Asymptotic methods for option pricing in finance

David Krief

Université Sorbonne Paris Cité
Université Paris Diderot

École Doctorale de Sciences Mathématiques de Paris Centre
Laboratoire de Probabilité Statistiques et Modélisation

Thèse de doctorat

Discipline: Mathématiques Appliquées

Présentée par
David Krief

Méthodes Asymptotiques pour la Valorisation d'Options en Finance

Asymptotic Methods for Option Pricing in Finance

Sous la direction de **Peter Tankov** et **Zorana Grbac**

Soutenue le **27 septembre 2018** devant le jury composé de:

Aurélien Alfonsi	Professeur, ENPC	examinateur
Noufel Frikha	Maître de Conf., Université Paris Diderot	examinateur
Zorana Grbac	Maître de Conf., Université Paris Diderot	directrice
Antoine Jacquier	Professeur, Imperial College	rapporteur
Benjamin Jourdain	Professeur, ENPC	rapporteur
Huyên Pham	Professeur, Université Paris Diderot	examinateur
Agnès Sulem	Professeur, INRIA	Présidente du jury
Peter Tankov	Professeur, ENSAE	directeur

Contents

Résumé

Dans cette thèse, nous étudions plusieurs problèmes de mathématiques financières liés à la valorisation des produits dérivés. Par différentes approches asymptotiques, nous développons des méthodes pour calculer des approximations précises du prix de certains types d'options dans des cas où il n'existe pas de formule explicite.

Dans le premier chapitre, nous nous intéressons à la valorisation des options dont le payoff dépend de la trajectoire du sous-jacent par méthodes de Monte-Carlo, lorsque le sous-jacent est modélisé par un processus affine à volatilité stochastique. Nous prouvons un principe de grandes déviations trajectoriel en temps long, que nous utilisons pour calculer, en utilisant le lemme de Varadhan, un changement de mesure asymptotiquement optimal, permettant de réduire significativement la variance de l'estimateur de Monte-Carlo des prix d'options.

Le second chapitre considère la valorisation par méthodes de Monte-Carlo des options dépendant de plusieurs sous-jacents, telles que les options sur panier, dans le modèle à volatilité stochastique de Wishart, qui généralise le modèle Heston. En suivant la même approche que dans le précédent chapitre, nous prouvons que le processus vérifie un principe de grandes déviations en temps long, que nous utilisons pour réduire significativement la variance de l'estimateur de Monte-Carlo des prix d'options, à travers un changement de mesure asymptotiquement optimal. En parallèle, nous utilisons le principe de grandes déviations pour caractériser le comportement en temps long de la volatilité implicite Black-Scholes des options sur panier.

Dans le troisième chapitre, nous étudions la valorisation des options sur variance réalisée, lorsque la volatilité spot est modélisée par un processus de diffusion à volatilité constante. Nous utilisons de récents résultats asymptotiques sur les densités des diffusions hypo-elliptiques pour calculer une expansion de la densité de la variance réalisée, que nous intégrons pour obtenir l'expansion du prix des options, puis de leur volatilité implicite Black-Scholes.

Le dernier chapitre est consacré à la valorisation des dérivés de taux d'intérêt dans le modèle Lévy de marché Libor qui généralise le modèle de marché Libor classique (log-normal) par l'ajout de sauts. En écrivant le premier comme une perturbation du second et en utilisant la représentation de Feynman-Kac, nous calculons explicitement l'expansion asymptotique du prix des dérivés de taux, en particulier, des caplets et des swaptions.

Mots clés : Valorisation d'options, Méthodes asymptotiques, Grandes déviations, Monte-Carlo, Échantillonnage préférentiel, Expansions asymptotiques, Volatilité stochastique, Processus à sauts, Modèles affines, Volatilité implicite

Abstract

In this thesis, we study several mathematical finance problems, related to the pricing of derivatives. Using different asymptotic approaches, we develop methods to calculate accurate approximations of the prices of certain types of options in cases where no explicit formulas are available.

In the first chapter, we are interested in the pricing of path-dependent options, with Monte-Carlo methods, when the underlying is modelled as an affine stochastic volatility model. We prove a long-time trajectorial large deviations principle. We then combine it with Varadhan's Lemma to calculate an asymptotically optimal measure change, that allows to reduce significantly the variance of the Monte-Carlo estimator of option prices.

The second chapter considers the pricing with Monte-Carlo methods of options that depend on several underlying assets, such as basket options, in the Wishart stochastic volatility model, that generalizes the Heston model. Following the approach of the first chapter, we prove that the process verifies a long-time large deviations principle, that we use to reduce significantly the variance of the Monte-Carlo estimator of option prices, through an asymptotically optimal measure change. In parallel, we use the large deviations property to characterize the long-time behaviour of the Black-Scholes implied volatility of basket options.

In the third chapter, we study the pricing of options on realized variance, when the spot volatility is modelled as a diffusion process with constant volatility. We use recent asymptotic results on densities of hypo-elliptic diffusions to calculate an expansion of the density of realized variance, that we integrate to obtain an expansion of option prices and their Black-Scholes implied volatility.

The last chapter is dedicated to the pricing of interest rate derivatives in the Levy Libor market model, that generalizes the classical (log-normal) Libor market model by introducing jumps. Writing the first model as a perturbation of the second and using the Feynman-Kac representation, we calculate explicit expansions of the prices of interest rate derivatives and, in particular, caplets and swaptions.

