sample of observations are weak dependent, the basic idea is to use a pseudo-likelihood approach. The computation of the likelihood is simplified by approximating the joint probabilities of all data by the product of marginal probabilities.

The data $(S^{i,0})_{i\geq 1}$ are collected on disjoint but consecutive time windows of constant size Δ. Proposition 2.2.1 suggests to sample the initial set of observations by keeping one observation over k to obtain a sequence of data with less dependence between each other. Let $k \in \mathbb{N}^\star$ and $r = (k-1)\Delta$, then the set of observations used in the estimation procedure is given by:

$$S^{i,r} = \sup_{s\in[t_{i-1}+(i-1)r, t_i+(i-1)r]} X_s.$$

The choice $k = 1$ is equivalent to deal with the initial data $(S^{i,0})_{i\geq 1}$.

Remark 2.3.1. *For all $r = (k-1)\Delta$ with $k \in \mathbb{N}^\star$, the sequence $(S^{i,r})_{i\geq 1}$ has the same properties as $(S^{i,0})_{i\geq 1}$.*

We consider a stationary OU process with parameter $\theta_0 = (a_0, b_0, \beta_0)$ and $\theta_0 \in \Theta$, a compact subset of $\mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$. We denote $\mathbb{P}_{\theta_0}$ the probability measure associated to marginals from the sequence $(S^{i,r})_{i\geq 1}$.

Let $N \in \mathbb{N}^\star$ and $(S^{1,r}, \cdots, S^{N,r})$ be a sample from the sequence $(S^{i,r})_{i\geq 1}$. The pseudo-likelihood $\mathcal{L}_N^r$ associated to the sub-sequence is given by:

$$\mathcal{L}_N^r(\theta) = \prod_{i=1}^{N} f_\Delta(S^{i,r}, \theta). \tag{2.3.1}$$

The OU process parameter $\theta_0 = (a_0, b_0, \beta_0)$ is estimated by:

$$\hat{\theta}_N = (\hat{a}_N, \hat{b}_N, \hat{\beta}_N) = \underset{\theta \in \Theta}{\text{Argmax}}\ \mathcal{L}_N^r(\theta). \tag{2.3.2}$$

The results on the sequence $(S^{i,r})_{i \geq 1}$ allow to get statistical properties on the estimator. A basic but necessary result is the identifiability of the statistical model.

Proposition 2.3.2. *The statistical model $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$ is identifiable.*

The following results state the statistical properties of the estimator.

Theorem 2.3.3. *Consider an OU process solution of (2.1.1), with parameter $\theta_0 = (a_0, b_0, \beta_0)$. Assume that θ_0 belongs to Θ a compact subset of $\mathbb{R} \times \mathbb{R}_+^\star \times \mathbb{R}_+^\star$. For any $N \in \mathbb{N}^\star$, the estimator $\hat{\theta}_N$ defined by (2.3.2) is consistent:*

$$\hat{\theta}_N \xrightarrow[N \to +\infty]{\mathbb{P}_{\theta_0}} \theta_0.$$

Using Central Limit Theorem on ρ-mixing sequence of random variables, the asymptotic normality of the estimator follows.

Theorem 2.3.4. *For any $N \in \mathbb{N}^\star$ and $r > \frac{ln\left(\frac{5}{3}\right)}{b_0}$, the following convergence is satisfied:*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow[N \to +\infty]{law} \mathcal{N}_3(0, I_{\theta_0}^{-1}), \tag{2.3.3}$$

where I_{θ_0} is the Fisher information matrix given by:

$$I_{\theta_0} = \mathbb{E}_{\theta_0}\left[\left(\nabla_\theta \log f_\Delta(., \theta)|_{\theta=\theta_0}\right)\left(\nabla_\theta \log f_\Delta(., \theta)|_{\theta=\theta_0}\right)^T\right].$$

Proofs of Proposition 2.3.2, Theorems 2.3.3 and 2.3.4 are postponed in Section 2.5.

2.4 Numerical experiment

In this section, we discuss the existence of an optimal gap r used in the estimation procedure. Then, we apply our estimation method to a simulated dataset and a real dataset. Part of of this section is dedicated to the comparaison between our method and the one proposed in [15].

2.4.1 Trade off between the number of observations and the gap r

For $N \in \mathbb{N}^\star$ and $r = (k-1)\Delta$ with $k \in \mathbb{N}^\star$, an important aspect of this study is the choice of the sample as we take the set of observations from the sequence $(S^{i,r})_{1\leq i\leq \lceil N/k\rceil}$ rather than $(S^{i,0})_{1\leq i\leq N}$, where $\lceil . \rceil$ is the ceiling function. The right choice of the gap r between supremum observations is given by controlling the upper bound of an appropriate inequality. Indeed for a fixed dataset, creating a gap between the observations of the supremum removes a quantity of observations in the sample used. For $r = (k-1)\Delta$ and $k \in \mathbb{N}^\star$, we have:

$$\mathbb{E}\left[\left(\frac{1}{\lceil N/k\rceil}\log\mathcal{L}^r_{\lceil N/k\rceil} - \frac{1}{\lceil N/k\rceil}\mathbb{E}\left[\log\mathcal{L}^r_{\lceil N/k\rceil}\right]\right)^2\right] \leq$$
$$\mathrm{Var}(\log(f_\Delta(S^{1,r},\theta)))\left(\frac{1}{\lceil N/k\rceil} + \frac{2}{\lceil N/k\rceil^2}\sum_{1\leq i<j\leq\lceil N/k\rceil}\mathrm{Cov}(\log(f_\Delta(S^{i,r},\theta),\log(f_\Delta(S^{j,r},\theta))\right). \tag{2.4.1}$$

The inequality (2.4.1) bounds the variance of the pseudo-likelihood. As $(S^{i,r})_{i\geq 1}$ is identically distributed, the quantity $\frac{1}{\lceil N/k\rceil}\mathbb{E}\left[\log\mathcal{L}^r_{\lceil N/k\rceil}\right]$ does not depend on k and N. For $N \in \mathbb{N}^\star$ observations and a time window of size Δ, there exists an r^* which minimizes the upper bound of this inequality.

Proposition 2.4.1. *Consider an OU process solution of* (2.1.1) *with parameter* $\theta_0 = (a_0, b_0, \beta_0)$. *The optimal upper bound of* (2.4.1) *is reached for* $r^* = (k^*-1)\Delta$ *with*

$$k^* = \underset{1\leq k\leq N}{Argmin}\ g(\Delta, N, k, \theta_0),$$

and

$$g(\Delta, N, k, \theta_0) = C\left(\frac{1}{\lceil N/k\rceil} + \frac{2}{\lceil N/k\rceil^2}\frac{e^{b_0\Delta}}{e^{b_0k\Delta}-1}\left[\lceil N/k\rceil + e^{b_0k\Delta}\left(\frac{e^{-b_0k\lceil N/k\rceil\Delta}-1}{e^{b_0\Delta k}-1}\right)\right]\right)$$

and C *a strictly positive constant.*

Proof. Using inequality (2.4.1) and the ρ-mixing property on the sequence $(\log(f_\Delta(S^{i,r},\theta)))_{1\leq i\leq\lceil N/k\rceil}$, we obtain the following inequality:

$$\mathbb{E}\left[\left(\frac{1}{\lceil N/k\rceil}\log\mathcal{L}^r_{\lceil N/k\rceil} - \frac{1}{\lceil N/k\rceil}\mathbb{E}\left[\log\mathcal{L}^r_{\lceil N/k\rceil}\right]\right)^2\right] \leq g(\Delta, N, k, \theta_0).$$

To obtain the optimal $r = (k-1)\Delta$, it is enough to minimize the function g:

$$k^* = \underset{1\leq k\leq N}{\mathrm{Argmin}}\ g(\Delta, N, k, \theta_0).$$

□

The appropriate gap for the estimation is $r^\star = (k^\star - 1)\Delta$.

Remark 2.4.2. *In the case where the minimal argument is not unique, we choose the largest one. As we will have fewer observations when $k^\star$ is larger, the minimisation algorithm will be numerically less expensive. Moreover, the larger $k^\star$, the greater the observations are decorrelated. Thus, the pseudo-likelihood will be more suitable for our model.*

Remark 2.4.3. *When we deal with a real dataset, we do not know the value of the parameter θ_0. To find $r^\star$, we first need to compute some bounds, see an example in [15].*

2.4.2 Numerical simulation

Numerical issues emerge from the μ-zeros of $D_\mu(c)$ involved in the probability density. For large values of $|c|$, the μ-zeros are no longer computable through dichotomy. We use the asymptotic expansions (2.6.8) and (2.6.9) to evaluate the density for large values of $|c|$. In the pseudo-likelihood maximization, a multi-start method with 10 repeats is also used to reduce the instability.

- Simulated data:

We simulate a stationary OU process with parameter $\theta_{0,1} = (a_{0,1}, b_{0,1}, \beta_{0,1}) = (0, 1, 1)$ using an Euler scheme with $T = 10^3$ and $dt = 10^{-3}$. We denote (N, r), the set of numerical parameters, with N the number of suprema observations and r the gap between these observations. We apply our estimation method for $\Delta = 1$. Repeating this process 100 times, we obtain a sample of our estimator. We also performed the estimation method for the parameter $\theta_{0,2} = (a_{0,2}, b_{0,2}, \beta_{0,2}) = (20.9, 0.95, 47.5)$. From Proposition 2.4.1, for $\theta_{0,1}$ and $\theta_{0,2}$ with $N = 1000$, we have $r^\star = 1$. Simulations will be carried out for three different sets of numerical parameters, $(1000, 0)$, $(500, 1)$ and $(250, 3)$. Each estimation will be compared using the relative root-mean-square error (RMSE) and the mean-error (ME).

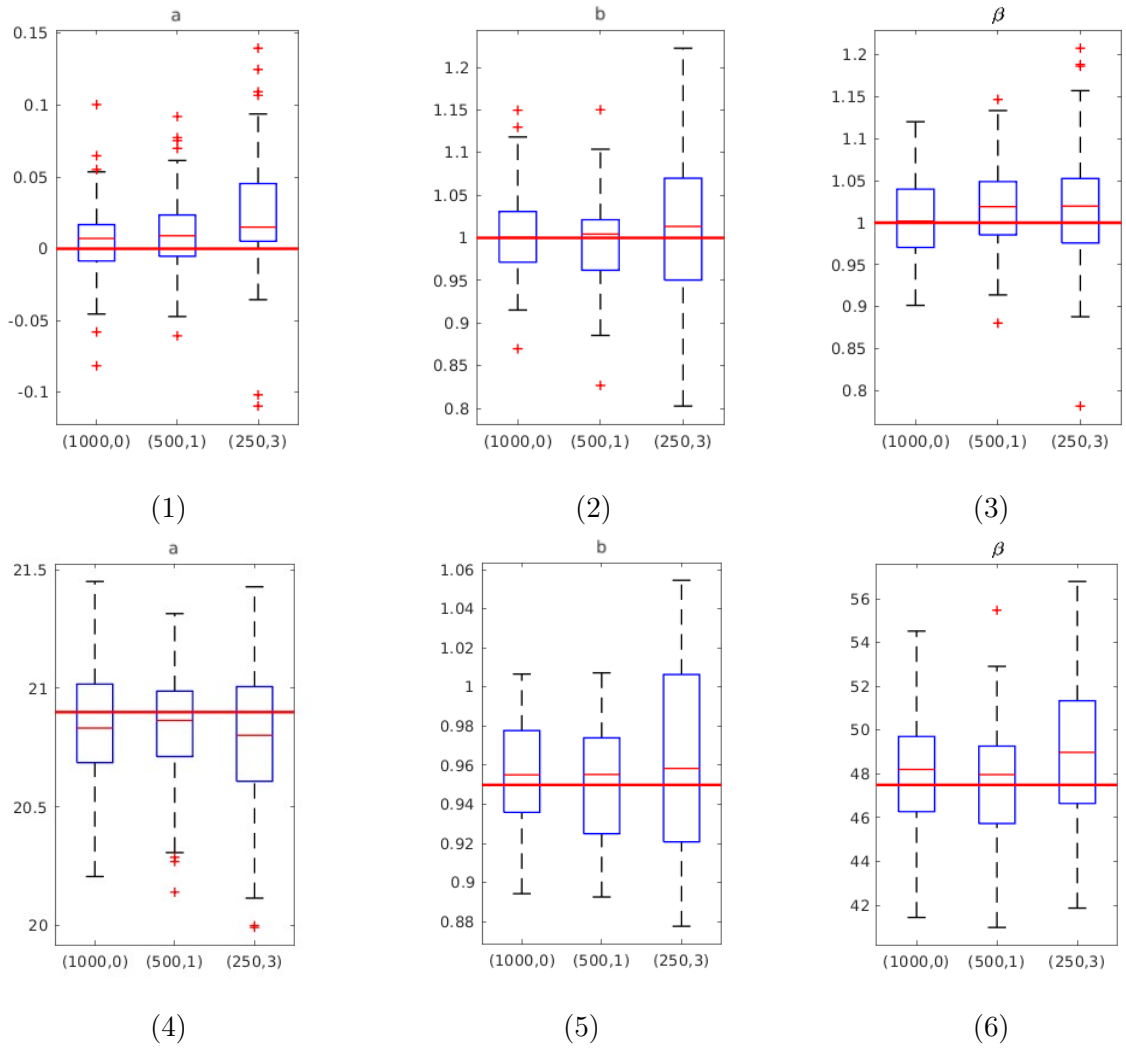


Figure 2.1: Boxplots of the estimated parameters: (1), (2), (3) $\theta_{0,1} = (a_{0,1}, b_{0,1}, \beta_{0,1}) = (0, 1, 1)$; (4), (5), (6) $\theta_{0,2} = (a_{0,2}, b_{0,2}, \beta_{0,2}) = (20.9, 0.95, 47.5)$. The red line corresponds to the theoretical value of the parameters.

Numerical parameters	Relative RMSE	ME
(250,3)	(0.0436, 0.0855, 0.0731)	(0.0240, 0.0114, 0.0201)
(500,1)	(0.0265, 0.0472, 0.0572)	(0.0081, -0.0048, 0.0198)
(1000,0)	(0.0279, 0.0507, 0.0473)	(0.0033, 0.0044, 0.0032)

Table 2.1: Table of the relative RMSE and ME for the estimator of $\theta_{0,1} = (a_{0,1}, b_{0,1}, \beta_{0,1}) = (0, 1, 1)$ with different numerical parameters.

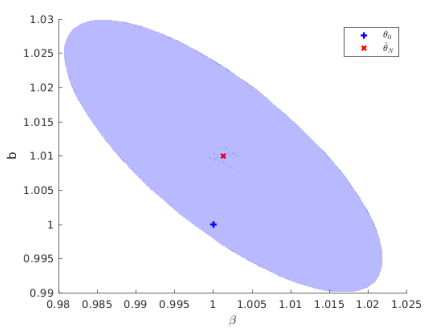
Numerical parameters	Relative RMSE	ME
(250,3)	(0.0154, 0.0615, 0.0701)	(-0.1072, -0.0063, -0.6946)
(500,1)	(0.0109, 0.0351 , 0.0557)	(-0.0538, -0.0096, -1.0821)
(1000,0)	(0.0113, 0.0348 , 0.0578)	(-0.0482, -0.0074, -1.0693)

Table 2.2: Table of the relative RMSE and ME for the estimator of $\theta_{0,2} = (a_{0,2}, b_{0,2}, \beta_{0,2}) = (20.9, 0.95, 47.5)$ with different numerical parameters.

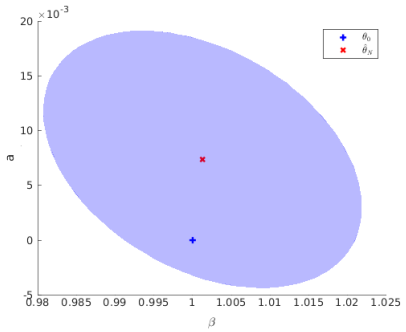
Relative RMSE are small enough and validate the results on the trade off between the number of observations and the time gap r.

The overestimation/bias on the β estimator comes from the decrease of $\beta \mapsto \mathcal{L}_N^r(a, b, \beta)$. As β becomes greater than b, the pseudo-likelihood function becomes flat in β. Consequently, the β estimator will have a big variation, which slowly decreases as the number of observations increases. Better results can be obtained by fixing β and performing the estimation method on parameters a and b (2D-estimation).

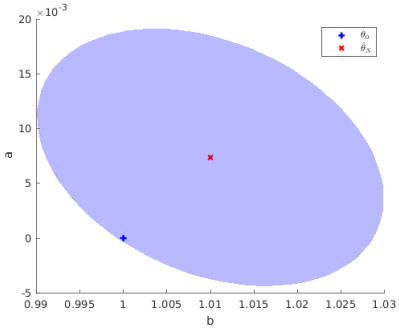
Using Theorem 2.3.3 in [93], we look at the probability that $\theta_{0,1}$ and $\theta_{0,2}$, fall into the 95% confidence ellispoïd for the set of numerical parameters $(500, 1)$.



(a) βb plane



(b) βa plane



(c) ba plane

Figure 2.2: Cut planes of the 95% confidence ellispoïd associated to the estimator of $\theta_{0,1} = (0, 1, 1)$ and the set of numerical parameters $(N, r) = (500, 1)$.

As a consequence of a small variance and a relatively high bias on the β estimator, the parameters $\theta_{0,1}$ and $\theta_{0,2}$ fall infrequently into the 95% confidence ellispoïd. For the 2D-estimation problem (β fixed), both parameters fall into the 95% confidence ellipse with a $\approx 93\%$ probability for the same set of numerical parameters.

- Weather data:

In [58], we can find the daily temperatures of Paris. This dataset is one of the longest we can find, as it started in 1900. In this dataset, we find the daily maximum and minimum temperature measurements as well as the average daily temperature. We choose to focus this study on the maximum temperatures from 15th of June to 14th of August from 1950 to 1984 (2135 days), using a gap of one day between each daily extreme values.

A multi-start method is also used for the estimation. We obtain the following estimates: $(a, b, \beta) = (18.0866, 0.9510, 36.0201)$.

We use the same approches as [15], to compare the two different estimation methods. The first validating method is the prediction. We take the mean temperature of 14/06/1985 as the starting point for the simulation of our 10-day process and we make confidence intervals on 1000 simulations of the maximum temperatures for each of these days and compare them with the true temperature values (between 15/06/1985 and 24/06/1985). The second method compares the theoretical quantiles with the empirical quantiles.

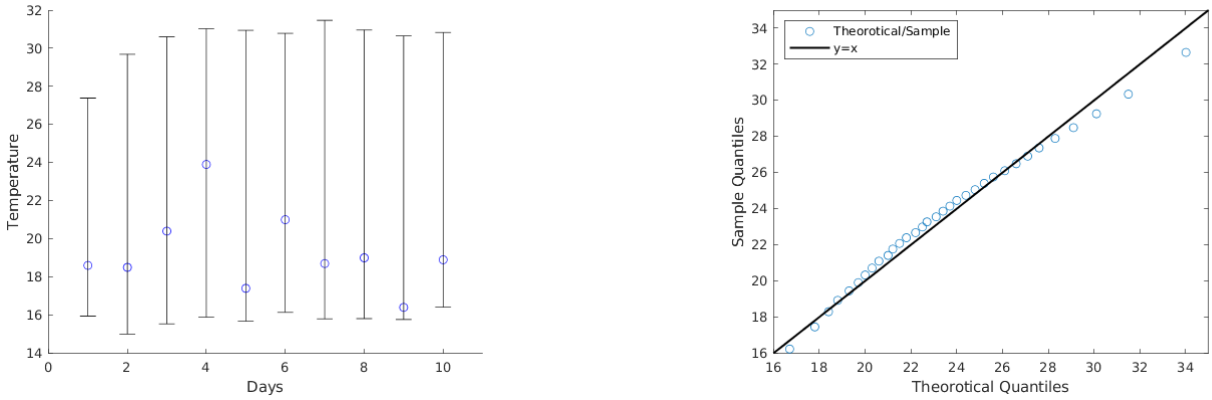


Figure 2.3: 95% confidence interval for daily extreme temperatures between 15/06/1985 and 24/06/1985, and Quantile-Quantile Graph.

The results are slightly better as [15] for the QQ-plot and the prediction method.

Our method is more efficient in computation time and in accuracy. In [15], the evaluation of the cumulative distribution function required a very expensive Monte Carlo method. Their estimation procedure took a week for the computation of the estimator, our method takes only a few minutes on the same machine.

2.5 Proofs

This section is dedicated to the proofs of Proposition 2.3.2, Theorem 2.3.3 and Theorem 2.3.4.

2.5.1 Proof of Proposition 2.3.2

Identifiability of the statistical model $\mathcal{P}$ depends on the injectivity of the supremum law. For the sake of proof, we denote $(X_t^\theta)_{t\geq 0}$ a stationary OU process with parameter $\theta = (a, b, \beta)$ and $S_t^\theta = \sup_{s\leq t} X_s^\theta$. Using (2.6.5), the cdf of the supremum (2.2.4) can be rewritten:

$$\mathbb{P}(S_t^\theta < c) = \frac{1}{2\sqrt{\pi}} e^{-(c-\frac{a}{b})^2 \frac{b}{\beta}} \sum_{n\geq 1} e^{-b\mu_{n,c,\theta} t} \frac{D^2_{\mu_{n,c,\theta}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\int_{-\infty}^{(c-\frac{a}{b})\sqrt{\frac{b}{\beta}}} D^2_{\mu_{n,c,\theta}}\left(-x\sqrt{2}\right) dx}. \tag{2.5.1}$$

We prove the injectivity of the measure $\mathbb{P}_\theta$ associated with the random variable S_t^θ for t fixed. We suppose, by absurd, that there exists $\theta_1 = (a_1, b_1, \beta_1)$ and $\theta_2 = (a_2, b_2, \beta_2)$ such that $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2}$ *i.e.*:

$$\forall c \in \mathbb{R}, \quad \mathbb{P}(S_t^{\theta_1} < c) = \mathbb{P}(S_t^{\theta_2} < c). \tag{2.5.2}$$

In particular; the equality is satisfied for $c \to \infty$. Using the asymptotic expansion (2.6.13) with the three parameters variable change, we get:

$$1 - \frac{e^{-(c-\frac{a_1}{b_1})^2 \frac{b_1}{\beta_1}}}{2\sqrt{\pi}\left(c-\frac{a_1}{b_1}\right)\sqrt{\frac{b_1}{\beta_1}}}(1+o(c^{-2+\delta})) = 1 - \frac{e^{-(c-\frac{a_2}{b_2})^2 \frac{b_2}{\beta_2}}}{2\sqrt{\pi}\left(c-\frac{a_2}{b_2}\right)\sqrt{\frac{b_2}{\beta_2}}}(1+o(c^{-2+\delta})),$$

with $0 < \delta < 2$. Therefore

$$\frac{\left(c-\frac{a_2}{b_2}\right)\sqrt{\frac{b_2}{\beta_2}}}{\left(c-\frac{a_1}{b_1}\right)\sqrt{\frac{b_1}{\beta_1}}} e^{-(c-\frac{a_1}{b_1})^2 \frac{b_1}{\beta_1} + (c-\frac{a_2}{b_2})^2 \frac{b_2}{\beta_2}} = 1 + o(c^{-2+\delta})$$

$$\iff \lim_{c\to+\infty} c^{2-\delta} \left[\frac{\left(c-\frac{a_2}{b_2}\right)\sqrt{\frac{b_2}{\beta_2}}}{\left(c-\frac{a_1}{b_1}\right)\sqrt{\frac{b_1}{\beta_1}}} e^{-c^2\left(\frac{b_1}{\beta_1}-\frac{b_2}{\beta_2}\right)+2c\left(\frac{a_1}{\beta_1}-\frac{a_2}{\beta_2}\right)+\frac{a_1^2}{b_1\beta_1}-\frac{a_2^2}{b_2\beta_2}} - 1 \right] = 0.$$

We deduce that $\frac{b_1}{\beta_1} = \frac{b_2}{\beta_2} = \frac{b}{\beta}$ and $\frac{a_1}{\beta_1} = \frac{a_2}{\beta_2} = \frac{a}{\beta}$ and thus $\mu_{n,c,\theta_1} = \mu_{n,c,\theta_2} = \mu_{n,c}$. Making use of (2.5.2) and (2.5.1), we have:

$$\sum_{n\geq 1} \left(e^{-b_1\mu_{n,c}t} - e^{-b_2\mu_{n,c}t}\right) \frac{D^2_{\mu_{n,c}-1}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\int_{-\infty}^{(c-\frac{a}{b})\sqrt{\frac{b}{\beta}}} D^2_{\mu_{n,c}}\left(-x\sqrt{2}\right) dx} = 0.$$

Since $\frac{D^2_{\mu_{n,c-1}}\left(-(c-\frac{a}{b})\sqrt{\frac{2b}{\beta}}\right)}{\int_{-\infty}^{(c-\frac{a}{b})\sqrt{\frac{b}{\beta}}} D^2_{\mu_{n,c}}\left(-x\sqrt{2}\right)dx} > 0$, then the previous sum is strictly positive if $b_1 < b_2$ (resp. strictly negative if $b_1 > b_2$) because all the terms are strictly positive (resp. strictly negative). We deduce that $b_1 = b_2$, so $\beta_1 = \beta_2$ and $a_1 = a_2$.

Therefore the measure $\mathbb{P}_\theta$ associated with the variable S_t^θ is identifiable.

2.5.2 Proof of Theorem 2.3.3

Thereafter, without loss of generality, we suppose $\theta = (0, b, 1)$. In the case of three parameters, the arguments are the same. In the following we write b instead of θ and Θ is then a compact subset of $\mathbb{R}_+^\star$. To prove the consistency of the estimator $\hat{\theta}_N$, we adapt the proof of Corollary 3.2.9 in [33]. We first prove some regularity properties on the density f_Δ with respect to the parameter b.

Lemma 2.5.1. *Let X be an OU process solution of (2.1.1), with parameter $b_0 \in \Theta$, the following conditions are satisfied:*

1. *For all $c \in \mathbb{R}$, $b \mapsto \log(f_\Delta(c, b))$ is continuous on Θ.*
2. *For all $b \in \Theta$, there exists a neighborhood V of b and $G \in L^1(\mathbb{P}_{b_0})$ such that:*

$$\sup_{\eta \in V} |\log(f_\Delta(., \eta)| \leq G, \tag{2.5.3}$$

with f_Δ the probability density function of the supremum.

Proof. 1. Continuity of $b \mapsto \log(f_\Delta(c, b))$ for all $c \in \mathbb{R}$ is proved using Proposition 2.6.1 in Section 2.6.

2. Using Corollary 2.6.5 and Corollary 2.6.9, there exists $C > 0$, $G_1 \in \mathrm{L}^1(\mathbb{P}_{b_0})$ and $G_2 \in \mathrm{L}^1(\mathbb{P}_{b_0})$ such that for all η in a neighborhood V of $b \in \Theta$:

$$\forall c \leq -C, \quad |\log(f_\Delta(c, \eta))| \leq G_1(c),$$

$$\forall c \geq C, \quad |\log(f_\Delta(c, \eta))| \leq G_2(c).$$

As $(c, b) \mapsto \log(f_\Delta(c, b))$ is continuous on $\mathbb{R} \times \mathbb{R}_+^\star$, we have:

$$\forall (c, \eta) \in [-C, C] \times V, \quad |\log(f_\Delta(c, \eta))| \leq K,$$

with $K > 0$. Then, for all $\eta \in V$ and $c \in \mathbb{R}$:

$$\sup_{\eta \in V} |\log(f_\Delta(c, \eta)| \leq \underbrace{G_1(c)\mathbb{1}_{]-\infty,-C[}(c) + K\mathbb{1}_{[-C,C]}(c) + G_2(c)\mathbb{1}_{]C,+\infty[}(c)}_{G(c)}.$$

We easily prove that $G \in \mathrm{L}^1(\mathbb{P}_{b_0})$ and the conclusion holds.

□

Combining the results of the previous lemma with Proposition 2.3.2 and the Ergodic Theorem (see [14] e.g.), we can prove Theorem 2.3.3. More precisely, for a sample $(S^{1,r}, \cdots, S^{N,r})$ with $N \in \mathbb{N}^\star$, the maximum pseudo-likelihood estimator is the maximal argument of the function:

$$M_N(b) = \frac{1}{N} \sum_{i=1}^{N} \log \left(\frac{f_\Delta(S^{i,r}, b)}{f_\Delta(S^{i,r}, b_0)} \right).$$

For all $b \in \Theta$, the strict concavity of the logarithm function and Jensen's inequality, give:

$$\mathbb{E}_{b_0} \left[\log \left(\frac{f_\Delta(S^{i,r}, b)}{f_\Delta(S^{i,r}, b_0)} \right) \right] < 0. \tag{2.5.4}$$

According to the Ergodic Theorem, the following convergence holds:

$$M_N(b) \xrightarrow[N \to +\infty]{\mathbb{P}_{b_0}\text{-a.s.}} M(b) = \mathbb{E}_{b_0} \left[\log \left(\frac{f_\Delta(., b)}{f_\Delta(., b_0)} \right) \right]. \tag{2.5.5}$$

Using Lemma 2.5.1, we prove that the convergence in Equation (2.5.5) is uniform.

To conclude, $M(b)$ is the Kullback-Leibler divergence and according to Proposition 2.3.2, it reaches its maximum for $b = b_0$, proving that the estimator is consistent.

2.5.3 Proof of Theorem 2.3.4

Asymptotic normality is a basic property used in statistics, details on the classical proof can be found in [105]. We mainly use the ρ-mixing property to obtain the asymptotic normality of the estimator. As in the previous section, we consider only the case $\theta = (0, b, 1)$, the three parameters case is obtained by similar reasoning. We introduce the following notations:

$$l_b(c) = \partial_b \log(f_\Delta(c, b)) \quad , \quad \dot{l}_b(c) = \partial_b^2 \log(f_\Delta(c, b)) \quad \text{and} \quad \ddot{l}_b(c) = \partial_b^3 \log(f_\Delta(c, b)),$$

$$\psi_N(b) = \frac{1}{N} \sum_{i=1}^{N} l_b(S^{i,r}) \quad , \quad \dot{\psi}_N(b) = \frac{1}{N} \sum_{i=1}^{N} \dot{l}_b(S^{i,r}) \quad \text{and} \quad \ddot{\psi}_N(b) = \frac{1}{N} \sum_{i=1}^{N} \ddot{l}_b(S^{i,r}).$$

To obtain the asymptotic normality, the first step is to perform a Taylor expansion of the function $\psi(\hat{b}_N)$ around b_0:

$$\sqrt{N}(\hat{b}_N - b_0) = \frac{-\sqrt{N}\psi_N(b_0)}{\dot{\psi}_N(b_0) + \frac{1}{2}(\hat{b}_N - b_0)\ddot{\psi}_N(\bar{b}_N)}, \tag{2.5.6}$$

with $\bar{b}_N$ a point located between $\hat{b}_N$ and b_0.

Lemma 2.5.2. *For $r > \frac{ln\left(\frac{5}{3}\right)}{b_0}$, the following convergence is satisfied:*

$$\sqrt{N}\psi_N(b_0) \xrightarrow[N \to +\infty]{law} \mathcal{N}\left(0, \mathbb{E}_{b_0}\left(l_{b_0}(.)^2\right)\right).$$

Proof. Convergence is obtained using a Central Limit Theorem for ρ-mixing sequence [39, Theorem 2]. Three conditions need to be checked to apply this theorem. The first condition is a direct consequence of the ρ-mixing property of $(S^{i,r})_{i\geq 1}$. The second one is satisfied from the asymptotic expansions and the smoothness of the cumulative distribution given in Section 2.6.
Concerning the last condition, it is necessary to check that $\lim_{N\to+\infty} \mathbb{V}\left(\sum_{i=1}^{N} l_{b_0}(S^{i,r})\right) = +\infty$. We notice that, by inclusion of the sigma-fields, the sequence $(l_{b_0}(S^{i,r}))_{i\geq 1}$ is ρ-mixing. Moreover, according to Lemma 2.5.4 below, for any $i \geq 1$ the random variable $l_{b_0}(S^{i,r})$ is centered. Using Proposition 1.5.1 in [39], we have:

$$\mathbb{V}\left(\sum_{i=1}^{N} l_{b_0}(S^{i,r})\right) = N\tilde{\sigma}^2 + O(1), \tag{2.5.7}$$

with $\tilde{\sigma}^2 = \mathbb{E}[l_{b_0}(S^{1,r})^2] + 2\sum_{i=1}^{+\infty} \mathbb{E}[l_{b_0}(S^{1,r}) l_{b_0}(S^{i,r})]$.
Using the ρ-mixing property, $\tilde{\sigma}^2 \geq \mathbb{E}[l_{b_0}(S^{i,r})^2]\frac{3-5e^{-b_0 r}}{1-e^{-b_0 r}}$. By assumption $r > \frac{\ln\left(\frac{5}{3}\right)}{b_0}$, then $\tilde{\sigma}^2 > 0$. So, according to (2.5.7):

$$\lim_{N\to+\infty} \mathbb{V}\left(\sum_{i=1}^{N} l_{b_0}(S^{i,r})\right) = +\infty.$$

Since all the conditions are satisfied, we can apply Theorem 2 in [39] and we obtain:

$$\sqrt{N}\psi_N(b_0) \xrightarrow[N\to+\infty]{law} \mathcal{N}\left(0, \mathbb{E}_{b_0}\left(l_{b_0}(.)^2\right)\right).$$

□

Remark 2.5.3. *The condition $r > \frac{ln\left(\frac{5}{3}\right)}{b_0}$ remains the same in the three parameters case.*

The required conditions on the other quantities involved in the Taylor expansion are listed in the following lemma:

Lemma 2.5.4. *For all $b \in \Theta$, the function $b \mapsto l_b(c)$ is twice continuously derivable for all $c \in \mathbb{R}$. Moreover, $\mathbb{E}_{\mathbb{P}_{b_0}} |l_{b_0}(.)| < \infty$ and $\mathbb{E}_{\mathbb{P}_{b_0}} \left|\dot{l}_{b_0}(.)\right|$ exists and is non-zero. The functions l_{b_0} and $\ddot{l}_{b_0}$ are respectively dominated by $G(.)$ and $F(.)$ in $L^1(\mathbb{P}_{b_0})$ for all b in a neighborhood of b_0.*

Proof. Integrability and differentiability conditions are satisfied using the asymptotic expansions and the smoothness of the cumulative distribution in Section 2.6. Domination conditions are verified using the same reasoning as the one in the proof of Lemma 2.5.1. □

Combining Lemmas 2.5.2, 2.5.4 and Theorem 2.3.3, we have the following convergence:

$$\sqrt{N}(\hat{b}_N - b_0) \xrightarrow[N\to+\infty]{law} \mathcal{N}(0, I_{b_0}^{-1}), \tag{2.5.8}$$

where I_{b_0} is the Fisher information.

2.6 Appendix

2.6.1 Parabolic Cylinder function

We recall the definition and some auxiliary results about the Parabolic Cylinder function. Some of these results can be found in [16, 68].

For all $x, \mu \in \mathbb{R}$, the Parabolic Cylinder function $D_\mu(x)$ is a solution of the differential equation:

$$\begin{cases} y''(z) + \left(\mu + \frac{1}{2} - \frac{1}{4}z^2\right) y(z) = 0, \\ y(z) \underset{z\to+\infty}{\sim} z^\mu e^{-z^2/4}, \end{cases} \tag{2.6.1}$$

Moreover for $\mu \in \mathbb{C}$ with $\mathrm{Re}(\mu) > 0$ and $z \in \mathbb{C}$, the function $(z, \mu) \mapsto D_\mu(z)$ is a holomorphic function ([68], Chapter 10).

The Parabolic Cylinder function satisfies the following relations:

$$\partial_x D_\mu(x) = \mu D_{\mu-1}(x) - \frac{x}{2} D_\mu(x). \tag{2.6.2}$$

For all $\mu, c \in \mathbb{R}$,

$$\int_{-\infty}^{c} e^{-\frac{x^2}{2}} D_\mu(-x\sqrt{2})dx = \frac{e^{-\frac{c^2}{2}}}{\sqrt{2}} D_{\mu-1}(-c\sqrt{2}). \tag{2.6.3}$$

and, from [68] p286:

$$D_\mu(x\sqrt{2})D_{\mu-1}(-x\sqrt{2}) + D_\mu(-x\sqrt{2})D_{\mu-1}(x\sqrt{2}) = \frac{\sqrt{2\pi}}{\Gamma(1-\mu)}. \tag{2.6.4}$$

We recall now some properties on the μ-zeros of the Parabolic Cylinder function. Let $n \in \mathbb{N}^\star$, we denote $\mu_{n,c}$, the positive (ordered) zeros of the function $\mu \mapsto D_\mu\left(-c\sqrt{2}\right)$. Then, thanks to [16] (Proposition 3.14), we have:

$$\int_{-\infty}^{c} D^2_{\mu_{n,c}}(-x\sqrt{2})dx = -\frac{\mu_{n,c}}{\sqrt{2}} D_{\mu_{n,c}-1}(-c\sqrt{2})\partial_\mu D_{\mu_{n,c}}(-c\sqrt{2}). \tag{2.6.5}$$

According to Proposition 3.2.1 and Proposition 3.3.2 in Chapter 3, we have:

$$\partial_c \mu_{n,c} = \sqrt{2}\frac{\partial_x D_{\mu_{n,c}}(-c\sqrt{2})}{\partial_\mu D_{\mu_{n,c}}(-c\sqrt{2})} \tag{2.6.6}$$

$$= -\frac{2}{\sqrt{\pi}\int_0^\infty e^{-(2\mu_{n,c}+1)u + c^2\tanh(u)} \mathrm{erfc}\left(-c\sqrt{\tanh(u)}\right) \frac{du}{\sqrt{\sinh(u)\cosh(u)}}}, \tag{2.6.7}$$

and the following asymptotic expansion is verified:

$$\mu_{n,c} \underset{c\to-\infty}{=} \frac{c^2}{2} - \frac{1}{2} - |c|^{\frac{2}{3}} 2^{-\frac{1}{3}} a_n + O\left(|c|^{-\frac{2}{3}}\right), \tag{2.6.8}$$

where a_n, $n \in \mathbb{N}^\star$ are the zeros of the first kind Airy function.

Furthermore, the following convergence is verified:

$$\mu_{n,c} \underset{c\to+\infty}{\longrightarrow} n-1. \tag{2.6.9}$$

Recall that the zeros a_n of the first kind Airy function are all real, negative and satisfy the following inequality ([90]), for $n \geq 1$:

$$-\left(\frac{3\pi}{8}(4n-1)\right)^{\frac{2}{3}}\left(1+\frac{5}{48\left(\frac{3\pi}{8}(4n-1)\right)^2}\right) < a_n \leq -\left(\frac{3\pi}{8}(4n-1)\right)^{\frac{2}{3}}. \tag{2.6.10}$$

2.6.2 Smoothness of the cumulative distribution function

Some smoothness properties are needed to prove the set of derivations and continuities of the functions presented in the proof of consistency and asymptotic normality. These properties will be proved using results on holomorphic functions. Making use of (2.6.6), we can rewrite the cumulative distribution function as:

$$\mathbb{P}(S^{1,0} < c) = -\frac{e^{-bc^2}}{2\sqrt{b\pi}}\sum_{n\geq 1} e^{-b\mu_{n,c,b}\Delta}\frac{\partial_c \mu_{n,c,b}}{\mu_{n,c,b}^2}. \tag{2.6.11}$$

Proposition 2.6.1. *The cumulative distribution function of the supremum verifies the following properties:*

1. $(c,b) \mapsto \mathbb{P}(S^{1,0} < c)$ *is a smooth function on* $\mathbb{R}\times\mathbb{R}_+^\star$.
2. *For any* $k, j \in \mathbb{N}^\star$*:*

$$\partial_b^k\partial_c^j \mathbb{P}(S^{1,0} < c) = \sum_{n\geq 1}\partial_b^k\partial_c^j f_n(\Delta, c, b),$$

with

$$f_n(\Delta,c,b) = -\frac{e^{-bc^2-b\mu_{n,c,b}\Delta}}{\sqrt{2\pi}}\frac{D_{\mu_{n,c,b}-1}\left(-c\sqrt{2b}\right)}{\mu_{n,c,b}\partial_\mu D_{\mu_{n,c,b}}\left(-c\sqrt{2b}\right)}. \tag{2.6.12}$$

Proof. We introduce the following notation, $\mathbb{C}_{C_1\leq|.|\leq C_2}^{\theta_1\leq\arg\leq\theta_2} = \{z \in \mathbb{C} \text{ s.t. } \theta_1 \leq \arg(z) \leq \theta_2\ , C_1 \leq |z| \leq C_2\}$. We denote $(z,b) \mapsto \tilde{F}_\Delta(z,b)$ the continuation of the cumulative distribution function on $\mathbb{C}\times\mathbb{C}_{|.|\neq 0}^{-\frac{\pi}{2}<\arg<\frac{\pi}{2}}$. According to the Implicit Function Theorem, the function $(z,b) \mapsto \mu_{n,z,b}$ is holomorphic. Then by composition of holomorphic functions, we deduce that (2.6.12) is holomorphic.

We can write $\tilde{F}_\Delta$ as follows:

$$\begin{aligned}\tilde{F}_\Delta(z,b) = {} & \tilde{G}_1(z,b)\mathbb{1}_{\mathbb{C}_{|.|>C}^{\frac{\pi}{2}<arg<\frac{3}{2\pi}}\times\mathbb{C}_{|.|>C}^{-\frac{\pi}{2}\leq\arg<\frac{\pi}{2}}}(z,b) + \tilde{G}_2(z,b)\mathbb{1}_{\mathbb{C}_{|.|<C}\times\mathbb{C}_{0<|.|<C}^{-\frac{\pi}{2}\leq arg<\frac{\pi}{2}}}(z,b) \\ & + \tilde{G}_3(z,b)\mathbb{1}_{\mathbb{C}_{|.|>C}^{-\frac{\pi}{2}<arg<\frac{\pi}{2}}\times\mathbb{C}_{|.|>C}^{-\frac{\pi}{2}\leq\arg<\frac{\pi}{2}}}(z,b),\end{aligned}$$

with C large enough.

For $\tilde{G}_1$ and $\tilde{G}_3$, using the asymptotic expansion (2.6.8) and the limit (2.6.9), we easily prove the normal convergence of the associate series.

For $\tilde{G}_2$, the normal convergence can be obtained using the following equivalence from [6]:

$$\mu_{n,z,b} \underset{n\to+\infty}{\sim} 2n-1+\frac{4bz^2}{\pi^2}-2\frac{\sqrt{bz^2}}{\pi}\sqrt{4n-1+4\frac{bz^2}{\pi^2}}$$

and

$$D_\mu(z) \underset{\mu\to+\infty}{=} \sqrt{2}\left(\mu+\frac{1}{2}\right)^{\frac{\mu}{2}} e^{-\left(\frac{\mu}{2}+\frac{1}{4}\right)} \cos\left(z\sqrt{\mu+\frac{1}{2}}-\frac{\pi\mu}{2}\right)\left(1+O(\mu^{-\frac{1}{2}})\right).$$

From Theorem 3.2 in [88], the conclusion holds. □

2.6.3 Asymptotic expansions

For the integrability conditions required in the Ergodic Theorem, some asymptotic expansions on the cumulative distribution and the probability density of the supremum $S^{1,0}$ are provided. In the following proofs, without loss of generality we assume $\theta=(0,b,1)$. To return to the three parameters case, we replace c, Δ, b respectively by $c-\frac{a}{b}$, $\beta\Delta$ and $\frac{b}{\beta}$.

For large positive c

Since the zeros $\mu_{n,c,b}$ tend to positive integers when c goes to infinity, then we are able to give an asymptotic expansion for (2.5.1).

Proposition 2.6.2. *For large positive c, the cumulative distribution function of $S^{1,0}$ has the following asymptotic expansion:*

$$\mathbb{P}(S^{1,0}<c) \underset{c\to+\infty}{=} 1-\frac{e^{-bc^2}}{2\sqrt{\pi bc}}\left(1+o(c^{-2+\delta})\right), \tag{2.6.13}$$

with $0<\delta<2$.