Key words : Option pricing, Asymptotic methods, Large deviations, Monte-Carlo, Optimal sampling, Asymptotic expansions, Stochastic volatility, Jump processes, Affine models, Implied volatility

Chapter 1

Introduction

1.1 Option pricing

1.1.1 The Black and Scholes model

The history of the mathematical pricing of financial derivatives begins with Louis Bachelier (Bachelier, 1900), who developed a pricing theory based on the assumption that stock prices evolve as a Brownian motion. In the 1970s, this theory made a huge leap forward with the long celebrated Black and Scholes (BS) model (Merton, 1973; Black and Scholes, 1973) that models the price of a stock as an exponential Brownian motion with SDE

$$dS_t = S_t \left(\mu \, dt + \sigma \, dW_t\right) \,, \tag{1.1.1}$$

where $\mu \in \mathbb{R}$ is the drift of the dynamics of the stock, $\sigma > 0$ is its volatility and $(W_t)_{t \geq 0}$ is a standard Brownian motion. Under absence of arbitrage opportunities and assuming zero interest rate for simplicity, the price of a derivative with payoff $p(S)$ is expressed as

$$\mathbb{E}^{\mathbb{Q}}[p(S)] \,,$$

where $\mathbb{Q}$ is a measure under which $(S_t)_{t \geq 0}$ is a martingale. The existence of such a measure is proved in (Dalang et al., 1990). Under this measure, $W_t^{\mathbb{Q}} := W_t + \frac{\mu}{\sigma} \, t$ is a standard Brownian motion and simple calculations show that

$$S_t = S_0 \, e^{-\frac{\sigma^2}{2} \, t + \sigma \, W_t} \overset{\mathbb{Q}}{\sim} \mathcal{LN}\left(\log(S_0) - \frac{\sigma^2}{2} \, t \,, \, \sigma^2 t\right) \,,$$

thus allowing to express the price of European derivatives as

$$\mathbb{E}^{\mathbb{Q}}\left[p(S_T)\right] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} p\left(S_0 \, e^{-\frac{\sigma^2}{2} T + \sigma \sqrt{T} w}\right) e^{-\frac{w^2}{2}} \, dw \,.$$

In particular, considering a call option with maturity T and strike K, we obtain the famous Black and Scholes formula

$$\mathbb{E}^{\mathbb{Q}}\left[(S_T - K)_+\right] = S_0 \, \Phi\left(d_+\right) - K \, \Phi\left(d_-\right) \,, \tag{1.1.2}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a Gaussian random variable and

$$d_{\pm} = \frac{\log(S_0/K)}{\sigma\sqrt{T}} \pm \frac{\sigma\sqrt{T}}{2} \,. \tag{1.1.3}$$

In the years following its development, the BS model became so popular that, on the markets, option prices are quoted in BS "implied volatility", that is the volatility σ, that yields the price of the option when inserted in eqs. (1.1.2) and (1.1.3).

1.1.2 Beyond the BS model

Even though highly tractable, the BS model suffers from several drawbacks. In particular, the model fails to replicate some of the behaviours observed on the financial markets, where the implied volatility of options as a function of the strike typically displays the shape of a "smile" (see Figure 1.1.1), whereas in the BS model, the implied volatility is trivially constant and therefore cannot fit the observed data.

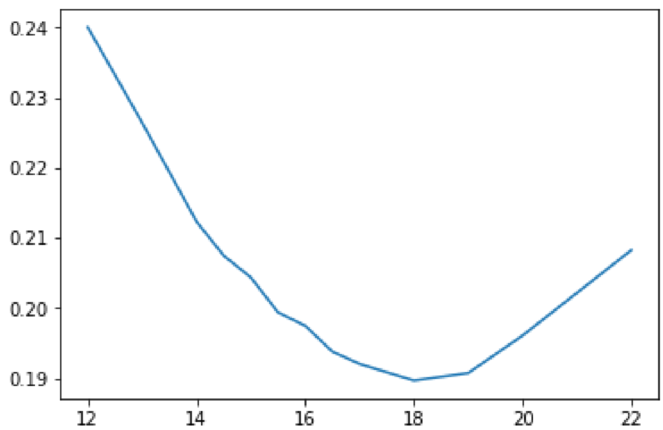

Figure 1.1.1: The implied volatility smile of call options on UBS stocks on the 31st of May 2018, with maturity 1 year.

In addition, the BS setting models the stock log-returns as Gaussian random variables, in which the probability of extreme events is excessively small, much smaller than what is observed on the market, as in Figure 1.1.2, where we can see three peaks between 6.5 and 7 σ, thus leading to an underestimation of the risk.

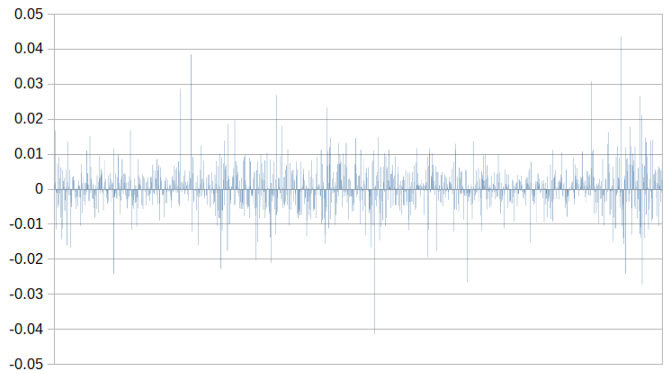

Figure 1.1.2: The daily log-returns of the Intel Corporation stock between June 2013 and May 2018. The standard deviation of the dataset is $\sigma = 6.35 \cdot 10^{-3}$.

These flaws led to the development of extensions of the BS model which are more flexible, thus allowing to better fit the market data. These models unfortunately induce new difficulties in the pricing and hedging of financial options, often requiring computationally expensive procedures.

Local and stochastic volatility

The first natural extension of the BS model is to allow the volatility to vary with time, that is

$$dS_t = S_t \, \sigma_t \, dW_t \,, \tag{1.1.4}$$

for some possibly stochastic volatility σ_t. Classically, one speaks about local volatility if the volatility $\sigma_t = \sigma(t, S_t)$ depends only on the price process and about stochastic volatility if σ_t depends on a new source of risk such as another Brownian motion. The class of local volatility models allows some flexibility to model the volatility smile, without introducing non-traded sources of risk, thus allowing to hedge financial options with the underlying stocks. Among this class, we find the popular "Constant elasticity of variance" (CEV) model (Cox, 1996) where $\alpha S_t^{\beta-1}$. In general, these models do not have closed-form formulas to price options. The "original" pricing method (Dupire, 1994) consists in solving numerically the PDE

$$\partial_t P(t,s) + \frac{1}{2}\sigma^2(t,s)\, s^2\, \partial_{ss} P(t,s) = 0 \,, \qquad P(T,s) = p(s) \,,$$

where $P(t,s) = \mathbb{E}^{\mathbb{Q}}\left[\, p(S_T) \,|\, S_t = s \,\right]$ is the price at time t of a derivative with payoff $p(S_T)$ if $S_t = s$. Later, in (Hagan and Woodward, 1999), the authors calculate an expansion of the implied volatility of options using singular perturbation techniques.

Even though fairly tractable, market evidence showed that local volatility models have limitations when it comes to replicating the behaviour of market data. Indeed, when comparing the evolution of the S&P500[1] index and the VIX[2] index (see Figure 1.1.3), we observe that they are not fully correlated, thus leading to the introduction of a new source of risk.

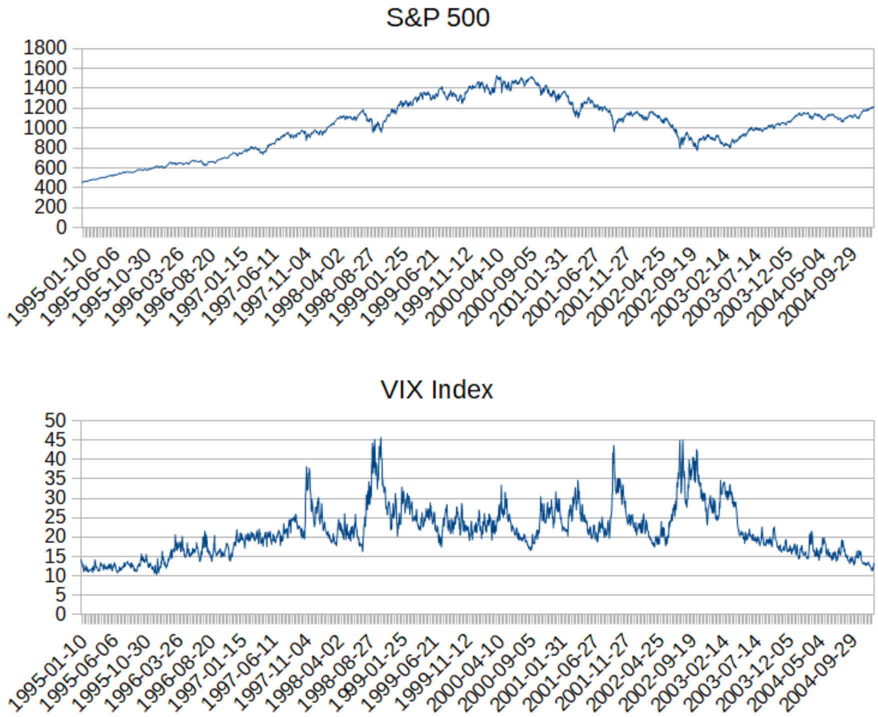

Figure 1.1.3: The daily data of the S&P 500 and the VIX index between 1995 and 2004.

Some of the noticeable examples of stochastic volatility models are: The Stein-Stein model (Stein and Stein, 1991), where σ_t is modelled as a mean-reverting Ornstein-Uhlenbeck process (Ornstein and Uhlenbeck, 1930), the Heston model (Heston, 1993), where σ_t^2 is modelled as a CIR process (Cox et al., 1985), the 3/2 stochastic volatility model (Carr and Sun, 2007), which uses the interest rate model from (Ahn and Gao, 1999) for σ_t^2 and the SABR (Stochastic α, β, ρ) model (Hagan et al., 2002), where $\sigma_t = x_t\, S_t^{\beta}$ and x_t is a log-normal process. These models allow for different levels of flexibility and no general method is available for the pricing of options. In some of these models, such as the class of affine models characterized in (Duffie et al., 2003), the Fourier transform of the stock log-price is known. The pricing problem is therefore generally tackled by inverting the Fourier transform (see also (Carr

[1] Index based on 500 large American companies.

[2] Volatility index based on the S&P 500 (SPX), that estimates expected volatility by aggregating, weighted prices of SPX puts and calls over a wide range of strikes. The detailed calculation methodology can be found in (Chicago Board Options Exchange, CBOE, 2018).

and Madan, 1999)). Others however require the use of asymptotic formulas or heavy numerical procedures to calculate option prices.

Jumps

Local and stochastic volatility models remedy the BS inability to model the volatility smile.

One of the main properties of the Brownian driven diffusion models is that their trajectories are continuous. However, structurally, market stock prices do not evolve continuously. They instead "jump" from one value to another, sometimes with a large difference. In addition, when time tends to zero, continuous local and stochastic volatility models tend to behave like the BS model, since the volatility does not have the possibility to vary a lot from its initial position. They have therefore limitations when it comes to pricing options with short maturities (see (Carr and Wu, 2003)). Such concerns lead, in many financial models, to the introduction of jumps. We address the reader to (Cont and Tankov, 2004) for a general reference on jump processes in finance. Following an idea of (Mandelbrot, 1963), several elements of the class of Lévy processes were proposed in the literature to model stock log-prices, such as the Merton model (Merton, 1976) with Gaussian jumps, the Kou model (Kou, 2002) with "double exponential" jumps, the Bates model (Bates, 1996) with jumping stochastic volatility, with finite jump activity, and the Normal-inverse-Gaussian model (Barndorff-Nielsen, 1997) and the CGMY model (Carr et al., 2002), with infinite jump activity. In Lévy models, when the jump activity is finite, the pricing can be done by conditioning on the jumps, whereas, when the jump activity is infinite, one classically uses the fact that the Fourier transform is known (Lévy-Khintchine formula) to use the fast Fourier transform method (Carr and Madan, 1999).