Proof. Recall that the cumulative distribution function is given by (2.5.1). Using Formula (2.6.9), we obtain:

$$\mathbb{P}(S^{1,0}<c) \underset{c\to+\infty}{=} \frac{1}{2\sqrt{\pi}} e^{-bc^2}\left[\frac{D_{-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_0^2(-x\sqrt{2})dx}+\sum_{n\geq 1} e^{-bn\Delta}\frac{D_{n-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_n^2(-x\sqrt{2})dx}\right]. \tag{2.6.14}$$

According to Formula (10.5.4) in [68]:

$$\frac{e^{-bc^2}}{2\sqrt{\pi}}\frac{D_{-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_0^2(-x\sqrt{2})dx}=\frac{1}{2}\left(1+erf(c\sqrt{b})\right),$$

where $erf(c) = \frac{2}{\sqrt{\pi}} \int_0^c e^{-t^2} dt$ is the Error function. Then applying Formulas (4.9.6) and (4.13.4) in [68], we get:

$$\sum_{n=1}^{\infty} e^{-bn\Delta} \frac{D_{n-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_n^2(-x\sqrt{2b})dx} \underset{c\to+\infty}{=} e^{-bc^2 \frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1).$$

As we combine these two results, we get:

$$\mathbb{P}(S^{1,0} < c) \underset{c\to+\infty}{=} \frac{1}{2}\left(1 + erf(c\sqrt{b})\right) + e^{-bc^2 \frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1).$$

Since $1 - erf(c) \underset{c\to+\infty}{=} \frac{e^{-c^2}}{\sqrt{\pi}c}(1 + O(c^{-2}))$, the conclusion holds. □

Proposition 2.6.3. *For large positive c, the asymptotic expansion of the probability density of $S^{1,0}$ is given by:*

$$f_\Delta(c,b) \underset{c\to+\infty}{=} \sqrt{\frac{b}{\pi}} e^{-bc^2} \left(1 + ce^{-bc^2 \frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1)\right), \tag{2.6.15}$$

with $|O(1)| \leq \frac{4\sqrt{b}e^{-b\Delta}}{\sqrt{\pi(1-e^{-2b\Delta})}}$.

Proof. When c goes to infinity, using (2.6.14) the derivative of the cumulative distribution function satisfies:

$$f_\Delta(c,b) \underset{c\to+\infty}{=} \sqrt{\frac{b}{\pi}} e^{-bc^2} + \frac{1}{2\sqrt{\pi}} \partial_c \left[e^{-bc^2} \sum_{n\geq 1} e^{-bn\Delta} \frac{D_{n-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_n^2(-x\sqrt{2})dx} \right].$$

Using Formulas (4.9.5) and (4.13.4) in [68], one can prove that:

$$\partial_c \left[e^{-bc^2} \sum_{n\geq 1} e^{-bn\Delta} \frac{D_{n-1}^2(-c\sqrt{2b})}{\int_{-\infty}^{c\sqrt{b}} D_n^2(-x\sqrt{2})dx} \right] \underset{c\to+\infty}{=} ce^{-\frac{2bc^2}{1+e^{-b\Delta}}} O(1),$$

with $|O(1)| \leq \frac{8be^{-b\Delta}}{\sqrt{\pi(1-e^{-2b\Delta})}}$. □

Remark 2.6.4. *We have:*

$$f_\Delta(c,b) \underset{c\to+\infty}{=} e^{-bc^2} \sqrt{\frac{b}{\pi}} \left(1 + o(c^{-\alpha})\right),$$

with $\alpha > 0$.

Corollary 2.6.5. *For large positive c,*

$$\log\left(f_\Delta(c,b)\right) \underset{c\to+\infty}{=} -bc^2 + \frac{1}{2}\log\left(\frac{b}{\pi}\right) + ce^{-c^2 b \frac{1-e^{-\Delta b}}{1+e^{-\Delta b}}} O(1), \tag{2.6.16}$$

with $|O(1)| \leq \frac{16\sqrt{b}e^{-b\Delta}}{3\sqrt{\pi(1-e^{-2b\Delta})}}$.

A similar reasoning as the one in the proof of Proposition 2.6.2 and Proprosition 2.6.3 may be applied to prove the following results:

Proposition 2.6.6. *For large positive c, the following asymptotic expansions are satisfied:*

1. $\partial_b \log\left(f_\Delta(c,b)\right) \underset{c\to+\infty}{=} -c^2 + \frac{1}{2b} + c^3 e^{-bc^2\frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1).$

2. $\partial_b^2 \log\left(f_\Delta(c,b)\right) \underset{c\to+\infty}{=} -\frac{1}{2b^2} + c^5 e^{-bc^2\frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1).$

3. $\partial_b^3 \log\left(f_\Delta(c,b)\right) \underset{c\to+\infty}{=} \frac{1}{b^3} + c^7 e^{-bc^2\frac{1-e^{-b\Delta}}{1+e^{-b\Delta}}} O(1).$

For large negative c

Using (2.6.7) and (2.6.8), we can give an asymptotic expansion for the cumulative distribution function of $S^{1,0}$ for large negative c.

Proposition 2.6.7. *For large negative c, the cumulative distribution function of $S^{1,0}$ has the following asymptotic expansion:*

$$\mathbb{P}(S^{1,0} < c) = 2|c\sqrt{b}|^{-3}\frac{e^{-bc^2-\left(\frac{c^2b}{2}-\frac{1}{2}\right)b\Delta+|c|^{\frac{2}{3}}b^{\frac{4}{3}}2^{-\frac{1}{3}}a_1\Delta}}{\sqrt{\pi}}(1+o(1)), \tag{2.6.17}$$

where a_1 is the first zero of the Airy function of the first kind.

Proof. Formula (2.6.11) gives:

$$\mathbb{P}\left(S^{1,0} < c\right) = -\frac{e^{-bc^2-b\Delta\mu_{1,c,b}}}{2\sqrt{\pi}}\frac{\partial_c\mu_{1,c,b}}{\sqrt{b}\mu_{1,c,b}^2}\left(1+\sum_{n\geq 2}e^{-b\Delta(\mu_{n,c,b}-\mu_{1,c,b})}\frac{\mu_{1,c,b}^2\partial_c\mu_{n,c,b}}{\mu_{n,c,b}^2\partial_c\mu_{1,c,b}}\right).$$

From the asymptotic expansion of μ-zeros for large negative c (2.6.7) and (2.6.8), it follows that:

$$\begin{aligned}
\frac{\partial_c\mu_{1,c,b}}{\sqrt{b}\mu_{1,c,b}^2} &\underset{c\to-\infty}{=} -4|c\sqrt{b}|^{-3}(1+o(1)),\\
\frac{\mu_{1,c,b}^2\partial_c\mu_{n,c,b}}{\mu_{n,c,b}^2\partial_c\mu_{1,c,b}} &\underset{c\to-\infty}{\longrightarrow} 1 \quad \text{if} \quad n < N(c,b) = \lfloor\frac{2bc^2}{3\pi}+\frac{1}{4}\rfloor + 1,\\
\frac{\mu_{1,c,b}^2\partial_c\mu_{n,c,b}}{\mu_{n,c,b}^2\partial_c\mu_{1,c,b}} &\underset{c\to-\infty,n\to\infty}{\longrightarrow} 0 \quad \text{if} \quad n > N(c,b).
\end{aligned}$$

Therefore for all $\epsilon > 1$, there exists $C > 0$ such that for all $c < -C$, we get:

$$\sum_{n\geq 2}e^{-b\Delta(\mu_{n,c,b}-\mu_{1,c,b})}\frac{\mu_{1,c,b}^2\partial_c\mu_{n,c,b}}{\mu_{n,c,b}^2\partial_c\mu_{1,c,b}} < \epsilon\sum_{n\geq 2}e^{-b\Delta(\mu_{n,c,b}-\mu_{1,c,b})}.$$

Using (2.6.10), we easily prove that for large negative c,

$$\sum_{n\geq 2} e^{-b\Delta(\mu_{n,c,b}-\mu_{1,c,b})} = e^{-b\Delta(\mu_{2,c,b}-\mu_{1,c,b})}O(1) = o(1).$$

We then conclude the proof with (2.6.8). □

Using similar arguments, we can prove the following result:

Proposition 2.6.8. *For large negative c, the asymptotic expansion of the probability density of $S^{1,0}$ is given by:*

$$f_\Delta(c,b) = 2|c|^{-2}\frac{e^{-bc^2-\left(\frac{c^2b}{2}-\frac{1}{2}\right)b\Delta+|c|^{\frac{2}{3}}b^{\frac{4}{3}}2^{-\frac{1}{3}}a_1\Delta}}{\sqrt{\pi}}(2+\Delta b)(1+O(|c|^{-\frac{2}{3}})). \tag{2.6.18}$$

Corollary 2.6.9. *For large negative c,*

$$\log f_\Delta(c,b) = -bc^2-\left(\frac{c^2b}{2}-\frac{1}{2}\right)b\Delta+|c|^{\frac{2}{3}}b^{\frac{4}{3}}2^{-\frac{1}{3}}a_1\Delta+\log\frac{2(2+\Delta b)}{\sqrt{\pi}}-2\log(|c|)+O(|c|^{-\frac{2}{3}}).$$

A similar reasoning as the one in the proof of Proposition 2.6.7 may be applied to prove the following:

Proposition 2.6.10. *For large negative c, the following asymptotic expansions are satisfied:*

1. $\partial_b \log(f_\Delta(c,b)) = -c^2 - \Delta\left(c^2b - \frac{1}{2} - |c|^{\frac{2}{3}}\frac{2^{\frac{2}{3}}b^{-\frac{2}{3}}}{3}a_1\right) + \frac{\Delta}{2+\Delta b} + |c|^{-\frac{2}{3}}O(1).$
2. $\partial_b^2 \log(f_\Delta(c,b)) = -\Delta\left(c^2 + |c|^{\frac{2}{3}}\frac{2^{\frac{5}{3}}b^{-\frac{5}{3}}}{9}a_1\right) - \frac{\Delta^2}{(2+\Delta b)^2} + |c|^{-\frac{2}{3}}O(1).$
3. $\partial_b^3 \log(f_\Delta(c,b)) = \Delta|c|^{\frac{2}{3}}\frac{5}{27}2^{\frac{5}{3}}b^{-\frac{8}{3}}a_1 + \frac{2\Delta^3}{(2+\Delta b)^3} + |c|^{-\frac{2}{3}}O(1).$

Chapter 3

Some properties for μ-zeros of Parabolic Cylinder functions

3.1 Introduction

Since the mid-twentieth century, real and complex zeros of special functions such as Bessel functions, Parabolic Cylinder functions, Hankel functions etc. have been intensively studied for various applications in physics, applied mathematics and engineering.

Studies on zeros for a special function of order μ and argument z have been performed by several authors. For example, Olver finds the z-zeros of Parabolic Cylinder functions [87] for large values of μ. The case of Bessel functions has been frequently studied (see for example Olver [86], Watson [109], Laforgia and Natalini [61]). In [40], the author presents a selection of results on the zeros of Bessel functions. Other authors have been interested in the z-zeros of Hermite functions or Confluent Hypergeometric functions (see for example Elbert and Muldoon [41], [42]). In [41], the authors study the variation of the z-zeros of the Hermite function and establish a formula for the derivative of a zero with respect to the parameter μ.

Fewer studies have been published on the μ-zeros. In [78] or [56], the authors study the behavior of the μ-zeros of the Hankel function of the first kind. Later on, these results were improved by Cochran [29]. Conde and Kalla [30] compute the μ-zeros of the Bessel function. Slater [94] gives an asymptotic formula for large μ-zeros of the Parabolic Cylinder function when z is fixed. Besides that, little is known about the μ-zeros of the Parabolic Cylinder function.

In this chapter we study the μ-zeros of the Parabolic Cylinder function, solution of the differential equation:

$$\begin{cases} y''(z) + \left(\mu + \frac{1}{2} - \frac{1}{4}z^2\right) y(z) = 0, \\ y(z) \underset{z\to+\infty}{\sim} z^{\mu} e^{-z^2/4}, \end{cases}$$

where the Parabolic Cylinder function, denoted $D_\mu(z)$, is to be considered as function of its order μ.

The aim of this chapter is to complete Slater's study and to propose a formula for μ-zeros for large values of z. We also establish a formula for the derivative of a μ-zero and deduce some monotonicity results. Since the z-zeros of Hermite functions are linked to those of Parabolic Cylinder functions, our analysis is based on the results of [41]. Asymptotic expressions for the μ-zeros are derived from the expansion of Olver [87]. Our analysis is similar to that of [29] for Hankel functions. This study is motivated by applications in Statistics. Indeed the μ-zeros of the Parabolic Cylinder function then appear in the law of the first time the Ornstein Uhlenbeck process hits a given level or in the law of its running supremum ([6], [16]). The estimation of the Ornstein Uhlenbeck parameters is a subject of interest in Statistics. The purpose of the parameter estimation is to provide a basis from which this process can be used in empirical applications (in finance, temperature modeling, medicine, physics...). Some authors proposed an estimation based on the observations of the first passages times ([64]) or of the supremum of this process ([18]). To obtain statistical properties on the estimators such as consistency and asymptotic normality, we need to check some integrability properties on the law of these observations. To do this, the asymptotic behavior of these μ-zeros is required.

The chapter is organized as follows: in Section 3.2 we present some properties for the μ-zeros of the Parabolic Cylinder function $D_\mu(z)$. Section 3.3 focuses on the behavior of the μ-zeros for large z. Moreover, numerical verifications of the asymptotic expansion are displayed. Only real parameters are considered in this chapter.

3.2 Variation of zeros

In this section we present some properties for the μ-zeros of the Parabolic Cylinder function $D_\mu(z)$ with respect to the real variable z. Since the function is holomorphic (see [68], ch. 10) in the complex plane, the set of μ-zeros has no accumulation points and there is a countably infinite number of zeros. Moreover, in real cases they are strictly positive [16]. In the following, we denote by $(\mu_n(z))_{n\geq 1}$ the ordered sequence of zeros of the function $\mu \mapsto D_\mu(z)$.

The following proposition gives some monotonicity properties of the zeros.

Proposition 3.2.1. *For all $n \in \mathbb{N}^*$:*

1. *The first derivative of the $n-th$ μ-zero is given by:*

$$\partial_z \mu_n(z) = \frac{2}{\sqrt{\pi}\int_0^\infty e^{-(2\mu_n(z)+1)u+\frac{z^2}{2}\tanh(u)} erfc\left(z\sqrt{\frac{\tanh(u)}{2}}\right)\frac{du}{\sqrt{\sinh(u)\cosh(u)}}} \tag{3.2.1}$$

2. *The function $z \mapsto \mu_n(z)$ is strictly increasing and convex.*

Proof. 1. Let $z(\mu,\alpha)$ be a zero of the function $z \mapsto \cos(\alpha)H_\mu(z)+\sin(\alpha)G_\mu(z)$ where α is fixed and $H_\mu(z)$ and $G_\mu(z)$ are linear independent solutions of $y''-2zy'+2\mu y=0$

with $H_\mu(z) \sim (2z)^\mu$ and $G_\mu(z) \sim \frac{1}{\sqrt{\pi}}\Gamma(1+\mu)z^{-\mu-1}e^{z^2}$ when $z \to +\infty$. In [41], the authors compute the derivative with respect to μ and obtain

$$\partial_\mu z(\mu,\alpha) = \frac{\sqrt{\pi}}{2}\int_0^\infty e^{-(2\mu+1)u+z(\mu,\alpha)^2\tanh(u)} erfc\left(z(\mu,\alpha)\sqrt{\tanh(u)}\right)\frac{du}{\sqrt{\sinh(u)\cosh(u)}}.$$

Since $D_\mu(z) = 2^{-\frac{\mu}{2}}e^{-\frac{z^2}{4}}H_\mu\left(\frac{z}{\sqrt{2}}\right)$, then choosing $\alpha = 0$ the result is a direct consequence of the local inversion theorem.

2. Elbert and Muldoon [41] (Corollary 7.2) prove that $\mu \mapsto z(\mu,\alpha)$ is completely monotonic : $\partial_\mu z(\mu,\alpha) > 0$, $(-1)^k\partial_\mu^{k+1}z(\mu,\alpha) \geq 0$, $k = 1,2,\ldots,n$. The conclusion follows from the local inversion theorem.

□

If $z = 0$, Formula (3.2.1) can be simplified. Indeed, the zeros $(\mu_n(0))_{n\geq 1}$ of $\mu \mapsto D_\mu(0)$ are the positive odd integers, $\mu_n(0) = 2n-1$. In this particular case, (3.2.1) becomes:

$$\partial_z\mu_n(z)\,|_{z=0} = \frac{2}{\sqrt{\pi}\int_0^\infty \frac{e^{-(4n-1)u}du}{\sqrt{\sinh(u)\cosh(u)}}} = \begin{cases} \frac{2}{\sqrt{\pi}} & \text{if} \quad n=1, \\ \frac{2}{\sqrt{\pi}(n-1)B\left(n-1,\frac{3}{2}\right)} & \text{if} \quad n \geq 2. \end{cases}$$

Remark 3.2.2. *We can prove that the function $z \mapsto \mu_n(z)$ is strictly increasing without using the form (3.2.1) of the derivative $\partial_z\mu_n(z)$. Indeed, on the one hand, thanks to [16] (Proposition 3.14), we have:*

$$\int_z^\infty D^2_{\mu_n(z)}(x)\,dx = -\mu_n(z)\,D_{\mu_n(z)-1}(z)\,\partial_\mu D_{\mu_n(z)}(z).$$

On the other hand, by differentiating $D_{\mu_n(z)}(z) = 0$ with respect to z, we get:

$$\partial_\mu D_{\mu_n(z)}(z)\,\partial_z\mu_n(z) + \mu_n(z)\,D_{\mu_n(z)-1}(z) = 0.$$

Therefore

$$\partial_\mu D_{\mu_n(z)}(z) = -\frac{\mu_n(z)\,D_{\mu_n(z)-1}(z)}{\partial_z\mu_n(z)}.$$

So that we finally get

$$\partial_z\mu_n(z) = \frac{\mu_n(z)^2\,D^2_{\mu_n(z)-1}(z)}{\int_z^\infty D^2_{\mu_n(z)}(x)\,dx} > 0.$$

As a consequence of (3.2.1), we obtain some bounds on the derivative of a μ-zero with respect to z.

Corollary 3.2.3. *The following inequalities hold:*

1. *If $z > 0$, then $\mu_n(z) > 1$ for all $n \in \mathbb{N}^*$ and*

$$\frac{4}{\sqrt{\pi}\,(\mu_n(z)-1)\,B\left(\frac{\mu_n(z)-1}{2},\frac{3}{2}\right)} \leq \partial_z\mu_n(z) \leq \frac{4e^{-\frac{z^2}{2}}}{\sqrt{\pi}\,(\mu_n(z)-1)\,erfc\left(\frac{z}{\sqrt{2}}\right)B\left(\frac{\mu_n(z)-1}{2},\frac{3}{2}\right)}.$$

2. If $z<0$, then $0<\mu_1(z)<1$ and $\mu_n(z)>1$ for $n\geq 2$. We also have

$$\frac{2e^{-\frac{z^2}{2}}}{c_{\mu_1(z)}\sqrt{\pi}erfc\left(\frac{z}{\sqrt{2}}\right)}\leq \partial_z\mu_1(z)\leq \frac{2}{\sqrt{\pi}c_{\mu_1(z)}},$$

$$\frac{4e^{-\frac{z^2}{2}}}{\sqrt{\pi}\left(\mu_n(z)-1\right)erfc\left(\frac{z}{\sqrt{2}}\right)B\left(\frac{\mu_n(z)-1}{2},\frac{3}{2}\right)}\leq \partial_z\mu_n(z)\leq \frac{4}{\sqrt{\pi}\left(\mu_n(z)-1\right)B\left(\frac{\mu_n(z)-1}{2},\frac{3}{2}\right)},\quad n\geq 2,$$

where $c_{\mu_1(z)}\in\left[1\ \frac{\pi}{2}\right)$ is a constant depending on $\mu_1(z)$.

Proof. 1. If $z=0$, the zeros $(\mu_n(0))_{n\geq 1}$ are the positive odd integers. Since the $z\mapsto \mu_n(z)$ is strictly increasing (see Proposition 3.2.1), then for $z>0$, we get:

$$\mu_n(z)>\mu_n(0)\geq\mu_1(0)=1,\quad n\in\mathbb{N}^*.$$

Since u is positive, then $\tanh(u)\in[0,1]$ and

$$e^{\frac{z^2}{2}}erfc\left(\frac{z}{\sqrt{2}}\right)\leq e^{\frac{z^2}{2}\tanh(u)}erfc\left(z\sqrt{\frac{\tanh(u)}{2}}\right)\leq 1.$$

Moreover, $\mu_n(z)>1$ for all $n\in\mathbb{N}^*$, then

$$\int_0^\infty \frac{e^{-(2\mu_n(z)+1)u}}{\sqrt{\sinh(u)\cosh(u)}}du=\frac{\mu_n(z)-1}{2}B\left(\frac{\mu_n(z)-1}{2},\frac{3}{2}\right).$$

2. On the one hand, in the case of negative z, the strictly monotonicity property of $z\mapsto\mu_n(z)$ gives $\mu_n(z)<\mu_n(0)=2n-1$ for all $n\in\mathbb{N}^*$.

On the other hand, the behavior of $D_\mu(z)$ for large negative z is ([68]) :

$$D_\mu(z)=z^\mu e^{-\frac{z^2}{4}}\left(1+O\left(|z|^{-2}\right)\right)-\frac{\sqrt{2\pi}e^{-\mu\pi i}}{\Gamma(-\mu)}z^{-\mu-1}e^{\frac{z^2}{4}}\left(1+O\left(|z|^{-2}\right)\right).\qquad(3.2.2)$$

If $\mu\in\mathbb{N}$, the dominant part (second term) in (3.2.2) vanishes and $D_\mu(z)\underset{z\to-\infty}{\longrightarrow}0$. Therefore $\mu_n(z)\underset{z\to-\infty}{\longrightarrow}n-1$ for all $n\in\mathbb{N}^*$. We deduce that for $n\in\mathbb{N}^*$, we have:

$$n-1<\mu_n(z)<2n-1.$$

If $z<0$, then $1\leq e^{\frac{z^2}{2}\tanh(u)}erfc\left(z\sqrt{\frac{\tanh(u)}{2}}\right)\leq e^{\frac{z^2}{2}}erfc\left(\frac{z}{\sqrt{2}}\right)$.

Moreover

$$\int_0^\infty \frac{e^{-(2\mu+1)u}}{\sqrt{\sinh(u)\cosh(u)}}du=\begin{cases} c_\mu\in\left[1\ \frac{\pi}{2}\right) & \text{if}\quad 0<\mu\leq 1,\\ \frac{\mu-1}{2}B\left(\frac{\mu-1}{2},\frac{3}{2}\right) & \text{if}\quad \mu>1,\end{cases}$$

where c_μ is a constant depending on μ.

□

Remark 3.2.4. *In the case $z<0$, by using the inequality $n-1<\mu_n(z)<2n-1$, we obtain less accurate bounds depending only on n.*

3.3 Asymptotic expansions of μ-zeros for large z

We are now interested in the behavior of μ-zeros for large positive values of z. Since the μ-zeros are positive, we restrict ourselves to the case of real positive μ.

Asymptotic expansion of Parabolic Cylinder function

Recall that the Parabolic Cylinder function $D_\mu(z)$ is solution of the differential equation:

$$y''(z) + \left(\mu + \frac{1}{2} - \frac{1}{4}z^2\right) y(z) = 0, \quad z \in \mathbb{C}. \tag{3.3.1}$$

The behavior of $D_\mu(z)$ for large positive z and $z >> |\mu|$ is ([68]):

$$D_\mu(z) = e^{-\frac{z^2}{4}} z^\mu \left[1 + O\left(z^{-2}\right)\right]. \tag{3.3.2}$$

Equation (3.3.1) has two turning points at $\sqrt{4\mu+2}$ and $-\sqrt{4\mu+2}$.The asymptotic behavior of $D_\mu(z)$ changes significantly depending on the relative position of z with respect to the turning points. The asymptotic behavior (3.3.2) is not valid if z runs through an interval containing one of the turning points. In this case, an Airy type expansion is needed to obtain those of the Parabolic Cylinder function. Its expression is ([87], [101]) :

$$D_{\frac{1}{2}\nu^2-\frac{1}{2}}\left(\nu t\sqrt{2}\right) \underset{\nu\to+\infty}{=} 2\sqrt{\pi}\nu^{\frac{1}{3}} g(\nu) \left(\frac{\xi(t)}{t^2-1}\right)^{\frac{1}{4}} \left[Ai\left(\nu^{\frac{4}{3}}\xi(t)\right) A_\nu(\xi) + \frac{Ai'\left(\nu^{\frac{4}{3}}\xi(t)\right)}{\nu^{\frac{8}{3}}} B_\nu(\xi)\right], \tag{3.3.3}$$

where Ai is the Airy function of the first kind,

$$\frac{2}{3}(-\xi(t))^{\frac{3}{2}} = \int_t^1 \sqrt{1-u^2}du, \quad -1 < t \leq 1 \, (\xi \leq 0),$$
$$\frac{2}{3}(\xi(t))^{\frac{3}{2}} = \int_1^t \sqrt{u^2-1}du, \quad t \geq 1 \, (\xi \geq 0),$$
$$g(\nu) \underset{\nu\to+\infty}{\sim} 2^{-\frac{1}{4}\nu^2-\frac{1}{4}} e^{-\frac{1}{4}\nu^2} \nu^{\frac{1}{2}\nu^2-\frac{1}{2}} \left(1 + \frac{1}{2}\sum_{s\geq 1} \frac{2^s\gamma_s}{\nu^{2s}}\right) \quad \text{and the coefficients } \gamma_s \text{ are defined by}$$
$$\Gamma\left(\frac{1}{2}+z\right) \underset{z\to+\infty}{\sim} \sqrt{2\pi} e^{-z} z^z \sum_{s\geq 0} \frac{\gamma_s}{z^s}.$$

More details on these coefficients γ_s and their computation can be found in [87], pages 134-135. For example, Olver finds for $s \leq 4$:

$$\gamma_0 = 1, \ \gamma_1 = -\frac{1}{24}, \ \gamma_2 = \frac{1}{1152},$$
$$\gamma_3 = \frac{1003}{414720}, \ \gamma_4 = \frac{4027}{39813120}.$$

The functions A_ν and B_ν satisfy $A_\nu(\xi) \underset{\nu\to+\infty}{\sim} \sum_{s\geq 0} \frac{A_s(\xi(t))}{\nu^{4s}}$, $B_\nu(\xi) \underset{\nu\to+\infty}{\sim} \sum_{s\geq 0} \frac{B_s(\xi(t))}{\nu^{4s}}$, where the coefficients $A_s(\xi(t))$ and $B_s(\xi(t))$ are given by:

$$A_s(\xi(t)) = \xi(t)^{-3s} \sum_{m=0}^{2s} \beta_m \left(\frac{\xi(t)}{t^2-1}\right)^{\frac{3}{2}} (2s-m) u_{2s-m}(t),$$

$$B_s(\xi(t)) = -\xi(t)^{-3s} \sum_{m=0}^{2s+1} \alpha_m \left(\frac{\xi(t)}{t^2-1}\right)^{\frac{3}{2}} (2s-m+1) u_{2s-m+1}(t),$$

$$\alpha_0 = 1 \text{ and } \alpha_m = \frac{(2m+1)(2m+3)\dots(6m-1)}{m!(144)^m}, \beta_m = -\frac{6m+1}{6m-1}\alpha_m$$

and $u_s(t)$ are polynomials in t of degrees $3s$ (s odd), $3s-2$ (s even, $s\geq 2$) and they satisfy the recurrence relation

$$(t^2-1)u'(t) - 3stu_s(t) = r_{s-1}(t),$$

where

$$8r_s(t) = (3t^2+2)u_s(t) - 12(s+1)tr_{s-1}(t) + 4(t^2-1)r'_{s-1}(t).$$

Formula (3.3.3) gives the asymptotic behavior of $D_\mu(z)$ if z runs through an interval containing the turning point $\sqrt{4\mu+2}$. Near the other turning point $-\sqrt{4\mu+2}$ (so when $z<0$), the asymptotic behavior of $D_\mu(z)$ is given by another formula (see [87], (9.7.)). As in this section we study the μ-zeros only in the case of large positive z, this second formula will not be useful here.

Remark 3.3.1. *If z belongs to an interval containing the other turning point $-\sqrt{4\mu+2}$, the study of the asymptotic behavior of the zeros is easier. The zeros tend to positive integers.*

Indeed, in this case the asymptotic behavior of the Parabolic Cylinder function is given by ([87], [101]):

$$D_{\frac{1}{2}\nu^2-\frac{1}{2}}\left(-\nu t\sqrt{2}\right) \underset{\nu\to+\infty}{=} 2\sqrt{\pi}\nu^{\frac{1}{3}} g(\nu) \left(\frac{\xi(t)}{t^2-1}\right)^{\frac{1}{4}} \left[\sin\left(\tfrac{1}{2}\pi\nu^2\right)\left(Ai\left(\nu^{\frac{4}{3}}\xi(t)\right)A_\nu(\xi) + \frac{Ai'\left(\nu^{\frac{4}{3}}\xi(t)\right)}{\nu^{\frac{8}{3}}}B_\nu(\xi)\right)\right.$$

$$\left. + \cos\left(\frac{1}{2}\pi\nu^2\right)\left(Bi\left(\nu^{\frac{4}{3}}\xi(t)\right)A_\nu(\xi) + \frac{Bi'\left(\nu^{\frac{4}{3}}\xi(t)\right)}{\nu^{\frac{8}{3}}}B_\nu(\xi)\right)\right], \tag{3.3.4}$$

where Bi is the Airy function of the second kind. Recall that ([86]):

$$Ai(z) \underset{z\to+\infty}{\sim} \frac{1}{2\sqrt{\pi}} z^{-\frac{1}{4}} e^{-\frac{2}{3}z^{\frac{3}{2}}} \qquad Ai'(z) \underset{z\to+\infty}{\sim} -\frac{1}{2\sqrt{\pi}} z^{\frac{1}{4}} e^{-\frac{2}{3}z^{\frac{3}{2}}},$$

$$Bi(z) \underset{z\to+\infty}{\sim} \frac{1}{\sqrt{\pi}} z^{-\frac{1}{4}} e^{\frac{2}{3}z^{\frac{3}{2}}} \qquad Bi'(z) \underset{z\to+\infty}{\sim} \frac{1}{\sqrt{\pi}} z^{\frac{1}{4}} e^{\frac{2}{3}z^{\frac{3}{2}}}.$$

As the factor $Bi\left(\nu^{\frac{4}{3}}\xi(t)\right)A_\nu(\xi) + \frac{Bi'\left(\nu^{\frac{4}{3}}\xi(t)\right)}{\nu^{\frac{8}{3}}}B_\nu(\xi)$ goes to infinity when $\nu\to+\infty$, to obtain the zeros of $D_{\frac{1}{2}\nu^2-\frac{1}{2}}$ we must cancel this term. If $\frac{1}{2}\nu^2 = n+\frac{1}{2}$ in (3.3.4), the cosine vanishes and, hence, the dominant part vanishes.

Asymptotic expansion of μ-zeros

For large positive values of z, the μ-zeros of $D_\mu(z)$ are linked to a_n, $n = 1, 2, \ldots$, the zeros of the Airy function ($Ai(a_n) = 0$). The zeros of Ai have been studied ([86]). Olver shows that they are all real and negative. They may be expressed asymptotically as

$$a_n \underset{n\to+\infty}{\sim} -\left(\frac{3\pi}{2}\left(n-\frac{1}{4}\right)\right)^{\frac{2}{3}}.$$

The following proposition gives the behavior of $\mu_n(z)$ when $z \to \infty$.

Proposition 3.3.2. *For large positive z, we have:*

$$\mu_n(z) = \frac{z^2}{4} - \frac{1}{2} - z^{\frac{2}{3}}2^{-\frac{2}{3}}a_n + O\left(z^{-\frac{2}{3}}\right), \tag{3.3.5}$$

where a_n, $n = 1, 2, \ldots$ are the zeros of the Airy function.

Proof. We apply the method given in [29] to compute the zeros of the Hankel function. Taking $\nu = \sqrt{2\mu+1}$ in (3.3.3), the original argument z of $D_\mu(z)$ has temporarily been replaced by $t(\xi)\sqrt{2}\sqrt{2\mu+1}$. The μ-zeros of $D_\mu\left(t(\xi)\sqrt{2}\sqrt{2\mu+1}\right)$ are given asymptotically by the μ-zeros of the right hand side of (3.3.3). These zeros in turn appear to be given by the μ-solutions of $Ai\left((2\mu+1)^{\frac{2}{3}}\xi(t)\right) = 0$, from which we deduce that

$$(2\mu+1)^{\frac{2}{3}}\xi(t) \sim a_n,$$

as $\mu \to +\infty$. Since a_n is negative, then $\xi(t) < 0$. Hence we obtain the asymptotic relation between zeros of the Airy function and μ-zeros of the Parabolic Cylinder function (we restore the original argument $z = t(\xi)\sqrt{2}\sqrt{2\mu+1}$):

$$\left(\frac{a_n}{\xi(t)}\right)^{\frac{3}{2}} \sim 2\mu_n + 1 = \frac{z^2}{2t^2(\xi)},$$

where ξ and t are related by the relation $\frac{2}{3}(-\xi(t))^{\frac{3}{2}} = \int_t^1 \sqrt{1-u^2}du$. We deduce that the limiting case that gives rise to large values of z (so large values of μ_n) is $\xi(t) \to 0$. As ξ is negative, so the case $\xi(t) \uparrow 0$ is associated with the behavior of the μ-zeros of $D_\mu(z)$ for large positive z. We easily deduce that if $\xi(t) \uparrow 0$, then $t \uparrow 1$ and $t(\xi) = 1 + 2^{-\frac{1}{3}}\xi - \frac{1}{10}2^{-\frac{2}{3}}\xi^2 + O(\xi^3)$.

Thus, for $z \to \infty$,

$$\begin{aligned}
2\mu_j + 1 = \frac{z^2}{2t^2(\xi)} &= \frac{z^2}{2}\frac{1}{\left(1 + 2^{-\frac{1}{3}}\xi - \frac{1}{10}2^{-\frac{2}{3}}\xi^2 + O(\xi^3)\right)^2}, \\
&= \frac{z^2}{2}\left(1 - 2^{\frac{2}{3}}\xi + \frac{16}{5}2^{-\frac{2}{3}}\xi^2 + O\left(\xi^3\right)\right), \\
&= \frac{z^2}{2} - z^{\frac{2}{3}}2^{\frac{1}{3}}a_j + \frac{8}{5}2^{\frac{2}{3}}z^{-\frac{2}{3}}a_j^2 + z^{-2}O(1).
\end{aligned}$$

□

Remark 3.3.3. *The expansions* (3.3.5) *are still valid for complex values of the parameters.*

Numerical verification

The values of the function $\mu \mapsto D_\mu(z)$ exceed the computer capabilities, the zeros are no longer observable and computable. For large values of z, the function oscillates between $+\infty$ and $-\infty$, so numerical verifications will be performed for moderate values of z. To check the quality of our results, we compare the μ-zeros given by the dichotomy method with those obtained by the asymptotic expansion (3.3.5). The computations are done in Python.

n	1	2	3	4	5
Dichotomy	143.8036	153.0062	160.6533	167.4948	173.8159
Asymptotic expansion	143.6623	152.5775	159.8764	166.3272	172.2242
Relative error	0.0009	0,0028	0,0048	0,0069	0,0091

Table 3.1: Numerical comparison of the first five μ-zeros of $\mu \mapsto D_\mu(23)$ obtained by the dichotomy method and by the asymptotic expansion formula (3.3.5).

In Table 1, we observe that, although the asymptotic formula (3.3.5) is for large z, for $z = 23$ we already obtain acceptable estimates compared to the dichotomy method. We clearly see the loss of accuracy, but numerical right shift of the μ-zeros estimated with (3.3.5) can be explained as follows. Since simulations cannot be performed for very large z, as n increase, the zeros of Airy function become dominant on (3.3.5). The quantity $-z^{\frac{2}{3}}2^{-\frac{2}{3}}a_n > 0$ becomes too small, which involves a right shift on the simulation. This example shows that our formula allows to evaluate the first μ-zeros even for moderates values of z.

Conclusion and Perspectives

Several perspectives are possible in this study. For example, questions may arise regarding the behavior of the pseudo-likelihood estimator compared to the likelihood estimator. Also, we believe that the condition $r > \frac{\log(\frac{5}{3})}{b_0}$ in the asymptotic normality property is only sufficient. We think that this property holds for $r > 0$.

Furthermore, in the data set [58], we can find the daily supremum, infimum, and mean of the temperature. One perspective would be to consider an estimation based on both infimum and supremum observations of a stationary OU process. Specifically, revisiting the previous notation, we would have a sample:

$$S^{i,0} = \sup_{s\in[t_{i-1},t_i]} X_s \quad \text{and} \quad I^{i,0} = \inf_{s\in[t_{i-1},t_i]} X_s.$$

The sample of the supremum and infimum would follow the same properties as stated in Proposition 2.2.1. The estimation would be constructed by making use of the joint law between the supremum and infimum of the OU process, which is directly related to the law of the first exit time from an interval of the OU process. In fact, for $x_0, c_1, c_2 \in \mathbb{R}$ such that $c_1 < x_0 < c_2$ we have:

$$\mathbb{P}_{x_0}\left(t < \tau_{[c_1,c_2]}\right) = \mathbb{P}_{x_0}\left(c_1 < \inf_{s<t} X_s,\ \sup_{s<t} X_s < c_2\right).$$

We denote $p(x_0,t) = \mathbb{P}_{x_0}\left(t < \tau_{[c_1,c_2]}\right)$. By applying Proposition 1.3.11 together with Remark 1.3.12, the function p is solution of:

$$\begin{cases} \partial_t p(x_0,t) = \frac{\beta}{2}\partial^2_{{x_0}^2} p(x_0,t) + (a - bx_0)\partial_{x_0} p(x_0,t) & (t,x_0) \in \mathbb{R}^\star_+ \times (c_1,c_2), \\ p(c_1,t) = p(c_2,t) = 0 & t \in \mathbb{R}^\star_+, \\ p(x_0,0) = 1 & x_0 \in [c_1,c_2]. \end{cases}$$

Making use of a spectral decomposition method (see Chapter 4), the function p can be explicitly computed.
In this case, we could propose a pseudo-likelihood method. We think that statistical properties for this estimator would follow from similar reasoning as in Section 2.5. In particular, the difficulty would lie in studying the series for asymptotic expansion and regularity. Similarly, we could propose an estimation method based on observations of the infimum, supremum, and mean of an OU process. In this case, the joint distribution has not been explicitly computed. Classical methods such as spectral decomposition used would not work because the distribution would be a solution of a partial differential equation whose spatial operator is not Hermitian.

Part II

Threshold diffusion processes

The analysis of discontinuous coefficients for the drift and volatility of SDEs represents a novel approach in the domain of stochastic modeling. These models are often categorized as threshold diffusion models, where the dynamics of the system change abruptly when some of the thresholds are crossed. This approach offers an alternative to the continuous-time models such as the Self-Exciting Threshold Autoregressive (SETAR), which belongs to the broader class of Threshold Autoregressive (TAR) models (see [102]). While SETAR models provide a framework for capturing non-linear dynamics in time series data, the introduction of discontinuous coefficients in SDEs offers a different perspective on modeling dynamic phenomena. These processes find applications across various domains, such as finance [34] and population ecology [24] for example. A representation for such processes can be given by the following SDE:

$$X_t = X_0 + \int_0^t \mu(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad t \geq 0, \tag{3.3.6}$$

where the functions μ and σ exhibit discontinuities. Equations of the class (3.3.6) are closely linked, through Itô-Tanaka formula, to processes solution of:

$$Y_t = Y_0 + \int_0^t \mu(Y_s)\,\mathrm{d}s + \int_0^t \sigma(Y_s)\,\mathrm{d}B_s + \int_{\mathbb{R}} L_t^x(Y)\nu(\mathrm{d}x), \quad t \geq 0, \tag{3.3.7}$$

where $L_t^x(X)$ is the symmetric local time, ν is a finite measure, singular with respect to the Lebesgue measure, with $|\nu(x)| \leq 1$ for all $x \in \mathbb{R}$.

The computation of explicit laws for SDEs is of paramount importance due to its wide-ranging applications in various fields. There is also significant interest in computing explicit laws for the process solution of thresholds SDE (3.3.6). For example, in [34], the authors compute the transition probability density of threshold Ornstein-Uhlenbeck (T-OU) and threshold Cox-Ingersoll-Ross (T-CIR) processes. Also, in [35], the authors compute the law of a drifted Skew Brownian motion using a method involving the inversion of the resolvent.

Several studies have been conducted on the parametric estimation of these type of processes. In [72], the authors use a maximum likelihood estimator for diffusions with piecewise constant drift and diffusion coefficients. Consistency and asymptotic normality results are provided for both continuous and discrete observations in the case of the ergodic and non-ergodic regimes of the process. In another study, [81] examines the estimation of drift parameters for a T-OU process. Additionally, [70] introduces a method for estimating the diffusion parameter of a threshold Brownian motion.

In Chapter 4, we propose an approach to studying the T-OU process. We provide explicit expressions for transition probability density and first hitting time density for the process killed at a constant boundary. Extensions to multi-threshold cases are explored, and an explicit expression for the Laplace transform of the first hitting time is proposed. In Chapter 5, we consider a thresholds Chan–Karolyi–Longstaff–Sanders (CKLS) process. This study focuses on the asymptotic behavior of maximum and quasi-maximum likelihood

estimators for drift parameters, considering both continuous and discrete time observations. We demonstrate that under high-frequency observations and infinite horizon, these estimators exhibit the same asymptotic normality property as in continuous time observations. Additionally, we discuss the estimation of diffusion coefficients. Finally, the application of these estimators to simulated and real data motivates the consideration of multiple thresholds. We conclude this part by suggesting directions for further research and highlighting areas for improvement.

Chapter 5 is in collaboration with Sara Mazzonetto and it is submitted, and available on Hal [80].

Chapter 4

The killed Threshold Ornstein-Uhlenbeck process

4.1 Introduction

Diffusion processes play a crucial role in modeling dynamic phenomena across various fields. A novel approach proposes investigating diffusion processes with piecewise constant coefficients. These processes find applications in various domains, such as population ecology and finance [24, 34]. For example, in [98, 81], the authors demonstrate through statistical tests that US Treasury Bills data exhibit threshold behavior, showing that this data can be modeled over some time windows by a threshold Ornstein-Uhlenbeck process.

This chapter deals with the computation of explicit laws of the process $X = (X_t)_{t\geq 0}$, which is solution of the following one-dimensional SDE:

$$X_t = x_0 + \int_0^t a(X_s) - b(X_s)X_s \,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad t \geq 0, \tag{4.1.1}$$

with $x_0 \in \mathbb{R}$, a, b and σ piecewise constants functions at levels called "*thresholds*".
Following the nomenclature in [81], we refer to the process X solution of (4.1.1) as the threshold Ornstein-Uhlenbeck (T-OU). Sometimes, we use (d)-T-OU to refer to a T-OU process with d thresholds. We use this alternative notation, when we wish to emphasize the number of thresholds.

Several studies have focused on the explicit computation of the laws of such processes. For instance, in [34], the transition probability densities of the (1)-T-OU and (1)-T-Cox-Ingersoll-Ross processes are computed using a spectral decomposition method. We can also mention [55], which computes laws associated to a threshold Brownian motion.