1.1.3 Non-equity derivatives

We have so far discussed the pricing of stock options. In this thesis, we also discuss the pricing of options that are not based on equity but also on interest rates and realized variance.

Interest rate derivatives and the Libor Market Model

According to the Bank of International Settlements (BIS), derivatives on interest rates amount for more than three quarters of the total volume of derivatives traded in the OTC market. The development of efficient pricing methods for such products is therefore a very important issue for financial institutions, but also for insurance companies and pension funds. Indeed, these companies structurally have to pay deterministic cash flows at future times and interest rate derivatives are a useful tool to hedge interest rate risk. Besides interest rate swaps, that have a linear payoff, the most common interest rate derivatives are caplets and swaptions. In order to model interest rates, several approaches were proposed in the last 50 years. Starting in the 1970's, the first approach models the "short rate", i.e. the interest rate that one needs to pay to borrow cash on the markets. Among the processes that are used to model the short rate, we cite the Vasicek model (Vasicek, 1977) and the CIR model (Cox et al., 1985). The second approach proposed by (Heath et al., 1992) models the instantaneous forward interest rate, i.e. the current interest rate applicable to a future transaction. The third approach called "Libor market model" (LMM) was developed in (Brace et al., 1997) as a justification of the Black formula, a version of the BS formula, that was commonly used by practitioners to price caplets, which can be seen as call options on the Libor rates.

Consider a set of times $0 \leq T_0 < ... < T_n$ for which zero-coupon bonds $B_t(T_k)$ are available. Denoting $\delta_j = T_j - T_{j-1}$, the forward Libor rate for the period $[T_{j-1}, T_j]$ is

$$L_t^j = \frac{1}{\delta_j}\left(\frac{B_t(T_{j-1})}{B_t(T_j)} - 1\right) .$$

Since $\frac{B_t(T_{j-1})}{B_t(T_j)}$ is the price of a traded asset $B_t(T_{j-1})$ discounted by $B_t(T_j)$, the theory of absence of arbitrage states that L_t^j should be a martingale under the measure $\mathbb{Q}^{T_j}$ that uses $B_t(T_j)$ as numéraire. The LMM therefore models L_t^j as a log-normal BS process under $\mathbb{Q}^{T_j}$. In the LMM, the price of caplets is calculated using the Black formula, thus justifying the market practice. Swaptions however cannot be priced exactly, but freezing some parameters provides a fairly accurate "Black-type" closed-form pricing approximation. For an overview on the LMM, we refer the reader to (Brigo and Mercurio, 2001).

As in the case of stocks, following the development of the LMM, several authors have enhanced the log-normal LMM with stochastic volatility (Gatarek, 2003; Piterbarg, 2003; Hagan and Lesniewski, 2008) and with jumps (Glasserman and Kou, 2003; Eberlein and Özkan, 2005). These modifications naturally induce new complications in the option pricing problems and give rise to an important literature. In (Eberlein and Kluge, 2007), the authors tackle the calibration of the Lévy Libor Market model (LLMM) of (Eberlein and Özkan, 2005) and several papers (Eberlein and Özkan, 2005; Kluge, 2005; Belomestny and Schoenmakers, 2011) propose methods to compute numerically caplet prices using Fourier transform inversion.

Worth noting are also the other challenges, such as negative interest rates and multi-curve modelling, that have appeared in the interest rate markets after the sub-prime crisis of 2008. We do not discuss those problems in the scope of this thesis and refer the reader to (Grbac and Runggaldier, 2015).

Realized variance and volatility derivatives

The trading of realized variance supposedly first happened at UBS in 1993, but really took off only in 1998, probably due to the elevated volatilities observed this year, (see (Carr and Lee, 2009)). Back then, the market offered mostly variance and volatility swaps. It was only in 2005, that the market of derivatives started to offer a wider range of realized-variance-based derivatives, such as options on realized variance. An option on realized variance with strike K is a non-linear derivative with payoff

$$h(Z) = \left(\frac{1}{T}\int_0^T Z_t^2\, dt - K\right)_+ \tag{1.1.5}$$

where Z_t is an instantaneous volatility process. The calculation of the expectation of (1.1.5) is generally a complicated task, as (1.1.5) is highly non-linear and the distribution of the squared time-integral of Z is generally unknown. A number of papers have been written on the subject, in particular (Carr et al., 2005) for pure jump processes, (Sepp, 2008) for the Heston model with jumps, (Kallsen et al., 2011) in an affine setting, (Drimus, 2012) for the 3/2 model and (Keller-Ressel and Muhle-Karbe, 2013) for exponential Lévy models, almost always using Laplace/Fourier transform methods. In particular, (Keller-Ressel and Muhle-Karbe, 2013) discuss the difference between the discrete and the continuous versions of realized variance. A non-parametric approach to price volatility derivatives is proposed in (Carr and Lee, 2008). Based on the replication of the derivative with a portfolio of the stock and vanilla options, the method is exact when asset and volatility are not correlated and is “immune” against non-zero correlation at first order. This is also discussed in (Henry-Labordere, 2017).

1.1.4 Pricing options with asymptotic methods

In general, as soon as we depart from the BS model, we cannot use closed-form formulas, thus the need to find alternative approaches. In some models, pricing is done through numerical integration of a known function, in some others, by solving numerically partial (integro-)differential equations. However, these methods are sometimes insufficient, in particular, when pricing path dependent options, when the dimension is large or when we need to compute a large number of prices, as in model calibration procedures. Indeed, in such cases, the amount of required computation can become too large to be executed in a reasonable time. When this is the case, the use of asymptotic methods is often a good solution to obtain accurate approximations of option prices in a decent time. Below, we review some of the important asymptotic methods.

Perturbation of differential equations

The perturbation of differential equation is a classical method, widely used in physics and engineering, which consists in writing an unsolvable differential equation as a perturbation of a solvable differential equation in order to approximate the solutions of the unsolvable differential equations by a truncated power series of the perturbation parameter. Perturbations methods are of two kinds, regular and singular.

Assume that $f_\lambda(x)$ is the solution of a differential equation

$$\mathcal{A}_\lambda f_\lambda(x) = 0\,, \qquad f_\lambda(a) = h\,, \tag{1.1.6}$$

where $\mathcal{A}_\lambda$ is an operator and $0 \leq \lambda \leq \bar{\lambda} \ll 1$ is a parameter. Classically, the problem of interest corresponds to the case with $\lambda = \bar{\lambda}$, whereas the case $\lambda = 0$ is a simple problem whose solution is explicitly known. Assume that we can write

$$f_\lambda(x) = \sum_{j=0}^{\infty} \lambda^j\, f^j(x) \qquad \text{and} \qquad \mathcal{A}_\lambda = \sum_{j=0}^{\infty} \lambda^j\, \mathcal{A}^j\,.$$

Then (1.1.6) becomes

$$\sum_{k=0}^{\infty} \lambda^k \left(\sum_{j=0}^{k} \mathcal{A}^j f^{k-j}(x) \right) = 0\,, \qquad f_\lambda(a) = h\,.$$

Hence

$$\mathcal{A}^0 f^0(x) = 0\,, \qquad f^0(a) = h \tag{1.1.7}$$

and

$$\mathcal{A}^0 f^k(x) = -\sum_{j=1}^{k} \mathcal{A}^j f^{k-j}\,, \qquad f^k(a) = 0\,, \quad k \geq 1\,. \tag{1.1.8}$$

Equation (1.1.7) corresponds to (1.1.6) for $\lambda = 0$, whose solution is explicitly known and (1.1.8) is of the same "type" as (1.1.7). Provided (1.1.8) can be solved up to order n, it is therefore straightforward to truncate the f_λ series to obtain an accurate approximation of $f_{\bar{\lambda}}$.

Many differential equations arise in finance. As an example, the price P of a European financial derivative with payoff $h(X_T)$ is generally expressed as the conditional expectation

$$P(t,x) = \mathbb{E}^{\mathbb{Q}}\left[h(X_T)\,|\,X_t = x\,\right]$$

under a certain probability measure $\mathbb{Q}$. By martingale property of $P(t, X_t)$, the price function verifies the Kolmogorov backward equation

$$0 = \partial_t P(t,x) + \mathcal{L}_t P(t,x)\,, \tag{1.1.9}$$

where $\mathcal{L}_t$ is the infinitesimal generator of X_t, under terminal condition $P(T, x) = h(x)$. When pricing options under a model where closed-form formulas do not exist, a possible approach is therefore to write the complicated model as a perturbation of a simple one, such as the BS model. The approach of regular perturbations is used in several papers in the literature. In (Sircar and Papanicolaou, 1999), the authors use regular perturbations to obtain an asymptotic expansion of call prices when the stochastic volatility varies quickly and (Lee, 2001) extends this methodology to expand the implied volatility of vanilla options for small variations of the volatility and slow variation of the variance. Regular perturbations techniques are used to obtain asymptotic expansions for option prices in local volatility models (Benhamou et al., 2008) and in the time-dependent Heston model (Benhamou et al., 2010), as perturbations of the BS model, and in jump-diffusion models (Benhamou et al., 2009), as a perturbation of the Merton model (Merton, 1976), combining the approach with Malliavin calculus to calculate the coefficients of the expansions. In (Jacquier and Lorig, 2015), a regular perturbation approach is used to obtain an expansion of implied volatility of vanilla options in models for which the characteristic function is known explicitly.

When f_λ can be expressed as an infinite series on the whole domain, regular perturbations are an effective way to approximate the solution of differential equations. When this is not possible however, some scaling is required in order to obtain expansions. These methods are called "singular perturbations" and are typically useful when the small parameter multiplies the highest derivative. We present, as an example, the case developed in (Widdicks et al., 2005), where the authors calculate an expansion of the BS price when volatility is small, using singular perturbations. Let $P(t, s; \sigma)$, be the price of a call option with maturity T and strike K, on a BS stock with volatility σ. Then $P(t, s)$ verifies the PDE

$$\partial_t P(t,s) + \frac{1}{2}\sigma^2 s^2 \partial_{ss} P(t,s) = 0 \,, \qquad P(T,s) = (s-K)_+ \,.$$

Since, when setting $\sigma = 0$, the second order derivative vanishes and remains only $P(t, s) = (s - K)_+$, let us define the scaling

$$\hat{P}(t,\hat{s}) = \frac{P(t,s)}{\sigma} \,, \qquad \text{where } \hat{s} = \frac{s-K}{\sigma} \,.$$

A simple change of variable in the BS equation then shows that $\hat{P}(t, \hat{s})$ verifies the PDE

$$\mathcal{A}_\sigma \hat{P}(t,\hat{s}) = 0 \,, \qquad \hat{P}(T,\hat{s}) = \hat{s}_+ \,,$$

where

$$\mathcal{A}_\sigma = \partial_t + \frac{1}{2} K^2 \partial_{\hat{s}\hat{s}} + \sigma\, K \hat{s}\, \partial_{\hat{s}\hat{s}} + \frac{1}{2}\sigma^2 \hat{s}^2 \partial_{\hat{s}\hat{s}} \,.$$

Since the second order term no longer vanishes when $\sigma = 0$, $\hat{P}(t, \hat{s})$ can now simply be expanded using regular perturbation techniques.