As noted in [36], threshold SDEs such as (4.1.1) are closely related to the solution of the

following SDE:

$$Y_t = Y_0 + \int_0^t a(Y_s) - b(Y_s)Y_s \, \mathrm{d}s + \int_0^t \sigma(Y_s) \, \mathrm{d}B_s + \int_{\mathbb{R}} L_t^y(Y)\nu(\mathrm{d}y), \quad t \geq 0, \tag{4.1.2}$$

where, $L_t^y(Y)$ is the symmetric local time, ν is a finite measure, singular with respect to the Lebesgue measure, with $|\nu(y)| \leq 1$ for all $y \in \mathbb{R}$, and a, b, and σ are measurable functions that may be discontinuous. The solutions to such equations are commonly referred to as singular diffusions (see [49]). From this equation, one can, for example, derive the SDEs associated with well-known processes in the literature, such as the Skew Brownian Motion (see [69]) or the Skew Ornstein-Uhlenbeck process (see [108]).

Computing the distributions for process solutions to (4.1.2) has been a key focus in the literature. For instance, in [35], the authors compute the law of a drifted Skew Brownian motion by inverting the resolvent. Also, in [108], the authors propose a spectral decomposition method to compute the law of the Skew Ornstein-Uhlenbeck process.

The computation of density distributions associated with killed processes over constant boundaries is a well-explored topic. In a series of works (see [76] and [75]), V. Linetsky computes the density of first hitting times for well-known processes such as the Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes. The author uses a spectral decomposition method, as presented in [49], which forms the foundation of their study.

In this chapter, we focus on computing the law associated with the killed T-OU process at a constant boundary. Specifically, we provide existence and explicit expressions for the transition probability density and the density of the first hitting time of the process. The transition probability density of the killed process is computed using a spectral decomposition method applied to the Fokker-Planck equation. Moreover, we can compute the density of the first hitting time using the transition probability density. Essentially, the hitting time density is related to the spatial derivative of the transition probability density. Since the transition probability density is expressed as a series of functions, the challenge lies in proving the term-by-term differentiation of this series. In [108], the authors study the Skew OU process and raise this challenging point. In particular, they note the difficulty in finding an associated asymptotic expansion for the eigenvalues of the spectral decomposition. In their article, the authors focus on a (1)-T-OU process, while in our case, we are investigating the multi-threshold case. In this chapter, we prove their intuition about the density (see Section 5 in [108]) and extend it to the multi-threshold case. Note that from our results, one can also recover the results in [6] and [8]. To our knowledge, this is the first time that the density of the first hitting time has been explicitly provided for a multiple thresholds process. Furthermore, we also propose an explicit expression for the Laplace transform of the first hitting time. This enables us to numerically test various methods for evaluating the hitting time density.

Outline. In Section 4.2, we introduce the T-OU by defining the drift and volatility as piecewise continuous functions. We also provide some preliminary results on the generator of the T-OU killed at a constant boundary. In Section 4.3, an explicit formulation is given for the transition probability density of the killed T-OU. Section 4.4 deals with the distribution of the first hitting time density of the T-OU process. Furthermore, in Section 4.5, we establish the connection between the T-OU and processes solutions to SDEs

of the class (4.1.2). In Section 4.6, some numerical experiments are performed. Proofs are collected in Section 4.7. Further results and discussion are provided in Section 4.8.

4.2 T-OU process

Let B be a standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We define X, the process solution of (4.1.1), with a, b, and σ as piecewise constant coefficients that are discontinuous at levels $-\infty = r_0 < r_1 < \cdots < r_d < r_{d+1} = +\infty$, where $d \in \mathbb{N}^\star$. Let $\bar{I}_d = (r_d, +\infty)$ and $\bar{I}_j = (r_j, r_{j+1}]$, for $j \in \{0, \cdots, d-1\}$. The drift coefficients and the volatility coefficients are given by:

$$a(x) = \sum_{j=0}^{d} a_j \mathbb{1}_{\bar{I}_j}(x) \quad , \quad b(x) = \sum_{j=0}^{d} b_j \mathbb{1}_{\bar{I}_j}(x) \text{ and } \sigma(x) = \sum_{j=0}^{d} \sigma_j \mathbb{1}_{\bar{I}_j}(x), \tag{4.2.1}$$

where $a_j \in \mathbb{R}$, $b_j > 0$ and $\sigma_j > 0$.

Proposition 4.2.1. *There exists a unique strong solution of* (4.1.1).

Proof. We easily verify that (ND) and (LI) hold for (4.1.1) (see Chapter 1). Then, from Theorem 1.2.13, there exists a unique weak solution of (4.1.1) up to an explosion time τ_∞. As the drift and volatility have at most linear growth, we have $\tau_\infty = +\infty$ a.s. (see Remark 1.2.16). Thus, the solution of (4.1.1) does not explode. Let us introduce the function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = \sum_{j=0}^{d} f_j(x) \mathbb{1}_{\bar{I}_j}(x)$ where for $j \in \{0, \cdots, d\}$, $f_j(x) = \sum_{k=0}^{j} \sum_{i=0}^{d} (\sigma_k - \sigma_i)^2$. Then, using Item 2 in Theorem 1.2.22, pathwise uniqueness holds for (4.1.1). Making use of Theorem 1.2.7, there exists a unique strong solution of (4.1.1). □

Let $c \in \mathbb{R}$ be fixed, we introduce the first hitting time of X of the level c:

$$\tau_c = \inf\{t > 0, X_t = c\},$$

Next, we sometimes denote $\tau_c(X)$ to refer to the first hitting of the constant boundary c of the process X. Without loss of generality, we suppose that $x_0 < c$ and $r_d < c$.

Remark 4.2.2. *Note that for $j \in \{0, \cdots, d\}$ and $c \in \bar{I}_j$, if $x_0 < c$, then τ_c is the first hitting time of a (j)-T-OU process. In the case $c < x_0$, $\tau_c(X)$ has the same law as $\tau_{-c}(-X)$, where, from Itô formula, $-X$ is still a T-OU process.*

Let us introduce $\bar{X}$, the T-OU killed at the boundary c, *i.e.*,

$$\bar{X}_t = \begin{cases} X_t \text{ on } \{t < \tau_c\}, \\ \Delta \quad \text{otherwise.} \end{cases} \tag{4.2.2}$$

where Δ is the cemetery. Note that from classical results (see [49], for example), the process $\bar{X}$ remains a diffusion process. The following proposition states the existence of a transition probability density with respect to the Lebesgue measure for the process $\bar{X}$.

Proposition 4.2.3. *For all $t \in \mathbb{R}_+^\star$, the probability $\mathbb{P}_{x_0}(\bar{X}_t \in dx)$ admits a density with respect to the Lebesgue measure.*

Proof. For every $B \in \mathcal{B}(\mathbb{R})$ such that B is a Lebesgue-negligible set, we have:

$$\mathbb{P}_{x_0}(\bar{X}_t \in B) = \mathbb{P}_{x_0}(X_t \in B, t \leq \tau_c) \leq \mathbb{P}_{x_0}(X_t \in B). \tag{4.2.3}$$

In particular, if $\mathbb{P}_{x_0}(X_t \in dx)$ admits a density with respect to the Lebesgue measure, then $\mathbb{P}_{x_0}(\bar{X}_t \in dx)$ also admits a density with respect to the Lebesgue measure.
The idea is to apply a well-chosen change of measure such that under the new measure $\mathbb{Q} \sim \mathbb{P}$, we have $\mathbb{Q}_{x_0}(X_t \in B) = 0$. Using Girsanov Theorem, we have:

$$\forall t \in [0,T], \quad \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t, \tag{4.2.4}$$

where $T \in \mathbb{R}_+$ is fixed and $Z_t = \exp\left(-\int_0^t \frac{a(X_s)-b(X_s)X_s}{\sigma(X_s)} dB_s - \frac{1}{2}\int_0^t \left(\frac{a(X_s)-b(X_s)X_s}{\sigma(X_s)}\right)^2 ds\right)$ is a $\mathbb{P}$-martingale. The martingale property of Z follows by applying a reasoning similar to Theorem 1 in [28]. In fact, using Theorem 1.7 and Proposition 1.2 in [91, Chapter VIII], we can prove that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent and $Z_t > 0$ for all $t \in [0,T]$, $\mathbb{P}$-almost-surely. Under the measure $\mathbb{Q}$, X is solution of:

$$X_t = x_0 + \int_0^t \sigma(X_t) dB_t^{\mathbb{Q}}, \quad t \geq 0$$

where, $B_t^{\mathbb{Q}} = B_t + \int_0^t \frac{a(X_s)-b(X_s)X_s}{\sigma(X_s)} ds$ is a $\mathbb{Q}$-Brownian motion. Using Theorem 9.1.9 in [96], $\mathbb{Q}_{x_0}(X_t \in dx)$ admits a density with respect to the Lebesgue measure. This implies that $\mathbb{P}_{x_0}(X_t \in B) = 0$. Then, from (4.2.3), for every Lebesgue-negligible set B, we have $\mathbb{P}_{x_0}(\bar{X}_t \in B) = 0$ and the conclusion holds. □

In the following, we denote $\mathbb{P}_{x_0}(\bar{X}_t \in dx) = p(x,t)dx$ and for $j \in \{0, \ldots, d-1\}$, we denote $I_j = \bar{I}_j$ and $I_d = (r_d, c]$.

It is a well-known result that a diffusion process is defined by its speed measure and its scale function (see Chapter 1 or Chapter VII in [91]). These quantities can be computed explicitly by using (1.3.4) and (1.3.6).

Proposition 4.2.4. *Let $\bar{X}$ be defined by (4.2.2). For $x \in (-\infty, c)$, the derivative of the scale function, denoted by S, and the speed measure m are defined as follows:*

$$S(x) = \frac{1}{C_0 S_0(x, r_1)} \mathbb{1}_{I_0}(x) + \sum_{j=1}^{d} \frac{1}{C_j S_j(x, r_j)} \mathbb{1}_{I_j}(x),$$

with, for $i \in \{0, \cdots, d\}$, $S_i(x,r) = \exp\left(\frac{2a_i(x-r)-b_i(x^2-r^2)}{\sigma_i^2}\right)$ and

$$m(dx) = m(x)dx = \frac{2}{\sigma^2(x)S(x)} dx,$$

where $C_0 = C_1 = 1$ and for $j \in \{2, \cdots, d\}$, $C_j = \prod_{i=1}^{j-1} S_i(r_{i+1}, r_i)$.

We now introduce the forward operator of the process $\bar{X}$. This will allow us to define the Fokker-Planck equation for the transition probability density p. From Definition 1.3.7 and Proposition 1.3.10, the infinitesimal generator $(\mathcal{L}, \mathrm{Dom}(\mathcal{L}))$ of the process $\bar{X}$, solution of (4.2.2), may be written as follows:

$$\mathcal{L}f = \frac{1}{m(.)}\left(S^{-1}(.)f'(.)\right)', \tag{4.2.5}$$

for all $f \in \mathrm{Dom}(\mathcal{L})$. The forward operator $\mathcal{L}^\star$ is defined as the adjoint operator of $\mathcal{L}$ in $L^2((-\infty, c])$:

$$\langle \mathcal{L}f, g\rangle_{L^2((-\infty,c])} = \langle f, \mathcal{L}^\star g\rangle_{L^2((-\infty,c])},$$

where $f \in \mathrm{Dom}(\mathcal{L})$ and $g \in \mathrm{Dom}(\mathcal{L}^\star)$ verifies:

$$\mathcal{L}^\star g = \left(S^{-1}(.)\left(\frac{g}{m}(.)\right)'\right)', \tag{4.2.6}$$

for all $g \in \mathrm{Dom}(\mathcal{L}^\star)$.

The domain of the adjoint generator can be constructed such that $\mathcal{L}^\star$ is Hermitian (see Section 4.7 for further details). In order for $\mathcal{L}^\star$ to be well-defined, each derivative needs to make sense in (4.2.6). So, the function $g \in \mathrm{Dom}(\mathcal{L}^\star)$ satisfies the transmission conditions: $\frac{g}{m}(.)$ and $S^{-1}(.)\left(\frac{g}{m}(.)\right)'$ are continuous in space. Using the explicit expressions for the derivative of the scale function and the speed measure in Proposition 4.2.4, $g \in \mathrm{Dom}(\mathcal{L}^\star)$ satisfies, for all $j \in \{1, \cdots, d\}$:

$$\begin{cases} \sigma_j^2 g(r_j^+) = \sigma_{j-1}^2 g(r_j^-), \\ \frac{\sigma_j^2}{2} g'(r_j^+) + (b_j r_j - a_j) g(r_j^+) = \frac{\sigma_{j-1}^2}{2} g'(r_j^-) + (b_{j-1} r_j - a_{j-1}) g(r_j^-). \end{cases} \tag{4.2.7}$$

The probability density p is solution of the following Fokker-Planck equation:

$$\begin{cases} \partial_t p(x,t) = \frac{1}{2}\partial_{x^2}^2(\sigma^2(x)p(x,t)) + \partial_x[(b(x)x - a(x))p(x,t)], \quad (x,t) \in (-\infty, c) \times \mathbb{R}_+^\star, \\ \text{for all } t \in \mathbb{R}_+^\star, \quad p(t,.) \text{ verifies (4.2.7)}, \\ p(c,t) = 0, \quad t \in \mathbb{R}_+^\star, \\ \lim_{t\to 0} p(x,t)dx = \delta_{x_0}(dx). \end{cases} \tag{4.2.8}$$

Note that, for all $j \in \{1, \cdots, d\}$, the function p is discontinuous at r_j. This discontinuity arises from the discontinuity of the volatility function σ. The discontinuities in the drift and the volatility imply discontinuities in the derivative of the transition probability density.

4.3 Transition probability density of the killed T-OU process

By solving the Fokker-Planck equation (4.2.8) using a spectral decomposition method, we provide an explicit expression of the transition probability density of the killed process $\bar{X}$.

We introduce the functions involved in formulating the solution of the Fokker-Planck equation (4.2.8). For all $j \in \{0, \cdots, d\}$, we define the function $\phi_\mu^{(j)}$ as follows:

$$\phi_\mu^{(j)}(x) = \sqrt{m(x)} \left[\alpha_\mu^{(j)} D_{\frac{\mu}{b_j}} \left(- \left(x - \frac{a_j}{b_j} \right) \frac{\sqrt{2b_j}}{\sigma_j} \right) + \beta_\mu^{(j)} D_{\frac{\mu}{b_j}} \left(\left(x - \frac{a_j}{b_j} \right) \frac{\sqrt{2b_j}}{\sigma_j} \right) \right],$$

where, for all $j \in \{1, \cdots, d\}$, the pairs $(\alpha_\mu^{(j)}, \beta_\mu^{(j)})$ are solutions of the linear system:

$$\begin{cases} \sigma_j^2 \phi_\mu^{(j)}(r_j) = \sigma_{j-1}^2 \phi_\mu^{(j-1)}(r_j), \\ \frac{\sigma_j^2}{2} (\phi_\mu^{(j)})'(r_j) + (b_j r_j - a_j) \phi_\mu^{(j)}(r_j) = \frac{\sigma_{j-1}^2}{2} (\phi_\mu^{(j-1)})'(r_j) + (b_{j-1} r_j - a_{j-1}) \phi_\mu^{(j-1)}(r_j), \end{cases} \tag{4.3.1}$$

with $(\alpha_\mu^{(0)}, \beta_\mu^{(0)}) = (1, 0)$. Here, D_μ is the Parabolic Cylinder function of parameter μ introduced in Chapter 2 and Chapter 3. The following theorem provides an explicit formulation for the transition probability density of the T-OU process killed at the constant c.

Theorem 4.3.1. *For $i, j \in \{0, \cdots, d\}$, suppose that $x \in I_i$ and $x_0 \in I_j$. The transition probability density of the T-OU process killed at the constant boundary c is given by:*

$$p(x,t) = \frac{1}{m(x_0)} \sum_{n=1}^{\infty} \frac{e^{-\mu_n t}}{N(\mu_n)} \phi_n^{(i)}(x) \phi_n^{(j)}(x_0), \tag{4.3.2}$$

where:

$$N(\mu_n) = \frac{\sigma_d^2}{2m(c)} \partial_x \left[\phi_n^{(d)}(x) \right]_{x=c} \partial_\mu \left[\phi_n^{(d)}(c) \right]_{\mu=\mu_n}, \tag{4.3.3}$$

and $\phi_n = \phi_{\mu_n}$, with $(\mu_n)_{n \geq 1}$ the ordered sequence of positive zeros of $\mu \mapsto \phi_\mu^{(d)}(c)$.

4.4 First hitting times distributions of the T-OU process

In this section, we focus on the first hitting time distribution of a constant boundary c for a T-OU process. First, we provide the Laplace transform, which is known to be the solution of a Sturm-Liouville problem. Next, we prove the existence of the density of the first hitting time. Then we give an explicit formulation for this density by differentiating with respect to the spatial variable the transition probability density (4.3.2).

4.4.1 Laplace transform of the first hitting times

Let $u_c(\mu, x_0) = \mathbb{E}_{x_0}[e^{-\mu \tau_c}]$, the Laplace transform of the first hitting time τ_c. Following the results in [49], the Laplace transform τ_c is the $\mathcal{C}^1$ solution of the following Sturm-Liouville equation:

$$\begin{cases} \mathcal{L} u_c(\mu, x_0) = \mu u_c(\mu, x_0), \quad x_0 \in (-\infty, c), \\ \lim\limits_{x_0 \to -\infty} u_c(\mu, x_0) = 0, \\ u_c(\mu, c) = 1. \end{cases} \tag{4.4.1}$$

Theorem 4.4.1. *For all $j \in \{0, \cdots, d\}$ and $x_0 \in I_j$, the Laplace transform of τ_c is given by:*

$$u_c(\mu, x_0) = \frac{\phi_{-\mu}^{(j)}(x_0)}{\phi_{-\mu}^{(d)}(c)}. \tag{4.4.2}$$

Furthermore, the function $u_c(., x)$, is holomorphic on $\{\mu \in \mathbb{C} | Re(\mu) > 0\}$.

4.4.2 Hitting times density

From the explicit representation (4.3.2) of the transition probability density of the killed T-OU process, we can derive a series representation for the first hitting time density. The following theorem provides this explicit formulation.

Theorem 4.4.2. *For all $x_0 < c$, we have $\mathbb{P}_{x_0}(\tau_c < +\infty) = 1$. Consequently, the distribution function of τ_c admits a density with respect to the Lebesgue measure, denoted p_{τ_c}. Furthermore, for all $t \in \mathbb{R}_+^\star$, $j \in \{0, \cdots, d\}$, and for $x_0 \in I_j$, the density p_{τ_c} is given by:*

$$p_{\tau_c}(t) = -\frac{m(c)}{m(x_0)} \sum_{n=1}^{\infty} e^{-\mu_n t} \frac{\phi_n^{(j)}(x_0)}{\partial_\mu \left[\phi_n^{(d)}(c)\right]_{\mu=\mu_n}}, \tag{4.4.3}$$

where $\phi_n^{(j)}(.)$ and $(\mu_n)_{n\geq 1}$ are defined as above.

Let us comment on the methodology to obtain the density (4.4.3) in Theorem 4.4.2. The idea is to use the transition probability density of the killed process computed in Theorem 4.3.1. Using the Fokker-Planck equation (4.2.8), we can establish a direct link between p_{τ_c} and p (see [8] for an example). Indeed, we have:

$$p_{\tau_c}(t) = -\frac{\sigma_d^2}{2} \partial_x p(x, t)|_{x=c}.$$

To differentiate the series term by term, it is necessary to verify that the series in (4.4.3) converges uniformly. To do this, for all $j \in \{0, \cdots, d\}$, we compute the asymptotic expansions of $\phi_\mu^{(j)}$ for μ large enough and of $(\mu_n)_{n\geq 1}$ for n large enough. Then Lemma 4.4.3 and Corollary 4.4.4 below are the key elements to prove Theorem 4.4.2 (for further details see Section 4.7 and Section 4.8).

Lemma 4.4.3. *For μ large enough and $x \in I_0$, we have: [1]:*

$$\frac{\phi_\mu^{(0)}(x)}{\sqrt{m(x)}} = \frac{2^{\frac{\mu}{2b_0}}}{\sqrt{\pi}} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \cos\left(x\frac{\sqrt{b_0 + 2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0}\right) \left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right], \tag{4.4.4}$$

and

$$\left(\frac{\phi_\mu^{(0)}(x)}{\sqrt{m(x)}}\right)' = -\frac{2^{\frac{\mu}{2b_0}}}{\sigma_0}\sqrt{\frac{b_0 + 2\mu}{\pi}} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \sin\left(x\frac{\sqrt{b_0 + 2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0}\right) \left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right]. \tag{4.4.5}$$

[1] Here, each asymptotic notation O and o depends on x.

For $j \in \{1, \cdots, d\}$ and $x \in I_j$:

$$\frac{\phi_\mu^{(j)}(x)}{\sqrt{m(x)}} = \frac{2^{\frac{\mu}{2b_0}} K_j(\mu)}{\sqrt{\pi} \prod_{i=1}^j \sigma_i} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \cos\left(\frac{\pi\mu}{2b_0}(1 + o(1))\right) \left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right],$$

and

$$\left(\frac{\phi_\mu^{(j)}(x)}{\sqrt{m(x))}}\right)' = -\frac{2^{\frac{\mu}{2b_0}} \tilde{K}_j(\mu)}{\sigma_j \prod_{i=1}^j \sigma_i} \sqrt{\frac{b_j + 2\mu}{\pi}} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \sin\left(\frac{\pi\mu}{2b_0}(1 + o(1))\right) \left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right].$$

Additionally, $\prod_{i=1}^j \sigma_i \wedge \sigma_{i-1} \leq K_j(\mu) \leq \prod_{i=1}^j \sigma_i \vee \sigma_{i-1}$, and the same bounds hold for $\tilde{K}_j$.

Corollary 4.4.4. *For n large enough, we have $\mu_n = b_0(2n+1)(1+o(1))$ and*

$$\left|\partial_\mu \left[\phi_\mu^{(d)}(c)\right]_{\mu=\mu_n}\right| = \sqrt{\pi m(c)} \frac{2^{\frac{\mu_n}{2b_0}-1} K_d(\mu_n)}{b_0 \prod_{i=1}^d \sigma_i} \Gamma\left(\frac{1}{2} + \frac{\mu_n}{2b_0}\right)(1 + o(1)). \tag{4.4.6}$$

4.5 Interplay between the T-OU and the threshold Skew OU

Let us introduce the process Y solution of the following SDE:

$$Y_t = y_0 + \int_0^t \tilde{a}(Y_s) - \tilde{b}(Y_s) Y_s \, \mathrm{d}s + \int_0^t \tilde{\sigma}(Y_s) \, \mathrm{d}B_s + \sum_{j=1}^d \tilde{\beta}_j L_t^{\tilde{r}_j}(Y), \quad t \geq 0, \tag{4.5.1}$$

where $\tilde{a}$, $\tilde{b}$, and $\tilde{\sigma}$ are piecewise constant coefficients, discontinuous at $\tilde{r}_j$, $j \in \{1, \cdots, d\}$, defined as (4.2.1) and $\tilde{\beta}_j \in (-1, 1)$. We refer to this process as the Threshold-Skew-OU (TS-OU) process. Strong existence and uniqueness of the SDE (4.5.1) are induced by the results in [66].

We can find a function that establishes a relationship between the process Y and the process X solution of (4.1.1). We provide an explicit example illustrating the link between the Skew-OU process and X solution of (4.1.1) when $d = 1$ (see Remark 4.5.3 for further details on the case $d > 1$). Suppose that Z is a Skew-OU process, solution of:

$$Z_t = z_0 + \int_0^t \tilde{a} - \tilde{b} Z_s \, \mathrm{d}s + \tilde{\sigma} B_t + \tilde{\beta} L_t^{\tilde{r}}(Z), \quad t \geq 0,$$

with $(\tilde{a}, \tilde{b}, \tilde{\sigma}, \tilde{r}) \in \mathbb{R} \times (\mathbb{R}_+^\star)^2 \times \mathbb{R}$, $\tilde{\beta} \in (-1, 1)$ and $L_t^{\tilde{r}}(Z)$, the symmetric local time of the process Z at $\tilde{r}$.

Proposition 4.5.1. *Let H be the function given by $H(x) = \frac{x - \tilde{r}}{\alpha(x)} + \tilde{r}$, where $\alpha(x) = \mathbb{1}_{x > \tilde{r}} + \alpha_0 \mathbb{1}_{x \leq \tilde{r}}$, and such that $\alpha_0 = \frac{1+\tilde{\beta}}{1-\tilde{\beta}}$. There exists a process X such that $H^{-1}(Z_t) = X_t$ for all $t \in \mathbb{R}_+$, where X is the solution of (4.1.1) with $x_0 = \alpha(z_0)(z_0 - \tilde{r}) + \tilde{r}$, threshold $\tilde{r}$, and drift and volatility parameters:*

$$[a_0; a_1] = [\tilde{a}\alpha_0 + \tilde{b}\tilde{r} - \alpha_0 \tilde{b}\tilde{r}; \tilde{a} + \tilde{b}\tilde{r} - \tilde{b}\tilde{r}], \quad [b_0; b_1] = \tilde{b}[1; 1], \quad [\sigma_0; \sigma_1] = \tilde{\sigma}[\alpha_0; 1]. \tag{4.5.2}$$

Proof. Applying Itô-Tanaka formula in Corollary 1.1.8, we obtain:

$$\mathrm{d}H(X_t) = \frac{1}{\alpha(X_t)} \mathbb{1}_{(X_t \neq \tilde{r})} dX_t + \frac{1}{2}\left(1 + \frac{1}{\alpha_0}\right) \mathbb{1}_{(X_t = \tilde{r})} dX_t + \frac{1}{2}\left(1 - \frac{1}{\alpha_0}\right) \mathrm{d}L_t^{\tilde{r}}(X)$$

From Proposition 1.3.9, we have:

$$\mathrm{d}H(X_t) = \frac{1}{\alpha(X_t)} \left[\left(\tilde{a}\alpha(X_t) + \tilde{b}\tilde{r} - \alpha(X_t)\tilde{b}\tilde{r} - \tilde{b}X_t\right) \mathrm{d}t + \tilde{\sigma}\alpha(X_t)\, \mathrm{d}B_t\right] + \frac{\alpha_0 - 1}{2\alpha_0}\, \mathrm{d}L_t^{\tilde{r}}(X),$$

We also prove that $L_t^{\tilde{r}}(X) = \frac{\alpha_0+1}{2\alpha_0} L_t^{\tilde{r}}(Z)$ by using Itô-Tanaka formula. Then, by setting $Z_t = H(X_t)$ a.s., the conclusion holds. □

Using Proposition 4.5.1, one can easily prove the following statement:

Proposition 4.5.2. *The transition probability density p_Z of the process Z, killed at the constant boundary c, and the density of $\tau_c(Z)$ exist and are given by:*

$$p_Z(x,t) = \frac{p(H(x),t)}{\alpha(x)} \quad \textit{and} \quad p_{\tau_c(Z)}(t) = p_{\tau_{H^{-1}(c)}(X)}(t),$$

where X is the process solution of (4.1.1) *with parameters* (4.5.2)*, killed at the constant boundary $H^{-1}(c)$.*

Remark 4.5.3. *For further details on the construction of the function H, we refer to Section 2.1 in [36]. In a similar way, the transition probability density and hitting time density of the process Y solution of* (4.5.1) *can be obtained from Theorem 4.3.1 and Theorem 4.4.2.*

4.6 Numerical experiments

In this section, we illustrate the explicit formulations obtained for the transition probability density (4.3.2) and the first hitting time density (4.4.3). We compare these expressions with an approximation obtained using a Monte Carlo method, employing 10^6 simulations with time step $\Delta_t = 10^{-3}$ in the Euler Scheme. To numerically compare these expressions, we use the parameters set given in the table below. Numerical simulations are performed using Python on a processor running at 3.6 GHz with 8 cores.

a_0	b_0	σ_0
-1	1	1

a_1	b_1	σ_1
-1	2	1.5

a_2	b_2	σ_2
0	1	1

r_1	r_2	x_0	c
-1	0	-1.5	1

Table 4.1: Simulations parameters.

Since the expressions of these densities admit a series representation, we approximate these solutions by truncating this series at $N \in \mathbb{N}$, *i.e*,

$$p_N(x,t) = \frac{1}{m(x_0)} \sum_{n=1}^{N} \frac{e^{-\mu_n t}}{N(\mu_n)} \phi_n^{(j)}(x_0)\phi_n^{(i)}(x), \tag{4.6.1}$$

and

$$p_{\tau_c,N}(t) = -\frac{m(c)}{m(x_0)} \sum_{n=1}^{N} e^{-\mu_n t} \frac{\phi_n^{(j)}(x_0)}{\partial_\mu \left[\phi_n^{(d)}(c)\right]_{\mu=\mu_n}}. \tag{4.6.2}$$

Here, the zeros $(\mu_n)_{n\leq N}$ are obtained using a dichotomy method on the interval $[0, 30]$. The coefficients $(\alpha_\mu^{(j)}, \beta_\mu^{(j)})_{j=1}^d$ are computed by explicitly solving (4.3.1). Note that these coefficients can also be computed by solving this system numerically.

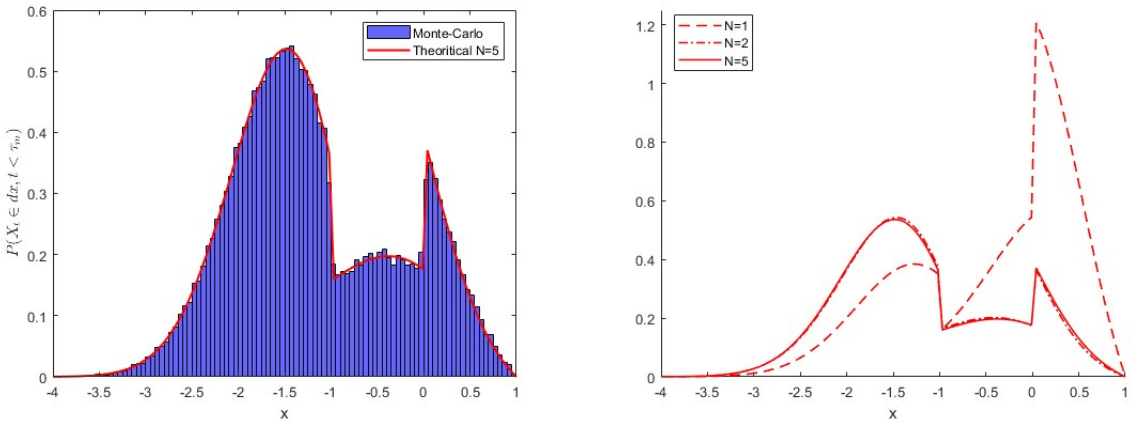


Figure 4.1: The left figure compares the truncated transition probability density (4.6.1) for $t = 1$ and $N = 5$ with a Monte Carlo method. The figure on the right compares the truncated density (4.6.1) for $N = 1$, $N = 2$, and $N = 5$.

The truncated density (4.6.1) fits well with the approximation obtained by the Monte Carlo method (see Figure 4.1). Evaluating the function p_N for $N = 5$ takes 95 ms (milliseconds). Furthermore, as seen in Figure 4.1, the truncated density $p_N(t, .)$ seems to approximate the actual density quite well, even for $N = 2$. It is worth noting that the approximation also depends on the parameters of the process listed in Table 5.1.

Alternative formulations for the transition probability density of the killed process can be explored numerically. For instance, in [26], the authors propose a semi-analytical formulation for the Fokker-Planck equation associated to a multi-Skew Brownian motion. Their method can be adapted to the solution of equation (4.2.8).

Remark 4.6.1. *Note that the parameter t influences the convergence speed of the series. Here, we have chosen $t = 1$. With a smaller t, the approximation is less accurate for small N. For example, for $t = 10^{-2}$, the approximation seems to approximate the actual density well for $N = 30$, and by increasing the search window for $(\mu_n)_{n\leq N}$ to $[0, 50]$.*

For the first hitting time density, we propose different approximation methods. First, we use the expression (4.6.2) for various values of N. Then, we suggest a method based on Laplace transform inversion of (4.4.2), using the Talbot method. For further details on this approach, see [1] and [100]. This method is already implemented in the mpmath library of Python.

The last method used is inspired by the results in [6] for the first hitting time of the OU process (see Definition 1.4.1). Making use of the holomorphic property of the Laplace transform of τ_c, we can prove that

$$p_{\tau_c}(t) = \frac{2}{\pi} \int_0^\infty \cos(\mu t) \mathbb{E}_{x_0} \left[\cos\left(\mu \tau_c\right) \right] d\mu, \tag{4.6.3}$$

where $\mathbb{E}_{x_0} \left[\cos\left(\mu \tau_c\right) \right] = \mathrm{Re}\left(\mathbb{E}_{x_0} \left[e^{i\mu\tau_c} \right] \right) = \mathrm{Re}\left(u_c(-i\mu, x_0) \right)$. This formulation provides an integral representation for the first hitting time density. In the same way as [6], we refer to this representation as the cosine transform. By using the results in [2], we approximate the integral, not by a Riemann sum, but by the following approximation:

$$p_{\tau_c, N_{\text{approx}}}(t) = \frac{e^{A/2}}{2t} u_c\left(\frac{A}{2t}, x_0\right) + \frac{e^{\frac{A}{2}}}{t} \sum_{k=1}^{N_{\text{approx}}} \mathrm{Re}\left[u_c\left(\frac{A}{2t} + \frac{k\pi i}{t}, x_0\right) \right], \tag{4.6.4}$$

with $A > 0$ and $N_{\text{approx}} \in \mathbb{N}^\star$. Furthermore, the question still arises of selecting suitable values for A and N. In the same way as in [6], we choose $A < 18.1$ to ensure that the discretization error is of the order 10^{-7}. We also set $N_{\text{approx}} = 500$. Suitable choices for the parameters N_{approx} and A are discussed in [2].

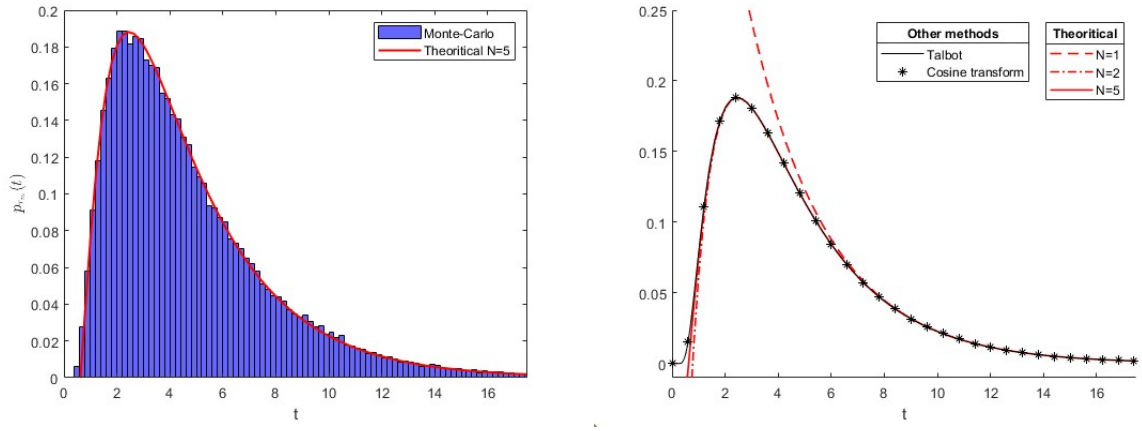


Figure 4.2: The figure on the left compares the expression (4.6.2) for $N = 5$ with a Monte Carlo method. The figure on the right compares the hitting time density obtained through cosine transform (4.6.3), Laplace transform inversion (Talbot method) of (4.4.2), and the truncated density (4.6.2) for $N = 1$, $N = 2$, and $N = 5$.

In the table below, we list the numerical evaluation times at a point of the hitting time density for each of the methods used.

Truncated series ($N = 5$)	Laplace inversion (Talbot)	Cosine transform
110 ms	2.69 s	31.3 s

Table 4.2: Computation time for the first hitting time density according to the different methods: truncated series, Laplace inversion (Talbot), cosine transform.

As with the transition probability density, the truncated series (4.6.2) for the first hitting time density closely matches the approximation obtained from the Monte Carlo method (see Figure 4.2). The cosine transform and Laplace transform inversion methods yield similar results. Note that the truncated series seems to approximate the density well for $N = 5$.
The primary distinction among these methods lies in their computation time (see Table 4.2). The truncated series method is notably faster than the other two. However, the Laplace transform inversion method also proves to be quite efficient in terms of computation time. Here, the truncated density is evaluated for $N = 5$. Note that, N has small effect on the computation speed of the method. Furthermore, what takes the most time here is the computation of μ-zeros of $\mu \mapsto \phi_\mu^{(d)}(c)$ by the Dichotomy method. But, these zeros are independent from the variable t and can be computed only once outside the function.

Note that for the approximation $p_{\tau_c,N}$, the series convergence breaks down for very small values of t, causing $p_{\tau_c,N}$ to tend towards $-\infty$ or $+\infty$ in such cases. In contrast, the Laplace inversion and cosine transform methods converge to 0 for small values of t. To address the issue of divergence for small t, as the Talbot method remains quick enough, we can use the Laplace transform inversion method for small values of t and the truncated series for higher values.

4.7 Proofs

This section is devoted to the proofs of Theorem 4.3.1, Theorem 4.4.1 and Theorem 4.4.2.

4.7.1 Proof of Theorem 4.3.1

The idea is to use a spectral decomposition method such that the explicit solution of the Fokker-Planck equation (4.2.8) can be expressed as follows:

$$p(x,t) = \sum_{n=1}^{\infty} c_n(t) e_n(x), \tag{4.7.1}$$

where $(e_n)_{n\geq 1}$ are eigenfunctions of $\mathcal{L}^\star$. To apply this spectral decomposition method, we use a technique from classical functional analysis (see [21]). We construct the operator $\mathcal{L}^\star$ over a well-chosen domain $\mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star)$, and then, based on the properties of

$(\mathcal{L}^\star, \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star))$, we prove that the solution of the equation (4.2.8) can be expressed as a spectral decomposition of eigenfunctions and eigenvalues.

Let us denote $\mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star)$ the set

$$\left\{ f \in L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right) | f(c) = 0, \frac{f}{m}(.) \in H^1\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right) \right.$$
$$\left. , S^{-1}(.)\left(\frac{f}{m}(.)\right)' \in H^1\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right)\right\}.$$

The operator $\mathcal{L}^\star$ is defined by

$$\mathcal{L}^\star : \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star) \to L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right)$$
$$f \to \left[\mathcal{L}^\star f : x \mapsto \left(S^{-1}(x)\left(\frac{f(x)}{m(x)}\right)'\right)' = \frac{1}{2}(\sigma(x)^2 f(x))'' + ((b(x)x - a(x))f(x))'\right].$$

Here, $\mathcal{L}^\star$ is constructed over a weighted space. Within this domain, the operator is Hermitian, *i.e.*,

$$\forall f, g \in \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star), \quad \langle \mathcal{L}^\star f, g\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)} = \langle f, \mathcal{L}^\star g\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)}.$$

The following lemma provides necessary properties on $(\mathcal{L}^\star, \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star))$ to justify the spectral decomposition method.

Lemma 4.7.1. *The space* $L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right)$ *admits a countable Hilbertian basis of eigenvectors of* $\mathcal{L}^\star$ *for the scalar product* $\langle ., .\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)}$. *Additionally, the spectrum of* $\mathcal{L}^\star$ *is contained in* $(-\infty, 0]$.

Proof. Since directly proving this result on $(\mathcal{L}^\star, \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star))$ is not straightforward, we use a common method by introducing a new operator $\mathbf{T}$ defined by

$$\mathbf{T} : L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right) \to L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right)$$
$$g \mapsto f_g$$

where f_g is solution of:

$$(\star) = \begin{cases} (\frac{\sigma(x)^2}{2} f_g(x))'' + ((b(x)x - a(x))f_g(x))' = g(x), \quad x \in (-\infty, c), \\ f_g \text{ verifies (4.2.7)}, \\ f_g(c) = 0, \\ f_g, f_g' \in L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right). \end{cases}$$

Using a similar reasoning as Lemma 2 in [8], we prove that $f_g(x) = 2\int_{-\infty}^{c} K(x,z)g(z)dz$ where $K \in L^2\left((-\infty,c]\times(-\infty,c], \frac{1}{\sqrt{m(.)}}\times\frac{1}{\sqrt{m(.)}}\right)$.

Then, the operator $\mathbf{T} : L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right) \to L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)$ is well defined, continuous and compact. Applying Theorem VIII.20 in [21], we conclude.

To complete the proof, we need to demonstrate that the spectrum is contained in $(-\infty, 0]$. Let us assume that $(f, -\mu)$ is a couple of eigenfunction and eigenvalue of $\mathcal{L}^\star$. Then we have:

$$\langle \mathcal{L}^\star f, f\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)} = -\mu||f||^2_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)}.$$

Using (4.2.6) and integration by part, we obtain:

$$\langle \mathcal{L}^\star f, f\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)} = -\int_{-\infty}^{c} S^{-1}(x)\left[\left(\frac{f(x)}{m(x)}\right)'\right]^2 dx.$$

Then, as the derivative of the scale function is strictly positive, $-\mu \leq 0$ and the conclusion holds. □

Let ϕ_μ be an a eigenvector of $\mathcal{L}^\star$ with eigenvalue $-\mu$. Then, for all $x \leq c$, the couple $(\phi_\mu, -\mu)$ is solution of the following equation on $\mathrm{Dom}^W(\mathcal{L}^\star)$:

$$\begin{aligned}\mathcal{L}^\star\phi_\mu(x) = -\mu\phi_\mu(x) &\Leftrightarrow \left(S^{-1}(x)\left(\frac{\phi_\mu(x)}{m(x)}\right)'\right)' = -\mu\phi_\mu(x),\\ &\Leftrightarrow \frac{1}{2}(\sigma(x)^2\phi_\mu(x))'' + ((b(x)x - a(x))\phi_\mu(x))' = -\mu\phi_\mu(x). \quad (4.7.2)\end{aligned}$$

We rewrite $\phi_\mu(x)$ as $\phi_\mu(x) = \sum_{j=0}^{d}\phi_\mu^{(j)}(x)\mathbb{1}_{I_j}(x)$. For all $j \in \{0,\cdots,d\}$ such that $x \in I_j$, the function $\tilde{\phi}_\mu^{(j)}(x) = \frac{1}{\sqrt{m(x)}}\phi_\mu^{(j)}(x) \in L^2((-\infty,c])$, satisfies the following equation:

$$(\tilde{\phi}_\mu^{(j)})''(x) + \left(\frac{b_j}{\sigma_j^2} - \left(\frac{(x-\frac{a_j}{b_j})b_j}{\sigma_j^2}\right)^2 + \frac{2\mu}{\sigma_j^2}\right)\tilde{\phi}_\mu^{(j)}(x) = 0, \quad \text{on } I_j. \quad (4.7.3)$$

Solutions of (4.7.3) are well known in the literature (see [68]). We obtain that:

$$\phi_\mu^{(j)}(x) = \sqrt{m(x)}\left[\alpha_\mu^{(j)} D_{\frac{\mu}{b_j}}\left(-\left(x-\frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right) + \beta_\mu^{(j)} D_{\frac{\mu}{b_j}}\left(\left(x-\frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right)\right], \quad \text{on } I_j,$$

where $(\alpha_\mu^{(j)}, \beta_\mu^{(j)}) \in \mathbb{R}^2$. As $\phi_\mu \in L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)$, we choose $(\alpha_\mu^{(0)}, \beta_\mu^{(0)}) = (1,0)$.

Since $\phi_\mu \in \mathrm{Dom}^{\mathrm{W}}(\mathcal{L}^\star)$, using the transmission conditions, for all $j \in \{1,\cdots,d\}$, the couple $(\alpha_\mu^{(j)}, \beta_\mu^{(j)})$ satisfies the following system of equations:

$$\begin{cases} \sigma_j^2\phi_\mu^{(j)}(r_j) = \sigma_{j-1}^2\phi_\mu^{(j-1)}(r_j), \\ \frac{\sigma_j^2}{2}(\phi_\mu^{(j)})'(r_j) + (b_j r_j - a_j)\phi_\mu^{(j)}(r_j) = \frac{\sigma_{j-1}^2}{2}(\phi_\mu^{(j-1)})'(r_j) + (b_{j-1}r_j - a_{j-1})\phi_\mu^{(j-1)}(r_j), \end{cases}$$

with $(\alpha_\mu^{(0)}, \beta_\mu^{(0)}) = (1, 0)$.