In the last 20 years, singular perturbations have been widely used to price options. In (Fouque et al., 2000; Fouque et al., 2003; Papanicolaou et al., 2003), the authors use singular perturbation to expand the price of options in the case of fast mean-reverting stochastic volatility model. Singular perturbations are also used in (Hagan and Woodward, 1999; Hagan et al., 2002) to develop asymptotics for the implied volatility of vanilla options in the famous SABR model. In addition to the presented example, (Widdicks et al., 2005) develop asymptotics for the price of American and barrier options. An expansion of the price of options in diffusive stochastic volatility models is obtained in (Antonelli and Scarlatti, 2009) as a power series of the correlation between the Brownian motions driving the asset and the volatility.

The applications of perturbation theory in finance are naturally not limited to option pricing, but other applications go beyond the scope of this thesis. We refer the reader to the introduction of (Černỳ et al., 2013) for an overview of other financial applications of perturbation theory, namely hedging and portfolio optimization.

Asymptotic expansions of the transition density

The first result aiming to understand the asymptotic behaviour of the transition density $p(s,t,x,y)$ of a multidimensional diffusion process

$$X_t = x + \int_s^t b(X_u)\,du + \int_s^t \sigma(X_u)\,dW_u$$

goes back to (Varadhan, 1967a) and (Varadhan, 1967b), who proved that, under uniform ellipticity condition, the solution $p(t,x,y)$ of the heat equation with variable coefficients

$$\partial_t p = \frac{1}{2}\sum_{i,j=1}^{n} \sigma_{ij}(x)\partial_{x_i x_j} p$$

with boundary condition $p(t,x,y) \to \delta_x(y)$ as $t \to 0$ satisfies

$$\lim_{t\to 0} 2t\,\log p(t,x,y) = -d^2(x,y)\,,$$

where $d(x,y)$ is the Riemannian distance induced by σ. This result was generalized to hypoelliptic diffusions under strong Hörmander condition by (Léandre, 1987). Asymptotic expansions of the type

$$p(s,t,x,y) = (t-s)^{-n/2}\, e^{-\frac{d_s^2(x,y)}{2(t-s)}} \left(\sum_{j=0}^{N} \alpha_j\,(t-s)^j + \mathcal{O}\big((t-s)^{N+1}\big) \right)$$

have been obtained by (Molčanov, 1975) and (Azencott, 1984) for elliptic operators. In (Bismut, 1984), a first Malliavin calculus approach to this problem is discussed to obtain results under the weaker than elliptic "H2" hypothesis. In (Watanabe, 1987), the author develops a theory of distributions on the Wiener space, thus allowing to study rigorously the density of a functional $F(W)$ of Brownian motion as $f(x) = \mathbb{E}(\delta_x F(W))$, where δ_x is a Dirac spike at x. Taking $F = F(\epsilon, W)$, where ϵ is a small parameter, this setting allows to study asymptotics of the law $f^\epsilon(x)$ of $F(\epsilon, W)$ as $f^\epsilon(x) = \mathbb{E}(\delta_x F(\epsilon, W))$ obtaining the transition density of diffusion processes as a particular case. Following (Azencott, 1984) and (Bismut, 1984), Ben Arous obtains in (Ben Arous, 1988a; Ben Arous, 1988b) asymptotic expansions for the density of hypoelliptic diffusion processes using the Laplace method on Wiener spaces and the Mallivin calculus. Later, (Deuschel et al., 2014a) consider the marginal diffusion of the density f^ϵ of the l-dimensional marginal $(X^\epsilon_{1t}, ..., X^\epsilon_{lt})$ of the n-dimensional diffusion process

$$dX^\epsilon_t = b(\epsilon, X^\epsilon_t)\, dt + \epsilon\, \sigma(X^\epsilon_t)\, dW_t\,,$$

where $b(\epsilon, X^\epsilon_t)$ is typically of the form $b(X^\epsilon_t)$ or $\epsilon^2\, b(X^\epsilon_t)$ and obtain the result

$$f^\epsilon = e^{-c_1/\epsilon^2}\, e^{c_2/\epsilon}\, \epsilon^{-l}\, \left(c_0 + \mathcal{O}(\epsilon)\right)\,, \qquad \text{as } \epsilon \to 0\,.$$

We finally cite a recent paper (Frikha and Kohatsu-Higa, 2016) which combines the parametrix technique with Malliavin calculus to prove an asymptotic expansion of the density of diffusions under weak Hörmander condition, assuming that b and σ are smooth functions with bounded derivatives of all orders and that X satisfies an integrability condition.

Around those results, many financial pricing methods were developed. Based on (Ben Arous, 1988a), (Bayer and Laurence, 2013a; Bayer and Laurence, 2013b) obtain an asymptotic expansion for the implied volatility of basket options in local volatility models out of and at the money. A small time expansion of implied volatility of options in stochastic volatility models is calculated in (Henry-Labordère, 2008). In (Deuschel et al., 2014b), the asymptotic implied volatility result for the Stein-Stein model obtained in (Gulisashvili and Stein, 2010) in the uncorrelated case is extended to the correlated case using the results of (Deuschel et al., 2014a).

In (Takahashi, 1999), the author expands the density of a diffusion, where the volatility is multiplied by a small parameter, as a "Gaussian" power series of the small parameter. The expansion of the density obtained in (Watanabe, 1987) is then integrated in order to obtain a small volatility expansion of the price of derivatives in local volatility models. The author calculates explicitly the expansion for vanilla and Asian options. The same methodology is used in (Kunitomo and Takahashi, 1995; Kunitomo and Takahashi, 2001) to price interest rate derivatives in the HJM framework (Heath et al., 1992). The authors provide an explicit expansion of the price of swaptions and Asian (interest rate) options. The validity of the latter expansions is finally proved in (Kunitomo and Takahashi, 2003).

A later series of papers, among which we can cite (Shiraya and Takahashi, 2011; Shiraya and Takahashi, 2014; Shiraya and Takahashi, 2016; Shiraya and Takahashi, 2017a; Shiraya and Takahashi, 2017b), extends these calculations to certain types of derivatives under stochastic volatility and jump diffusion models.

Methods based on large deviations and geometry

The methods based on large deviations are in the same spirit as the ones based on density expansions. The Freidlin-Wentzell theory (Freidlin and Wentzell, 2012) proves, under certain hypotheses, that the diffusion

$$dX_t^\epsilon = b(X_t^\epsilon)\,dt + \epsilon\,\sigma(X_t^\epsilon)\,dW_t\,, \qquad X_0^\epsilon = x\,,$$

where $(W_t)_{t\leq T}$ is Brownian motion, satisfies

$$\epsilon\,\log \mathbb{P}\left(X_T^\epsilon \in A\right) \sim I(A)\,, \qquad \text{as } \epsilon \to 0\,,$$

for the point-set distance

$$I(A) = \inf_{x_0=x,\,x_T\in A} \frac{1}{2}\int_0^T \left(\frac{\dot{x}_t - b(x_t)}{\sigma(x_t)}\right)^2 dt\,.$$

Based on the numerous versions and generalizations of this result, a wide range of pricing methods was developed.

In (Berestycki et al., 2004), the authors show, using large deviations, that the implied volatility in stochastic volatility models can be expressed as a function of a distance function connected to a Hamilton-Jacobi equation. Asymptotic expansions for the Heston implied volatility are calculed in (Forde and Jacquier, 2009) for the short-time case and in (Forde and Jacquier, 2011a; Jacquier and Mijatović, 2014) for the long-time case. In (Forde and Jacquier, 2011b), the authors use the Freidlin-Wentzell theory to add some rigour to the implied volatility expansions obtained in (Henry-Labordère, 2008) and (Paulot, 2015) for stochastic volatility models. In (Jacquier et al., 2013), a large deviation principle is shown for affine stochastic volatility models and asymptotics for the implied volatility are obtained.

Monte-Carlo methods, large deviations and optimal sampling

When one wishes to compute the expectation of a random variable, if no simpler method is available, the "last chance" solution is the use of Monte-Carlo methods. These methods were developed in the 1940s and popularised along the years as computer power increased. Let ξ be a square integrable random variable. Our goal is to calculate $p := \mathbb{E}(\xi)$. Assume that we can simulate efficiently from the distribution of ξ and let $\xi^{(1)}, ..., \xi^{(n)}$ be independent and identically distributed realisations of ξ. We define the Monte-Carlo estimator $\hat{p}$ of p as

$$\hat{p} = \frac{1}{n}\sum_{k=1}^{n} \xi^{(k)} .$$

The Central Limit Theorem stipulates that $\sqrt{n}(\hat{p} - p)$ converges in law to a Gaussian random variable with law $\mathbb{N}(0, \mathbb{V}\mathrm{ar}(\xi))$. Hence, when n is large enough, $\hat{p}$ behaves like a $\mathbb{N}(p,\, n^{-1}\,\mathbb{V}\mathrm{ar}(\xi))$ random variable and therefore, for a fixed probability α, we can approximate a confidence interval for p at level α by

$$I_\alpha = \left[\hat{p} + n^{-1/2}\hat{\sigma}\, q_{\alpha/2},\, \hat{p} + n^{-1/2}\hat{\sigma}\, q_{1-\alpha/2}\right] ,$$

where $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{k=1}^{n}\left(\xi^{(k)} - \hat{p}\right)^2$ and q_α is the quantile at the level α of a standard Gaussian random variable. By taking $\xi = P(X)$ as the payoff P of a derivative on a random log-price trajectory X, Monte-Carlo methods then become a powerful option pricing tool. Indeed, they allow to compute numerical estimates of the price of almost any derivative no matter how complicated its payoff is. The biggest drawback of Monte-Carlo methods is that, since the standard deviation of $\hat{p}$ is proportional to $n^{-1/2}$, one needs a large number of simulations to achieve the desired accuracy. To overcome this problem, a possible approach is, instead of simulating the path processes $X^{(1)}, ..., X^{(N)}$ under the actual risk-neutral probability measure $\mathbb{P}$, to simulate the path processes $X^{(1,\mathbb{Q})}, ..., X^{(N,\mathbb{Q})}$ under an equivalent probability measure $\mathbb{Q}$. The estimator of the price of an option with payoff $P(\cdot)$ on an underlying X then becomes

$$\hat{p}_{\mathbb{Q}} := \frac{1}{N}\sum_{k=1}^{N} P\left(X^{(k,\mathbb{Q})}\right) \frac{d\mathbb{P}}{d\mathbb{Q}}\left(X^{(k,\mathbb{Q})}\right) .$$

The estimators of the class $\{\hat{p}_{\mathbb{Q}} \,:\, \mathbb{Q} \sim \mathbb{P}\}$ are unbiased and have variance

$$\begin{aligned} \mathbb{V}\mathrm{ar}^{\mathbb{Q}}\left[\hat{p}_{\mathbb{Q}}\right] &= \frac{1}{N}\,\mathbb{V}\mathrm{ar}^{\mathbb{Q}}\left[P(X)\,\frac{d\mathbb{P}}{d\mathbb{Q}}(X)\right] \\ &= \frac{1}{N}\left(\mathbb{E}\left[P^2(X)\,\frac{d\mathbb{P}}{d\mathbb{Q}}(X)\right] - \mathbb{E}^2\left[P(X)\right]\right) , \end{aligned}$$