Moreover, from the boundary condition, we have $\phi_\mu(c) = 0$, *i.e.*, $\phi_\mu^{(d)}(c) = 0$. In the following we denote by $(\mu_n)_{n\geq 1}$ the ordered sequence of zeros of $\mu \mapsto \phi_\mu^{(d)}(c)$. Note that from Lemma 4.7.1, the eigenvalues $(-\mu_n)_{n\geq 1}$ are negative. Normalizing the eigenfunctions $(\phi_n)_{n\geq 1}$, we obtain:

$$e_n(x) = \frac{\sum_{j=0}^d \phi_n^{(j)}(x)\mathbb{1}_{I_j}(x)}{\sqrt{\sum_{j=0}^d \int_{I_j} \left[\phi_n^{(j)}(y)\right]^2 m^{-1}(y)dy}}.$$

Furthermore, classical results on the Parabolic Cylinder functions allow us to express the denominator without the integral.

Lemma 4.7.2. *Under the assumptions of Theorem 4.3.1, we have:*

$$N(\mu_n) = \sum_{j=0}^d \int_{I_j} \left(\phi_n^{(j)}(y)\right)^2 m^{-1}(y)dy = \frac{\sigma_d^2}{2m(c)} \partial_x \left[\phi_n^{(d)}(x)\right]_{x=c} \partial_\mu \left[\phi_n^{(d)}(c)\right]_{\mu=\mu_n}.$$

Proof. The proof follows by using Remark 4.8.1. It is identical to Lemma 3 in [8]. □

Then $(e_n)_{n\geq 1}$ forms an Hilbertian basis of $L^2\left((-\infty, c], \frac{1}{\sqrt{m(.)}}\right)$. Let us return to the initial problem of solving the Fokker-Planck equation (4.2.8). We can now decompose p in the form (4.7.1). Now, the only remaining task is to explicitly provide the function $t \mapsto c_n(t)$ for $n \geq 1$.
Using the initial condition in (4.2.8), we have $p(x, 0) = \sum_{n=1}^\infty c_n(0)e_n(x)$, then for $n \in \mathbb{N}^\star$:

$$\begin{aligned} c_n(0) &= \langle p(., 0), e_n\rangle_{L^2\left((-\infty,c], \frac{1}{\sqrt{m(.)}}\right)} = \sum_{j=0}^d \int_{I_j} \frac{p(z, 0)e_n(z)}{m(z)}dz \\ &= \frac{\sum_{j=0}^d \frac{\phi_n^{(j)}(x_0)}{m(x_0)}\mathbb{1}_{I_j}(x_0)}{\sqrt{N(\mu_n)}} \end{aligned}$$

It remains to find $c_n(t)$ for $t > 0$. By differentiating (4.7.1) and using $\partial_t p(x, t) = -\mathcal{L}^\star p(x, t) = -\sum_{n=1}^\infty c_n(t)\mu_n e_n(x)$, one can derive $c_n'(t) = -c_n(t)\mu_n$, leading to $c_n(t) = c_n(0)e^{-\mu_n t}$ due to the uniqueness of the decomposition, thus the conclusion holds.

4.7.2 Proof of Theorem 4.4.1

The method for computing the solution is inspired by the previous results obtained for the transition probability density of the killed process. Making use of the formulation of the infinitesimal generator $\mathcal{L}$ in (4.2.5), we have:

$$\mathcal{L}u_c(\mu, x_0) = \mu u_c(\mu, x_0) \iff \frac{1}{m(x_0)}\partial_{x_0}\left(S^{-1}(x_0)\partial_{x_0}(u_c(\mu, x_0))\right) = \mu u_c(\mu, x_0).$$

Then using (4.7.2), one can remark that $\phi_{-\mu}(x_0) = m(x_0)u_c(\mu, x_0)$ is solution of $\mathcal{L}^\star\phi_{-\mu}(x_0) = \mu\phi_{-\mu}(x_0)$. Let us write $u_c(\mu, x_0) = \sum_{j=0}^{d} u_c^{(j)}(\mu, x_0)\mathbb{1}_{I_j}(x_0)$, hence, for $x_0 \in I_j$, there exists $(\tilde{\alpha}_\mu^{(j)}, \tilde{\beta}_\mu^{(j)}) \in \mathbb{R}^2$ such that we have:

$$u_c^{(j)}(\mu, x_0) = \frac{1}{\sqrt{m(x_0)}}\left[\tilde{\alpha}_\mu^{(j)} D_{-\frac{\mu}{b_j}}\left(-\left(x - \frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right) + \tilde{\beta}_\mu^{(j)} D_{-\frac{\mu}{b_j}}\left(\left(x - \frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right)\right].$$

Since $\lim_{x_0\to-\infty} u_c(\mu, x_0) = 0$, we choose $\tilde{\beta}_\mu^{(0)} = 0$. For all $j \in \{1, \cdots, d\}$, the regularity of the Laplace transform leads to the system:

$$\begin{cases} \sigma_j^2 u_c^{(j)}(\mu, r_j) = \sigma_{j-1}^2 u_c^{(j-1)}(\mu, r_j), \\ \frac{\sigma_j^2}{2}\partial_{x_0} u_c^{(j)}(\mu, x_0)|_{x_0=r_j} + (b_j r_j - a_j)u_c^{(j)}(\mu, r_j) = \frac{\sigma_{j-1}^2}{2}\partial_{x_0} u_c^{(j-1)}(\mu, x_0)|_{x_0=r_j} + (b_{j-1}r_j - a_{j-1})u_c^{(j-1)}(\mu, r_j). \end{cases}$$

Note that this system is the same as (4.3.1). Then, for all $j \in \{0, \cdots, d\}$, we obtain $(\tilde{\alpha}_\mu^{(j)}, \tilde{\beta}_\mu^{(j)}) = \tilde{\alpha}_\mu^{(0)}(\alpha_{-\mu}^{(j)}, \beta_{-\mu}^{(j)})$. The solution still depends on $\tilde{\alpha}_\mu^{(0)}$ which needs to be determined. Furthermore, using that $u_c(\mu, c) = 1$, we obtain $\tilde{\alpha}_\mu^{(0)} = \frac{1}{\phi_{-\mu}(c)}$ and the conclusion holds.
Using Lemma 2.1 in [73], the holomorphic property of the Laplace transform follows directly.

4.7.3 Proof of Theorem 4.4.2

Using the scale function (see Proposition 1.3.18), defined by $s(x) = \int^x S(y)dy$, we can establish that the process $\bar{X}$ is recurrent. The regime of the process $\bar{X}$ only depends on its behavior over the interval I_0. As $b_0 > 0$, we can verify that $\lim_{x\to-\infty} s(x) = -\infty$, which implies that for all $x_0, c \in \mathbb{R}$ such that $x_0 < c$, we have $\mathbb{P}_{x_0}(\tau_c < +\infty) = 1$ (see the proof of Proposition 1.3.18). Let us now prove that $\mathbb{P}_{x_0}(\tau_c \in dt)$ is absolutely continuous with respect to the Lebesgue measure. For every $B \in \mathcal{B}(\mathbb{R})$ such that B is Lebesgue-negligible, we have:

$$\mathbb{P}_{x_0}(\tau_c(X) \in B) \leq \mathbb{P}_{x_0}(\tau_{[x_0-\epsilon,c]}(X) \in B) \tag{4.7.4}$$

where $\epsilon > 0$ and $\tau_{[x_0-\epsilon,c]}(X)$ is the first exit time of the process X from the interval $(x_0 - \epsilon, c)$.

We construct a process Y such that it admits a state space $[l_1, l_2]$ with $l_1 \leq x_0 - \epsilon < c \leq l_2$ and such that $Y_{.\wedge\tau_{[x_0-\epsilon,c]}(Y)} \overset{law}{=} X_{.\wedge\tau_{[x_0-\epsilon,c]}(X)}$. Here, $\{l_1\}$ and $\{l_2\}$ are chosen as non-singular boundaries (see Definition 1.3.13). As an example, Y can be instantaneously reflected on the boundaries $\{l_1\}$ and $\{l_2\}$:

$$Y_t = x_0 + \int_0^t a(Y_s) - b(Y_s)Y_s ds + \int_0^t \sigma(Y_s)dB_s + L_t^{l_1}(Y) - L_t^{l_2}(Y).$$

Strong existence and uniqueness for this SDE follows from Theorem 4.5 in [9]. For more details on reflected SDEs, we refer to [89]. Now using (4.7.4), we have:

$$\mathbb{P}_{x_0}(\tau_c(X) \in B) \leq \mathbb{P}_{x_0}(\tau_{x_0-\epsilon}(Y) \in B) + \mathbb{P}_{x_0}(\tau_c(Y) \in B). \tag{4.7.5}$$

Then, from Theorem 6.1 in [57], $\tau_{x_0-\epsilon}(Y)$ and $\tau_c(Y)$ admit a density with respect to the Lebesgue measure. Using the inequality (4.7.5), we prove the existence of the density p_{τ_c}.

We now focus on computing the explicit expression for p_{τ_c}. Recall that p and p_{τ_c} are linked by:

$$p_{\tau_c}(t) = -\frac{\sigma_d^2}{2}\partial_x p(x,t)|_{x=c}.$$

It suffices now to compute the derivative in space of the function p at the point c. Since the function p admits a series representation, obtaining formulation (4.4.3) requires proving that we can differentiate each term of the series in (4.3.2) separately.

Let $x \in [c-\epsilon, c]$, we want to show that:

$$\frac{1}{m(x_0)}\sum_{n=1}^{\infty}\frac{e^{-\mu_n t}}{N(\mu_n)}\partial_x\left[\phi_n^{(d)}(x)\right]\phi_n^{(j)}(x_0), \tag{4.7.6}$$

converges uniformly. Using (4.3.3) and Corollary 7 in [8], we have:

$$\frac{1}{m(x_0)}\sum_{n=1}^{\infty}\left|\frac{e^{-\mu_n t}}{N(\mu_n)}\partial_x\left[\phi_n^{(d)}(x)\right]\phi_n^{(j)}(x_0)\right| \leq \frac{\sigma_d^2 m(c)}{2m(x_0)}\sum_{n=1}^{\infty}\left|e^{-\mu_n t}\frac{\phi_n^{(j)}(x_0)}{\partial_\mu\left[\phi_n^{(d)}(c)\right]}\right|.$$

Then, from Corollary 4.4.4, we have, for large n:

$$\left|\frac{e^{-\mu_n t}}{N(\mu_n)}\partial_x\left[\phi_n^{(d)}(x)\right]\phi_n^{(j)}(x_0)\right| \leq Ce^{-b_0(2n+1)t},$$

with C a strictly positive constant. The series (4.7.6) converges uniformly and the conclusion holds.

4.8 Appendix

4.8.1 On the Parabolic Cylinder functions

In this section, for the convenience of the reader, we recall the definition and some properties of the Parabolic Cylinder function.

For all $x, \mu \in \mathbb{R}$, the Parabolic Cylinder function, denoted $D_\mu(.)$, is the solution of the differential equation:

$$\begin{cases} y''(x) + \left(\mu + \frac{1}{2} - \frac{1}{4}x^2\right) y(x) = 0, \\ y(x) \underset{x\to+\infty}{\sim} x^\mu e^{-x^2/4}. \end{cases} \tag{4.8.1}$$

In the case $\mu > -1$, the Parabolic Cylinder function can be expressed as follows (see [68] p290):

$$D_\mu(x\sqrt{2}) = \frac{2^{\frac{\mu}{2}+1}e^{\frac{x^2}{2}}}{\sqrt{\pi}}\int_0^\infty e^{-t^2}t^\mu \cos\left(-2xt + \frac{\mu\pi}{2}\right)dt. \tag{4.8.2}$$

Furthermore, we have the following relations:

$$\partial_x D_\mu(x) = \mu D_{\mu-1}(x) - \frac{x}{2} D_\mu(x). \tag{4.8.3}$$

and, from [68] p286:

$$D_\mu(x\sqrt{2})D_{\mu-1}(-x\sqrt{2}) + D_\mu(-x\sqrt{2})D_{\mu-1}(x\sqrt{2}) = \frac{\sqrt{2\pi}}{\Gamma(1-\mu)}. \tag{4.8.4}$$

For μ large enough, we have:

$$D_\mu(x) = \frac{2^{\frac{\mu}{2}}}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2}\right)\cos\left(x\sqrt{\mu+\frac{1}{2}} - \frac{\pi\mu}{2}\right)\left(1+O(\mu^{-\frac{1}{4}})\right). \tag{4.8.5}$$

Moreover, using (4.8.1) and integration by parts:

$$\int_{-\infty}^{x} D_\mu^2\left(-y\sqrt{2}\right)dy = -\frac{\sqrt{2}}{2}\left[D_\mu\left(-x\sqrt{2}\right)\partial_\mu\left(\mu D_{\mu-1}\left(-x\sqrt{2}\right)\right) - \mu D_{\mu-1}\left(-x\sqrt{2}\right)\partial_\mu D_\mu\left(-x\sqrt{2}\right)\right]. \tag{4.8.6}$$

Remark 4.8.1. *If $f:\mathbb{R}^2\to\mathbb{R}$ is a solution of $\partial_{x^2}^2 f(x,\mu) + \left(\mu+\frac{1}{2}-\frac{1}{4}x^2\right)f(x,\mu) = 0$, then for all $x,\mu,c_1,c_2\in\mathbb{R}$,*

$$\begin{aligned}\int_{c_1}^{c_2} f(x,\mu)^2\,dx = &-\frac{1}{2}\left[f(c_2,\mu)\partial_x\partial_\mu f(c_2,\mu) - \partial_x f(c_2,\mu)\partial_\mu f(c_2,\mu) - f(c_1,\mu)\partial_x\partial_\mu f(c_1,\mu)\right.\\ &\left.+\partial_x f(c_1,\mu)\partial_\mu f(c_2,\mu)\right].\end{aligned}$$

4.8.2 Homogeneous and particular solutions of the weighted Sturm-Liouville problem

For all $j\in\{0,\cdots,d\}$ and $x\in I_j$, we have $\phi_\mu^{(j)}(x) = \sqrt{m(x)}\tilde{\phi}_\mu^{(j)}(x)$, where:

$$\tilde{\phi}_\mu^{(j)}(x) = \alpha_\mu^{(j)} D_{\frac{\mu}{b_j}}\left(-\left(x-\frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right) + \beta_\mu^{(j)} D_{\frac{\mu}{b_j}}\left(\left(x-\frac{a_j}{b_j}\right)\frac{\sqrt{2b_j}}{\sigma_j}\right),$$

with $\tilde{\phi}$ solution of:

$$(\tilde{\phi}_\mu^{(j)})''(x) + \left(\frac{b_j}{\sigma_j^2} - \left(\frac{(x-\frac{a_j}{b_j})b_j}{\sigma_j^2}\right)^2 + \frac{2\mu}{\sigma_j^2}\right)\tilde{\phi}_\mu^{(j)}(x) = 0. \tag{4.8.7}$$

Additionally, for all $j\in\{1,\cdots,d\}$, the couple $(\alpha_\mu^{(j)},\beta_\mu^{(j)})$ verifies:

$$\begin{cases}\tilde{\phi}_\mu^{(j)}(r_j) = \frac{\sigma_{j-1}}{\sigma_j}\tilde{\phi}_\mu^{(j-1)}(r_j),\\ (\tilde{\phi}_\mu^{(j)})'(r_j) = \frac{\sigma_{j-1}}{\sigma_j}\left[(\tilde{\phi}_\mu^{(j-1)})'(r_j) + r_j\left(\frac{b_{j-1}}{\sigma_{j-1}^2} - \frac{b_j}{\sigma_j^2}\right)\tilde{\phi}_\mu^{(j-1)}(r_j) + \left(\frac{a_j}{\sigma_j^2} - \frac{a_{j-1}}{\sigma_{j-1}^2}\right)\tilde{\phi}_\mu^{(j-1)}(r_j)\right],\end{cases} \tag{4.8.8}$$

with $(\alpha_\mu^{(0)}, \beta_\mu^{(0)}) = (1, 0)$. Making use of the method on p284 in [68], we provide a different formulation for the solution of equation (4.8.7). For all $j \in \{1, \cdots, d\}$ and $x \in I_j$, we have:

$$\tilde{\phi}_\mu^{(j)}(x) = \tilde{\phi}_{H,\mu}^{(j)}(x) + \tilde{\phi}_{P,\mu}^{(j)}(x),$$

where:

$$\begin{aligned}\tilde{\phi}_{H,\mu}^{(j)}(x) =& \frac{\sigma_{j-1}}{\sigma_j}\tilde{\phi}_\mu^{(j-1)}(r_j)\cos\left((x-r_j)\frac{\sqrt{b_j+2\mu}}{\sigma_j}\right) + \sin\left((x-r_j)\frac{\sqrt{b_j+2\mu}}{\sigma_j}\right) \\ &\times\left(\frac{\sigma_{j-1}(\tilde{\phi}_\mu^{(j-1)})'(r_j) + r_{j-1}\sigma_{j-1}\left(\frac{b_{j-1}}{\sigma_{j-1}^2} - \frac{b_j}{\sigma_j^2}\right)\tilde{\phi}_\mu^{(j-1)}(r_j) + \sigma_{j-1}\left(\frac{a_j}{\sigma_j^2} - \frac{a_{j-1}}{\sigma_{j-1}^2}\right)\tilde{\phi}_\mu^{(j-1)}(r_j)}{\sqrt{b_j+2\mu}}\right),\end{aligned} \tag{4.8.9}$$

and

$$\tilde{\phi}_{P,\mu}^{(j)}(x) = \frac{b_j^2}{\sigma_j^3\sqrt{b_j+2\mu}}\int_{r_j}^x \left(y - \frac{a_j}{b_j}\right)^2 \tilde{\phi}_\mu^{(j)}(y)\sin\left((x-y)\frac{\sqrt{b_j+2\mu}}{\sigma_j}\right)dy. \tag{4.8.10}$$

Here, $\tilde{\phi}_{H,\mu}^{(j)}$ is the homogeneous solution and $\tilde{\phi}_{P,\mu}^{(j)}$ is the particular solution of (4.8.7).

4.8.3 Asymptotic expansions

This section is dedicated to the proof of Lemma 4.4.3 and Corollary 4.4.4. In this section, without loss of generality, we suppose that $a(.) = 0$. If $a(.) \neq 0$, the proof follows by straightforward computations.

Proof of the asymptotic expansions in Lemma 4.4.3:

This lemma is demonstrated by recurrence. In the case $j = 0$, (4.4.4) is given in [68], and (4.4.5) is easily obtained by using the method on p69 in [94]. The recurrence involves technical and laborious computations. However, when focusing on the case $j = 1$, the proof remains similar across all other instances where $j \in \{2, \cdots, d\}$. We focus on proving that, for $x \in I_1$ and μ large enough:

$$\frac{\phi_\mu^{(1)}(x)}{\sqrt{m(x)}} = \frac{2^{\frac{\mu}{2b_0}}K_1(\mu)}{\sigma_1\sqrt{\pi}}\Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right)\cos\left(\frac{\pi\mu}{2b_0}(1+o(1))\right)\left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right], \tag{4.8.11}$$

where $\sigma_1 \wedge \sigma_0 \leq K_1(\mu) \leq \sigma_1 \vee \sigma_0$.

Recall that for $x \in I_1$, $\frac{\phi_\mu^{(1)}(x)}{\sqrt{m(x)}} = \tilde{\phi}_\mu^{(1)}(x)$. We focus on the asymptotic expansion of $\tilde{\phi}_\mu^{(1)}$. In this case, for all $x \in I_1$, we have $\tilde{\phi}_\mu^{(1)}(x) = \alpha_\mu^{(1)} D_{\frac{\mu}{b_1}}\left(-x\frac{\sqrt{2b_1}}{\sigma_1}\right) + \beta_\mu^{(1)} D_{\frac{\mu}{b_1}}\left(x\frac{\sqrt{2b_1}}{\sigma_1}\right)$ where $(\alpha_\mu^{(1)}, \beta_\mu^{(1)})$ satisfies:

$$\begin{cases} \sigma_1 \left[\alpha_\mu^{(1)} D_{\frac{\mu}{b_1}}\left(-r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) + \beta_\mu^{(1)} D_{\frac{\mu}{b_1}}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right)\right] = \sigma_0 D_{\frac{\mu}{b_0}}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right), \\ \frac{1}{\sqrt{b_1}}\left[\beta_\mu^{(1)} D_{\frac{\mu}{b_1}-1}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) - \alpha_\mu^{(1)} D_{\frac{\mu}{b_1}-1}\left(-r_1\frac{\sqrt{2b_1}}{\sigma_1}\right)\right] = -\frac{1}{\sqrt{b_0}} D_{\frac{\mu}{b_0}-1}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right). \end{cases} \tag{4.8.12}$$

We assume that $\mu \notin b_1\mathbb{N}$ (see Remark 4.8.5 for more details). Solving the system (4.8.12) and using (4.8.4), we obtain:

$$\alpha_\mu^{(1)} = \Gamma\left(1-\frac{\mu}{b_1}\right)\frac{\left[\frac{\sigma_0}{\sigma_1} D_{\frac{\mu}{b_0}}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right) D_{\frac{\mu}{b_1}-1}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) + \sqrt{\frac{b_1}{b_0}} D_{\frac{\mu}{b_0}-1}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right) D_{\frac{\mu}{b_1}}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right)\right]}{\sqrt{2\pi}},$$

$$\beta_\mu^{(1)} = \Gamma\left(1-\frac{\mu}{b_1}\right)\frac{\left[\frac{\sigma_0}{\sigma_1} D_{\frac{\mu}{b_0}}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right) D_{\frac{\mu}{b_1}-1}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) - \sqrt{\frac{b_1}{b_0}} D_{\frac{\mu}{b_0}-1}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right) D_{\frac{\mu}{b_1}}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right)\right]}{\sqrt{2\pi}}.$$

From Section 4.8.2:

$$\tilde{\phi}_\mu^{(1)}(x) = \tilde{\phi}_{H,\mu}^{(1)}(x) + \tilde{\phi}_{P,\mu}^{(1)}(x).$$

We start by studying the asymptotic expansion of the particular and the homogeneous solution separately. We obtain the following asymptotic expansion for the particular solution.

Lemma 4.8.2. *For μ large enough and $x \in I_1$, we have:*

$$\tilde{\phi}_{P,\mu}^{(1)}(x) = 2^{\frac{\mu}{2b_0}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)O(\mu^{-\frac{1}{4}}).$$

Proof. Using (4.8.10), we have $\tilde{\phi}_{P,\mu}^{(1)}(x) = \alpha_\mu^{(1)} I_1 + \beta_\mu^{(1)} I_2$, where,

$$I_1 = \frac{b_1^2}{\sigma_1^3\sqrt{b_1+2\mu}}\int_{r_1}^{x} y^2 D_{\frac{\mu}{b_1}}\left(-y\frac{\sqrt{2b_1}}{\sigma_1}\right)\sin\left((x-y)\frac{\sqrt{b_1+2\mu}}{\sigma_1}\right)dy,$$
$$I_2 = \frac{b_1^2}{\sigma_1^3\sqrt{b_1+2\mu}}\int_{r_1}^{x} y^2 D_{\frac{\mu}{b_1}}\left(y\frac{\sqrt{2b_1}}{\sigma_1}\right)\sin\left((x-y)\frac{\sqrt{b_1+2\mu}}{\sigma_1}\right)dy.$$

Then, from Cauchy-Schwartz inequality together with (4.8.6) and Lemma 4.8.8, we obtain:

$$I_1 = 2^{\frac{\mu}{2b_1}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_1}\right)O(\mu^{-\frac{1}{4}}) \quad \text{and} \quad I_2 = 2^{\frac{\mu}{2b_1}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_1}\right)O(\mu^{-\frac{1}{4}}).$$

Using formula (4.8.5) together with Stirling formula (see (1.4.25) in [68]), we prove that:

$$\begin{aligned}\alpha_\mu^{(1)}O(\mu^{-\frac{1}{4}}) + \beta_\mu^{(1)}O(\mu^{-\frac{1}{4}}) =& 2^{\frac{\mu}{2b_0}+\frac{\mu}{2b_1}}\Gamma\left(1-\frac{\mu}{b_1}\right)\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_1}\right) \\ &\times\left[\frac{\Gamma\left(\frac{\mu}{2b_1}\right)}{\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_1}\right)}O(\mu^{-\frac{1}{4}}) + \frac{\Gamma\left(\frac{\mu}{2b_0}\right)}{\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)}O(\mu^{-\frac{1}{4}})\right],\end{aligned}$$

and $\frac{\Gamma\left(\frac{\mu}{2b_0}\right)}{\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)} = O(\mu^{-\frac{1}{2}})$. Then, we obtain:

$$\tilde{\phi}^{(1)}_{P,\mu}(x) = 2^{\frac{\mu}{2b_0}+\frac{\mu}{b_1}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)\Gamma\left(1-\frac{\mu}{b_1}\right)\Gamma^2\left(\frac{1}{2}+\frac{\mu}{2b_1}\right)O(\mu^{-\frac{3}{4}}).$$

Using again Stirling formula, we have $2^{\frac{\mu}{b_1}}\Gamma\left(1-\frac{\mu}{b_1}\right)\Gamma^2\left(\frac{1}{2}+\frac{\mu}{2b_1}\right)O(\mu^{-\frac{3}{4}}) = O(\mu^{-\frac{1}{4}})$ and the conclusion holds. □

We now focus on the asymptotic expansion of the homogeneous solution of (4.8.7).

Lemma 4.8.3. *For μ large enough and $x \in I_1$, we have:*

$$\tilde{\phi}^{(1)}_{H,\mu}(x) = \frac{2^{\frac{\mu}{2b_0}}K_1(\mu)}{\sigma_1\sqrt{\pi}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)\left[\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}+\theta(\mu)\right)+O(\mu^{-\frac{1}{4}})\right],$$

where $\sigma_1 \wedge \sigma_0 \leq K_1(\mu) \leq \sigma_1 \vee \sigma_0$ and $\frac{\theta(\mu)}{\sqrt{b_1+2\mu}} = O(1)$.

Proof. Using (4.4.4), (4.4.5), and (4.8.9), we have:

$$\begin{aligned}\tilde{\phi}^{(1)}_{H,\mu}(x) =& \frac{2^{\frac{\mu}{2b_0}-1}(\sigma_0+\sigma_1)}{\sigma_1\sqrt{\pi}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)\left[\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}+f(x,\mu)\right)\right.\\ &\left.+\frac{\sigma_0-\sigma_1}{\sigma_0+\sigma_1}\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}-f(x,\mu)\right)+O\left(\mu^{-\frac{1}{4}}\right)\right],\end{aligned}$$

where $f(x,\mu) = \frac{(x-r_1)\sqrt{b_1+2\mu}}{\sigma_1}$. Then, we get:

$$\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}+f(x,\mu)\right)+\frac{\sigma_0-\sigma_1}{\sigma_0+\sigma_1}\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}-f(x,\mu)\right) = \frac{2\sqrt{A^2(\mu)+B^2(\mu)}}{\sigma_0+\sigma_1}$$
$$\times\left[\frac{A(\mu)}{\sqrt{A^2(\mu)+B^2(\mu)}}\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}\right)-\frac{B(\mu)}{\sqrt{A^2(\mu)+B^2(\mu)}}\sin\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}\right)\right],$$

where $A(\mu) = \sigma_0\cos(f(x,\mu))$ and $B(\mu) = \sigma_1\sin(f(x,\mu))$.
Using the following variable change $\frac{A(\mu)}{\sqrt{A^2(\mu)+B^2(\mu)}} = \cos(\theta(\mu))$ and $\frac{B(\mu)}{\sqrt{A^2(\mu)+B^2(\mu)}} = \sin(\theta(\mu))$, we get:

$$\begin{aligned}&\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}+f(x,\mu)\right)+\frac{\sigma_0-\sigma_1}{\sigma_0+\sigma_1}\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}-f(x,\mu)\right)\\ &= \frac{2K_1(\mu)}{\sigma_0+\sigma_1}\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0}+\frac{\pi\mu}{2b_0}+\theta(\mu)\right).\end{aligned}$$

with $K_1(\mu) = \sqrt{A^2(\mu)+B^2(\mu)}$. Note that $\sigma_0^2 \wedge \sigma_1^2 \leq A^2(\mu)+B^2(\mu) \leq \sigma_0^2 \vee \sigma_1^2$, then $\sigma_0 \wedge \sigma_1 \leq K_1(\mu) \leq \sigma_0 \vee \sigma_1$. As $\theta(\mu) = \arctan\left(\frac{A(\mu)}{B(\mu)}\right)+n\pi$, by differentiation with respect to μ, we have:

$$\theta'(\mu) = \frac{\sigma_0(r_1-x)}{\sqrt{b_1+2\mu}(A^2(\mu)+B^2(\mu))}.$$

Then, we obtain $\frac{r_1-x}{(\sigma_0^2\wedge\sigma_1^2)\sqrt{b_1+2\mu}} \leq \theta'(\mu) \leq \frac{r_1-x}{(\sigma_0^2\vee\sigma_1^2)\sqrt{b_1+2\mu}}$. Integrating with respect to μ leads to $\frac{\theta(\mu)}{\sqrt{b_1+2\mu}} = O(1)$. □

Combining the two preceding lemmas, we provide an asymptotic expansion for the function $\tilde{\phi}_\mu^{(1)}$.

Lemma 4.8.4. *For μ large enough and $x \in I_1$, we have:*

$$\tilde{\phi}_\mu^{(1)}(x) = \frac{2^{\frac{\mu}{2b_0}} K_1(\mu)}{\sigma_1\sqrt{\pi}} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0} + \theta(\mu)\right)\left[1 + O(\mu^{-\frac{1}{4}})\right]. \tag{4.8.13}$$

Proof. Using Lemma 4.8.2 and Lemma 4.8.3, for μ large enough and $x \in I_1$, we have:

$$\tilde{\phi}_\mu^{(1)}(x) = \frac{2^{\frac{\mu}{2b_0}} K_1(\mu)}{\sigma_1\sqrt{\pi}} \Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right) \left[\cos\left(\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0} + \theta(\mu)\right) + O(\mu^{-\frac{1}{4}})\right].$$

Let $\mu_\star$ be large enough and such that $\frac{r_1\sqrt{b_0+2\mu_\star}}{\sigma_0} + \frac{\pi\mu_\star}{2b_0} + \theta(\mu_\star) = n\pi + \frac{\pi}{2}$. When μ is not in a neighborhood of a zero $\mu_\star$, we can factorize the cosine in our expression. It remains to prove that in a neighborhood of $\mu_\star$ we can also factorize the cosine. The idea is to use Lemma 4.8.7. By differentiation, we have:

$$\partial_\mu\left[\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0} + \theta(\mu)\right]_{\mu=\mu_\star} = \frac{r_1}{\sigma_0\sqrt{b_0+2\mu_\star}} + \frac{\pi}{2b_0} + \theta'(\mu_\star). \tag{4.8.14}$$

Then, as $\theta'(\mu_\star) = O(\mu_\star^{-\frac{1}{2}})$, we obtain:

$$\left|\partial_\mu\left[\frac{r_1\sqrt{b_0+2\mu}}{\sigma_0} + \frac{\pi\mu}{2b_0} + \theta(\mu)\right]_{\mu=\mu_\star}\right| > C,$$

with C a strictly positive constant. We conclude the proof by applying Lemma 4.8.7. □

This completes the proof of (4.8.11). Note that, by differentiating $\tilde{\phi}_{H,\mu}^{(1)}(x)$ and $\tilde{\phi}_{P,\mu}^{(1)}(x)$ with respect to x and by applying a similar reasoning we can prove that, for μ large enough and $x \in I_1$:

$$(\tilde{\phi}_\mu^{(1)})'(x) = -\frac{2^{\frac{\mu}{2b_0}} \tilde{K}_1(\mu)}{\sigma_1^2}\sqrt{\frac{b_1+2\mu}{\pi}}\Gamma\left(\frac{1}{2} + \frac{\mu}{2b_0}\right)\sin\left(\frac{\pi\mu}{2b_0}(1+o(1))\right)\left[1 + O\left(\mu^{-\frac{1}{4}}\right)\right],$$

where $\sigma_1 \wedge \sigma_0 \leq \tilde{K}_1(\mu) \leq \sigma_1 \vee \sigma_0$.

Remark 4.8.5 (Pathological case)**.** *Given that $\mu \in b_1\mathbb{N}$, the function $x \mapsto D_{\frac{\mu}{b_1}}(x)$ is either odd or even. In this case, we have: $\phi_{nb_1}^{(1)}(x) = \left((-1)^n\alpha_{nb_1}^{(1)} + \beta_{nb_1}^{(1)}\right) D_n\left(\frac{x\sqrt{2b_1}}{\sigma_1}\right)$ if a solution of the following system exists:*

$$\begin{cases} \left((-1)^n\alpha_{nb_1}^{(1)} + \beta_{nb_1}^{(1)}\right) D_n\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) = \frac{\sigma_0}{\sigma_1} D_{\frac{nb_1}{b_0}}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right), \\ \left((-1)^n\alpha_{nb_1}^{(1)} + \beta_{nb_1}^{(1)}\right) D_{n-1}\left(r_1\frac{\sqrt{2b_1}}{\sigma_1}\right) = \sqrt{\frac{b_1}{b_0}} D_{\frac{nb_1}{b_0}-1}\left(-r_1\frac{\sqrt{2b_0}}{\sigma_0}\right). \end{cases}$$

Therefore, we still obtain the same asymptotic expansion for $\phi_{nb_1}^{(1)}$ when n is large enough.

Remark 4.8.6 (Reccurence in the other state)**.** *Let us assume that the statement is true for $j-1$. In this case, we have $\tilde{\phi}_\mu^{(j)}(x) = \tilde{\phi}_{H,\mu}^{(j)}(x) + \tilde{\phi}_{P,\mu}^{(j)}(x)$. For $\mu \notin b_j\mathbb{N}$, using* (4.8.3) *and* (4.8.4)*, for all $x \in I_j$, $(\alpha_\mu^{(j)}, \beta_\mu^{(j)})$ are solutions of:*

$$\begin{cases} \alpha_\mu^{(j)} &= \Gamma\left(1-\frac{\mu}{b_j}\right) \frac{\frac{\sigma_{j-1}}{\sigma_j^2}\tilde{\phi}_\mu^{(j-1)}(r_j) D_{\frac{\mu}{b_j}-1}\left(r_j\frac{\sqrt{2b_j}}{\sigma_j}\right) - \frac{\sigma_{j-1}\sqrt{b_j}}{\mu\sqrt{2}} D_{\frac{\mu}{b_j}}\left(r_j\frac{\sqrt{2b_j}}{\sigma_j}\right)\left[(\tilde{\phi}_\mu^{(j-1)})'(r_j) + r_j\frac{b_{j-1}}{\sigma_{j-1}^2}\tilde{\phi}_\mu^{(j-1)}(r_j)\right]}{\sqrt{2\pi}}, \\ \beta_\mu^{(j)} &= \Gamma\left(1-\frac{\mu}{b_j}\right) \frac{\frac{\sigma_{j-1}}{\sigma_j^2}\tilde{\phi}_\mu^{(j-1)}(r_j) D_{\frac{\mu}{b_j}-1}\left(r_j\frac{\sqrt{2b_j}}{\sigma_j}\right) + \frac{\sigma_{j-1}\sqrt{b_j}}{\mu\sqrt{2}} D_{\frac{\mu}{b_j}}\left(r_j\frac{\sqrt{2b_j}}{\sigma_j}\right)\left[(\tilde{\phi}_\mu^{(j)})'(r_{j-1}) + r_j\frac{b_{j-1}}{\sigma_{j-1}^2}\tilde{\phi}_\mu^{(j-1)}(r_j)\right]}{\sqrt{2\pi}}. \end{cases}$$

The proof follows by similar reasoning based on the previous system and (4.8.8) *together with* (4.8.9)*. If $\mu \in b_j\mathbb{N}$, the asymptotic behavior remains the same by a similar reasoning as in Remark 4.8.5.*

Proof of the asymptotic expansions in Corollary 4.4.4:

Using Lemma 4.4.3, for all $x \in I_d$ and n large enough, we have:

$$\phi_n^{(d)}(x) = \frac{2^{\frac{\mu_n}{b_0}} K_d(\mu_n)}{\prod_{i=1}^d \sigma_i}\sqrt{\frac{m(x)}{\pi}}\Gamma\left(\frac{1}{2}+\frac{\mu_n}{2b_0}\right)\cos\left(\frac{\pi\mu_n}{2b_0}(1+o(1))\right)\left[1+O\left(\mu_n^{-\frac{1}{4}}\right)\right]. \tag{4.8.15}$$

Here $(\mu_n)_{n\geq 1}$ is the ordered sequence of zeros of $\mu \mapsto \phi_n^{(d)}(c)$. From (4.8.15), for n large enough, μ_n are the zeros of $\mu \mapsto \cos\left(\frac{\pi\mu}{2b_0}(1+o(1))\right)$. Then, we have:

$$\mu_n = b_0(2n+1)(1+o(1)).$$

It remains now to prove (4.4.6). By differentiating (4.8.15) with respect to μ, we have:

$$\begin{aligned} \partial_\mu\left[\phi_\mu^{(d)}(c)\right]_{\mu=\mu_n} =& \partial_\mu\left[\frac{2^{\frac{\mu_n}{b_0}} K_d(\mu_n)}{\prod_{i=1}^d \sigma_i}\sqrt{\frac{m(c)}{\pi}}\Gamma\left(\frac{1}{2}+\frac{\mu}{2b_0}\right)\left(1+O(\mu^{-\frac{1}{4}})\right)\right]_{\mu=\mu_n} \\ & \times \overbrace{\cos\left(\frac{\pi\mu_n}{2b_0}(1+o(1))\right)}^{=0} - \sqrt{\pi m(c)}\frac{2^{\frac{\mu_n}{2b_0}-1}K_d(\mu_n)}{b_0\prod_{i=1}^d\sigma_i} \\ & \times \Gamma\left(\frac{1}{2}+\frac{\mu_n}{2b_0}\right)\underbrace{\sin\left(\frac{\pi\mu_n}{2b_0}(1+o(1))\right)}_{=\pm 1}(1+o(1)). \end{aligned}$$

Applying the absolute values, we obtain:

$$\left|\partial_\mu\left[\phi_\mu^{(d)}(c)\right]_{\mu=\mu_n}\right| = \sqrt{\pi m(c)}\frac{2^{\frac{\mu_n}{2b_0}-1}K_d(\mu_n)}{b_0\prod_{i=1}^d\sigma_i}\Gamma\left(\frac{1}{2}+\frac{\mu_n}{2b_0}\right)(1+o(1)).$$

The conclusion holds, and Corollary 4.4.4 is proved.

4.8.4 Auxiliary results

Lemma 4.8.7. *Let $f, g : \mathbb{R}_+ \to \mathbb{R}$ such that f is differentiable, and for μ large enough, $g(\mu) = O\left(\mu^{-\frac{1}{4}}\right)$. Let $\mu_\star$ such that $f(\mu_\star) = n\pi + \frac{\pi}{2}$. If $|f'(\mu_\star)| > \epsilon > 0$, then for n large enough, we have:*

$$\lim_{\mu \to \mu_\star} \frac{g(\mu)}{\cos(f(\mu))} = O\left(\mu_\star^{-\frac{5}{4}}\right).$$

Proof. This result is a direct application of L'Hôpital's rule. □

Using (4.8.2) and (4.8.6), one can prove the following lemma.

Lemma 4.8.8. *For μ large enough,*

$$\left|\frac{D_\mu(-x\sqrt{2})\partial_\mu\left(\mu D_{\mu-1}(-x\sqrt{2})\right)}{2^\mu \Gamma^2\left(\frac{1}{2}+\frac{\mu}{2}\right)\sqrt{\mu}}\right| + \left|\frac{\mu D_{\mu-1}(-x\sqrt{2})\partial_\mu D_\mu(-x\sqrt{2})}{2^\mu \Gamma^2\left(\frac{1}{2}+\frac{\mu}{2}\right)\sqrt{\mu}}\right| = O(1). \tag{4.8.16}$$

Proof. Using (4.8.2), we have $|D_\mu(-x\sqrt{2})| \leq \frac{2^{\frac{\mu}{2}} e^{\frac{x^2}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}+\frac{\mu}{2}\right)$ and $|D_{\mu-1}(-x\sqrt{2})| \leq \frac{2^{\frac{\mu}{2}-\frac{1}{2}} e^{\frac{x^2}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{\mu}{2}\right)$. Then, by differentiating (4.8.2) with respect to μ, we obtain:

$$\begin{aligned}\partial_\mu D_\mu(-x\sqrt{2}) =& \frac{\ln(2)}{2} D_\mu(-x\sqrt{2}) - \sqrt{\pi} 2^{\frac{\mu}{2}} e^{\frac{x^2}{2}} \int_0^\infty e^{-t^2} t^\mu \sin\left(2xt + \frac{\mu\pi}{2}\right) dt \\ &+ \frac{2^{\frac{\mu}{2}+1} e^{\frac{x^2}{2}}}{\sqrt{\pi}} \int_0^\infty e^{-t^2} t^\mu \ln(t) \cos\left(2xt + \frac{\mu\pi}{2}\right) dt, \end{aligned} \tag{4.8.17}$$

and

$$\begin{aligned}\partial_\mu D_{\mu-1}(-x\sqrt{2}) =& \frac{\ln(2)}{2} D_{\mu-1}(-x\sqrt{2}) + \sqrt{\pi} 2^{\frac{\mu-1}{2}} e^{\frac{x^2}{2}} \int_0^\infty e^{-t^2} t^{\mu-1} \cos\left(2xt + \frac{\mu\pi}{2}\right) dt \\ &+ \frac{2^{\frac{\mu+1}{2}} e^{\frac{x^2}{2}}}{\sqrt{\pi}} \int_0^\infty e^{-t^2} t^{\mu-1} \ln(t) \sin\left(2xt + \frac{\mu\pi}{2}\right) dt. \end{aligned} \tag{4.8.18}$$

Using (4.8.2), (4.8.17) and (4.8.18), we have:

$$\begin{aligned} &- D_\mu(-x\sqrt{2})\partial_\mu\left(\mu D_{\mu-1}(-x\sqrt{2})\right) + \mu D_{\mu-1}(-x\sqrt{2})\partial_\mu D_\mu(-x\sqrt{2}) \\ &= - \underbrace{D_\mu(-x\sqrt{2}) D_{\mu-1}(-x\sqrt{2})}_{=A_1} \underbrace{- \mu 2^{\mu+\frac{1}{2}} e^{x^2} \int_0^\infty \int_0^\infty e^{-t^2-y^2} t^\mu y^{\mu-1} \cos\left(2x(t-y)\right) dtdy}_{=A_2} \\ &\underbrace{- \mu \frac{2^{\mu+\frac{3}{2}} e^{x^2}}{\pi} \int_0^\infty \int_0^\infty e^{-t^2-y^2} t^\mu y^{\mu-1} \cos\left(2xt + \frac{\mu\pi}{2}\right) \sin\left(2yt + \frac{\mu\pi}{2}\right) \left(\ln(y) - \ln(t)\right) dtdy}_{=A_3}. \end{aligned}$$

For $i = 1, 2, 3$, it remains to bound $|A_i|$. We easily get:

$$|A_1| \leq C_1 2^\mu \Gamma\left(\frac{1}{2} + \frac{\mu}{2}\right) \Gamma\left(\frac{\mu}{2}\right) \quad \text{and} \quad |A_2| \leq C_2 \mu 2^\mu \Gamma\left(\frac{1}{2} + \frac{\mu}{2}\right) \Gamma\left(\frac{\mu}{2}\right),$$

with C_1, C_2 two strictly positive constants. It remains now to bound A_3: Using that, for x strictly positive, we have $1 - \frac{1}{x} \leq \ln(x) \leq x - 1$, then for $(t, y) \in \mathbb{R}_+ \times [0, t]$:

$$\left|(\ln(y) - \ln(t)) \cos\left(2xt + \frac{\mu\pi}{2}\right) \sin\left(2xy + \frac{\mu\pi}{2}\right)\right| \leq \frac{t}{y} - 1,$$

and, for $(t, y) \in \mathbb{R}_+ \times [t, +\infty)$, we have:

$$\left|(\ln(y) - \ln(t)) \cos\left(2xt + \frac{\mu\pi}{2}\right) \sin\left(2xy + \frac{\mu\pi}{2}\right)\right| \leq \frac{y}{t} - 1.$$

We bound A_3 in the following way:

$$\begin{aligned} |A_3| &\leq C_3 \mu 2^\mu \left[\int_0^\infty \int_0^t e^{-t^2-y^2} t^{\mu+1} y^{\mu-2} dy dt + \int_0^\infty \int_t^\infty e^{-t^2-y^2} t^{\mu-1} y^\mu dy dt\right], \\ &\leq C_3 \mu 2^\mu \left[\Gamma\left(\frac{\mu}{2} + 1\right) \Gamma\left(\frac{\mu}{2} - \frac{1}{2}\right) + \Gamma\left(\frac{\mu}{2}\right) \Gamma\left(\frac{\mu}{2} + \frac{1}{2}\right)\right]. \end{aligned}$$

with C_3 a strictly positive constant. Using Stirling formula, for μ large enough, we have $\frac{\sum_{i=1}^3 |A_i|}{2^\mu \sqrt{\mu} \Gamma^2\left(\frac{\mu}{2} + \frac{1}{2}\right)} = O(1)$. Then the conclusion holds. □

Chapter 5

Parameters estimation of an ergodic Multi-threshold CKLS process from continuous and discrete observations

5.1 Introduction

We deal with parameter estimation of a stochastic process, which follows different Chan-Karolyi-Longstaff-Sanders (CKLS) dynamics separately on different intervals (see equation (5.2.1) for a definition), that we call *threshold CKLS* (T-CKLS) process. T-CKLS is in particular a Self Exciting Threshold model [34]. It excites itself by changing dynamics according to its own position. The class of T-CKLS models includes several threshold and non-threshold models such as Ornstein-Uhlenbeck (OU) or Vasicek, Cox-Ingersoll-Ross (CIR), Black-Scholes, Merton, Constant Elasticity of Variance (CEV) model. Let us recall the equation satisfied by (non-threshold) CKLS:

$$X_t = x_0 + \int_0^t (a - bX_s)\,\mathrm{d}s + \int_0^t \sigma |X_s|^\gamma \,\mathrm{d}B_s \tag{5.1.1}$$

where $\sigma \in (0, \infty)$, $b \in \mathbb{R}$, $\gamma \in [0, 1]$. Assume also that the deterministic initial condition x_0 and the coefficient a are strictly positive. CKLS was considered for interest rate modeling in [27]. Among statistical studies of non-threshold diffusions such as CKLS and CIR, let us mention [83] and [10, 11, 12]. The approaches exploit the knowledge of the law of some functionals of the process. Except for very special threshold diffusions (TDs), this law is not available nor easily exploitable. Therefore, when dealing with TDs, one must consider different techniques, which would also hold for non-threshold cases (see Section 5.5.2 for details).

Recently, several studies have been conducted on the parametric estimation of TDs. Just to mention some results, in the case of continuous time observations [59, 97, 72], high-frequency observations on finite or infinite horizon [70, 81, 82], low frequency ones [113, 112, 82]. TDs attract attention for applications, in financial modelling e.g. [34, 71, 85],

population ecology [24], etc. One of the features of the TDs is that they allow for mean reversions, even with several mean reversion levels.

In this chapter we focus on ergodic T-CKLS. Note that this includes ergodic T-CIR model and allows for different dynamics on fixed intervals: OU on a region, CIR on another, and possibly other special cases of CKLS on another interval. We consider both continuous time observations and discrete observations, which are not necessarily equally spaced.

We study both maximum likelihood estimation (MLE) and a quasi maximum likelihood estimation (QMLE) for the drift coefficients and we also propose an estimator for the diffusion coefficient σ based on quadratic variation. We study the asymptotic behavior of these estimators, obtaining a law of large numbers and a central limit theorem for the drift coefficients estimators. We discuss these asymptotic results in long time in the continuous time setting, see Theorem 5.3.5, and in high frequency and long time in the discrete time setting, see Theorem 5.4.4. We also study consistency and provide a lower bound for the speed of convergence for the diffusion coefficient estimator in high frequency and long time, see Theorem 5.4.3.

In the continuous time setting, Theorem 5.3.5 is new for MLE. The drift QMLE was considered in [97]. In the same paper, numerical studies were conducted on the discretization of the QMLE. In the discrete time setting, the statistical properties of the discretized MLE and QMLE are proven in Theorem 5.4.4, which is the main result of this chapter. Let us mention that, up to our knowledge, the results obtained in this chapter are also new when the process follows some special dynamics of T-CKLS process, such as T-CIR. Moreover, up to our knowledge, Theorem 5.4.3 about diffusion coefficient estimation is the first result of its kind in the context of TDs observed in high frequency and long time.

In the case of discrete observations, we assume that the process is in its stationary regime (X_t is distributed according to the stationary distribution for all $t \geq 0$) and it is observed on a time-grid of N observations $0 = t_0 < t_1 < \ldots < t_N = T_N$ with maximal lag between two consecutive observations, say Δ_N, such that $\lim_{N\to\infty} T_N = +\infty$ and $\lim_{N\to\infty} \Delta_N \to 0$. To prove the convergence speed, we require an additional condition of the form $\lim_{N\to\infty} \Delta_N^\alpha T_N = 0$, for some power $\alpha \in (0, 1]$ that depends on the parameter vector γ. This is consistent with the existing literature in the context of parameter estimation of diffusions from discrete observations with and without threshold, e.g. [7, 10, 81] (where $\alpha = 1$). Further comments on our results, their assumptions, possible extensions and comparisons with the literature are provided in Section 5.5.

Outline. In Section 5.2, we introduce the model and the quasi-likelihood and likelihood functions associated to the T-CKLS process. In Section 5.3, we deal with estimation of T-CKLS from continuous observations. We provide some statistical properties related to the drift estimator. The main results of this chapter are provided in Section 5.4, which deals with drift and volatility estimation from discrete observations. We study the asymptotic behavior in high frequency and long time for both estimators. In Section 5.5 we comment on the results. Numerical experiments are provided in Section 5.6, where the estimators are implemented and tested on simulated data and US interest rates data. Proofs are collected in Section 5.7. Further useful results are available in Section 5.8.

Throughout the chapter, we use the notion of stable convergence, denoted $\xrightarrow{stably}$. Further details on this type of convergence can be found in [50] and [51].

5.2 The framework: model and assumptions

In the entire chapter, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ denotes a filtered probability space, and B an $(\mathcal{F}_t)$-standard Brownian motion. The T-CKLS process solves the following one-dimensional SDE:

$$X_t = X_0 + \int_0^t (a(X_s) - b(X_s)X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)|X_s|^{\gamma(X_s)}\,\mathrm{d}B_s, \quad t \geq 0, \tag{5.2.1}$$

with $X_0 > 0$, X_0 (either deterministic, or X_0 independent of $(B_t)_{t\geq 0}$), piecewise constant coefficients a, b, σ and γ possibly discontinuous at levels $0 = r_0 < r_1 < \ldots < r_d < r_{d+1} = +\infty$, $d \in \mathbb{N}$. We focus on the case $\gamma(\mathbb{R}) \subseteq [0,1]$ and $\gamma(0) \in [1/2, 1] \cup \{0\}$. More precisely, let $I_j := [r_j, r_{j+1})$, for $j \in \{0, \ldots, d\}$, unless $\gamma(0) = 0$, in which case $I_0 = (-\infty, r_1)$. The drift coefficients are given by:

$$a(x) = \sum_{j=0}^{d} a_j \mathbb{1}_{I_j}(x) \in \mathbb{R} \quad \text{and} \quad b(x) = \sum_{j=0}^{d} b_j \mathbb{1}_{I_j}(x) \in \mathbb{R},$$

and similarly, the volatility coefficients are given by:

$$\sigma(x) = \sum_{j=0}^{d} \sigma_j \mathbb{1}_{I_j}(x) > 0 \quad \text{and} \quad \gamma(x) = \sum_{j=0}^{d} \gamma_j \mathbb{1}_{I_j}(x) \geq 0.$$

When $\gamma_0 \in [1/2, 1)$, we assume in addition that $a_0 > 0$. When $\gamma_0 = 1$, we allow for $a_0 \geq 0$. When $\gamma_0 = 0$, for the sake of simplicity, we keep the assumptions $r_1 > 0$ and, although unnecessary, $X_0 > 0$.

For existence of a pathwise unique strong solution to (5.2.1) under the assumption that $\gamma_0 \in [1/2, 1] \cup \{0\}$, we refer for instance to [66] for existence and uniqueness results for threshold diffusions. Moreover, T-CKLS is a Markov process (see e.g. [44]). When $\gamma_0 \in [1/2, 1]$ the process is always non-negative and 0 is either an unreachable point or a reflecting one (see Lemma 5.7.1 in Section 5.7.1).

We suppose that $(\gamma_j)_{j=0}^d$ and the thresholds $(r_j)_{j=0}^d$ are known, and we estimate drift and diffusion parameter vectors a, b, σ for continuous time observations and discrete (not necessarily equally spaced) high frequency observations and infinite horizon.

Definition 5.2.1 (Ergodicity)**.** *We say that the process is* ergodic*, if it is positive recurrent.*

In this chapter we consider only the case in which the process is ergodic. For instance, the process is ergodic if we restrict the drift coefficients of the first interval I_0 and the last interval I_d to satisfy

$$(a_0, b_0) \in (0, +\infty) \times \mathbb{R} \quad \text{and} \quad (a_d, b_d) \in \mathbb{R} \times (0, +\infty)$$

when $\gamma_0 \neq 0$. When $\gamma_0 = 0$, the process is ergodic if, for instance,

$$(a_0, b_0) \in \mathbb{R} \times (0, +\infty) \quad \text{and} \quad (a_d, b_d) \in \mathbb{R} \times (0, +\infty).$$

More precise restrictions to the parameters are given in Table 5.4 in Section 5.7.1.

When the process is ergodic, there exists a stationary distribution (invariant distribution for the transition semigroup), denoted by μ. An expression for the stationary distribution is given in Section 5.7.1.

Definition 5.2.2 (Stationarity)**.** *We say that the process is* stationary *if it is ergodic, X_0 is independent of the driving Brownian motion, and X_0 is distributed according to the stationary distribution.*

The maximum likelihood estimation method for the drift parameter requires some additional parameter restrictions that we are going to detail in the remainder of this section. Assume that we have access to an observation of an entire trajectory on the time interval $[0, T]$ of the T-CKLS. We denote by $\theta := (a, b) = (a_j, b_j)_{j=0}^d$ the drift parameters and we assume the thresholds r and the coefficients γ to be known. In the next sections, we assume σ to be unknown, and we propose an estimator. Yet, in the following lines, the reader should think as if σ is known (replaced by an estimator). We consider two different contrast functions: likelihood and quasi-likelihood. The likelihood function $\theta \mapsto \mathcal{L}_T(\theta; \sigma, \gamma)$ is related to the Girsanov weight:

$$\mathcal{L}_T(\theta; \sigma, \gamma) = \exp\left(\int_0^T \frac{a(X_s) - b(X_s)X_s}{\sigma(X_s)^2 (X_s)^{2\gamma(X_s)}} \, \mathrm{d}X_s - \frac{1}{2} \int_0^T \frac{\left(a(X_s) - b(X_s)X_s\right)^2}{\sigma(X_s)^2 (X_s)^{2\gamma(X_s)}} \, \mathrm{d}s \right). \quad (5.2.2)$$

Note that it is well defined if the integrals above are well defined (see Proposition 5.8.1 for details). Hence, we further restrict the parameter space $\Theta^{(\mathcal{L})}$ to the coefficients for which the integrals above are well defined: e.g. ,if $\gamma_0 = 1/2$, we have to restrict to $a_0 \geq \sigma_0^2/2$.

We consider the quasi-likelihood function $\theta \mapsto q\text{-}\mathcal{L}_T(\theta) := \ln \mathcal{L}_T(\theta; 1, 0)$ [97]. Hence,

$$q\text{-}\mathcal{L}_T(\theta) = \int_0^T \left(a(X_s) - b(X_s)X_s\right) \, \mathrm{d}X_s - \frac{1}{2} \int_0^T \left(a(X_s) - b(X_s)X_s\right)^2 \, \mathrm{d}s. \quad (5.2.3)$$

The advantage of this contrast function is the fact that it does not depend on the diffusion's coefficients γ, σ. We denote $\Theta^{(q\text{-}\mathcal{L})}$ the set of parameters such that the quasi-likelihood is well defined. Note that $q\text{-}\mathcal{L}_T$ is always well defined because the process we consider has continuous trajectories. So, it covers a wider range of parameters than the likelihood function.

Summarizing, in the next sections, we suppose that the parameter $r = (r_j)_{j=1}^d$ and $\gamma = (\gamma_j)_{j=0}^d$ are known. We denote $\theta_\star := (a, b) = (a_j, b_j)_{j=0}^d$ and $\sigma_\star$ the parameters to be estimated. We suppose that the parameters are in $\Theta^{(\mathcal{L})}$, resp. $\Theta^{(q\text{-}\mathcal{L})}$, when dealing with the likelihood, resp. quasi-likelihood function. Moreover, we assume that the process is ergodic (see Table 5.4 in Section 5.7.1 for the parameter restrictions ensuring ergodicity).

5.3 Estimation from continuous time observations

Let $T \in (0, \infty)$, and assume we have at our disposal continuous time observations on the interval $[0, T]$ of a trajectory of the process X solution to the SDE (5.2.1).

First, we provide estimators which maximize likelihood and quasi-likelihood. Next, we study the asymptotic behavior of the estimators in long time under the assumption that the process is ergodic (see Definition 5.2.1).

5.3.1 Estimators expressions

The drift parameters estimators, MLE and QMLE, are defined as the maximal argument of the log-likelihood (5.2.2) and the quasi-likelihood (5.2.3):

$$\theta_T^{(\mathcal{L})} := \underset{\theta \in \Theta^{(\mathcal{L})}}{\operatorname{argmax}} \ \ln \mathcal{L}_T(\theta; \sigma, \gamma) \quad \text{and} \quad \theta_T^{(q\text{-}\mathcal{L})} := \underset{\theta \in \Theta^{(q\text{-}\mathcal{L})}}{\operatorname{argmax}} \ q\text{-}\mathcal{L}_T(\theta).$$

We look for expressions for MLE and QMLE in terms of the following quantities:

$$Q_T^{j,m} := \int_0^T X_s^m \mathbb{1}_{I_j}(X_s) \, \mathrm{d}s \quad \text{and} \quad M_T^{j,m} := \int_0^T X_s^m \mathbb{1}_{I_j}(X_s) \, \mathrm{d}X_s \tag{5.3.1}$$

for $j \in \{0, \dots, d\}$ and $m \in \{-2\gamma_j, 1 - 2\gamma_j, 2 - 2\gamma_j, 2\gamma_j\} \cup \{0, 1, 2\}$. It is convenient to express log-likelihood and quasi-likelihood as follows. The log-likelihood satisfies

$$\ln \mathcal{L}_T(\theta; \sigma, \gamma) := \sum_{j=0}^{d} \frac{1}{\sigma_j^2} \left(a_j M_T^{j,-2\gamma_j} - b_j M_T^{j,1-2\gamma_j} - \frac{a_j^2}{2} Q_T^{j,-2\gamma_j} - a_j b_j Q_T^{j,1-2\gamma_j} - \frac{b_j^2}{2} Q_T^{j,2-2\gamma_j} \right), \tag{5.3.2}$$

and the quasi-likelihood $q\text{-}\mathcal{L}_T(\theta) := \ln \mathcal{L}_T(\theta; 1, 0)$, which rewrites as

$$q\text{-}\mathcal{L}_T(\theta) = \sum_{j=0}^{d} a_j M_T^{j,0} - b_j M_T^{j,1} - \frac{a_j^2}{2} Q_T^{j,0} - a_j b_j Q_T^{j,1} - \frac{b_j^2}{2} Q_T^{j,2}. \tag{5.3.3}$$

The following proposition provides explicit expression of MLE and QMLE, in terms of the quantities in (5.3.1).

Proposition 5.3.1. *Let $T \in (0, \infty)$, the maximum of the likelihood $\mathcal{L}_T(\theta; \sigma, \gamma)$ is achieved at $\theta_T^{(\mathcal{L})} := (a_T^{j,\gamma}, b_T^{j,\gamma})_{j=0}^d$ with*

$$(a_T^{j,\gamma}, b_T^{j,\gamma}) = \left(\frac{M_T^{j,-2\gamma_j} Q_T^{j,2-2\gamma_j} - Q_T^{j,1-2\gamma_j} M_T^{j,1-2\gamma_j}}{Q_T^{j,-2\gamma_j} Q_T^{j,2-2\gamma_j} - (Q_T^{j,1-2\gamma_j})^2}, \frac{M_T^{j,-2\gamma_j} Q_T^{j,1-2\gamma_j} - Q_T^{j,-2\gamma_j} M_T^{j,1-2\gamma_j}}{Q_T^{j,-2\gamma_j} Q_T^{j,2-2\gamma_j} - (Q_T^{j,1-2\gamma_j})^2} \right). \tag{5.3.4}$$

The maximum of the quasi-likelihood $q\text{-}\mathcal{L}_T(\theta)$ is achieved at $\theta_T^{(q\text{-}\mathcal{L})} := (a_T^{j,0}, b_T^{j,0})_{j=0}^d$, that is

$$(a_T^{j,0}, b_T^{j,0}) = \left(\frac{M_T^{j,0} Q_T^{j,2} - Q_T^{j,1} M_T^{j,1}}{Q_T^{j,0} Q_T^{j,2} - (Q_T^{j,1})^2}, \frac{M_T^{j,0} Q_T^{j,1} - Q_T^{j,0} M_T^{j,1}}{Q_T^{j,0} Q_T^{j,2} - (Q_T^{j,1})^2} \right).$$

Proof. We sketch the proof for MLE. The same works for QMLE. One shows that (5.3.4) is the unique singular point of the gradient (vector of the derivatives with respect to a_j and b_j for all $j \in \{0, \dots, d\}$) of (5.3.2) and the Hessian is negative definite. □

Remark 5.3.2. *If $\gamma \equiv 0$, then the diffusion coefficient is piecewise constant, so T-CKLS is a threshold OU (T-OU) and QMLE and MLE coincide, as noticed in [81].*

Remark 5.3.3. *For every $j = 0, \dots, d$, $(a_T^{j,\gamma}, b_T^{j,\gamma})$ only depend on the observations of the trajectory $t \mapsto X_t$ which belong to I_j. Of course, the same holds for $(a_T^{j,0}, b_T^{j,0})$.*

MLE and QMLE do not depend on $\sigma_\star$ explicitly, but only on the quantities in (5.3.1). The following result ensures that $\sigma_\star$ is a.s. equal to an estimator expressed in terms of $Q_T^{j,\cdot}, M_T^{\cdot,0}, M_T^{j,1}, X_T, X_0$.

Proposition 5.3.4. *Let $T \in (0, \infty)$ and $j \in \{0, \dots, d\}$. Then*

$$\sigma_j = \sqrt{\frac{\mathbf{Q}_T^j}{Q_T^{j,2\gamma_j}}} \quad \text{a.s. on the event } \{Q_T^{j,0} > 0\}$$

where

$$\begin{aligned} \mathbf{Q}_T^0 &:= (f_0(X_T))^2 - (f_0(X_0))^2 + 2\left(r_1 M_T^{0,0} - M_T^{0,1}\right) - 2r_1 \mathfrak{f}_0 \\ \mathbf{Q}_T^d &:= (f_d(X_T))^2 - (f_d(X_0))^2 + 2\left(r_d M_T^{d,0} - M_T^{d,1}\right), \end{aligned} \tag{5.3.5}$$

with $f_0(x) = x\mathbb{1}_{I_0}(x) + r_1\mathbb{1}_{[r_1,+\infty)}(x)$, $f_d(x) = \mathbb{1}_{I_d}(x)(x - r_d)$, and $\mathfrak{f}_0 = \min(X_T, r_1) - \min(X_0, r_1)$, and for $j \in \{1, \dots, d-1\}$:

$$\mathbf{Q}_T^j := (f_j(X_T))^2 - (f_j(X_0))^2 - 2M_T^{j,1} + 2r_j M_T^{j,0} + 2(r_{j+1} - r_j)\big(\mathfrak{f}_{j+1} - \sum_{i=j+1}^{d} M_T^{i,0}\big) \tag{5.3.6}$$

with $f_j(x) = \mathbb{1}_{I_j}(x)(x - r_j) + (r_{j+1} - r_j)\mathbb{1}_{[r_{j+1},+\infty)}(x)$ and $\mathfrak{f}_j = \max(X_T, r_j) - \max(X_0, r_j)$.

Proof. We only consider the case $j \in \{1, \dots, d-1\}$. When $j \in \{0, d\}$ the proof works analogously. Considering the event $\{Q_T^{j,0} > 0\}$ corresponds to take trajectories, which spend some time in I_j and so $Q_T^{j,2\gamma_j}$ does not vanish. Applying Itô-Tanaka formula (see Corollary 1.1.8 in Chapter 1 or [91, Chapter VI, exercice 1.25]) ensures that

$$\mathrm{d}f_j(X_s) = \mathbb{1}_{I_j}(X_s)\,\mathrm{d}X_s + 2^{-1}\,\mathrm{d}\left(L_s^{r_j}(X) - L_s^{r_{j+1}}(X)\right).$$

The quadratic variation of $f_j(X)$ satisfies a.s. the equality:

$$\langle f_j(X)\rangle_T = \sigma_j^2 \int_0^T (X_s)^{2\gamma_j}\mathbb{1}_{I_j}(X_s)\,\mathrm{d}s = \sigma_j^2 Q_T^{j,2\gamma_j}.$$

Since for every semi-martingale Y Itô formula ensures that a.s. $\mathrm{d}Y^2 = 2Y\,\mathrm{d}Y + \mathrm{d}\langle Y\rangle$, Itô formula applied to the semi-martingale $(f_j(X_T))^2$, yields that a.s.

$$\mathrm{d}\langle f_j(X)\rangle_s = \mathrm{d}\left(f_j(X_s)\right)^2 + 2r_j\,\mathrm{d}M_s^{j,0} - 2\,\mathrm{d}M_s^{j,1} + (r_{j+1} - r_j)\,\mathrm{d}L_s^{r_{j+1}}(X).$$

In order to check that $\langle f_j(X)\rangle_T$ is a.s. equal to $\mathbf{Q}_T^j$, we exploit Itô-Tanaka formula to rewrite the local times in terms of $M^{j,0}, M^{j,1}$. Itô-Tanaka formula applied to $\max(X_T, r_j)$ yields $L_T^{r_j}(X) = 2\mathfrak{f}_j - 2\sum_{i=j}^d M_T^{i,0}$ a.s.. The proof is thus completed. □

5.3.2 Asymptotic properties: long time

In this section, we explore the statistical properties as $T \to \infty$ of the MLE and QMLE from continuous time observations of a trajectory of the T-CKLS process.

We assume that the process is ergodic, μ is the stationary distribution given in Section 5.7.1, and we introduce the following hypotheses:

- $\mathbf{H}_{\mathcal{L}}$: μ admits finite $(-2\gamma_0)$-th and $(2-2\gamma_d)$-th moment,
- $\mathbf{H}_{q\text{-}\mathcal{L}}$: μ admits finite $(2+2\gamma_d)$-th moment.

See Section 5.5.3 for comments on these assumptions. However, often the above required moments are finite. For instance, if $b_d \neq 0$, then μ admits positive moments of all order and for $\gamma_0 > 1/2$ admits negative moments of all order without any further condition. If $\gamma_0 = 1/2$ then parameters restrictions may be necessary (see Proposition 5.8.2).

The asymptotic behavior of the MLE and QMLE are provided in the following theorem, which states that MLE and QMLE are strongly consistent and asymptotically normal estimators of the drift parameter $\theta_\star$.

Theorem 5.3.5. *For $\ell \in \{\mathcal{L}, q\text{-}\mathcal{L}\}$, under Hypothesis $\mathbf{H}_\ell$, the MLE and QMLE are strongly consistent estimators of $\theta_\star$ i.e.*

$$\theta_T^{(\ell)} \xrightarrow[T\to+\infty]{a.s.} \theta_\star.$$

Furthermore, the following convergence is satisfied:

$$\sqrt{T}\left(\theta_T^{(\ell)} - \theta_\star\right) \xrightarrow[T\to+\infty]{stably} \mathbf{N}^{(\ell)},$$

where $\mathbf{N}^{(\ell)} = (N_{j,a}^{(\ell)}, N_{j,b}^{(\ell)})_{j=0}^{d}$ are $d+1$ independent, independent of X, two-dimensional centered Gaussian random variables with covariance matrices respectively given by $\sigma_j^2\Gamma_j^{(\ell)}$ such that

$$\Gamma_j^{(\mathcal{L})} := \Gamma_j^{(\mathcal{L},\gamma_j)} := \begin{pmatrix} Q_\infty^{j,-2\gamma_j} & -Q_\infty^{j,1-2\gamma_j} \\ -Q_\infty^{j,1-2\gamma_j} & Q_\infty^{j,2-2\gamma_j} \end{pmatrix}^{-1} \quad \text{and} \quad \Gamma_j^{(q\text{-}\mathcal{L})} = \left(\Gamma_j^{(\mathcal{L},0)}\right)^{-1} \Gamma_j^{(\mathcal{L},-\gamma_j)} \left(\Gamma_j^{(\mathcal{L},0)}\right)^{-1}$$

where $Q_\infty^{j,\cdot}$ are real constants defined in Lemma 5.3.6.

Proof. The proof is similar to the one in [72, 81], nevertheless we summarize the steps and stress the specificity of the case we consider in this chapter. Note that MLE and QMLE rewrite as follows:

$$\theta_j^{(\mathcal{L})} \overset{a.s.}{=} \theta_j + \sigma_j \left(\mathfrak{M}_T^{j,-\gamma_j}, -\mathfrak{M}_T^{j,1-\gamma_j}\right) \Gamma^{(\mathcal{L})} \quad \text{and} \quad \theta_j^{(q\text{-}\mathcal{L})} \overset{a.s.}{=} \theta_j + \sigma_j \left(\mathfrak{M}_T^{j,\gamma_j}, -\mathfrak{M}_T^{j,1+\gamma_j}\right) \Gamma^{(\mathcal{L},0)}$$

where $\mathfrak{M}_T^{j,k} = \int_0^T (X_s)^k \mathbb{1}_{I_j}(X_s)\,\mathrm{d}B_s$ for $k \in \{-\gamma_j, 1-\gamma_j, \gamma_j, 1+\gamma_j\}$ are martingales. Indeed, note that $Q_T^{j,2k}$ is the quadratic variation of $M_T^{j,k-\gamma_j}$ which is the one of $\mathfrak{M}_T^{j,k}$ up to a

multiplicative factor. We can now exploit martingale theorems. The consistency of the estimator (MLE and QMLE) follows directly from [74, Theorem 1] and the ergodicity of the process which implies, for instance, Lemma 5.3.6. The asymptotic normality property follows from [32, Theorem 2.2]. Hypotheses $\mathbf{H}_{\mathcal{L}}$ and $\mathbf{H}_{q\text{-}\mathcal{L}}$ are necessary for the application of [74, Theorem 1]. □

Lemma 5.3.6 (Ergodic properties)**.** *For $j \in \{0, \dots, d\}$ and $m \in \mathbb{R}$, if the m-th moment of μ is finite on the set I_j, then*

$$Q_\infty^{j,m} \overset{a.s.}{:=} \lim_{T\to\infty} \frac{Q_T^{j,m}}{T} = \int_{I_j} x^m \mu(\,\mathrm{d}x),$$

are non-vanishing constants.

Remark 5.3.7. *The above asymptotic normality of the estimators implies the local asymptotic normality (LAN) property (see [65]). The LAN property, is a fundamental concept in the asymptotic theory of statistics. For instance, when it is satisfied, it can be combined with the Minimax theorem to establish a lower bound for the asymptotic variance of estimators.*

Since statement and proof are analogous to the one of the T-OU process (corresponding to $\gamma = 0$) given in [81, Theorem 1.(iv)], we just provide a short statement.

Let $\ell \in \{\mathcal{L}, q\text{-}\mathcal{L}\}$ and assume that $\mathbf{H}_\ell$ holds. The LAN property holds for the ℓ-function with rate of convergence $1/\sqrt{T}$. Furthermore, the asymptotic Fisher information is given by:

$$\Gamma^{(\ell)} = \begin{bmatrix} \frac{1}{\sigma_0^2}\left(\Gamma_0^{(\ell)}\right)^{-1} & 0_{\mathbb{R}^{2\times 2}} & \dots & 0_{\mathbb{R}^{2\times 2}} \\ 0_{\mathbb{R}^{2\times 2}} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{1}{\sigma_{d-1}^2}\left(\Gamma_{d-1}^{(\ell)}\right)^{-1} & 0_{\mathbb{R}^{2\times 2}} \\ 0_{\mathbb{R}^{2\times 2}} & \dots & 0_{\mathbb{R}^{2\times 2}} & \frac{1}{\sigma_d^2}\left(\Gamma_d^{(l)}\right)^{-1} \end{bmatrix}.$$

5.4 Estimation from discrete observations

In this section, we assume to observe the process on a discrete time grid $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T_N < \infty$, for $N \in \mathbb{N}$.

First, we provide estimators, which maximize a discretized versions of likelihood and quasi-likelihood. Next, we study the asymptotic behavior of the estimators in high frequency and long time under the assumption that the process is stationary (see Definition 5.2.2).

5.4.1 Estimators expressions

There is no exploitable explicit expression for the transition densities of the T-CKLS process, nor for the finite dimensional distributions. This is also true in the well known special case of threshold Brownian motion with piecewise constant drift. Hence, instead

of considering the likelihood function associated to the sample $(X_{t_k})_{k=0}^N$, we considered a discretization of the likelihood $\mathcal{L}_T$ (5.2.2) and quasi-likelihood $q\text{-}\mathcal{L}_T$ (5.2.3). Further comments on this choice are given in Section 5.5. Once these discretizations introduced, we denote them respectively by $\mathcal{L}_{T_N,N}$ and $q\text{-}\mathcal{L}_{T_N,N}$ and we compute the estimators

$$\theta^{(\mathcal{L})}_{T_N,N} = \underset{\theta\in\Theta^{(\mathcal{L})}}{\operatorname{argmax}}\ \ln \mathcal{L}_{T_N,N}(\theta;\sigma,\gamma) \quad \text{and} \quad \theta^{(q\text{-}\mathcal{L})}_{T_N,N} = \underset{\theta\in\Theta^{(q\text{-}\mathcal{L})}}{\operatorname{argmax}}\ q\text{-}\mathcal{L}_{T_N,N}(\theta).$$

Let us denote by $Q^{j,m}_{T_N,N}$ and $M^{j,m}_{T_N,N}$ the discrete versions of $Q^{j,m}_{T_N}$ and $M^{j,m}_{T_N}$ in (5.3.1):

$$Q^{j,m}_{T_N,N} := \sum_{i=0}^{N-1} X_{t_i}^m \mathbb{1}_{I_j}(X_{t_i})(t_{i+1}-t_i) \quad \text{and} \quad M^{j,m}_{T_N,N} := \sum_{i=0}^{N-1} X_{t_i}^m \mathbb{1}_{I_j}(X_{t_i})(X_{t_{i+1}}-X_{t_i}) \tag{5.4.1}$$

for $j \in \{0,\dots,d\}$ and $m \in \{-2\gamma_j, 1-2\gamma_j, 2-2\gamma_j, 2\gamma_j\} \cup \{-1,0,1,2\}$. To obtain the discretized likelihood and quasi-likelihood, it would be natural to replace the above quantities in the continuous-time observations likelihood and quasi-likelihood functions given in (5.3.2)-(5.3.3). The discretized quasi-likelihood is then

$$q\text{-}\mathcal{L}_{T_N,N}(\theta) := \sum_{j=0}^{d} a_j M^{j,0}_{T_N,N} - b_j M^{j,1}_{T_N,N} - \frac{a_j^2}{2} Q^{j,0}_{T_N,N} - a_j b_j Q^{j,1}_{T_N,N} - \frac{b_j^2}{2} Q^{j,2}_{T_N,N}, \tag{5.4.2}$$

and we *could* do similarly for the log-likelihood. We would get

$$\ln \mathcal{L}_{T_N,N}(\theta;\sigma,\gamma) = \sum_{j=0}^{d} \frac{1}{\sigma_j^2}\left(a_j M^{j,-2\gamma_j}_{T_N,N} - b_j M^{j,1-2\gamma_j}_{T_N,N} - \frac{a_j^2}{2} Q^{j,-2\gamma_j}_{T_N,N} - a_j b_j Q^{j,1-2\gamma_j}_{T_N,N} - \frac{b_j^2}{2} Q^{j,2-2\gamma_j}_{T_N,N}\right). \tag{5.4.3}$$

Actually, we do not choose the latter quantity as discretized-log-likelihood. Instead, we consider a different discretization of $M^{j,m}$ based on an alternative expression, which depends on $M^{j,0}$ and $Q^{j,0}$ and $Q^{j,-1}$ (see Lemma 5.5.1 in Section 5.5, where we also explain this choice for the discretization). For every $j \in \{0,\dots,d\}$ we define $\mathcal{M}^{j,0}_{T_N,N} := M^{j,0}_{T_N,N}$ and for $m \in \{-2\gamma_j, 1-2\gamma_j\} \setminus \{0\}$, then

$$\mathcal{M}^{0,m}_{T_N,N} = f_{0,m+1}(X_0) - f_{0,m+1}(X_{T_N}) - \frac{m}{2}\sigma_0^2 Q^{0,m+2\gamma_0-1}_{T_N,N} + r_1^m\left(M^{0,0}_{T_N,N} + \mathfrak{f}_0\right),$$

$$\mathcal{M}^{d,m}_{T_N,N} = f_{d,m+1}(X_{T_N}) - f_{d,m+1}(X_0) - \frac{m}{2}\sigma_d^2 Q^{d,m+2\gamma_d-1}_{T_N,N} + r_d^m\left(M^{d,0}_{T_N,N} - \mathfrak{f}_d\right),$$

where $f_{0,m}(x) = \int_x^{r_1} y^{m-1}dy\,\mathbb{1}_{I_0}(x)$, $f_{d,m}(x) = \int_{r_d}^{x} y^{m-1}dy\,\mathbb{1}_{I_d}(x)$, and if $j \in \{1,\dots,d-1\}$, then

$$\begin{aligned}\mathcal{M}^{j,m}_{T_N,N} = {} & f_{j,m+1}(X_{T_N}) - f_{j,m+1}(X_0) - \frac{m}{2}\sigma_j^2 Q^{j,m+2\gamma_j-1}_{T_N,N} \\ & + r_j^m M^{j,0}_{T_N,N} + r_{j+1}^m \mathfrak{f}_{j+1} - r_j^m \mathfrak{f}_j - (r_{j+1}^m - r_j^m)\sum_{k=j+1}^{d} M^{k,0}_{T_N,N}\end{aligned} \tag{5.4.4}$$

where $f_{j,m}(x) = \int_{r_j}^{x\wedge r_{j+1}} y^{m-1}dy\,\mathbb{1}_{(r_j,+\infty)}(x)$.

Finally, we consider the following discretized log-likelihood:

$$\ln \mathcal{L}_{T_N,N}(\theta;\sigma,\gamma) = \sum_{j=0}^{d} \frac{1}{\sigma_j^2}\left(a_j \mathcal{M}_{T_N,N}^{j,-2\gamma_j} - b_j \mathcal{M}_{T_N,N}^{j,1-2\gamma_j} - \frac{a_j^2}{2} Q_{T_N,N}^{j,-2\gamma_j} - a_j b_j Q_{T_N,N}^{j,1-2\gamma_j} - \frac{b_j^2}{2} Q_{T_N,N}^{j,2-2\gamma_j} \right).$$

The following proposition establishes an explicit expression of the discretized MLE and discretized QMLE.

Proposition 5.4.1. *Let $(T_N)_{N\in\mathbb{N}}$ be a sequence in $(0,\infty)$ and let $N \in \mathbb{N}$. The maximum of the discretized likelihood is achieved at $\theta_{T_N,N}^{(\mathcal{L})} = (a_{T_N,N}^{j,\gamma_j}, b_{T_N,N}^{j,\gamma_j})_{j=0}^d$ with*

$$(a_{T_N,N}^{j,\gamma_j}, b_{T_N,N}^{j,\gamma_j}) := \left(\frac{\mathcal{M}_{T_N,N}^{j,-2\gamma_j} Q_{T_N,N}^{j,2-2\gamma_j} - Q_{T_N,N}^{j,1-2\gamma_j} \mathcal{M}_{T_N,N}^{j,1-2\gamma_j}}{Q_{T_N,N}^{j,-2\gamma_j} Q_{T_N,N}^{j,2-2\gamma_j} - (Q_{T_N,N}^{j,1-2\gamma_j})^2}, \frac{\mathcal{M}_{T_N,N}^{j,-2\gamma_j} Q_{T_N,N}^{j,1-2\gamma_j} - Q_{T_N,N}^{j,-2\gamma_j} \mathcal{M}_{T_N,N}^{j,1-2\gamma_j}}{Q_{T_N,N}^{j,-2\gamma_j} Q_{T_N,N}^{j,2-2\gamma_j} - (Q_{T_N,N}^{j,1-2\gamma_j})^2} \right). \tag{5.4.5}$$

The maximum of discretized quasi-likelihood is achieved at $\theta_{T_N,N}^{(q\text{-}\mathcal{L})} = (a_{T_N,N}^{j,0}, b_{T_N,N}^{j,0})_{j=0}^d$ with

$$(a_{T_N,N}^{j,0}, b_{T_N,N}^{j,0}) = \left(\frac{M_{T_N,N}^{j,0} Q_{T_N,N}^{j,2} - Q_{T_N,N}^{j,1} M_{T_N,N}^{j,1}}{Q_{T_N,N}^{j,0} Q_{T_N,N}^{j,2} - (Q_{T_N,N}^{j,1})^2}, \frac{M_{T_N,N}^{j,0} Q_{T_N,N}^{j,1} - Q_{T_N,N}^{j,0} M_{T_N,N}^{j,1}}{Q_{T_N,N}^{j,0} Q_{T_N,N}^{j,2} - (Q_{T_N,N}^{j,1})^2} \right). \tag{5.4.6}$$

The proof is omitted because it is analogous to the one of Proposition 5.3.1.

Note that the QMLE does not depend on the parameter vector $\sigma_\star = (\sigma_j)_{j=0}^d$, instead the MLE does because so do the expressions $\mathcal{M}_{T_N,N}^{j,m}$. Since we assume $\sigma_\star$ is not known, we replace it by an estimator. For $j \in \{0,\dots,d\}$,

$$\sigma_{T_N,N}^j = \sqrt{\frac{\mathbf{Q}_{T_N,N}^j}{Q_{T_N,N}^{j,2\gamma_j}}}, \tag{5.4.7}$$

where $\mathbf{Q}_{T_N,N}^j$ is obtained by discretizing the right hand side of formula (5.3.5) and (5.3.6) in Proposition 5.3.4. Note that $\mathbf{Q}_{T_N,N}^j$ depends on $M_{T_N,N}^{j,1}$ and $M_{T_N,N}^{j,0}$ defined in (5.4.1). For instance,

$$\mathbf{Q}_{T_N,N}^d := \mathbb{1}_{I_d}(X_T)(X_T - r_d)^2 - \mathbb{1}_{I_d}(X_0)(X_0 - r_d)^2 + 2\left(r_d M_{T_N,N}^{d,0} - M_{T_N,N}^{d,1} \right).$$

5.4.2 Asymptotic properties: high frequency - long time

In this section, we state the statistical properties of the discretized MLE and QMLE.

Let $\Delta_N := \max_{k=1,\dots,N}(t_k - t_{k-1})$ denote the maximal lag between two consecutive observations. We assume that the observation time window goes to infinity (long time) and the maximal lag between consecutive observations vanishes (high frequency):

$$\lim_{N\to+\infty} T_N = +\infty \quad \text{and} \quad \lim_{N\to+\infty} \Delta_N = 0. \tag{5.4.8}$$

Moreover, we assume that the process is stationary, see Definition 5.2.2.

The results of this section require additional assumptions on the moments of the stationary distribution and on Δ_N. Let $\mathcal{H}_{\mathcal{L}}$ and $\mathcal{H}_{q\text{-}\mathcal{L}}$ be the following assumptions:

- $\mathcal{H}_{\mathcal{L}}$: μ admits finite p-th moment and $(-q)$-th moments with
  - $p = 2 + 2\gamma_d$ and $q = 2\gamma^{\max}$ if $\gamma_0 = 0$ and $\gamma^{\max} \in [0, 1/2)$,
  - $p = \max(2 + 2\gamma_d, p_{\mathcal{L}})$ and $q = q_{\mathcal{L}}$ if $\gamma^{\max} \in (0, 1/2)$ and $\gamma_0 \neq 0$ or $\gamma^{\max} = 1/2$,
  - $p = \max(2 + 2\gamma_d, p_{\mathcal{L}}, p'\gamma^{\max})$ and $q = \max(q_{\mathcal{L}}, 2q'\gamma^{\max})$ if $\gamma^{\max} \in (1/2, 1]$,
- $\mathcal{H}_{q\text{-}\mathcal{L}} = \mathbf{H}_{q\text{-}\mathcal{L}}$: μ admits finite $(2 + 2\gamma_d)$-th moment,

where $\gamma^{\max} := \max\{\gamma_j \colon j = 0, \dots, d\}$, and $p_{\mathcal{L}}, q_{\mathcal{L}}, p', q' \geq 1$ such that $1 = 1/p_{\mathcal{L}} + 2/q_{\mathcal{L}} = 1/p' + 3/(2q')$. Observe that, if $p' = 4$ then $q' = 2$ and $p'\gamma^{\max} = 2q'\gamma^{\max} = 4\gamma^{\max}$. Note that, the conditions on moments of μ may lead to parameters restrictions for our asymptotic results (see Proposition 5.8.2), e.g. if $\gamma_0 = 1/2$. Nevertheless, if we restrict to $b_d \neq 0$ and $\gamma_0 > 1/2$ then μ admits positive and negative moments of all order. The interplay between $p_{\mathcal{L}}, q_{\mathcal{L}}, p', q'$ may come into play to reduce the required negative moments if, for instance, μ admits positive moments of all orders. See Section 5.5.3 for more details and comments on these assumptions.

Without loss of generality we assume $\Delta_N \in (0, 1)$ for all N. Then, we also introduce the following quantities $g_N^{(q\text{-}\mathcal{L})} = \Delta_N$,

$$g_N^{(\mathcal{L})} = \max_{j=0,\dots,d} \begin{cases} \Delta_N^{2-2\gamma_j} & \text{if } \gamma_j \in (\frac{3}{4}, 1), \\ \Delta_N^{2\gamma_j - 1} & \text{if } \gamma_j \in (\frac{1}{2}, \frac{3}{4}], \\ \Delta_N^{1-2\gamma_j} & \text{if } \gamma_j \in (\frac{1}{4}, \frac{1}{2}), \\ \Delta_N^{2\gamma_j} & \text{if } \gamma_j \in (0, \frac{1}{4}], \\ \Delta_N & \text{if } \gamma_j \in \{0, \frac{1}{2}, 1\}, \end{cases} \qquad g_N^{(\sigma)} := \max_{j=0,\dots,d} \begin{cases} \Delta_N^{\gamma_j} & \text{if } \gamma_j \in (\frac{1}{2}, 1), \\ \Delta_N^{2\gamma_j} & \text{if } \gamma_j \in (0, \frac{1}{2}), \\ \Delta_N & \text{if } \gamma_j \in \{0, \frac{1}{2}, 1\}. \end{cases}$$

Remark 5.4.2. $\Delta_N = g_N^{(q\text{-}\mathcal{L})} \leq g_N^{(\sigma)} \leq g_N^{(\mathcal{L})}$. *If* $\gamma^{\max} \in \{0, 1/2, 1\}$, *then* $g_N^{(\mathcal{L})} = \Delta_N$.

We are now ready to provide our first convergence result in the discrete setting. We consider the volatility estimator in (5.4.7) and we prove consistency, and show that the speed of convergence is larger than $\sqrt{T_N}$.

Theorem 5.4.3. *Assume that* (5.4.8) *holds, that the T-CKLS X is stationary and that Hypothesis* $\mathcal{H}_{q\text{-}\mathcal{L}}$ *holds. Then, the estimator* $\sigma^2_{T_N,N} = ((\sigma^j_{T_N,N})^2)_{j=0}^d$ *in* (5.4.7) *is a consistent estimators of the diffusion coefficient vector* $\sigma^2_\star = ((\sigma_j)^2)_{j=0}^d$, *i.e.*

$$\sigma^2_{T_N,N} \xrightarrow[N\to+\infty]{\mathbb{P}} \sigma^2_\star.$$

Under the additional assumption that $\lim_{N\to+\infty} T_N g_N^{(\sigma)} = 0$, *it holds that:*

$$\sqrt{T_N}\left(\sigma^2_{T_N,N} - \sigma^2_\star\right) \xrightarrow[N\to+\infty]{\mathbb{P}} 0_{\mathbb{R}^{d+1}}.$$

Since we assume that the diffusion coefficient $\sigma_\star$ vector is unknown, we replace it by estimator (5.4.7) in the expression of $\mathcal{M}^{j,\cdot}_{T_N,N}$.

The following theorems state the asymptotic properties in high frequency and long time observations of the discretized MLE and QMLE of $\theta_\star$.

Theorem 5.4.4. *Assume that* (5.4.8) *holds and that the T-CKLS X is stationary. For $\ell \in \{\mathcal{L}, q\text{-}\mathcal{L}\}$, under Hypothesis $\mathcal{H}_\ell$, the MLE and QMLE are consistent estimators of $\theta_\star$ i.e.*

$$\theta^{(\ell)}_{T_N,N} \xrightarrow[N\to+\infty]{\mathbb{P}} \theta_\star.$$

Furthermore, if $\lim_{N\to+\infty} T_N g_N^{(\ell)} = 0$, then under the same hypothesis, we have:

$$\sqrt{T_N}\left(\theta^{(\ell)}_{T_N,N} - \theta_\star\right) \xrightarrow[N\to+\infty]{stably} \mathbf{N}^{(\ell)},$$

where $\mathbf{N}^{(.)}$ is defined in Theorem 5.3.5.

The proof of Theorem 5.4.3 and Theorem 5.4.4 are postponed to Section 5.7. Nevertheless we give the main ideas and tools in Section 5.5.

Remark 5.4.5. *For $\ell \in \{\mathcal{L}, q\text{-}\mathcal{L}\}$, it follows from the previous results that the estimator $(\theta^{(\ell)}_{T_N,N}, \sigma_{T_N,N})$ is a consistent estimator of $(\theta_\star, \sigma_\star)$. By the stable convergence properties, $\sqrt{T_N}(\theta^{(\ell)}_{T_N,N}, \sigma_{T_N,N})$ converges stably to the vector $(\mathbf{N}^{(\ell)}, 0_{\mathbb{R}^{d+1}})$. If $\sigma_\star$ is known, the analogous to the LAN property in Remark 5.3.7 holds for the discretized MLE and QMLE.*

5.5 Comments on the results and their proofs

In this section, we comment the results of the previous section and we summarize the key elements of the proofs of consistency and asymptotic normality of the continuous time and discretized MLE and QMLE. Moreover, we discuss assumptions and extensions, and compare with related literature.

The main results of this chapter are the asymptotic results in the context of a T-CKLS process observed in high frequency and long time: Theorems 5.4.3 and 5.4.4. As we already mentioned, these results are also new in the context of other threshold diffusions such as T-CIR, and for mixed dynamics: CIR on a space interval, OU on another, CKLS on another one. We exploit this feature in Section 5.6. In Section 5.5.1 we comment on the estimators. In particular, we discuss the discretization choice $\mathcal{M}$ and the novelty on estimation of $\sigma_\star$.

Since the process is ergodic, the proofs of the asymptotic results in the context of continuous time observations follow from standard martingale central limit theorems. This is the case in other results for TDs, e.g. [97, 72, 81]. Therefore, in this section, we only focus on the proofs in the case of discrete observations. This is done in Section 5.5.2. Nevertheless, in Section 5.5.3, we provide some comments on the assumptions of the results both in the context of discrete and continuous time observations and we compare our assumptions with those considered in the literature.

5.5.1 On the estimators

On the volatility estimator. Estimator (5.4.7) is inspired by the results in [70]. The latter reference studies some estimators for the volatility of a threshold Brownian motion

(null drift, $\gamma = 0$, one threshold) from high frequency discrete observations over a finite time horizon. The estimators are based on quadratic variation, but on two separate intervals over which the volatility is constant. Our estimator exploits a different discretization choice for the quadratic variation over the intervals I_j, which allows to obtain information about its behavior in high frequency and long time. Up to our knowledge, this is the first time an estimator of the volatility for TDs (with $\gamma \equiv 0$ as well) is analyzed in high frequency and long time.

Likelihood discretization: $M^{\cdot,m}_{T_N,N}$ versus $\mathcal{M}^{\cdot,m}_{T_N,N}$. Let us first introduce an expression $\mathbb{P}$-a.s. equal to $M^{j,m}_T$, which would not involve any term $M^{j,k}$ except if $k = 0$. The notation, in particular the functions $f_{j,m}$, has been introduced for equation (5.4.4).

Lemma 5.5.1. *Let $T \in (0,\infty)$, $j \in \{1\dots,d\}$ and $m \in \{-2\gamma_j, 1-2\gamma_j\} \setminus \{0\}$. It holds $\mathbb{P}$-a.s. that:*

$$M^{0,m}_T = f_{0,m+1}(X_0) - f_{0,m+1}(X_T) - \frac{m}{2}\sigma_0^2 Q^{0,m+2\gamma_0-1}_T + r_1^m \left(M^{0,0}_T + \mathfrak{f}_0\right), \tag{5.5.1}$$

$$M^{d,m}_T = f_{d,m+1}(X_T) - f_{d,m+1}(X_0) - \frac{m}{2}\sigma_d^2 Q^{d,m+2\gamma_d-1}_T + r_d^m \left(M^{d,0}_T - \mathfrak{f}_d\right), \tag{5.5.2}$$

and for $j \in \{1,\dots,d-1\}$:

$$\begin{aligned} M^{j,m}_T = {} & f_{j,m+1}(X_T) - f_{j,m+1}(X_0) - \frac{m}{2}\sigma_j^2 Q^{j,m+2\gamma_j-1}_T \\ & + r_j^m M^{j,0}_T + r_{j+1}^m \mathfrak{f}_{j+1} - r_j^m \mathfrak{f}_j - (r_{j+1}^m - r_j^m) \sum_{k=j+1}^{d} M^{k,0}_T, \end{aligned} \tag{5.5.3}$$

where $\mathfrak{f}_0 = \min(X_T, r_1) - \min(X_0, r_1)$ and $\mathfrak{f}_j = \max(X_T, r_j) - \max(X_0, r_j)$ for $j \geq 1$.

Proof. We prove only the case $j \in \{1,\dots,d-1\}$, and the case $j \in \{0, d\}$ works similarly. Applying Itô-Tanaka formula (see Corollary 1.1.8 in Chapter 1 or [91, Chapter VI, exercice 1.25]), yields the a.s. equality

$$f_{j,m+1}(X_T) = f_{j,m+1}(X_0) + \frac{m}{2}\sigma_j^2 Q^{j,m+2\gamma_j-1} + M^{j,m}_T + \frac{1}{2}\left(r_j^m L^{r_j}_T(X) - r_{j+1}^m L^{r_{j+1}}_T(X)\right). \tag{5.5.4}$$

Applying a second time Itô-Tanaka formula to $\max(X_T, r_j)$ shows that a.s. $L^{r_j}_T(X) = 2\mathfrak{f}_j - 2\sum_{i=j}^d M^{i,0}_T$. Combining the latter equation with (5.5.4) completes the proof. □

We used the latter result to obtain $\mathcal{M}^{j,m}_{T_N,N}$: we just considered the discretized versions of the right hand side of (5.5.1), (5.5.2) and (5.5.3) in Lemma 5.5.1 by replacing the quantities $M^{j,0}$ and $Q^{j,m+2\gamma_j-1}$ by their discretized versions (5.4.1). Without this choice, proceeding as in the next sections, we would have obtained a more restrictive hypothesis $\mathcal{H}_{\mathcal{L}}$.

5.5.2 Key elements of proofs, comments and extensions.

We now comment on the proofs of consistency and asymptotic normality for drift and diffusion coefficients, namely Theorems 5.4.3 and 5.4.4. They rely on the next Lemma 5.5.2 and on the continuous time results Proposition 5.3.4 and Theorem 5.3.5. More precisely, one shows that the rescaled difference between the discrete-time and continuous-time estimators (e.g. $\sqrt{T_N}\left(\theta^{(\mathcal{L})}_{T_N,N} - \theta^{(\mathcal{L})}_{T_N}\right)$) vanishes as $N \to \infty$. Since all estimators depend on $M^{j,m}_{T_N,N}$ and $Q^{j,k}_{T_N,N}$ for suitable k, m, the proof is based on the following result.

Lemma 5.5.2. *Let $\lambda \in \{1, 2\}$. Assume that (5.4.8) holds and that the T-CKLS X is stationary (see Definition 5.2.2). Then*

$$\lim_{N\to\infty} T_N^{-1/\lambda}\mathbb{E}\left[|Q^{j,k}_{T_N} - Q^{j,k}_{T_N,N}|\right] = 0 \quad \textit{and} \quad \lim_{N\to\infty} T_N^{-1/\lambda}\mathbb{E}\left[|M^{j,m}_{T_N} - M^{j,m}_{T_N,N}|\right] = 0 \qquad (5.5.5)$$

for all $j \in \{0, \dots, d\}$ in each one of the following cases:

(a) (drift QMLE) for every $k \in \{0, 1, 2\}$ and $m \in \{0, 1\}$, under the assumptions: $\lim_{N\to+\infty} T_N^{\lambda-1} g_N^{(q\text{-}\mathcal{L})} = 0$ and hypothesis $\mathcal{H}_{q\text{-}\mathcal{L}}$.

(b) (volatility estimation) for every $k = 2\gamma_j$ and $m \in \{0, 1\}$, under the assumptions: $\lim_{N\to+\infty} T_N^{\lambda-1} g_N^{(\sigma)} = 0$ and hypothesis $\mathcal{H}_{q\text{-}\mathcal{L}}$.

(c) (drift MLE, $\sigma_\star$ unknown) for every $k \in \{-2\gamma_j, 1-2\gamma_j, 2-2\gamma_j, 2\gamma_j\} \cup \{-1, 0\}$ and $m \in \{0, 1\}$, under the assumptions: $\lim_{N\to+\infty} T_N^{\lambda-1} g_N^{(\mathcal{L})} = 0$ and hypothesis $\mathcal{H}_{\mathcal{L}}$.

Remark 5.5.3. *If $\lambda = 1$, then $\lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\cdot)} = 0$ is equivalent to $\lim_{N\to\infty} \Delta_N = 0$.*

The proof of Lemma 5.5.2 is provided in Section 5.7.4. It is quite technical so we give the main ideas and tools here. It relies on two auxiliary results for T-CKLS processes: Propositions 5.5.4-5.5.5, whose proofs are postponed to Section 5.7.5 and Section 5.7.6 respectively. Proposition 5.5.4 is a property very commonly used in statistics for diffusion processes.

Proposition 5.5.4. *Assume that the T-CKLS X is stationary. Let $j \in \{0, \dots, d\}$ and $m \geq 1$. Assume that the m-th moment of μ is finite. Then there exists a constant $C \in (0, \infty)$ such that for all $0 \leq s < t$ it holds $\mathbb{E}\left[|X_t - X_s|^m\right] \leq C(t-s)^{m/2}$.*

Proposition 5.5.5, instead, is the remedy to the lack of knowledge of the finite dimensional distributions of TDs. Indeed, the quantities $M^{j,\cdot}, Q^{j,\cdot}$ consider only observations taking values on I_j together with their following observation. Hence, in the proof of Lemma 5.5.2, one needs to bound the probability that the process crossed a threshold between two consecutive observations.

Proposition 5.5.5. *Assume that the T-CKLS X is stationary. Let $j \in \{0, \dots, d\}$ and $m \in \mathbb{R}$, such that the m-th moment of μ is finite on the set I_j. Then there exists a constant $C \in (0, \infty)$ such that for all $0 \leq s < t$, we have:*

$$\mathbb{E}\left[|X_s|^m \mathbb{P}_{X_s}\left(\tau^{\xi^{(j)}}_{I_j} < t - s\right) \mathbb{1}_{I_j}(X_s)\right] \leq C(t-s)^{1/2},$$

where $\xi^{(j)}$ is a CKLS process with parameters $(a_j, b_j, \sigma_j, \gamma_j)$ starting at X_s and driven by a Brownian motion independent of $\mathcal{F}_s$ (denoted by B as well), and $\tau_{I_j}^{\xi^{(j)}}$ is the first hitting time of the interval I_j for the process $\xi^{(j)}$.

Differences with respect to related literature. Let us first consider standard diffusions related to T-CKLS, for instance CIR process. The proofs in [10, 11] rely on the knowledge of the law of the process, and some other quantities such as the integral $\int_0^t (X_s)^{-1}\,\mathrm{d}s$. We do not have access to the law of T-CKLS, nor even to that of threshold Brownian motion (with multiple thresholds). Under the ergodicity assumption, in the context of discrete observations our Theorem 5.4.4, recovers and improves the existing result for CIR.

The proof strategy that we just illustrated above, is analogous to the one exploited in [81] for drift estimation of T-OU. Nevertheless, we believe that the way we deal with controlling the probability of crossing a threshold between two consecutive observations (proof of Proposition 5.5.5) is the key to extend the results of this chapter to more general TDs. Hence, in our opinion, Proposition 5.5.5 is one of the most relevant results of this chapter.

5.5.3 Comments on assumptions

Assumptions in the continuous setting: $\mathbf{H}_{q\text{-}\mathcal{L}}$ and $\mathbf{H}_{\mathcal{L}}$. In the case of T-OU process (solution to (5.2.1) with $\gamma \equiv 0$), $\mathbf{H}_{q\text{-}\mathcal{L}}$ and $\mathbf{H}_{\mathcal{L}}$ rewrite as follows: μ admits finite second moment. When T-OU is ergodic (see Table 5.4), μ admits moments of all order (see Proposition 5.8.2), so $\mathbf{H}_{q\text{-}\mathcal{L}}$ and $\mathbf{H}_{\mathcal{L}}$ hold. Indeed, for the asymptotic results in [81], no additional assumptions have been introduced.

For CIR process (solution to the SDE (5.1.1), with $\gamma = 1/2$). The asymptotic behavior of the MLE of an ergodic standard CIR process with $a, b > 0$ is studied in [11] under the additional condition $a > \sigma^2/2$. This condition ensures that μ admits finite (-1)-moment (see Proposition 5.8.2) and so $\mathbf{H}_{\mathcal{L}}$ is satisfied. Indeed, since $b > 0$ the stationary measure of the CIR process admits positive moments of all order (see Proposition 5.8.2).

In a similar way, in [83], the author studies the asymptotic behavior of the MLE of a CKLS process (solution to the SDE (5.1.1)) in the ergodic case with $a, b > 0$ and $\gamma \in (1/2, 1)$. In this case, the stationary distribution of the CKLS process admits moments of all order (see [83, Proposition 2.1] or next Proposition 5.8.2). Therefore, $\mathbf{H}_{q\text{-}\mathcal{L}}$ and $\mathbf{H}_{\mathcal{L}}$ are satisfied.

We improve the conditions in [97], where consistency and asymptotic normality of the QMLE have been proven. More precisely, in [97], the process is supposed to be stationary and geometrically ergodic and they require finiteness of the fourth order moment. We only assume ergodicity, and, for the QMLE, assumption $\mathbf{H}_{q\text{-}\mathcal{L}}$ is less restrictive than the existence of the fourth order moment.

Assumptions in the discrete setting: the discretization choice. Let us note that in the likelihood function (5.4.3) there could be a term such as $M_T^{0,-1}$ (take $\gamma_0 = 1/2$). If one takes $M_{T,N}^{0,-1}$ instead of $\mathcal{M}_{T_N,N}^{0,-1}$ for the discretization of the likelihood function, then one should prove the convergence in Lemma 5.5.2 for it. Following our proof, this requires

stronger hypothesis than the ones we consider in $\mathcal{H}_{\mathcal{L}}$. Let us mention that the idea of considering a different expression is inspired by [11], which deals with non-threshold CIR process.

In the case of the QMLE estimator, the replacement of $M^{j,1}_{T_N,N}$ by $\mathcal{M}^{j,1}_{T_N,N}$ makes no sense if $\sigma_\star$ is unknown because the estimator (5.4.7) of $\sigma_\star$ depends on $M^{j,1}_{T_N,N}$, and, if $\sigma_\star$ is known, the replacement would not allow for better assumptions (see Remark 5.7.5).

Assumptions in the discrete setting: asymptotic normality. In the case $\gamma \equiv 0$ (T-OU) or $\gamma \equiv \frac{1}{2}$ (T-CIR), the hypotheses $\lim_{N\to\infty} T_N\Delta_N = 0$ has been considered for the asymptotic normality of estimators for CIR in [11], T-OU process and threshold drifted Brownian motion in [81, 82]. The condition degrades when the diffusion coefficient is non-linear in some interval, *i.e.* when there exists $j \in \{0, \dots, d\}$ such that γ_j is not equal to $0, 1/2$, or 1. Indeed, $\Delta_N = g_N^{(q\text{-}\mathcal{L})} \leq g_N^{(\sigma)} \leq g_N^{(\mathcal{L})}$.

The quantities $g_N^{(\mathcal{L})}$ and $g_N^{(\sigma)}$ involve the maximum over all j of some powers of Δ_N depending on γ_j. Indeed, for each interval I_j we get a condition for convergence on Δ_N depending on γ_j (for a rigorous proof, see Section 5.7.4), and, taking the most restrictive condition, corresponds to taking the maximum.

Let us comment on the Hypothesis $\mathcal{H}_{\mathcal{L}}$. Finiteness of moments for the stationary measure μ is summarized in Proposition 5.8.2. When $\gamma_0 = 0$ and $\gamma^{\max} < 1/2$, $\mathcal{H}_{\mathcal{L}}$ is satisfied. When $\gamma_0 = 1/2$ and $b_d > 0$, $\mathcal{H}_{\mathcal{L}}$ is satisfied if $a_0 > \sigma_0^2$. Indeed, $q_{\mathcal{L}}$ in hypothesis $\mathcal{H}_{\mathcal{L}}$ can be taken such that $2 < q_{\mathcal{L}} < 2a_0/\sigma_0^2$. This condition is the one in [11, Proposition 5] in the case of the CIR process (solution to (5.1.1) with $\gamma = 1/2$).

Theorem 5.4.4 when $\sigma_\star$ is known. When $\sigma_\star$ is known, there is no need to replace it by its estimator. Theorem 5.4.4 holds with weaker assumptions, we show this in Section 5.7.4. The case of the QMLE is discussed in Remark 5.7.5. Nevertheless, the assumption $\mathcal{H}_{q\text{-}\mathcal{L}}$ have to be more restrictive than $\mathbf{H}_{q\text{-}\mathcal{L}}$ and it is already equal to it. The case of the MLE is developed in details in the paragraph for the proof of Item (c) (see Section 5.7.4).

5.6 Numerical Experiments

In this section, we implement the MLE and QMLE based on discrete observations on simulated and US interest rates data.

5.6.1 Simulated Data

In this section, we investigate the efficiency of our estimators on simulated data. We simulate the T-CKLS process combining known Euler-Type schemes on different intervals such as the schemes in [5, equation (3)] when the process is a CIR or a drifted version of the scheme in [99] when the diffusion coefficient is non-linear. More precisely we use the

following scheme. Given $X_0 \in (0,\infty)$ and $(G_k)_{k\in\mathbb{N}}$ a sequence of i.i.d. standard Gaussian random variables, next we set $X_0^{(n)} = X_0$ and, we define for all $k \in \mathbb{N}$

$$X^{(n)}_{(k+1)/n} := \left| X^{(n)}_{k/n} + \frac{1}{n}\left(a(X^{(n)}_{k/n}) - b(X^{(n)}_{k/n})X^{(n)}_{k/n}\right) + \frac{1}{\sqrt{n}}\sigma(X^{(n)}_{k/n})(X^{(n)}_{k/n})^{\gamma(X^{(n)}_{k/n})}G_k \right|.$$

Further discussion about the most suitable numerical scheme for T-CKLS is beyond the purpose of this chapter. To estimate the parameters from the simulated data, we use the estimators from discrete observations in Section 5.4.1. The implementation has been done using Matlab and the parameters are as in Table 5.1.

a_0	b_0	σ_0	γ_0
0.3	0.2	0.2	0.5

a_1	b_1	σ_1	γ_1
0	0	0.4	0

a_2	b_2	σ_2	γ_2
0.3	0.2	0.2	0.5

r_1	r_2
1	1.5

Table 5.1: Simulations parameters.

Firstly, we illustrate Theorem 5.4.4 for the drift parameters $\theta_\star = (a_i, b_i)_{i=0}^{d=2}$, diffusion parameters $\sigma_\star$, γ and two thresholds r_1, r_2 given in Table 5.1. We consider a process, which follows a CIR dynamic close to 0 and far away from 0 and is a BM on an intermediate bounded interval. We simulate 10^4 trajectories of the T-CKLS with two threshold. The set of numerical parameters is $(T, N) = (10^3, 10^6)$ with starting condition determined as follows. As the process is supposed to be stationary, we first simulate one trajectory starting from $X_0 > 0$ chosen arbitrarily, say $X_0 = 1$, with the scheme $X^{(n)}$ given above with $n = kN/T$ for some $k \in \mathbb{N} \setminus \{0\}$ (we took $k = 1$). Then, we consider the final value of the latter trajectory as initial condition of the 10^4 trajectories.

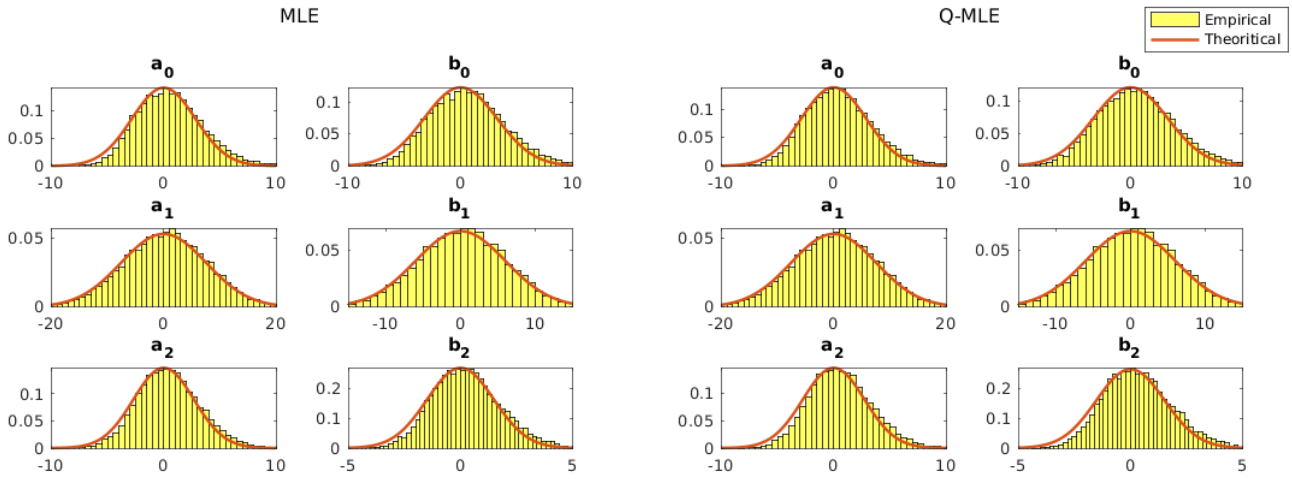


Figure 5.1: Asymptotic normality property in Theorem 5.4.4, with parameters as in Table 5.1. We plot the theoritical distribution using (5.7.1) and compare with the empirical distribution on 10^4 trajectories.

Remark that, despite the fact that the set of numerical parameters (T, N) does not satisfy the conditions for the asymptotic normality in Theorem 5.4.4 ($T^2 = N$ instead of $T^2 \ll N$), numerics show good results.

We compare the estimators by means of relative root-mean-square error (RMSE) and mean-error (ME) in Table 5.2.

Estimator	Relative RMSE	ME
$(a^{0,(\mathcal{L})}_{T_N,N}, a^{1,(\mathcal{L})}_{T_N,N}, a^{2,(\mathcal{L})}_{T_N,N})$	(0.3206, 0.2341, 0.2921)	(0.0154, 0.0092, 0.0156)
$(b^{0,(\mathcal{L})}_{T_N,N}, b^{1,(\mathcal{L})}_{T_N,N}, b^{2,(\mathcal{L})}_{T_N,N})$	(0.5500, 0.1857, 0.2439)	(0.0173, 0.0069, 0.0093)
$(a^{0,(q\text{-}\mathcal{L})}_{T_N,N}, a^{1,(q\text{-}\mathcal{L})}_{T_N,N}, a^{2,(q\text{-}\mathcal{L})}_{T_N,N})$	(0.3233, 0.2341, 0.2965)	(0.0170, 0.0092, 0.0191)
$(b^{0,(q\text{-}\mathcal{L})}_{T_N,N}, b^{1,(q\text{-}\mathcal{L})}_{T_N,N}, b^{2,(q\text{-}\mathcal{L})}_{T_N,N})$	(0.5555, 0.1857, 0.2478)	(0.0182, 0.0069, 0.0112)
$(\sigma^0_{T_N,N}, \sigma^1_{T_N,N}, \sigma^2_{T_N,N})$	(0.0088, 0.0087, 0.0015)	(0.0087, -0.0034, 0.001)

Table 5.2: Table of the relative RMSE and ME for the estimator of $(\theta_\star, \sigma_\star)$ using the MLE (5.4.5), QMLE (5.4.6) and volatility estimator (5.4.7).

Observe that the MLE gives a better estimation of the drift parameters, which is easily explained by the fact that the likelihood contains more information about the model. In general the MLE tends to have a better RMSE and ME than the QMLE. Applying the estimator on several data sets, we remark that the QMLE has a greater variance than the MLE.

Remark 5.6.1 (Threshold estimation)**.** *The thresholds can be estimated using the method proposed in [97] based on continuous observations. We use this procedure in the next section. It is a QMLE-based method without an explicit expression for the threshold estimator. Consequently, the numerical cost of this method increases with the number of thresholds. However, satisfactory results can still be achieved in the case of two thresholds. Mathematical study of the estimator for discrete observations should be conducted.*

Remark 5.6.2 (Diffusion exponent estimation)**.** *The parameter γ was estimated in [83] for non-threshold CKLS. A study on the estimator properties for discrete observations in the context of CKLS or its generalisation to T-CKLS is necessary. However, we expect that the estimator converges quite slowly, especially around 0.*

5.6.2 Interest rates analysis

In this section, we apply our estimators to the ten year US treasury rate based on the Federal Reserve Bank's H15 data set. We apply a discretized version of the test to evaluate the existence of one threshold (see [97], see Remark 5.6.1) or more thresholds (see [98]) in this dataset. In [98], the authors introduce a test statistic to identify thresholds in the drift term of a diffusion model: detect their presence and estimate them. They also develop a computationally efficient approach to calibrate the p-value and extend the test

to detect multiple thresholds. Inspired by [113], we exploit here a discretized version of the method presented in [98] and combine with the drift MLE and $\sigma_\star$ estimator considered in this chapter.

General step. Let us describe the test to be applied at a general step and then the procedure describing the steps. Suppose that there are m thresholds, and that we look for the presence of an additional threshold on the k-th interval $I_k = (r_{k-1}, r_k)$ (we know the value of r_{k-1}, r_k: either known or estimated in previous steps). We consider the hypothesis:

$$\begin{cases} H_0: \text{ Null hypothesis} & m \text{ thresholds;} \\ H_1: \text{ Alternative hypothesis} & (m+1) \text{ thresholds.} \end{cases} \tag{5.6.1}$$

Under the null hypothesis H_0, the model has m thresholds. Under hypothesis H_1, there is an additional threshold $\bar{r}$ in the k-th regime, meaning that the sequence of thresholds becomes $-\infty = r_0 < r_1 < \ldots < r_{k-1} < \bar{r} < r_k < \ldots < r_m$. The quasi-likelihood ratio test statistic is given by:

$$T = \sup_{\bar{r} \in [a,b]} T(\bar{r})$$

with

$$T(\bar{r}) := 2\left(q\text{-}\mathcal{L}_{T_N,N}\left(\theta^{(H_1)}_{T_N,N}(m+1,\bar{r})\right) - q\text{-}\mathcal{L}_{T_N,N}\left(\theta^{(H_0)}_{T_N,N}(m)\right)\right) \tag{5.6.2}$$

where a and b are 20 and 80 percentiles of the data in $[r_{k-1}, r_k]$, the value $\theta^{(H_1)}_{T_N,N}(m+1,\bar{r})$ is the drift MLE under hypothesis H_1 with the additional threshold given by $\bar{r}$, and $\theta^{(H_0)}_{T_N,N}(m)$ is the drift MLE of the model under hypothesis H_0. The MLE is given in Proposition 5.4.1, with volatility parameter $\sigma_\star$ estimated by (5.4.7).

We compute the statistics for some values of $\bar{r}$, say $\bar{r}_j := a(1-j/n)+bj/n$, $j \in \{0, 1, \ldots, n\}$ (we choose $n = 10^3$). Next, we take as an estimator for the threshold r_j the $\bar{r}_j$ which maximises $T(\bar{r})$ and the observed test statistics T_{data} is then the quantity $T(\bar{r}_j)$.

The distribution of the test statistic (5.6.2) is obtained using a bootstrap method. To compute the p-value we simulate 10^3 trajectories of the process with the parameters under H_0, we compute T_j the statistics on the j-th trajectory. Then the p-value is given by $\#\{j\colon T_{data} < T_j\}/10^3$.

We fix the significance level at the conventional 5%.

Procedure. We apply the test above as follows, in a sequential procedure. We first test for the presence of a threshold on the data: applying this test for $m = 0$. If the test is significant, then we take as an estimator for the threshold $\widehat{r}_1$ the $\bar{r}$ which realises the maximum in the test statistics. This threshold divides the state space into two intervals. We then test the presence of thresholds on each of the two intervals, starting from the left to the right. On each interval, if the test is significant, we keep dividing the interval into two sub-intervals and so on. Once we do not have evidence of new thresholds in the interval we are considering, we go to the next interval.

Application to ten year US treasury rate. We consider the ten year US Treasury rate and we adopt the convention that the daily time intervals is $\mathrm{d}t = 0.046$, where one unit of time represents one month. We assume that the data follow a T-CIR dynamics, *i.e.* $\gamma \equiv 1/2$.

We consider the ten year US Treasury rate for two different time window: Jan 2016 - Dec 2019, and Jan 2020 - Jan 2024 represented in Figure 5.2.

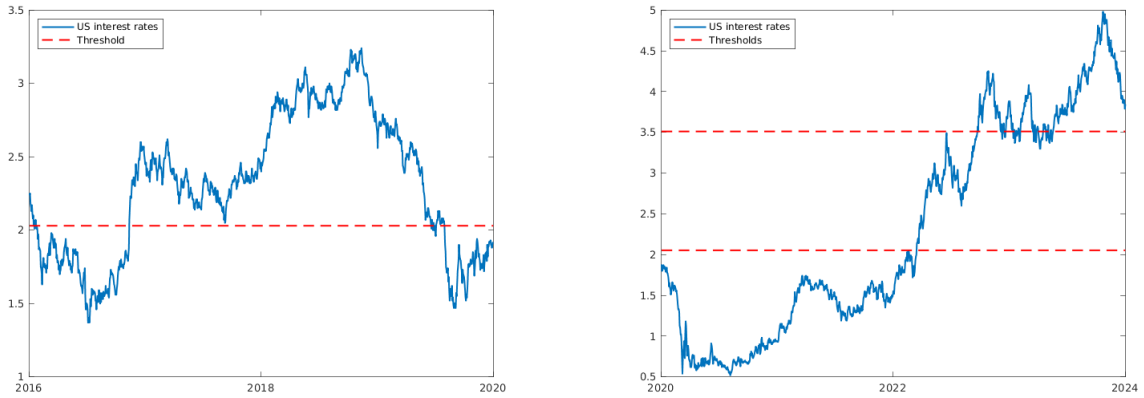


Figure 5.2: The figure shows the interest rate daily data (solid line) for the time window Jan 2016 - Dec 2019, and Jan 2020 - Jan 2024. The fitted thresholds are represented by the dashed lines.

Let us consider the time window, Jan 2016 - Dec 2019. The threshold test (5.6.1) for $m = 0$ is significant and the threshold estimation is $r_1 = 2.0303$. We apply the test (5.6.1) to detect a threshold on $(0, r_1)$. It is not significant. The same conclusion holds testing for threshold presence on $(r_1, +\infty)$.

On the time window, Jan 2020 - Jan 2024, the threshold test (5.6.1) for $m = 0$ is significant and we estimate the threshold $r_1 = 2.0507$. There is no evidence of additional thresholds on $(0, r_1)$, and the null hypothesis is rejected for the existence of a threshold $r_2 = 3.5112$ on $(r_1, +\infty)$. Instead, for the tests (5.6.1) with $m = 2$ for finding a threshold on (r_1, r_2) and $(r_2, +\infty)$, the null hypothesis H_0 is not-rejected.

Jan 2016 - Dec 2019	
Estimator	Value
$(a^{0,(\mathcal{L})}_{T_N,N}, a^{,(\mathcal{L})}_{T_N,N})$	(1.6434, 0.1713)
$(b^{0,(\mathcal{L})}_{T_N,N}, b^{1,(\mathcal{L})}_{T_N,N})$	(0.9410, 0.0723)
$(\sigma^0_{T_N,N}, \sigma^1_{T_N,N})$	(0.1616, 0.1053)
$r^1_{T_N,N}$	(2.0303)

Jan 2020 - Jan 2024	
Estimator	Value
$(a^{0,(\mathcal{L})}_{T_N,N}, a^{1,(\mathcal{L})}_{T_N,N}, a^{2,(\mathcal{L})}_{T_N,N})$	(0.2013, 0.5826, -0.0207)
$(b^{0,(\mathcal{L})}_{T_N,N}, b^{1,(\mathcal{L})}_{T_N,N}, b^{2,(\mathcal{L})}_{T_N,N})$	(0.1556, 0.0670, 0.0236)
$(\sigma^0_{T_N,N}, \sigma^1_{T_N,N}, \sigma^2_{T_N,N})$	(0.2129, 0.2091, 0.1807)
$(r^1_{T_N,N}, r^2_{T_N,N})$	(2.0507, 3.5112)

Table 5.3: Estimated parameters corresponding to Figure 5.2.

Therefore, we conclude that there is a single threshold in the time window, Jan 2016 - Dec 2019, and two thresholds in Jan 2020 - Jan 2024. In Table 5.3, we summarize the values obtained for each of the fitted parameters using the estimators.

5.7 Proofs

In this section, recalling some well known results, we show under which conditions, on the parameters, the process admits a stationary distribution. Then, we prove Theorem 5.4.3, Theorem 5.4.4, Lemma 5.5.2, Proposition 5.5.4 and Proposition 5.5.5.

5.7.1 The process: properties of solutions and conditions for the stationary distribution

T-CLKS shows several behaviors, it may behave as an OU process on some intervals, a CIR or CKLS in others. The state space of T-CKLS is determined by the process behavior around 0, in particular at I_0 where it behaves as a standard CKLS process. The regime of the process (transient, recurrent, positive-recurrent) is also determined by the behavior at I_0 and I_d.

Let us introduce the scale function s and the speed measure $m(x)\,\mathrm{d}x$. The interested reader could refer to Chapter 1 or [19, II.4] for a summary or find more details e.g. in [91, Chapter VII, Section 3]. The scale function is continuous, unique up to a multiplicative constant, and its derivative satisfies $s'(x) = \exp\left(-\int_{r_1}^{x} \frac{2(a(y)-b(y)y)}{\sigma(y)^2 y^{2\gamma(y)}}\,\mathrm{d}y\right)$. The speed measure is given by $m(x)\,\mathrm{d}x = \frac{2}{(\sigma(x))^2|x|^{2\gamma(x)}s'(x)}\,\mathrm{d}x$. The state of space of the T-CKLS process, denoted $I = \cup_{j=0}^{d} I_j$, depends on the value of the parameters in I_0 (a_0, σ_0 and γ_0). When $\gamma_0 \in [1/2, 1]$, a suitable comparison theorem for SDEs ensures that the process is non-negative. This, together with an adaptation of the proof of Proposition 1.5 in [91, Chapter XI] and the Feller boundary classification criteria (see Section 1.3.3 in Chapter 1 or [91]), imply the following lemma.

Lemma 5.7.1. *Let X be the solution to the SDE* (5.2.1).

- *If $\gamma_0 = 0$, the state of space of the process is $I = \mathbb{R}$.*
- *If $\gamma_0 = 1/2$ and $0 < a_0 < \sigma_0^2/2$, the state space is $I = [0, +\infty)$ and the point 0 is instantaneously reflecting.*
- *If $\gamma_0 \in (1/2, 1]$ or if $\gamma_0 = 1/2$ and $a_0 \geq \sigma_0^2/2$, then $I = (0, +\infty)$ and 0 is an unattainable boundary.*

The regime of the process can be obtained by properties on the scale function and the spead measure (see Section 1.3.4, [91, Exercice 3.15 in Chapter X] and [53, Theorem 20.15]). We recall that, in the recurrent case the measure $m(x)\,\mathrm{d}x$ is a stationary measure

and the fact that $m(x)\,\mathrm{d}x$ is a finite measure corresponds to positive recurrence (*ergodicity*) of the process. In the ergodic case, the stationary measure can be renormalized to the stationary distribution:

$$\mu(\mathrm{d}x) = \frac{m(x)}{\int_I m(y)\,\mathrm{d}y}\,\mathrm{d}x. \tag{5.7.1}$$

The recurrent positivity property of the process only depends on the parameters below the first threshold (on I_0) and above the last threshold (on I_d). In the following table, we give conditions on the parameters a_0, a_d, b_0, b_d and σ_0, σ_d depending on the value of γ_0, γ_d such that the process is ergodic (admits a stationary distribution).

$\gamma_0 = 0$	$a_0 \in \mathbb{R}$ and $b_0 > 0$, or $a_0 > 0$ and $b_0 = 0$
$\gamma_0 \in [1/2, 1)$	$a_0 > 0$ and $b_0 \in \mathbb{R}$
$\gamma_0 = 1$	$a_0 > 0$ and $b_0 \in \mathbb{R}$, or $a_0 = 0$ and $b_0 < -\sigma_0^2/2$

	$a_d \in \mathbb{R}$ and $b_d > 0$, or
$\gamma_d \in [0, 1/2]$	$a_d < 0$ and $b_d = 0$
$\gamma_d \in (1/2, 1)$	$a_d \in \mathbb{R}$ and $b_d = 0$
$\gamma_d = 1$	$a_d \in \mathbb{R}$ and $b_d \in (-\sigma_d^2/2, 0]$

Table 5.4: Parameter conditions for ergodicity of T-CKLS process X solution to (5.2.1).

5.7.2 Proof of Theorem 5.4.3

This proof relies on Lemma 5.5.2 whose proof is provided in Section 5.7.4.

We study the asymptotic behavior of $(\sigma_{T_N,N})^2 - \sigma_\star^2$. For $j \in \{0, \dots, d\}$, by Proposition 5.3.4, on the event $\{Q_T^{j,0} > 0\}$, we have:

$$(\sigma^j_{T_N,N})^2 - \sigma_j^2 = \sigma_j^2 \frac{Q^{j,2\gamma_j}_{T_N} - Q^{j,2\gamma_j}_{T_N,N}}{Q^{j,2\gamma_j}_{T_N,N}} + \frac{\mathbf{Q}^j_{T_N,N} - \mathbf{Q}^j_{T_N}}{Q^{j,2\gamma_j}_{T_N,N}}.$$

By equations (5.3.5)-(5.3.6), we have:

$$\left|\mathbf{Q}^j_{T_N} - \mathbf{Q}^j_{T_N,N}\right| \le 2\left|M^{j,1}_{T_N} - M^{j,1}_{T_N,N}\right| + 2\max(|r_j|, |r_{j+1}|) \sum_{i=j}^{d} \left|M^{i,0}_{T_N} - M^{i,0}_{T_N,N}\right|,$$

for $j \in \{1, \dots, d-1\}$ and for $j \in \{0, d\}$

$$\left|\mathbf{Q}^j_{T_N} - \mathbf{Q}^j_{T_N,N}\right| \le 2\left|M^{j,1}_{T_N} - M^{j,1}_{T_N,N}\right| + 2|r_{(j+1)\wedge d}|\left|M^{j,0}_{T_N} - M^{j,0}_{T_N,N}\right|.$$

Then, we conclude by making use of item (b) in Lemma 5.5.2 and the fact that $\mathbb{P}\left(\lim_{T\to+\infty} Q_T^{j,0} > 0\right) = 1$.

5.7.3 Proof of Theorem 5.4.4

This proof relies on Lemma 5.5.2 that we prove in the next section. For all $N \in \mathbb{N}$ and $\ell \in \{\mathcal{L}, q\text{-}\mathcal{L}\}$ it holds that:

$$\left(\theta^{(\ell)}_{T_N,N} - \theta_\star\right) = \left(\theta^{(\ell)}_{T_N,N} - \theta^{(\ell)}_{T_N}\right) + \left(\theta^{(\ell)}_{T_N} - \theta_\star\right).$$

The second term, on the right hand side of the equality, provides the asymptotic behavior by applying Theorem 5.3.5.
In the case $\ell = \mathcal{L}$, for $j \in \{0, \dots, d\}$, using equations (5.3.4) and (5.4.5), each component of the first term can be rewritten as follows:

$$\left(\frac{Q^{j,k}_{T_N,N}}{Q^{j,-2\gamma_j}_{T_N,N} Q^{j,2-2\gamma_j}_{T_N,N} - (Q^{j,1-2\gamma_j}_{T_N,N})^2} - \frac{Q^{j,k}_{T_N}}{Q^{j,-2\gamma_j}_{T_N} Q^{j,2-2\gamma_j}_{T_N} - (Q^{j,1-2\gamma_j}_{T_N})^2}\right) \mathcal{M}^{j,m}_{T_N} + \frac{Q^{j,k}_{T_N,N}(\mathcal{M}^{j,m}_{T_N,N} - \mathcal{M}^{j,m}_{T_N})}{Q^{j,-2\gamma_j}_{T_N,N} Q^{j,2-2\gamma_j}_{T_N,N} - (Q^{j,1-2\gamma_j}_{T_N,N})^2},$$

with $k \in \{-2\gamma_j, 1-2\gamma_j, 2-2\gamma_j\}$ and $m \in \{-2\gamma_j, 1-2\gamma_j\}$. Let $\bar{r} := \max\{|r_j| \colon j = 1, \dots, d\}$. Then, using formula (5.5.1), (5.5.2) and (5.5.3), for $j \in \{1, \dots, d-1\}$, we have:

$$|\mathcal{M}^{j,m}_{T_N} - \mathcal{M}^{j,m}_{T_N,N}| \le \bar{r}^m \sum_{i=j}^{d} |M^{i,0}_{T_N} - M^{i,0}_{T_N,N}| + \frac{|m|}{2} \left|\left(\sigma^j_{T_N,N}\right)^2 - \sigma_j^2\right| \left|Q^{j,m+2\gamma_j-1}_{T_N} - Q^{j,m+2\gamma_j-1}_{T_N,N}\right|,$$

and for $j \in \{0, d\}$:

$$|\mathcal{M}^{j,m}_{T_N} - \mathcal{M}^{j,m}_{T_N,N}| \le \bar{r}^m |M^{j,0}_{T_N} - M^{j,0}_{T_N,N}| + \frac{|m|}{2} \left|\left(\sigma^j_{T_N,N}\right)^2 - \sigma_j^2\right| \left|Q^{j,m+2\gamma_j-1}_{T_N} - Q^{j,m+2\gamma_j-1}_{T_N,N}\right|.$$

This, Item (c) in Lemma 5.5.2, and Theorem 5.4.3 ensure that

$$T_N^{-1/\lambda} \left|Q^{j,k}_{T_N,N} - Q^{j,k}_{T_N}\right| \xrightarrow[N\to+\infty]{\mathbb{P}} 0 \quad \text{and} \quad T_N^{-1/\lambda} \left|\mathcal{M}^{j,m}_{T_N,N} - \mathcal{M}^{j,m}_{T_N}\right| \xrightarrow[N\to+\infty]{\mathbb{P}} 0$$

with $\lambda = 1$ to get the consistency of the MLE and $\lambda = 2$ for the speed of convergence.

The case $\ell = q\text{-}\mathcal{L}$, the first term works analogously, exploiting Item (a) in Lemma 5.5.2.

5.7.4 Proof of Lemma 5.5.2

We prove (5.5.5) under the assumptions corresponding to the different possible values of k and m appearing in the different estimators we consider.

Let us recall that (5.4.8) holds and that the process is stationary, where μ denotes the stationary distribution. Let us introduce the round ground notation $\lfloor t \rfloor_{\Delta_N} := t_k$ for $t \in [t_k, t_{k+1}) \subseteq [t_k, t_k + \Delta_N]$. Moreover, without loss of generality, we assume $T_N \le N$ and $\Delta_N \le 1$ for all $N \in \mathbb{N}$.

Analogously to [81], the proof relies on Proposition 5.5.4 and 5.5.5, that we prove in the next sections. Basically, for $\lambda \in \{1, 2\}$, the proof of Lemma 5.5.2 reduces to prove that the following integrals are $o(T_N^{1/\lambda})$:

$$\int_0^{T_N} \mathbb{E}\left[|X_t - X_{\lfloor t \rfloor_{\Delta_N}}|^p\right] \mathrm{d}t, \tag{5.7.2}$$

$$\int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{1}_{X_t \notin I_j, X_{\lfloor t \rfloor_{\Delta_N}} \in I_j}\right] \mathrm{d}t \quad \text{and} \quad \int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{1}_{X_t \in I_j, X_{\lfloor t \rfloor_{\Delta_N}} \notin I_j}\right] \mathrm{d}t \tag{5.7.3}$$

for some suitable $p > 0$ and $q \in \mathbb{R}$ and specific assumptions, all depending on the value of m and k appearing in $M^{j,m}$ and $Q^{j,k}$.

Proposition 5.5.4 and Proposition 5.5.5, once they can be applied (i.e. if μ admits finite moments of order $\max(p, 1)$ and q respectively), yield upper bounds involving T_N and Δ_N for (5.7.2) and (5.7.3) :

$$\int_0^{T_N} \mathbb{E}\left[|X_t - X_{\lfloor t \rfloor_{\Delta_N}}|^p\right] \mathrm{d}t \leq \int_0^{T_N} \mathbb{E}\left[|X_t - X_{\lfloor t \rfloor_{\Delta_N}}|^{\max(p,1)}\right]^{\min(p,1)} \mathrm{d}t \leq C T_N \Delta_N^{p/2} \tag{5.7.4}$$

and

$$\int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q (\mathbb{1}_{X_t \notin I_j, X_{\lfloor t \rfloor_{\Delta_N}} \in I_j} + \mathbb{1}_{X_t \in I_j, X_{\lfloor t \rfloor_{\Delta_N}} \notin I_j})\right] \mathrm{d}t \leq C T_N \sqrt{\Delta_N} \tag{5.7.5}$$

for some positive constant C independent of N. The assumptions on $g_N^{(\cdot)}$ ensure that these quantities are $o(T^{1/\lambda})$.

The proof of each item of Lemma 5.5.2 is then reduced to determine the corresponding values of p and q in (5.7.4) and (5.7.5) and so deduce the assumptions $\mathcal{H}_{\mathcal{L}}$ and $\mathcal{H}_{q\text{-}\mathcal{L}}$ on moments of μ and the conditions on $g_N^{(\cdot)}$. This is done in the next sections. We consider separately the two quantities in (5.5.5) and then we summarize the items of Lemma 5.5.2 in Section 5.7.4.

To conclude the section, for the reader's convenience, let us rewrite equation (5.7.3) in such a way that it can be more easily recognised that Proposition 5.5.5 can be applied. The tower property of conditional expectation and Markov property imply that:

$$\mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{1}_{X_t \notin I_j, X_{\lfloor t \rfloor_{\Delta_N}} \in I_j}\right] \leq \mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{P}_{X_{\lfloor t \rfloor_{\Delta_N}}}\left(\tau_{I_j}^{\xi^{(j)}} < \Delta_N\right) \mathbb{1}_{I_j}(X_{\lfloor t \rfloor_{\Delta_N}})\right],$$

and

$$\mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{1}_{X_t \in I_j, X_{\lfloor t \rfloor_{\Delta_N}} \notin I_j}\right] \leq \sum_{\substack{i=0 \\ i \neq j}}^{d} \mathbb{E}\left[|X_{\lfloor t \rfloor_{\Delta_N}}|^q \mathbb{P}_{X_{\lfloor t \rfloor_{\Delta_N}}}\left(\tau_{I_i}^{\xi^{(i)}} < \Delta_N\right) \mathbb{1}_{I_i}(X_{\lfloor t \rfloor_{\Delta_N}})\right],$$

where $\xi^{(j)}$ is a CKLS process with parameter $(a_j, b_j, \sigma_j, \gamma_j)$ starting at $\xi_0^{(j)} = X_{\lfloor t \rfloor_{\Delta_N}}$ but driven by a BM, (denoted by B as well), independent of $\mathcal{F}_{\lfloor t \rfloor_{\Delta_N}}$ and $\tau_{I_j}^{\xi_j}$ is the first hitting time of the interval I_j for the process $\xi^{(j)}$. The right-hand-side of the latter equations is of the same kind of (5.7.3).

Dealing with $\mathbb{E}\left[|Q^{j,k}_{T_N} - Q^{j,k}_{T_N,N}|\right]$.

In this section, exploiting (5.7.4)-(5.7.5), we precise under which assumptions on the moments of μ and on Δ_N, it holds that $\mathbb{E}\left[|Q^{j,k}_{T_N} - Q^{j,k}_{T_N,N}|\right]$ is $o(T_N^{1/\lambda})$ for $k \in [-2,2]$, $j \in \{0,\dots,d\}$. Let us remind that we are interested in $k \in \{-2\gamma_j, 1-2\gamma_j, 2-2\gamma_j, 2\gamma_j\} \cup \{-1,0,1,2\}$.

Let us first note that

$$\begin{aligned} Q^{j,k}_{T_N} - Q^{j,k}_{T_N,N} &= \int_0^{T_N} (X_t^k - X^k_{\lfloor t\rfloor_{\Delta_N}})\mathbb{1}_{I_j}(X_t)\,\mathrm{d}t + \int_0^{T_N} X^k_{\lfloor t\rfloor_{\Delta_N}}(\mathbb{1}_{I_j}(X_t) - \mathbb{1}_{I_j}(X_{\lfloor t\rfloor_{\Delta_N}}))\,\mathrm{d}t \\ &= \int_0^{T_N} X^k_{\lfloor t\rfloor_{\Delta_N}}(\mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j} - \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j})\,\mathrm{d}t + \int_0^{T_N} (X_t^k - X^k_{\lfloor t\rfloor_{\Delta_N}})\mathbb{1}_{I_j}(X_t)\,\mathrm{d}t, \end{aligned}$$

therefore

$$\begin{aligned} \mathbb{E}\left[|Q^{j,k}_{T_N} - Q^{j,k}_{T_N,N}|\right] &\le \int_0^{T_N} \mathbb{E}\left[|X_t^k - X^k_{\lfloor t\rfloor_{\Delta_N}}|\mathbb{1}_{I_j}(X_t)\right]\,\mathrm{d}t \\ &+ \int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t\rfloor_{\Delta_N}}|^k \mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j}\right]\,\mathrm{d}t + \int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t\rfloor_{\Delta_N}}|^k \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j}\right]\,\mathrm{d}t. \end{aligned} \tag{5.7.6}$$

Inequality (5.7.6) involves only terms of the kind (5.7.3) with $q = k$ (leading to the bounds in (5.7.5)) and a term similar to (5.7.2):

$$\int_0^{T_N} \mathbb{E}\left[|X_t^k - X^k_{\lfloor t\rfloor_{\Delta_N}}|\mathbb{1}_{I_j}(X_t)\right]\,\mathrm{d}t. \tag{5.7.7}$$

This term is trivial if $k = 0$, it is bounded from above by (5.7.2) if $k = 1$, and, in all other cases, we manage to bound (5.7.7) with functions as (5.7.2). We then derive the desired conditions by (5.7.4)-(5.7.5). Our bounds are not necessarily optimal.

The case $k = 0, 1$. μ admits finite k-th moment and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N = 0$.

The case $k \in (0,1)$. μ admits finite first moment and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N^k = 0$.

Note that $k \in (0,1)$ if $\gamma_j \in (0,1/2)$ for $k = 2\gamma_j$, $k = 1-2\gamma_j$ and if $\gamma_j \in (1/2,1)$ for $k = 2(1-\gamma_j)$. Sub-additivity of $x \mapsto x^k$ yields (5.7.4).

Remark 5.7.2. *The inequalities in the next paragraph work for all $k \in (0,2]$. Nevertheless, when $k \in (0,1)$ the previous paragraph, requires less restrictive assumptions: $\Delta_N^k \le \Delta_N^{k/2}$.*

The case $k \in (1,2]$. μ admits finite k-th moment and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N^{k/2} = 0$.

This is the case if $\gamma_j \in (1/2,1]$ when $k = 2\gamma_j$, and if $\gamma_j \in [0,1/2)$ when $k = 2(1-\gamma_j)$. Cauchy-Schwarz inequality implies

$$\mathbb{E}\left[|X_t^k - X_{\lfloor t\rfloor_{\Delta_N}}^k|\mathbb{1}_{I_j}(X_t)\right] \leq \sqrt{\mathbb{E}\left[\left(X_t^{k/2} - X_{\lfloor t\rfloor_{\Delta_N}}^{k/2}\right)^2\right]\left(\mathbb{E}\left[X_t^k\right] + \mathbb{E}\left[X_{\lfloor t\rfloor_{\Delta_N}}^k\right]\right)}.$$

Moreover, the fact that $x \mapsto x^{k/2}$ is sub-additive ensures that

$$\mathbb{E}\left[\left(X_t^{k/2} - X_{\lfloor t\rfloor_{\Delta_N}}^{k/2}\right)^2\right] \leq \mathbb{E}\left[|X_t - X_{\lfloor t\rfloor_{\Delta_N}}|^k\right].$$

The case $k \in [-1,0)$. μ admits finite moments of order $-q_{\mathcal{L}}$ and $p_{\mathcal{L}}$ with $p_{\mathcal{L}}, q_{\mathcal{L}} \geq 1$ such that $1 = 1/p_{\mathcal{L}} + 2/q_{\mathcal{L}}$, and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N^{|k|} = 0$.

This is the case if $k = -1$ or $k = 1 - 2\gamma_j \in [-1,0)$ with $\gamma_j \in (1/2,1]$ or $k = -2\gamma_j$ with $\gamma_j \in (0,1/2)$. By sub-additivity of $x \mapsto x^{-k}$ and Jensen's inequality, we have:

$$\mathbb{E}\left[|X_t^k - X_{\lfloor t\rfloor_{\Delta_N}}^k|\mathbb{1}_{I_j}(X_t)\right] \leq \mathbb{E}\left[|X_t^{-1} - X_{\lfloor t\rfloor_{\Delta_N}}^{-1}|^{|k|}\mathbb{1}_{I_j}(X_t)\right] \leq \mathbb{E}\left[|X_t^{-1} - X_{\lfloor t\rfloor_{\Delta_N}}^{-1}|\mathbb{1}_{I_j}(X_t)\right]^{|k|}.$$

Hölder's inequality with $1 = 1/p_{\mathcal{L}} + 2/q_{\mathcal{L}}$ gives

$$\mathbb{E}\left[|X_t^{-1} - X_{\lfloor t\rfloor_{\Delta_N}}^{-1}|\mathbb{1}_{I_j}(X_t)\right] \leq \mathbb{E}\left[|X_t - X_{\lfloor t\rfloor_{\Delta_N}}|^{p_{\mathcal{L}}}\right]^{1/p_{\mathcal{L}}} \mathbb{E}\left[|X_t^{-1}|^{q_{\mathcal{L}}}\right]^{1/q_{\mathcal{L}}} \mathbb{E}\left[|X_{\lfloor t\rfloor_{\Delta_N}}^{-1}|^{q_{\mathcal{L}}}\right]^{1/q_{\mathcal{L}}}.$$

The case $k \in [-2,-1)$. μ admits finite moments of order $-q'|k|$ and $\max(1, p'|k|/2)$ with $p', q' \geq 1$ such that $1 = 1/p' + 1/q' + 1/(2q')$ and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N^{|k|/2} = 0$.

This is the case when $k = -2\gamma_j$ with $\gamma_j \in (1/2,1]$. Note that what follows holds for all $k \in [-2,0)$ as well. Observe that

$$\mathbb{E}\left[|X_t^k - X_{\lfloor t\rfloor_{\Delta_N}}^k|\right] \leq \mathbb{E}\left[|X_t^{k/2} - X_{\lfloor t\rfloor_{\Delta_N}}^{k/2}||X_t^{k/2} + X_{\lfloor t\rfloor_{\Delta_N}}^{k/2}|\right]$$

This rewrites $\mathbb{E}\left[|X_t^{|k|/2} - X_{\lfloor t\rfloor_{\Delta_N}}^{|k|/2}||X_t^k X_{\lfloor t\rfloor_{\Delta_N}}^{k/2} + X_t^{k/2} X_{\lfloor t\rfloor_{\Delta_N}}^k|\right]$. The fact that $x \mapsto x^{|k|/2}$ is sub-additive and Hölder's inequality (with $p', q' \geq 1$ such that $1 = 1/p' + 1/q' + 1/(2q')$, e.g. $p' = 4$, $q' = 2$) yield

$$\begin{aligned}\mathbb{E}\left[|X_t^k - X_{\lfloor t\rfloor_{\Delta_N}}^k|\right] \leq \mathbb{E}\left[|X_t - X_{\lfloor t\rfloor_{\Delta_N}}|^{p'|k|/2}\right]^{1/p'} \Bigg(&\mathbb{E}\left[X_t^{q'k}\right]^{1/q'} \mathbb{E}\left[X_{\lfloor t\rfloor_{\Delta_N}}^{2q'k/2}\right]^{1/(2q')} \\ &+ \mathbb{E}\left[X_t^{2q'k/2}\right]^{1/(2q')} \mathbb{E}\left[X_{\lfloor t\rfloor_{\Delta_N}}^{q'k}|\right]^{1/q'}\Bigg).\end{aligned}$$

Remark 5.7.3. *If $k \in [-1,0)$, considering the same bounds as in the last paragraph, we get a more restrictive condition on Δ_N than with the previously considered bounds in the case $k \in [-1,0)$ but a possibly less restrictive condition on moments. For instance, taking $p' = 4$ and $q' = 2$, if $k = -1$ then finite moments of order -2 and 2 are finite for μ suffice, while $q_{\mathcal{L}} \geq 2$.*

Remark 5.7.4. *Let $k \in [-2, 0]$ and let $j \neq 0$. Note that $X_t^k \mathbb{1}_{I_j}(X) \leq r_1^k \mathbb{1}_{I_j}(X)$. Hence, the integrand of (5.7.7) in k can be bounded from above by*

$$r_1^{2k}\mathbb{E}\left[|X_t^{|k|} - X_{\lfloor t\rfloor_{\Delta_N}}^{|k|}|\mathbb{1}_{I_j}(X_t)\mathbb{1}_{I_j}(X_{\lfloor t\rfloor_{\Delta_N}})\right] + \mathbb{E}\left[|X_t^k - X_{\lfloor t\rfloor_{\Delta_N}}^k|\mathbb{1}_{I_j}(X_t)\mathbb{1}_{I_j^c}(X_{\lfloor t\rfloor_{\Delta_N}})\right].$$

The first term is the integrand of (5.7.7) in $|k| \in [0, 2]$ and the second is bounded by

$$r_1^k\mathbb{E}\left[\mathbb{1}_{I_j}(X_t)\mathbb{1}_{I_j^c}(X_{\lfloor t\rfloor_{\Delta_N}})\right] + \mathbb{E}\left[X_{\lfloor t\rfloor_{\Delta_N}}^k \mathbb{1}_{I_j}(X_t)\mathbb{1}_{I_j^c}(X_{\lfloor t\rfloor_{\Delta_N}})\right]$$

which is sum of terms of the kind (5.7.3) with $q = 0$ and $q = k$. If, e.g., $k \in [0, 1]$ then the conditions are μ admits finite first and $(-|k|)$-th moment and $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N^{|k|} = 0$.

Dealing with $\mathbb{E}\left[|M_{T_N}^{j,m} - M_{T_N,N}^{j,m}|\right]$

As in the previous section, we exploit (5.7.4)-(5.7.5) to obtain the assumptions on the moments of μ and on Δ_N under which $\mathbb{E}\left[|M_{T_N}^{j,m} - M_{T_N,N}^{j,m}|\right]$ is $o(T_N^{1/\lambda})$ for $m \in \{0, 1\}$, $j \in \{0, \dots, d\}$.

As for $Q^{j,k}$, we can rewrite $M_{T_N,N}^{j,m} - M_{T_N}^{j,m}$ as follows:

$$\begin{aligned} M_{T_N,N}^{j,m} - M_{T_N}^{j,m} = &\int_0^{T_N} \left(X_t^m \mathbb{1}_{I_j}(X_t) - X_{\lfloor t\rfloor_{\Delta_N}}^m \mathbb{1}_{I_j}(X_{\lfloor t\rfloor_{\Delta_N}})\right)(a(X_t) - b(X_t)X_t)\,\mathrm{d}t \\ &+ \int_0^{T_N} \left(X_t^m \mathbb{1}_{I_j}(X_t) - X_{\lfloor t\rfloor_{\Delta_N}}^m \mathbb{1}_{I_j}(X_{\lfloor t\rfloor_{\Delta_N}})\right)\sigma(X_t)(X_t)^{\gamma(X_t)}\,\mathrm{d}B_t. \end{aligned}$$

Using Triangular inequality, Hölder's inequality, and Itô-isometry, we obtain:

$$\begin{aligned} &\mathbb{E}\left[|M_{T_N}^{j,m} - M_{T_N,N}^{j,m}|\right] \leq \int_0^{T_N} \mathbb{E}\left[|X_t^m - X_{\lfloor t\rfloor_{\Delta_N}}^m|(|a_j| + b_j|X_{\lfloor t\rfloor_{\Delta_N}}| + b_j|X_t - X_{\lfloor t\rfloor_{\Delta_N}}|)\right]\,\mathrm{d}t \\ &+ \int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t\rfloor_{\Delta_N}}|^m \mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j}\left(\max_{i=0,\dots,d}|a_i| + \max_{i=0,\dots,d}|b_i|\left(|X_{\lfloor t\rfloor_{\Delta_N}}| + |X_t - X_{\lfloor t\rfloor_{\Delta_N}}|\right)\right)\right]\,\mathrm{d}t \\ &+ \int_0^{T_N} \mathbb{E}\left[|X_{\lfloor t\rfloor_{\Delta_N}}|^m \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j}\left(\max_{i=0,\dots,d}|a_i| + \max_{i=0,\dots,d}|b_i|\left(|X_{\lfloor t\rfloor_{\Delta_N}}| + |X_t - X_{\lfloor t\rfloor_{\Delta_N}}|\right)\right)\right]\,\mathrm{d}t \\ &+ \sqrt{2}\max_{i=0,\dots,d}(\sigma_i)\left(\int_0^{T_N} \mathbb{E}\left[(X_t^m - X_{\lfloor t\rfloor_{\Delta_N}}^m)^2(X_t)^{2\gamma(X_t)} + X_{\lfloor t\rfloor_{\Delta_N}}^{2m}(X_t)^{2\gamma(X_t)}\mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j}\right.\right. \\ &\left.\left. + X_{\lfloor t\rfloor_{\Delta_N}}^{2m}(X_t)^{2\gamma(X_t)}\mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j}\right]\,\mathrm{d}t\right)^{1/2}. \end{aligned} \tag{5.7.8}$$

The case $m = 0$. μ has finite $\max(1, 2\gamma_d)$-th moment and $\lim_{N\to 0} T_N\Delta_N = 0$.

Inequality (5.7.8) involves terms of the kind (5.7.2) with $p = 1$ and (5.7.3) with $q \in \{0, 1\}$ but also

$$\left(\int_0^{T_N} \mathbb{E}\left[(X_t)^{2\gamma(X_t)}(\mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j} + \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j})\right]\,\mathrm{d}t\right)^{1/2}.$$

Since I_j is not bounded only if $j = d$ or $j = 0$ with $\gamma_0 = 0$, the above quantity is bounded from above for all $j \neq d$ and for $j = d$ if $\gamma_d = 0$ by terms (5.7.3) with $q = 0$. If $j = d$ and $\gamma_d \neq 0$, its square is bounded by a sum of terms as (5.7.2) with $p = 2\gamma_d$ and (5.7.3) with $q = 2\gamma_d$:

$$\mathbb{E}\left[(X_t)^{2\gamma_d}\mathbb{1}_{X_t\in I_d, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_d}\right] \leq C\mathbb{E}\left[|X_t - X_{\lfloor t\rfloor_{\Delta_N}}|^{2\gamma_d} + X^{2\gamma_d}_{\lfloor t\rfloor_{\Delta_N}}\mathbb{1}_{X_t\in I_d, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_d}\right]$$

for some positive constant C. The latter inequality is derived from sub-additivity (if $\gamma_d \leq 1/2$) or Jensen's inequality (if $\gamma_d > 1/2$).

Inequalities (5.7.4)-(5.7.5) ensure that

$$\mathbb{E}\left[|M^{j,0}_{T_N} - M^{j,0}_{T_N,N}|\right] \leq C_1 T_N\sqrt{\Delta_N} + C_2\sqrt{T_N\Delta_N^{\max(1,2\gamma_d)/2}} \leq (C_1 + C_2)T_N\sqrt{\Delta_N}$$

for some positive constants C_1, C_2.

The case $m = 1$. μ admits finite $2(1 + \gamma_d)$-th moment and $\lim_{N\to 0} T_N\Delta_N = 0$.

Inequality (5.7.8) involves terms of the kind (5.7.2) with $p \in \{1, 2\}$ and (5.7.3) with $q \in \{1, 2\}$ but also the square root of the following term:

$$\int_0^{T_N}\mathbb{E}\left[(X_t - X_{\lfloor t\rfloor_{\Delta_N}})^2X_t^{2\gamma(X_t)} + X^2_{\lfloor t\rfloor_{\Delta_N}}X_t^{2\gamma(X_t)}(\mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j} + \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j})\right]\,\mathrm{d}t.$$

Similarly to the case $m = 0$, by the fact that $X_s^{2\gamma_j}\mathbb{1}_{I_j}(X_s)$ is not bounded from above (up to multiplicative constant) by $\mathbb{1}_{I_j}(X_s)$ only if $j = d$, the fact that the same holds for $X_s^2\mathbb{1}_{I_j}(X_s)$ unless $j = d$ or $j = 0$ and $\gamma_0 = 0$, and by sub-additivity of the square-root, we can reduce to

$$\begin{aligned}
&\left(\int_0^{T_N}\mathbb{E}\left[(X_t - X_{\lfloor t\rfloor_{\Delta_N}})^2 + \mathbb{1}_{X_t\in I_j, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_j} + \mathbb{1}_{X_t\notin I_j, X_{\lfloor t\rfloor_{\Delta_N}}\in I_j}\right]\,\mathrm{d}t\right)^{1/2}\\
&+\left(\int_0^{T_N}\mathbb{E}\left[(X_t - X_{\lfloor t\rfloor_{\Delta_N}})^2X_t^{2\gamma_d}\mathbb{1}_{I_d}(X_t) + X^2_{\lfloor t\rfloor_{\Delta_N}}X_t^{2\gamma_d}\mathbb{1}_{X_t\in I_d, X_{\lfloor t\rfloor_{\Delta_N}}\in I_0}\right]\right)^{1/2}\\
&+\left(\int_0^{T_N}\mathbb{E}\left[X^2_{\lfloor t\rfloor_{\Delta_N}}\mathbb{1}_{X_t\notin I_d, X_{\lfloor t\rfloor_{\Delta_N}}\in I_d} + X_t^{2\gamma_d}\mathbb{1}_{X_t\in I_d, X_{\lfloor t\rfloor_{\Delta_N}}\notin I_d\cup I_0}\right]\right)^{1/2}.
\end{aligned}$$

The first line involves terms like (5.7.2) with $p = 2$ and (5.7.3) with $q = 0$. The last line shows a term (5.7.3) with $q = 2$ and a term appearing for the case $m = 0$. The second line, by subadditivity and Hölder's inequality ($(1 + \gamma_d)^{-1} + (1 + \gamma_d)^{-1}\gamma_d = 1$), is bounded from above by

$$\begin{aligned}
&\left(\int_0^{T_N}\mathbb{E}\left[(X_t - X_{\lfloor t\rfloor_{\Delta_N}})^{2(1+\gamma_d)}\right]^{1/(1+\gamma_d)}\mathbb{E}\left[X_t^{2(1+\gamma_d)}\right]^{\gamma_d/(1+\gamma_d)}\,\mathrm{d}t\right)^{1/2}\\
&+\left(\int_0^{T_N}\mathbb{E}\left[X^{2(1+\gamma_d)}_{\lfloor t\rfloor_{\Delta_N}}\mathbb{1}_{X_t\in I_d, X_{\lfloor t\rfloor_{\Delta_N}}\in I_0}\right]^{1/(1+\gamma_d)}\mathbb{E}\left[X_t^{2(1+\gamma_d)}\right]^{\gamma_d/(1+\gamma_d)}\,\mathrm{d}t\right)^{1/2}
\end{aligned}$$

where we recognise (5.7.2) and (5.7.3) with $p = q = 2(1 + \gamma_d)$.

By (5.7.4) and (5.7.5) we deduce $\mathbb{E}\left[|M_{T_N}^{j,1} - M_{T_{N,N}}^{j,1}|\right]$ is bounded from above by

$$C_1(T_N\sqrt{\Delta_N} + T_N\Delta_N + \sqrt{T_N\Delta_N} + \sqrt{T_N\Delta_N^{1/2(1+\gamma_j)}}) \leq 4C_1T_N\sqrt{\Delta_N}$$

for some positive constant C_1.

End of the proof of Lemma 5.5.2

In this section, we complete the proof of Lemma 5.5.2 by summarising the assumptions of the previous sections for each item.

Proof of Item (a) (QMLE $\theta_{T_N,N}^{(q\text{-}\mathcal{L})}$). The QMLE (5.4.6) involves the statistics $Q^{j,k}$ and $M^{j,m}$, where $m \in \{0,1\}$, $k \in \{0,1,2\}$, and $j \in \{0,\dots,d\}$. The most restrictive assumptions on the moments of μ is obtained for $m = 1$: μ has finite moment of order $2(1+\gamma_d)$, that is $\mathcal{H}_{q\text{-}\mathcal{L}}$. Instead, the unique condition on Δ_N is $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N = 0$. Hence, $g_N^{(q\text{-}\mathcal{L})} = \Delta_N$.

Remark 5.7.5. *If $\sigma_\star$ is known, one could replace $M_{T_N,N}^{j,1}$ by $\mathcal{M}_{T_N,N}^{j,1}$ in* (5.4.2) *obtaining the analogous of* (5.4.6)*. However, in order to exploit the asymptotic properties of the estimator from continuous time observations, $\mathbf{H}_{q\text{-}\mathcal{L}}$ has to be included in $\mathcal{H}_{\mathcal{L}}$, hence it cannot be relaxed. Moreover, the most restrictive condition on Δ_N would come from $k = 2\gamma_j$ when $\gamma_j \in (0,1/2)\cup(1/2,1)$, for all $j \in \{0,\dots,d\}$, leading to the condition $\lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\sigma)} = 0$, which is worse than $\lim_{N\to\infty} T_N^{\lambda-1}\Delta_N = 0$.*

Proof of Item (b) (Volatility estimation). Only the statistics $Q^{j,k}$, $M^{j,m}$, and $M^{i,0}$ with $m \in \{0,1\}$, $k = 2\gamma_j$, and $i,j \in \{0,\dots,d\}$ appear in the volatility estimator (5.4.7). We observe that the most restrictive assumptions on the moments of μ is obtained for $m = 1$: μ has finite moment of order $2(1+\gamma_d)$, that is $\mathcal{H}_{q\text{-}\mathcal{L}}$. Instead the unique condition on Δ_N more restrictive than $\lim_{N\to\infty} T_N\Delta_N = 0$ comes from $k = 2\gamma_j$ when $\gamma_j \in (0,1/2)\cup(1/2,1)$, for all $j \in \{0,\dots,d\}$, leading to the condition $\lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\sigma)} = 0$.

Proof of Item (c) (MLE $\theta_{T_N,N}^{(\mathcal{L})}$ with $\sigma_\star$ unknown). First, assume that $\sigma_\star$ is known. To deal with the case $\sigma_\star$ is unknown, one shall then take the most restrictive conditions among the ones in Item (b) and the ones obtained in the case $\sigma_\star$ known.

The MLE (5.4.5) involves the statistics $Q^{j,k}$, $M^{i,0}$, where $k \in \{-2\gamma_j, 1-2\gamma_j, 2-2\gamma_j\}\cup\{-1,0\}$, and $i,j \in \{0,\dots,d\}$. This case is more subtle than the others. The more restrictive conditions come from the different values of k. Let us consider separately the cases $\gamma_j \in (0,1/2]$ and $\gamma_j \in (1/2,1]$. Let $p_{\mathcal{L}}, q_{\mathcal{L}}, p', q'$ given in $\mathcal{H}_{\mathcal{L}}$.

If $\gamma_j \in (0,1/2]$, the unique negative moment condition comes from $k = -1$ (and $k = -2\gamma_j$), while for the positive moments the most restrictive conditions come from $k = -1$ and $k = 2(1-\gamma_j)$: finite moments of order $-q_{\mathcal{L}}$ and $\max(p_{\mathcal{L}}, 2(1-\gamma_j))$. The condition on Δ_N follows from the fact that $\lim_{N\to\infty} T_N^{\lambda-1}(\Delta_N + \Delta_N^{2\gamma_j} + \Delta_N^{1-2\gamma_j} + \Delta_N^{1-\gamma_j}) \leq \lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\mathcal{L})} = 0$.

If $\gamma_j \in (1/2, 1]$, the most restrictive moment conditions comes from $k = -1$ and $k = -2\gamma_j$: $\max(p_{\mathcal{L}}, p'\gamma_j)$ and $-\max(q_{\mathcal{L}}, 2q'\gamma_j)$. The most restrictive condition on Δ_N is given by $\lim_{N\to\infty} T_N^{\lambda-1}(\Delta_N + \Delta_N^{\gamma_j} + \Delta_N^{2\gamma_j-1} + \Delta_N^{2(1-\gamma_j)}) \leq \lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\mathcal{L})} = 0$.

In conclusion, since $j \in \{0, \dots, 1\}$, we have $\lim_{N\to\infty} T_N^{\lambda-1} g_N^{(\mathcal{L})} = 0$ and μ has finite moments of orders

$$-\max(q_{\mathcal{L}}, 2q'\gamma^{\max}\mathbb{1}_{(1/2,1]}(\gamma^{\max})) \quad \text{and} \quad \max(p_{\mathcal{L}}, p'\gamma^{\max}\mathbb{1}_{(1/2,1]}(\gamma^{\max}), 2(1-\gamma^{\min})\mathbb{1}_{(0,1/2]}(\gamma^{\min})). \tag{5.7.9}$$

Remark 5.7.6. *If $\gamma^{\max} \leq 1/2$, taking $p' = 4$ and $q' = 2$, the orders are $-q_{\mathcal{L}}$ and $p_{\mathcal{L}}$.*

We consider separately the case $\gamma_0 = 0$ and $\gamma^{\max} \leq 1/2$. If $\gamma_0 = 0$ then we deal with positive moments up to the one of order 2, and with negative moments when $k = -2\gamma_j$. By Remark 5.7.4, the most restrictive moment conditions are finite 2-nd and $-2\gamma_j$-th moment and the following one on Δ_N: $\lim_{N\to 0} T_N^{\lambda-1} g_N^{(\mathcal{L})} = 0$.

5.7.5 Proof of Proposition 5.5.4

Given s and t such that $0 \leq s < t$, we show that for every $m \geq 1$ such that μ admits finite m-th moment, there exists a constant $C \in (0, \infty)$ depending only on m and the parameters of the process such that $\mathbb{E}\left[|X_t - X_s|^m\right] \leq C(t-s)^{m/2}$.

By the triangular inequality,

$$\begin{aligned}|X_t - X_s| &\leq \int_s^t |a(X_u) - b(X_u)X_u| \, \mathrm{d}u + \left| \int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)} \, \mathrm{d}B_u \right| \\ &\leq (t-s) \max_{i=0,\dots,d} |a_i| + \max_{i=0,\dots,d} |b_i| \int_s^t |X_u| \, \mathrm{d}u + \left| \int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)} \, \mathrm{d}B_u \right|.\end{aligned}$$

Then, Jensen's inequality ensures that for $m \geq 1$ it holds that:

$$\begin{aligned}\mathbb{E}\left[|X_t - X_s|^m\right] \leq{}& 2^{2m-2} \max_{i=0,\dots,d} |a_i|(t-s)^m + 2^{2m-2} \max_{i=0,\dots,d} |b_i|(t-s)^{m-1} \int_s^t \mathbb{E}\left[|X_u|^m\right] \mathrm{d}u \\ &+ 2^{m-1}\mathbb{E}\left[\left| \int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)} \, \mathrm{d}B_u \right|^m \right].\end{aligned}$$

Since X_0 is distributed as the stationary distribution μ, which admits finite m-th moment, then $\sup_{u\in[s,t]} \mathbb{E}\left[|X_u|^m \mathbb{1}_{I_d}(X_u)\right] < \infty$.

Burkholder-Davis-Gundy inequality implies that

$$\mathbb{E}\left[\left| \int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)} \, \mathrm{d}B_u \right|^m \right] \leq \mathbb{E}\left[\left(\int_s^t \max_{i=0,\dots,d} |\sigma_i|(X_u)^{2\gamma(X_u)} \, \mathrm{d}u \right)^{m/2} \right].$$

Then, we distinguish the case $m \geq 2$ and $m \in [1, 2)$. In both cases we apply Hölder's inequality but in a different way. If $m \geq 2$, we obtain:

$$\begin{aligned}\mathbb{E}\left[\left| \int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)} \, \mathrm{d}B_u \right|^m \right] &\leq (t-s)^{m/2-1} (\max_{i\in\{0,\dots,d\}} |\sigma_i|)^{m/2} \int_s^t \mathbb{E}\left[|X_u|^{m\gamma_d}\right] \mathrm{d}u \\ &\leq C(t-s)^{m/2},\end{aligned}$$

since, $\sup_{u\in[s,t]} \mathbb{E}\left[|X_u|^{m\gamma_d}\mathbb{1}_{I_d}(X_u)\right] < \infty$. If $m \in [1,2)$, we reduce to the previous case ($m \geq 2$):

$$\mathbb{E}\left[\left|\int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)}\,\mathrm{d}B_u\right|^m\right] \leq \left(\mathbb{E}\left[\left|\int_s^t \sigma(X_u)(X_u)^{\gamma(X_u)}\,\mathrm{d}B_u\right|^{2m}\right]\right)^{1/2} \leq C(t-s)^{m/2}.$$

The proof is thus completed.

5.7.6 Proof of the key result: Proposition 5.5.5

For all $j \in \{0,\dots,d\}$, let $\xi^{(j)}$ denote a standard CKLS process with parameters $(a_j, b_j, \sigma_j, \gamma_j)$ starting at X_s. Let $s,t \in [0,\infty)$ be fixed such that $0 \leq s < t$. Note that:

$$\begin{aligned}&\mathbb{E}\left[|X_s|^m \mathbb{P}_{X_s}\left(\tau_{I_j}^{\xi^{(j)}} < t-s\right)\right]\\ &\leq \mathbb{E}\left[|X_s|^m\left(\mathbb{P}_{X_s}\left(\tau_{r_j,\searrow}^{\xi^{(j)}} < t-s\right) + \mathbb{P}_{X_s}\left(\tau_{r_{j+1},\nearrow}^{\xi^{(j)}} < t-s\right)\right)\mathbb{1}_{I_j}(X_s)\right],\end{aligned}$$

where $\tau_{r_j,\searrow}^{\xi^{(j)}}$ is the first hitting time from above of the level r_j, $\tau_{r_{j+1},\nearrow}^{\xi^{(j)}}$ is the first hitting time from below of the level r_{j+1} of the process $\xi^{(j)}$.

Without loss of generality, we reduce to show that:

$$\mathbb{E}\left[|X_s|^m \mathbb{P}_{X_s}\left(\tau_{r_d,\searrow}^{\xi^{(d)}} < t-s\right)\mathbb{1}_{I_d}(X_s)\right] \leq C_2(t-s)^{1/2}, \tag{5.7.10}$$

and

$$\mathbb{E}\left[|X_s|^m \mathbb{P}_{X_s}\left(\tau_{r_1,\nearrow}^{\xi^{(0)}} < t-s\right)\mathbb{1}_{I_0}(X_s)\right] \leq C_1(t-s)^{1/2}, \tag{5.7.11}$$

where C_1 and C_2 are strictly positive constant.

Indeed, in the other cases, X_s belongs in I_j for $j \notin \{0,d\}$, which is compact, and the desired inequality can be deduced using a similar reasoning.

Bounds on the first hitting time from above of the level r_d: (5.7.10)

We focus on the case $\gamma_d \in [0,1)$. The case $\gamma_d = 1$ can be proven using a similar reasoning, the proof is thus omitted. Let us recall that the parameters $(a_d, b_d, \sigma_d, \gamma_d)$ satisfy the ergodicity conditions in Table 5.4 in Section 5.7.1, in particular $b_d \geq 0$.
The main idea of this proof is to bound the first hitting time by the hitting times of some drifted Brownian motions. To do so, we apply the Lamperti transform and we bound the process, over a well chosen time interval.

We define the process $(Y_u)_{u\geq 0}$ as follows. For all $u \geq 0$ let $Y_u = \psi(\xi_u^{(d)})$ where $\psi(x) = \int_0^x \frac{1}{\sigma_d y^{\gamma_d}}\,\mathrm{d}y = \frac{x^{1-\gamma_d}}{\sigma_d(1-\gamma_d)}$ (Lamperti transform). We denote ψ^{-1} as the inverse function of ψ, then Y is solution to the following SDE:

$$\mathrm{d}Y_u = \frac{a_d}{\sigma_d}(\psi^{-1}(Y_u))^{-\gamma_d} - b_d(1-\gamma_d)Y_u - \frac{1}{2}\sigma_d\gamma_d(\psi^{-1}(Y_u))^{\gamma_d-1}\,\mathrm{d}u + \mathrm{d}B_u.$$

Let $\varepsilon > 0$ be fixed and let $\tau^Y_{[\psi(r_d),\psi(X_s)+\varepsilon]}$ denote the first hitting time of the boundary of $[\psi(r_d), \psi(X_s) + \varepsilon]$ of the process Y. Then, we have

$$\mathbb{P}_{X_s}\left(\tau^{\xi^{(d)}}_{r_d,\searrow} < t - s\right) = \mathbb{P}_{\psi(X_s)}\left(\tau^Y_{\psi(r_d),\searrow} < t - s\right) \leq \mathbb{P}_{\psi(X_s)}\left(\tau^Y_{[\psi(r_d),\psi(X_s)+\varepsilon]} < t - s\right).$$

By applying the Comparison Theorem (e.g. [48, p352]) until time $\tau^Y_{[\psi(r_d),\psi(X_s)+\varepsilon]}$, $Y - \psi(X_s)$ is bounded from above by a drifted BM B^{ν_+} starting from 0 and from below by a drifted BM B^{ν_-} starting from 0 with parameters:

$$\begin{cases} \nu_+ = \frac{|a_d|}{\sigma_d} r_d^{-\gamma_d} \\ \nu_- = -\frac{|a_d|}{\sigma_d} r_d^{-\gamma_d} - b_d(1-\gamma_d)(\psi(X_s)+\varepsilon) - \frac{1}{2}\sigma_d\gamma_d r_d^{\gamma_d - 1}. \end{cases}$$

Hence, the following inequality holds:

$$\mathbb{P}_{X_s}\left(\tau^{\xi^{(d)}}_{r_d,\searrow} < t - s\right) \leq \mathbb{P}_0\left(\tau^{B^{\nu_-}}_{\psi(r_d)-\psi(X_s),\searrow} < t - s\right) + \mathbb{P}_0\left(\tau^{B^{\nu_+}}_{\varepsilon,\nearrow} < t - s\right),$$

and classical results on the first hitting of a drifted Brownian motion (see [19]) yield

$$\mathbb{P}_0\left(\tau^{B^{\nu_+}}_{\varepsilon,\nearrow} < t - s\right) \leq K_1 e^{-\frac{\varepsilon^2}{2(t-s)}},$$

and,

$$\mathbb{P}_0\left(\tau^{B^{\nu_-}}_{\psi(r_d)-\psi(X_s),\searrow} < t - s\right) \leq K_2 e^{-\frac{|\psi(r_d)-\psi(X_s)|^2}{2(t-s)} + b_d(1-\gamma_d)\psi(X_s)^2} f(X_s),$$

where K_1 and K_2 are two strictly positive constants, f is an explicit function which depends only on X_s and such that $\lim_{x\to+\infty} f(x)e^{-\psi(x)^2} = 0$.

Let us note that

$$\mu(x)\mathbb{1}_{x\geq r_d} = K_3 \frac{2}{\sigma_d^2 x^{1-2\frac{a_d}{\sigma_d^2}}} \exp\left(-b_d(1-\gamma_d)\psi(x)^2\right)\mathbb{1}_{x\geq r_d},$$

with K_3 a strictly positive constant. Since μ admits finite m-th moment by assumption, inequality (5.7.10) holds. We avoid details here, but the interested reader could appreciate the following remarks. When $t - s$ is close to 0, $e^{-K/(t-s)}$ decreases faster than any polynomial of $(t - s)$. So, by splitting the integral to distinguish between when $\psi(x) - \psi(r_d)$ is small (e.g. $O(\sqrt{t-s})$) and when it isn't, we can easily compute the bounds. To avoid repetitions, we do not mention anymore this remark.

Bounds on the first hitting time from below of the level r_1: (5.7.11)

The parameters $(a_0, b_0, \sigma_0, \gamma_0)$ satisfy the conditions ensuring ergodicity in Table 5.4. We remark that, on a suitable time interval, the process $(\xi^{(0)})^{2(1-\gamma_0)}$ can be bounded from above by the norm of a multi-dimensional Brownian motion. This leads to obtaining (5.7.11) under the assumption that $\gamma_0 \in \{0, 1/2\}$. Instead, if $\gamma_0 \in (1/2, 1]$, this bound is not enough. Thus, we additionally bound from below the Lamperti transform of the process $\xi^{(0)}$ by a drifted Brownian motion.

The case $\gamma_0 = 1/2$. We remind the following results, for $i \in \{1, \dots, n\}$, we denote $\tilde{B} := (\tilde{B}^i)_{i \leq n}$ a n-dimensional Brownian motion. Itô formula and Levy characterization imply that

$$\forall u \geq 0, \quad \left\| \frac{\sigma_0}{2} \tilde{B}_u - \sqrt{\frac{X_s}{n}} \right\|_2^2 = Y_{u,n} = X_s + n\frac{\sigma_0^2}{4} u + \int_0^u \sigma_0 \sqrt{Y_{v,n}} \, \mathrm{d}W_v,$$

where W is a Brownian motion. Then $Y_{\cdot,n}$ is a CIR process whose coefficients satisfy the conditions for ergodicity in Table 5.4. Moreover we take n such that $0 < a_0 < n\sigma_0^2/4$. So, by the Comparison Theorem, it holds a.s. that $\xi_u^{(0)} \leq Y_{u,n}$ for all $0 \leq u \leq \tau_{r_1,\nearrow}^{\xi^{(0)}}$ and then

$$\mathbb{P}_{X_s}\left(\tau_{r_1,\nearrow}^{\xi^{(0)}} \leq t - s\right) \leq \mathbb{P}_{X_s}\left(\tau_{r_1,\nearrow}^{Y_{\cdot,n}} \leq t - s\right).$$

Moreover, we have

$$\{\tau_{r_1,\nearrow}^{Y_{\cdot,n}} \leq t - s\} \subseteq \bigcup_{i=1}^{n} \left\{ \forall u \leq t - s, -\sqrt{\frac{r_1}{n}} < \left(\frac{\sigma_0}{2} \tilde{B}_u^i - \sqrt{\frac{X_s}{n}} \right) < \sqrt{\frac{r_1}{n}} \right\}^c.$$

Then, by the Comparison Theorem and the symmetry of BM, we obtain:

$$\mathbb{P}_{X_s}\left(\tau_{r_1,\nearrow}^{\xi^{(0)}} \leq t - s\right) \leq 2n\mathbb{P}_0\left(\tau_{2\frac{\sqrt{r_1}-\sqrt{X_s}}{\sigma_0\sqrt{n}},\nearrow}^{\tilde{B}^1} \leq t - s\right) \leq K_4 e^{-\frac{2(\sqrt{r_1}-\sqrt{X_s})^2}{\sigma_0^2 n(t-s)}},$$

with K_4 a strictly positive constant. This, and the stationary distribution μ (5.7.1), if μ admits finite m-th moment, yield inequality (5.7.11).

The case $\gamma_0 = 0$. By applying Itô formula on $Y := (\xi^{(0)})^2$, for all $u \geq 0$, we have:

$$\mathrm{d}Y_u = 2\left(\mathrm{sgn}(\xi^{(0)}) a_0 \sqrt{Y_u} - b_0 Y_u + \frac{\sigma_0^2}{2}\right) \mathrm{d}u + 2\sigma_0 \sqrt{Y_u} \, \mathrm{d}\tilde{B}_u$$

with $Y_0 = X_s^2$ and $\tilde{B}_u$ another Brownian motion. Since $2\,\mathrm{sgn}(\xi^{(0)}) a_0 \sqrt{Y_u} - 2b_0 Y_u + \sigma_0^2 \leq 2|a_0| r_1 + \sigma_0^2 \leq n\sigma_0^2$ for some $n \in \mathbb{N}$, the Comparison Theorem ensures that for all $u \leq \tau_{r_1^2,\nearrow}^{Y}$ the process Y is bounded from above by the norm of a n-dimensional Brownian motion. Hence, similarly to the case $\gamma_0 = \frac{1}{2}$, we have:

$$\mathbb{P}_{X_s}\left(\tau_{r_1,\nearrow}^{\xi^{(0)}} \leq t - s\right) \leq 2n\mathbb{P}_0\left(\tau_{\frac{r_1 - X_s}{\sigma_0\sqrt{n}},\nearrow}^{\tilde{B}^1} \leq t - s\right) \leq K_5 e^{-\frac{(r_1 - X_s)^2}{2\sigma_0^2 n(t-s)}},$$

with K_5 a strictly positive constant. We conclude analogously to the case $\gamma_0 = 1/2$.

The case $\gamma_0 \in (1/2, 1)$. By Lamperti transform, $Y_u := \mathrm{d}\psi(\xi_u^{(0)})$ with $\psi(x) = \frac{x^{1-\gamma_0}}{\sigma_0(1-\gamma_0)}$ (in particular $Y_0 = \frac{X_s^{1-\gamma_0}}{\sigma_0(1-\gamma_0)}$). So,

$$\mathrm{d}Y_u = \left[\frac{a_0}{\sigma_0}\left((1-\gamma_0)Y_u\right)^{1-\frac{1}{(1-\gamma_0)}} - b_0(1-\gamma_0)Y_u - \frac{\gamma_0}{2(1-\gamma_0)} Y_u^{-1}\right] \mathrm{d}u + \mathrm{d}B_u.$$

By the Comparison Theorem, it holds a.s. for all $u \in [0, \tau_{\left[\frac{\psi(X_s)}{2}, \psi(r_1)\right]})$ that

$$B_u^{\nu_-} \leq Y_u \quad \text{and} \quad Y_u^2 \leq \left|\left| \frac{1}{2}\tilde{B}_u - \frac{\psi(X_s)}{\sqrt{n}} \right|\right|_2^2,$$

where B^{ν_-} is drifted Brownian motion and $\tilde{B} := (\tilde{B}^i)_{i \leq n}$ a n-dimensional Brownian motion. Here the drift parameter ν_- and the dimension n are given by

$$\begin{cases} \nu_- = -\frac{|b_0| r_1^{1-\gamma_0}}{\sigma_0} - \frac{\gamma_0 \sigma_0}{X_s^{1-\gamma_0}}, \\ n(X_s) = \left\lceil \frac{a_0}{\sigma_0^2(1-\gamma_0)} X_s^{1-2\gamma_0} + |b_0|(1-\gamma_0) \frac{r_1^{2(1-\gamma_0)}}{\sigma_0^2(1-\gamma_0)^2} \right\rceil. \end{cases}$$

Finally, we obtain the following inequality:

$$\begin{aligned} \mathbb{P}_{X_s}\left(\tau^{\xi^{(0)}}_{r_1,\nearrow} < t-s\right) &\leq \mathbb{P}_0\left(\tau^{B^{\nu_-}}_{-\frac{\psi(X_s)}{2},\searrow} < t-s\right) + 2n(X_s)\mathbb{P}_0\left(\tau^{\tilde{B}^1}_{2\frac{\psi(r_1)^2-\psi(X_s)^2}{\sqrt{n(X_s)}},\nearrow} < t-s\right) \\ &\leq K_6\left(e^{-\frac{|\psi(r_1)-\psi(X_s)|^2}{2(t-s)}} + n(X_s)e^{-\frac{2\left(\psi(r_1)^2-\psi(X_s)^2\right)^2}{n(X_s)(t-s)}}\right), \end{aligned}$$

with K_6 a strictly positive constant.

Similarly to the previous cases, this, and the stationary distribution μ (5.7.1), if μ admits finite m-th moment, yield inequality (5.7.11).

The case $\gamma_0 = 1$. One can conclude by bounding directly the Lamperti transform from above by a drifted Brownian motion as it has been done for the bounds on the first hitting time from above of the level r_d.

5.8 Appendix: Auxiliary results

In this section, we provide some auxiliary results on well posedness of some integrals appearing in the likelihood in Proposition 5.8.1 and on the finiteness of the moments of the stationary measure μ in Proposition 5.8.2. More precisely, we give some properties of the moments of the T-CKLS process. Some are straightforward applications of the ergodic properties.

The following proposition describes the behavior of various integrals of the T-CKLS process. It establishes whether the likelihood (5.2.2) is well defined or not. Properties of this kind have also been considered in [84].

Proposition 5.8.1. *Let X be solution to the SDE* (5.2.1).

(a) If $\gamma_0 = 1/2$ and $a_0 \geq \sigma_0^2/2$ or if $\gamma_0 \in (1/2, 1] \cup \{0\}$, then

$$\forall t \geq 0, \quad \int_0^t \frac{1}{X_s^{2\gamma_0}} \mathbb{1}_{I_0}(X_s)\, \mathrm{d}s < \infty \quad \mathbb{P}_{x_0} - \ a.s..$$

(b) If $\gamma_0 = 1/2$ *and* $a_0 < \sigma_0^2/2$*, then*

$$\forall t \geq 0, \quad \mathbb{P}_{x_0}\left(\int_0^t \frac{1}{X_s^{2\gamma_0}} \mathbb{1}_{I_0}(X_s)\,\mathrm{d}s = \infty\right) > 0.$$

Proof. If $\gamma_0 = 0$, it is trivial. When $\gamma_0 \neq 0$, the first item follows from the fact that 0 is an unattainable boundary and continuity of the trajectories: the image of $[0, t]$ through each trajectory $s \mapsto X_s(\omega)$ is a compact of $]0, \infty[$. The second item, for $\gamma_0 = 1/2$ has been proven in [11] by using properties of the Laplace transform. □

The following proposition describes the behaviour of the moments from the stationary distribution of the T-CKLS process in the ergodic regime. We recall that the stationary distribution μ is given by (5.7.1).

Proposition 5.8.2. *Let* $m \in (0, \infty)$ *and assume that the conditions in Table 5.4 (ensuring that* μ *is the stationary distribution) hold. Then* μ *admits finite* m*-th moment unless it holds simultaneously* $b_d = 0$ *and* $\gamma_d \in [1/2, 1]$ *in which case the* m*-th moment is finite if*

- $\gamma_d = 1/2$*,* $b_d = 0$ *and* $a_d < -m\,\sigma_d^2/2$*.*
- $\gamma_d \in (1/2, 1)$*,* $b_d = 0$ *and* $m < \gamma_d - 1/2 < 1$*.*
- $\gamma_d = 1$ *and* $m \leq 1$*.*

The measure μ *admits finite* $-m$*-th moment unless* $\gamma_0 \in \{0, 1/2\}$ *in which case the* $(-m)$*-th moment is finite if*

- $\gamma_0 = 1/2$ *and* $a_0 > m\,\sigma_0^2/2$*.*
- $\gamma_0 = 0$ *and* $m < 1$*.*

Conclusion and Perspectives

In Chapter 4, we explicitly compute some densities related to the killed Ornstein-Uhlenbeck process. We focused on the transition probability density of the killed process over a constant boundary and the density of the first hitting time. To compute the transition probability density, we use a spectral decomposition method. This approach allows us to derive an explicit formulation of the transition probability density in the form of a function series. Based on this transition probability density, we have provided an explicit expression for the density of the first hitting time of this process. We justified the term-by-term derivation of the series using asymptotic expansion on the Parabolic Cylinder functions. We also computed the Laplace transform of the first hitting time. Then, we numerically compared each explicit formulation of these distributions with some well-known methods used in the literature. In terms of computational time, we have demonstrated the numerical advantage of our formulations. In this chapter, several observations can be made, and various perspectives can be explored. Firstly, for $j \in \{0, \cdots, d\}$, we computed the transition probability density and the first hitting time density only in the case of a T-OU that admits a mean-reverting mechanism on each subinterval I_j (i.e., $b_j > 0$). Note that, under this hypothesis, the process X is ergodic. Thus, one perspective would be to remove the assumption $b_j > 0$ on I_j. For example, we could consider a Brownian motion on the intervals I_j (i.e., $a_j, b_j = 0$) or a drifted Brownian motion (i.e., $b_j = 0$).

Regarding the transition probability density of the killed process, we have only proposed a formulation based on the spectral decomposition method. We believe that the method in [26] can be readapted for (4.2.8). This method would provide us with a semi-analytical formulation. Note that their formulation avoids the computation of eigenvalues that depend on the process parameters on each I_j. Furthermore, this method is presented in a more analytical manner, and we believe that a probabilistic interpretation is possible. Additionally, in [25], the same authors also provide this semi-analytical formulation for non-homogeneous boundary conditions.

The transition probability density of the killed process and the density of the first hitting times are closely related to the Parabolic Cylinder functions. These functions belong to a special class of Confluent Hypergeometric functions (see [68] for further discussion on this particular function). Moreover, some properties of the Parabolic Cylinder functions represent particular cases of properties of the Confluent Hypergeometric function. Here, the key element to deriving an explicit formulation for the density of the first hitting times lies in the asymptotic expansion associated with the Parabolic Cylinder functions. We believe that in some cases, our results can be extended to hitting time problems for other processes. More precisely, killed processes that have transition probability densities

expressed as Confluent Hypergeometric functions. For example, in the case of the T-CIR process (see [34]), the expressions for the first hitting time density are in the form of series of functions that involve Bessel functions. Furthermore, note that these series of functions remain complex to manipulate, especially when considering estimation problems as discussed in Chapter 2. An alternative perspective would be to explore other types of formulations that would facilitate a more straightforward derivation of properties regarding the density.

In Chapter 5, we propose a method for estimating the drift and volatility parameters of the T-CKLS process based on observations from a trajectory. We focus on the T-CKLS under the ergodic regime and the framework of continuous and high-frequency long-term observations. We propose drift estimators based on both likelihood and quasi-likelihood approaches. The volatility parameters are estimated using a quadratic estimator. We have established the consistency and asymptotic normality of our estimators. Following that, we compare our convergence hypotheses with those already existing in the literature for specific cases of the T-CKLS. Finally, we apply our estimators to a set of simulated data and to an interest rate dataset.

Note that the study of the T-CKLS process highlighted the problem of existence and uniqueness of some threshold SDEs. One can easily verify that the volatility function in (5.2.1) does not satisfy the hypotheses of Theorem 1.2.22 in Chapter 1. The following discussions are a work in progress with Sara Mazzonetto and focus on extending the existence and uniqueness results to a class of SDEs with discontinuous coefficients, which satisfies:

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad t \geq 0 \tag{5.8.1}$$

where B is a one-dimensional Brownian motion and b and σ are measurable functions which may be discontinuous at points called *"thresholds"*. We can see the discontinuous coefficients as follows:

$$b(x) = \sum_{i=1}^{N} b_i(x)\mathbb{1}_{I_i}(x) \quad \text{and} \quad \sigma(x) = \sum_{i=1}^{N} \sigma_i(x)\mathbb{1}_{I_i}(x),$$

for functions b_i and σ_i which are the restrictions of b and σ on the interval I_i. In Chapter 1, we remind the key results on pathwise uniqueness for SDEs under classical assumptions on the drift and volatility functions. These results are outlined in Theorem 1.2.22. The benefit of these results lies in their simple proof, which depends on the fact that the local time between two solutions of the same SDE is a.s. zero at zero. For specific threshold SDEs, these results can be extended. The main approach involves localizing each assumption outlined in Theorem 1.2.22. In our current work, we have extended these assumptions to the framework of processes solving the equation (5.8.1). Specifically, our results enable us to establish the existence and uniqueness of a strong solution of (5.2.1). Additionally, we can provide a comparison theorem for this type of SDE.

In Chapter 5, several remarks can be made, and various perspectives can be explored. Firstly, we believe that our assumptions regarding the moments of μ for consistency and asymptotic normality are sufficient in high-frequency long-time data. It would be interesting to explore other proof methods that could potentially weaken these assumptions.

Additionally, further study of the estimator for the parameter σ could aim to achieve a more optimal speed than $\sqrt{T_N}$. The estimation of the parameters γ and r remains an open problem, and the main goal would be to establish joint properties for the estimation of all parameters.

To explore further, investigating the statistical properties of our estimators for other regimes of the process would be valuable. We believe that obtaining these properties in the null recurrent regime should be achievable without significant difficulty. However, studying these estimators in the transient regime is an ongoing task.

Obtaining results for high-frequency data and fixed time under the assumption of ergodicity remains unresolved. In [81], for the T-OU process, the authors use a change of measure and the transition probability density of a threshold Brownian motion to establish statistical properties in this setting. The goal here would be to find a process with a reasonable transition probability density to apply similar reasoning (see [62] for the case of a CKLS process).

Based on our results, we can construct statistical tests for the existence of one or more thresholds in a CKLS process. We believe that this follows quite easily from the asymptotic normality properties stated in Theorem 5.4.4.

Additionally, as mentioned previously, the proofs of Proposition 5.5.4 and Proposition 5.5.5 are applicable to other processes. A research direction would be to extend our results on drift parameter estimation of a more general SDE:

$$X_t = X_0 + \int_0^t \mu(X_s, \theta_0)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad t \geq 0$$

where μ, σ are given functions and discontinuous at certain levels and θ_0, the parameter to be estimated. In this case, proofs will be slightly different as we cannot explicitly compute the MLE and QMLE. We would need to study the statistical properties on the likelihood and quasi-likelihood functions. In some cases, we believe that we could also obtain results in the high-frequency and fixed time data case using the local time approximation [79], as demonstrated in [82] for the threshold drifted Brownian motion.

Furthermore, note that we currently have no results concerning the existence of the transition probability density of the T-CKLS process. Indeed, the process does not satisfy the assumptions outlined in [45]. Partly because the volatility function is no longer elliptical around zero nor locally Hölder continuous near a threshold.

Bibliography

[1] J. Abate and P. P. Valkó. Multi-precision Laplace transform inversion. *Int. J. Numer. Methods Eng.*, 60(5):979–993, 2004. ISSN 0029-5981. doi: 10.1002/nme.995.

[2] J. Abate and W. Whitt. The Fourier-series method for inverting transforms of probability distributions. *Queueing Syst.*, 10(1-2):5–87, 1992. ISSN 0257-0130. doi: 10.1007/BF01158520.

[3] Y. Aït-Sahalia. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262, 2002. ISSN 0012-9682. doi: 10.1111/1468-0262.00274.

[4] P. Alaton, B. Djehiche, and D. Stillberger. On modelling and pricing weather derivatives. *Appl. Math. Finance*, 9(1):1–20, 2002. ISSN 1350-486X. doi: 10.1080/13504860210132897.

[5] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384, 2005. ISSN 0929-9629. doi: 10.1515/156939605777438569.

[6] L. Alili, P. Patie, and J. L. Pedersen. Representations of the first hitting time density of an Ornstein-Uhlenbeck process. *Stoch. Models*, 21(4):967–980, 2005. ISSN 1532-6349. doi: 10.1080/15326340500294702.

[7] C. Amorino and A. Gloter. Contrast function estimation for the drift parameter of ergodic jump diffusion process. *Scandinavian Journal of Statistics*, 47(2):279–346, 2020.

[8] S. Ankirchner, C. Blanchet-Scalliet, D. Dorobantu, and L. Gay. First passage time density of an Ornstein-Uhlenbeck process with broken drift. *Stoch. Models*, 38(2): 308–329, 2022. ISSN 1532-6349. doi: 10.1080/15326349.2022.2026790.

[9] R.F. Bass and Z. Chen. One-dimensional stochastic differential equations with singular and degenerate coefficients. *Sankhyā*, 67(1):19–45, 2005.

[10] M. Ben Alaya and A. Kebaier. Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases. *Stoch. Models*, 28(4):609–634, 2012. ISSN 1532-6349. doi: 10.1080/15326349.2012.726042.

[11] M. Ben Alaya and A. Kebaier. Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stochastic Anal. Appl.*, 31(4):552–573, 2013. ISSN 0736-2994. doi: 10.1080/07362994.2013.798175.

[12] M. Ben Alaya, T. Ngo, and S. Pergamenchtchikov. Optimal guaranteed estimation methods for the Cox - Ingersoll - Ross models. working paper or preprint, January 2023.

[13] J. Besag. Spatial interaction and the statistical analysis of lattice systems. Discussion. *J. R. Stat. Soc., Ser. B*, 36:192–236, 1974. ISSN 0035-9246.

[14] P. Billingsley. *Ergodic theory and information*. Wiley Ser. Probab. Math. Stat. John Wiley & Sons, Hoboken, NJ, 1965.

[15] C. Blanchet-Scalliet, D. Dorobantu, L. Gay, V. Maume-Deschamps, and P. Ribereau. Risk assessment using suprema data. *Stochastic Environmental Research and Risk Assessment*, 32(10):2839–2848, 2018.

[16] C. Blanchet-Scalliet, D. Dorobantu, and L. Gay. Joint law of an Ornstein-Uhlenbeck process and its supremum. *J. Appl. Probab.*, 57(2):541–558, 2020. ISSN 0021-9002. doi: 10.1017/jpr.2020.22.

[17] C. Blanchet-Scalliet, D. Dorobantu, and B. Nieto. Some properties for ν-zeros of Parabolic Cylinder functions. *Matematiche*, 78(2):277–287, 2023. ISSN 0373-3505. doi: 10.4418/2023.78.2.1.

[18] C. Blanchet-Scalliet, D. Dorobantu, and B. Nieto. A pseudo-likelihood estimator of the Ornstein-Uhlenbeck parameters from suprema observations. *Stat. Inference Stoch. Process.*, 27(2):407–425, 2024. ISSN 1387-0874. doi: 10.1007/s11203-024-09307-4.

[19] A. N. Borodin and P. Salminen. *Handbook of Brownian motion: Facts and formulae.* Probab. Appl. Basel: Birkhäuser, 2nd ed. edition, 2002. ISBN 3-7643-6705-9.

[20] R. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144, 2005. ISSN 1549-5787. doi: 10.1214/154957805100000104.

[21] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. New York, NY: Springer, 2011. ISBN 978-0-387-70913-0.

[22] D. Brody, J. Syroka, and M. Zervos. Dynamical pricing of weather derivatives. *Quant. Finance*, 2(3):189–198, 2002. ISSN 1469-7688. doi: 10.1088/1469-7688/2/3/302.

[23] R. Brown. Xxvii. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The philosophical magazine*, 4(21):161–173, 1828.

[24] R. Cantrell and C. Cosner. Diffusion models for population dynamics incorporating individual behavior at boundaries: Applications to refuge design. *Theor. Popul. Biol.*, 55(2):189–207, 1999. ISSN 0040-5809. doi: 10.1006/tpbi.1998.1397.

[25] E. Carr and N. March. Semi-analytical solution of multilayer diffusion problems with time-varying boundary conditions and general interface conditions. *Appl. Math. Comput.*, 333:286–303, 2018. ISSN 0096-3003. doi: 10.1016/j.amc.2018.03.095.

[26] E. Carr and I. Turner. A semi-analytical solution for multilayer diffusion in a composite medium consisting of a large number of layers. *Appl. Math. Modelling*, 40(15-16):7034–7050, 2016. ISSN 0307-904X. doi: 10.1016/j.apm.2016.02.041.

[27] K. C. Chan, G. Karolyi, F. Longstaff, and A. Sanders. The volatility of short-term interest rates: an empirical comparison of alternative models of the term structure of interest rates. In *The new interest rate models. Recent developments in the theory and application of yield curve dynamics*, pages 87–100. London: Risk Books, 2000. ISBN 1-899332-97-9.

[28] P. Cheridito, D. Filipović, and R. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83(1):123–170, 2007.

[29] J. A. Cochran. The zeros of Hankel functions as functions of their order. *Numer. Math.*, 7:238–250, 1965. ISSN 0029-599X. doi: 10.1007/BF01436080.

[30] S. Conde and S. L. Kalla. The v-zeros of $J_{-\nu}(x)$. *Math. Comput.*, 33:423–426, 1979. ISSN 0025-5718. doi: 10.1090/S0025-5718-1979-0514838-0.

[31] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985. ISSN 0012-9682. doi: 10.2307/1911242. URL `semanticscholar.org/paper/853c1f9b36a574432f112d72b473a68f377f410e`.

[32] I. Crimaldi and L. Pratelli. Convergence results for multivariate martingales. *Stochastic Processes Appl.*, 115(4):571–577, 2005. ISSN 0304-4149. doi: 10.1016/j.spa.2004.10.004.

[33] D. Dacunha-Castelle and M. Duflo. Probability and statistics. Volume II. Transl. from the French by David McHale. New York etc.: Springer-Verlag. XIV, 410 p. DM 86.00 (1986)., 1986.

[34] M. Decamps, M. Goovaerts, and W. Schoutens. Self exciting threshold interest rates models. *Int. J. Theor. Appl. Finance*, 9(7):1093–1122, 2006. ISSN 0219-0249. doi: 10.1142/S0219024906003937.

[35] D. Dereudre, S. Mazzonetto, and S. Roelly. An explicit representation of the transition densities of the skew Brownian motion with drift and two semipermeable barriers. *Monte Carlo Methods Appl.*, 22(1):1–23, 2016. ISSN 0929-9629. doi: 10.1515/mcma-2016-0100.

[36] K. Ding and Z. Cui. A general framework to simulate diffusions with discontinuous coefficients and local times. *ACM Transactions on Modeling and Computer Simulation*, 32(4):1–29, 2022.

[37] B. Dischel. At last: A model for weather risk. *Energy and Power Risk Management*, 11(3):20–21, 1998.

[38] F. Dornier and M. Queruel. Caution to the wind. *Energy & Power Risk Management*, 13(8):30–32, 2000.

[39] P. Doukhan. *Mixing: Properties and examples*, volume 85 of *Lect. Notes Stat.* New York: Springer-Verlag, 1994. ISBN 0-387-94214-9.

[40] A. Elbert. Some recent results on the zeros of Bessel functions and orthogonal polynomials. *J. Comput. Appl. Math.*, 133(1-2):65–83, 2001. ISSN 0377-0427. doi: 10.1016/S0377-0427(00)00635-X.

[41] A. Elbert and M. Muldoon. Inequalities and monotinicity properties for zeros of Hermite functions. *Proc. R. Soc. Edinb., Sect. A, Math.*, 129(1):57–75, 1999. ISSN 0308-2105. doi: 10.1017/S0308210500027463.

[42] A. Elbert and M. Muldoon. Approximations for zeros of Hermite functions. In *Special functions and orthogonal polynomials. AMS special session, Tucson, AZ, USA, April 21–22, 2007*, pages 117–126. Providence, RI: American Mathematical Society (AMS), 2008. ISBN 978-0-8218-4650-6.

[43] H. J. Engelbert and W. Schmidt. On one-dimensional stochastic differential equations with generalized drift. Stochastic differential systems, Proc. IFIP-WG 7/1 Work. Conf., Marseille-Luminy/France 1984, Lect. Notes Control Inf. Sci. 69, 143-155 (1985)., 1985.

[44] H. J. Engelbert and W. Schmidt. Strong Markov continuous local martingales and solutions of one- dimensional stochastic differential equations. III. *Math. Nachr.*, 151:149–197, 1991. ISSN 0025-584X. doi: 10.1002/mana.19911510111.

[45] N. Fournier and J. Printems. Absolute continuity for some one-dimensional processes. *Bernoulli*, 16(2):343–360, 2010. ISSN 1350-7265. doi: 10.3150/09-BEJ215.

[46] E. Gobet. *Monte-Carlo methods and stochastic processes. From linear to non-linear.* Boca Raton, FL: CRC Press, 2016. ISBN 978-1-4987-4622-9; 978-1-315-36875-7.

[47] E. Gobet and G. Matulewicz. Parameter estimation of Ornstein-Uhlenbeck process generating a stochastic graph. *Stat. Inference Stoch. Process.*, 20(2):211–235, 2017. ISSN 1387-0874. doi: 10.1007/s11203-016-9142-4.

[48] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.*, volume 24 of *North-Holland Math. Libr.* Amsterdam etc.: North-Holland; Tokyo: Kodansha Ltd., 2nd ed. edition, 1989. ISBN 0-444-87378-3; 4-06-203231-7.

[49] K. Itô and H. McKean. *Diffusion processes and their sample paths.* Class. Math. Berlin: Springer-Verlag, repr. of the 1974 ed. edition, 1996. ISBN 3-540-60629-7.

[50] J. Jacod and P. Protter. *Discretization of processes.*, volume 67 of *Stoch. Model. Appl. Probab.* Berlin: Springer, 2012. ISBN 978-3-642-24126-0; 978-3-642-24127-7. doi: 10.1007/978-3-642-24127-7.

[51] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes.*, volume 288 of *Grundlehren Math. Wiss.* Berlin: Springer, 2nd ed. edition, 2003. ISBN 3-540-43932-3.

[52] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets.* Springer Finance. London: Springer, 2009. ISBN 978-1-85233-376-8; 978-1-4471-2524-2; 978-1-84628-737-4. doi: 10.1007/978-1-84628-737-4.

[53] O. Kallenberg. *Foundations of modern probability.* Probab. Appl. New York, NY: Springer, 2nd ed. edition, 2002. ISBN 0-387-95313-2.

[54] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Grad. Texts Math.* New York etc.: Springer-Verlag, 1988. ISBN 0-387-96535-1.

[55] J. Keilson and J. Wellner. Oscillating Brownian motion. *J. Appl. Probab.*, 15: 300–310, 1978. ISSN 0021-9002. doi: 10.2307/3213403.

[56] J. B. Keller, S. I. Rubinow, and M. Goldstein. Zeros of Hankel functions and poles of scattering amplitudes. *J. Math. Phys.*, 4:829–832, 1963. ISSN 0022-2488. doi: 10.1063/1.1724325.

[57] J. Kent. Eigenvalue expansions for diffusion hitting times. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 52:309–319, 1980. ISSN 0044-3719. doi: 10.1007/BF00538895.

[58] A.M.G. Klein Tank, J.B. Wijngaard, G.P. Können, R. Böhm, G. Demarée, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, et al. Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 22(12):1441–1453, 2002.

[59] Y. Kutoyants. On identification of the threshold diffusion processes. *Ann. Inst. Statist. Math.*, 64(2):383–413, 2012.

[60] K. Laaidi, M. Pascal, M. Ledrans, A. Le Tertre, S. Medina, and C. Schonemann. Le système français d'alerte canicule et santé 2004 (Sacs 2004). un dispositif intégré au plan national canicule. *Bulletin épidémiologique hebdomadaire*, (30-31):134–136, 2004.

[61] A. Laforgia and P. Natalini. Zeros of Bessel functions: monotonicity, concavity, inequalities. *Matematiche*, 62(2):255–270, 2007. ISSN 0373-3505.

[62] G. Lan, Y. Hu, and C. Zhang. The explicit solution and precise distribution of CKLS model under girsanov transform, 2014.

[63] P. Lansky, P. Sanda, and J. He. The parameters of the stochastic leaky integrate-and-fire neuronal model. *J. Comput. Neurosci.*, 21(2):211–223, 2006. ISSN 0929-5313. doi: 10.1007/s10827-006-8527-6.

[64] P. Lansky, P. Sanda, and J. He. The parameters of the stochastic leaky integrate-and-fire neuronal model. *J. Comput. Neurosci.*, 21(2):211–223, 2006. ISSN 0929-5313. doi: 10.1007/s10827-006-8527-6.

[65] L. Le Cam and G. Lo Yang. *Asymptotics in statistics: some basic concepts*. Springer Ser. Stat. New York etc.: Springer-Verlag, 1990. ISBN 0-387-97372-9.

[66] J. F. Le Gall. One-dimensional stochastic differential equations involving the local times of the unknown process. Stochastic analysis and applications, Proc. int. Conf., Swansea 1983, Lect. Notes Math. 1095, 51-82 (1984)., 1984.

[67] J. F. Le Gall. *Brownian motion, martingales, and stochastic calculus*, volume 274 of *Grad. Texts Math.* Cham: Springer, 2016. ISBN 978-3-319-31088-6; 978-3-319-31089-3. doi: 10.1007/978-3-319-31089-3.

[68] N. Lebedev. Special functions and their applications. *Physics Today*, 18(12):70, 1965.

[69] A. Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.*, 3: 413–466, 2006. ISSN 1549-5787. doi: 10.1214/154957807000000013.

[70] A. Lejay and P. Pigato. Statistical estimation of the oscillating Brownian motion. *Bernoulli*, 24(4B):3568–3602, 2018. ISSN 1350-7265. doi: 10.3150/17-BEJ969.

[71] A. Lejay and P. Pigato. A threshold model for local volatility: evidence of leverage and mean reversion effects on historical data. *International Journal of Theoretical and Applied Finance*, 22(4):1950017, 2019.

[72] A. Lejay and P. Pigato. Maximum likelihood drift estimation for a threshold diffusion. *Scand. J. Stat.*, 47(3):609–637, 2020. ISSN 0303-6898. doi: 10.1111/sjos.12417.

[73] A. Lejay, L. Lenôtre, and G. Pichot. *One-dimensional skew diffusions: explicit expressions of densities and resolvent kernels*. PhD thesis, Inria Rennes-Bretagne Atlantique; Inria Nancy-Grand Est, 2015.

[74] D. Lepingle. Sur le comportement asymptotique des martingales locales. Semin. Probab. XII, Univ. Strasbourg 1976/77, Lect. Notes Math. 649, 148-161 (1978)., 1978.

[75] V. Linetsky. Computing hitting time densities for CIR and OU diffusions: applications to mean-reverting models. *Journal of Computational Finance*, 7:1–22, 2004.

[76] V. Linetsky. Lookback options and diffusion hitting times: a spectral expansion approach. *Finance Stoch.*, 8(3):373–398, 2004. ISSN 0949-2984. doi: 10.1007/s00780-003-0120-5.

[77] E. Löcherbach. Ergodicity and speed of convergence to equilibrium for diffusion processes. *Unpublished manuscript, June*, 2013.

[78] W. Magnus and L. Kotin. The zeros of the Hankel function as a function of its order. *Numer. Math.*, 2:228–244, 1960. ISSN 0029-599X. doi: 10.1007/BF01386226.

[79] S. Mazzonetto. Rates of convergence to the local time of Oscillating and Skew Brownian Motions. *arXiv preprint arXiv:1912.04858*, 2019.

[80] S. Mazzonetto and B. Nieto. Parameters estimation of a threshold ckls process from continuous and discrete observations. 2024.

[81] S. Mazzonetto and P. Pigato. Drift estimation of the threshold Ornstein-Uhlenbeck process from continuous and discrete observations. *Stat. Sin.*, 34(1):313–336, 2024. ISSN 1017-0405.

[82] S. Mazzonetto and P. Pigato. Estimation of parameters and local times in a discretely observed threshold diffusion model. *Preprint arXiv:2403.06858*, 2024.

[83] Y. Mishura, K. Ralchenko, and O. Dehtiar. Parameter estimation in CKLS model by continuous observations. *Stat. Probab. Lett.*, 184:10, 2022. ISSN 0167-7152. doi: 10.1016/j.spl.2022.109391. Id/No 109391.

[84] Y. Mishura, A. Pilipenko, and A. Yurchenko-Tytarenko. Low-dimensional cox-ingersoll-ross process. *Stochastics*, pages 1–21, 2024.

[85] P. Mota and M. Esquível. On a continuous time stock price model with regime switching, delay, and threshold. *Quant. Finance*, 14(8):1479–1488, 2014. ISSN 1469-7688. doi: 10.1080/14697688.2013.879990.

[86] F.W.J Olver. The asymptotic expansion of Bessel functions of large order. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 247(930):328–368, 1954.

[87] F.W.J Olver. Uniform asymptotic expansions for Weber Parabolic Cylinder functions of large orders. *J. Res. Nat. Bur. Standards Sect. B*, 63(2):131–169, 1959.

[88] B. Palka. *An introduction to complex function theory.* Undergraduate Texts Math. New York etc.: Springer-Verlag, 1991. ISBN 0-387-97427-X.

[89] A. Pilipenko. *An introduction to stochastic differential equations with reflection*, volume 1 of *Lect. Pure Appl. Math.* Potsdam: Potsdam University Press, 2014. ISBN 978-3-86956-297-1.

[90] G. Pittaluga and L. Sacripante. Inequalities for the zeros of the Airy functions. *SIAM J. Math. Anal.*, 22(1):260–267, 1991. ISSN 0036-1410. doi: 10.1137/0522015.

[91] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren Math. Wiss.* Berlin: Springer, 3rd ed., 3rd. corrected printing edition, 2005. ISBN 3-540-64325-7; 978-3-642-08400-3.

[92] J. Robert, V. John, and P. William. Pairs trading. *Quantitative Finance*, 5(3): 271–276, 2005. doi: 10.1080/14697680500149370.

[93] G. Saporta. *Probabilités, analyse des données et statistique. (Probabilities, data analysis and statistics).* Paris: Éditions Technip, 1990. ISBN 2-7108-0565-0.

[94] L. J. Slater. Confluent hypergeometric functions. Moskau: Rechenzentrum der Akad. der Wiss. der UdSSR 1966, 250 S. (1966)., 1966.

[95] W.F. Stout. Almost sure convergence. Probability and Mathematical Statistics. 24. New York - San Francisco - London: Academic Press, a subsidiary of Harcourt Brace Jovanovich., 1974.

[96] D. Stroock and S. Varadhan. *Multidimensional diffusion processes.* Class. Math. Berlin: Springer, reprint of the 2nd correted printing (1997) edition, 2006. ISBN 3-540-28998-4.

[97] F. Su and K. Chan. Quasi-likelihood estimation of a threshold diffusion process. *J. Econom.*, 189(2):473–484, 2015. ISSN 0304-4076.

[98] F. Su and K. Chan. Testing for threshold diffusion. *Journal of Business & Economic Statistics*, 35(2):218–227, 2017.

[99] D. Taguchi and A. Tanaka. On the Euler-Maruyama scheme for degenerate stochastic differential equations with non-sticky condition. In *Séminaire de probabilités L*, pages 165–185. Cham: Springer, 2019. ISBN 978-3-030-28534-0; 978-3-030-28535-7. doi: 10.1007/978-3-030-28535-7_9.

[100] A. Talbot. The accurate numerical inversion of Laplace transform. *J. Inst. Math. Appl.*, 23:97–120, 1979. ISSN 0020-2932. doi: 10.1093/imamat/23.1.97.

[101] N. Temme. Numerical and asymptotic aspects of Parabolic Cylinder functions. *J. Comput. Appl. Math.*, 121(1-2):221–246, 2000. ISSN 0377-0427. doi: 10.1016/S0377-0427(00)00347-2.

[102] H. Tong. Threshold models in time series analysis – 30 years on. *Stat. Interface*, 4 (2):107–118, 2011. ISSN 1938-7989. doi: 10.4310/SII.2011.v4.n2.a1.

[103] D.C. Trost, E. A. Overman, J.H. Ostroff, W. Xiong, and P. March. A Model for Liver Homeostasis Using Modified Mean-Reverting Ornstein–Uhlenbeck Process. *Computational and Mathematical Methods in Medicine*, 11(3):21, 2010.

[104] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the Brownian motion. *Phys. Rev., II. Ser.*, 36:823–841, 1930. ISSN 0031-899X. doi: 10.1103/PhysRev.36.823.

[105] A. W. Van der Vaart. *Asymptotic statistics*, volume 3 of *Camb. Ser. Stat. Probab. Math.* Cambridge: Cambridge Univ. Press, 1998. ISBN 0-521-49603-9. doi: 10.1017/CBO9780511802256.

[106] O. Vasicek. An equilibrium characterization of the term structure. *J. Financ. Econ.*, 5(2):177–188, 1977. ISSN 0304-405X. doi: 10.1016/0304-405X(77)90016-2.

[107] O. Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2):177–188, 1977.

[108] S. Wang, S. Song, and Y. Wang. Skew Ornstein-Uhlenbeck processes and their financial applications. *J. Comput. Appl. Math.*, 273:363–382, 2015. ISSN 0377-0427. doi: 10.1016/j.cam.2014.06.023.

[109] G. N. Watson. A treatise on the theory of Bessel functions. Cambridge: University Press, viii, 804 S. (1922)., 1922.

[110] M. Wilkinson and A. Pumir. Spherical Ornstein-Uhlenbeck processes. *J. Stat. Phys.*, 145(1):113–142, 2011. ISSN 0022-4715. doi: 10.1007/s10955-011-0332-6.

[111] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971. ISSN 0023-608X. doi: 10.1215/kjm/1250523691.

[112] H. Yaozhong and X. Yuejuan. Parameter estimation for threshold Ornstein–Uhlenbeck processes from discrete observations. *Journal of Computational and Applied Mathematics*, 411:114264, 2022.

[113] T. Yu, H. Tsai, and H. Rachinger. Approximate maximum likelihood estimation of a threshold diffusion process. *Comput. Stat. Data Anal.*, 142:14, 2020. ISSN 0167-9473. doi: 10.1016/j.csda.2019.106823. Id/No 106823.

Multi-Mean Reverting Processes: Analytical and Statistical Approaches

Abstract : This thesis explores the theory and application of stochastic differential equations (SDEs), with a particular focus on parameter estimation and the behavior of processes with discontinuous coefficients. The first part focuses on the Ornstein-Uhlenbeck (OU) process. We introduce an estimator for the OU parameters based on supremum observations of one trajectory. We derive an analytic expression for the supremum density and build an estimator using a pseudo-likelihood method. The statistical properties of this estimator, including consistency and asymptotic normality, are established using weak-dependence results. The effectiveness of our estimator is demonstrated through its application to both simulated and real data. Additionally, we explore the behavior of Parabolic Cylinder functions, which are involved in the law of the OU supremum. Specifically, we investigate the μ-zeros of the function $\mu \mapsto D_\mu(z)$ with respect to the real variable z, establishing a formula for the derivative of a zero and providing an asymptotic expansion for large positive z. The second part investigates processes governed by SDEs with discontinuous coefficients. Introducing the threshold Ornstein-Uhlenbeck (T-OU) process, we provide explicit expressions for transition probability density and first hitting time density for the killed process. Then, we introduce the threshold CKLS (T-CKLS) process and focus on estimating its parameters using observations of a single trajectory. We study the asymptotic behavior of maximum likelihood and quasi-maximum likelihood estimators for drift parameters, as well as a volatility estimator. Some statistical properties under continuous and high-frequency observations over long time horizons are derived. Finally, we highlight the relevance of multiple thresholds through applications to simulated and real data.

Keywords: OU, CIR, CKLS, supremum law, parameters estimation, pseudo-likelihood, asymptotic normality, hitting time, self-exciting process , maximum likelihood, thresholds diffusion.

Processus à plusieurs réversions à la moyenne : Approches analytiques et statistiques

Résumé : Cette thèse traite de la théorie et des applications des équations différentielles stochastiques (EDS), en se concentrant particulièrement sur l'estimation des paramètres et le comportement des processus aux coefficients discontinus. La première partie introduit un estimateur pour les paramètres du processus d'Ornstein-Uhlenbeck (OU), construit à partir d'observations du supremum d'une unique trajectoire. Une fois l'expression analytique pour la densité du supremum établie, nous procédons à la construction d'un estimateur en utilisant une méthode de pseudo-vraisemblance. Les propriétés statistiques de cet estimateur, à savoir la consistance et la normalité asymptotique, sont établies en utilisant des propriétés de faible dépendance de l'échantillon d'observations. L'efficacité de notre estimateur est démontrée à travers son application à des données simulées et réelles. De plus, nous étudions le comportement des fonctions Paraboliques Cylindrique, qui sont impliquées dans la loi du supremum de l'OU. Plus précisément, nous étudions les μ-zéros de la fonction $\mu \mapsto D_\mu(z)$ par rapport à la variable réelle z. Nous établissons une formule pour la dérivée d'un zéro et fournissons un développement asymptotique pour de grands z positifs. La deuxième partie développe la théorie des processus solutions des EDS à coefficients discontinus. Après avoir introduit le processus d'Ornstein-Uhlenbeck à seuil (T-OU), nous établissons des expressions analytiques pour la densité de probabilité de transition et la densité de premier temps d'atteinte pour le processus tué. Ensuite, le processus CKLS avec seuil (T-CKLS) est introduit et nous nous concentrons sur l'estimation de ses paramètres de dérive et de volatilité en utilisant des observations d'une seule trajectoire. L'analyse du comportement asymptotique des estimateurs de maximum de vraisemblance et de quasi-maximum de vraisemblance pour les paramètres de dérive, ainsi qu'un estimateur de volatilité, est effectuée. Les propriétés statistiques sont obtenues à partir d'observations continues et à haute fréquence en temps long. Enfin, la pertinence d'une modélisation à plusieurs seuils est mise en évidence à travers des applications à des données simulées et réelles.

Mots clefs. OU, CIR, CKLS, loi du supremum, estimation des paramètres, pseudo-vraisemblance, normalité asymptotique, premier temps d'atteinte, processus auto-excitant, maximum de vraisemblance, diffusion à seuils.

UNIVERSITÉ DE LYON