thus the interest to find the measure $\mathbb{Q} \sim \mathbb{P}$ that minimizes $\mathbb{E}\left[P^2(X)\,\frac{d\mathbb{P}}{d\mathbb{Q}}(X)\right]$. Since this minimization problem is generally complex, some authors, starting with (Siegmund, 1976), considered the minimization of an asymptotic version of $\mathbb{E}\left[P^2(X)\,\frac{d\mathbb{P}}{d\mathbb{Q}}(X)\right]$ based on the theory of large deviations. Following this idea, (Dupuis and Wang, 2004) discuss the use of adaptive control-theoretic measure changes. Adaptive schemes are further discussed in the context of Gaussian functionals in (Jourdain and Lelong, 2009). In (Guasoni and Robertson, 2008), the authors combine Laplace method with Schilder's Theorem and Varadhan's Lemma (Dembo and Zeitouni, 1998, Thms. 5.2.3

and 4.3.1) to calculate explicitly the small noise large deviation proxy to the optimal measure to price Asian options when the underlying behaves as a geometric Brownian motion. Then, (Robertson, 2010) generalizes the Freidlin-Wentzell theory (Dembo and Zeitouni, 1998, Section 5.6) proving a trajectorial small-noise large deviations principle for a diffusive stochastic volatility model, thus extending the methodology of (Guasoni and Robertson, 2008) to the stochastic volatility framework. Finally, in a more recent work, (Genin and Tankov, 2016) combine a similar approach with the long-time large deviations results of (Léonard, 2000) for processes with independent increments to develop an optimal sampling method for path dependent options when the underlying is modelled as an exponential Lévy process.

1.2 Summary of the thesis

In this thesis, we tackle four financial pricing problems using asymptotic approaches.

In Chapter 2, we consider the problem of pricing potentially path-dependent derivatives via Monte-Carlo methods when the stock price is given by an affine stochastic volatility model. We start by proving a large-time large deviations principle for the paths of such processes under strong but verifiable hypotheses. We then consider the class of measure changes defined by the time-dependent Esscher transform. By simulating under a measure change, the variance of the Monte-Carlo estimator changes. We write the variance reduction problem, which happens to be unsolvable explicitly. Using the large deviations result, we formulate an asymptotic approximation of the variance reduction problem that we solve to compute an asymptotically optimal measure change and obtain a significant variance reduction when using Monte-Carlo simulations. We test the method on the Heston model with and without jumps to demonstrate its numerical efficiency.

In Chapter 3, we consider the problem of pricing European basket options when stock prices are modelled as a Wishart stochastic volatility process, which is an n-dimensional generalization of the very popular Heston model, using Monte-Carlo methods. Following the approach of Chapter 2, we start by proving that the Wishart stochastic volatility process satisfies a large deviations principle when time tends to infinity. We then write the Monte-Carlo variance minimization problem induced by the Esscher transform class of measure changes and use the large deviation principle to write a solvable approximate minimization problem. We finally test the method numerically and see that the variance reduction obtained allows a significant reduction of the required number of Monte-Carlo simulations. In addition and independently, we use the large deviation result to calculate the asymptotic implied volatility of basket options when time is long and test the convergence numerically using Monte-Carlo simulations.

In Chapter 4, we study the pricing of options on realized variance when the instantaneous volatility is modelled as a diffusion process with general drift and constant volatility. We calculate a first order expansion of the density of the integral of the squared volatility process using recent results on asymptotic expansions for marginal densities of hypo-elliptic diffusions. We then integrate the option payoff against the density to expand the option price and compare the result to the expansion of the BS price to calculate an expansion of the associated BS implied volatility.

In Chapter 5, we tackle the issue of pricing interest rates derivatives under the Lévy Libor market model, that generalises the popular log-normal Libor market model by introducing jumps. We expand the generator of the Lévy Libor market model around the generator of the log-normal Libor market model and use the Feynman-Kac formula to calculate an explicit asymptotic expansion of the price of "European-type" options at second order. We compare the numerical results with an accurate Monte-Carlo simulation to demonstrate the efficiency of the method.

1.3 The main results of the thesis

Let us now present in more detail the main results of this thesis.

1.3.1 Affine stochastic volatility models, large deviations and optimal sampling (Chapters 2 and 3)

The class of affine models, whose characterization can be found in (Duffie et al., 2003) is a wide class of models whose Laplace transform is known up to the resolution of an ODE. A particularly interesting subclass of affine models is the class of affine stochastic volatility (ASV) models (Keller-Ressel, 2011). It was developed as a generalization of the very popular Heston model (Heston, 1993) and combines the tractability of affine models with the flexibility of stochastic volatility models to reflect stylized facts observed on the markets. An ASV process $(X_t, V_t)_{t\geq 0}$, where X_t is the asset log-price and V_t is the instantaneous variance process, is a bivariate affine model whose Laplace transform is of the form

$$\mathbb{E}\left[e^{uX_t+wV_t}\,\middle|\,\mathcal{F}_s\right] = e^{\phi(t-s,u,w)+\psi(t-s,u,w)\,V_s+u\,X_s}\,,$$

where ϕ and ψ satisfy the generalized Riccati equations

$$\partial_t\phi(t,u,w) = F(u,\psi(t,u,w))\,, \qquad \phi(0,u,w)=0 \tag{1.3.1a}$$
$$\partial_t\psi(t,u,w) = R(u,\psi(t,u,w))\,, \qquad \psi(0,u,w)=w\,, \tag{1.3.1b}$$

and F and R have Lévy-Khintchine forms. This feature granted ASV models an increasing popularity in the industry, owing to the fact that the prices of single asset vanilla options can be efficiently computed using Laplace transform inversion methods (Carr and Madan, 1999; Duffie et al., 2000).

Pricing path-dependent derivatives in ASV models

Pricing path-dependent derivatives in ASV models is a more complicated task and often requires Monte-Carlo simulations. We consider the pricing of such derivatives and, in particular, Asian option with payoff

$$\left(\frac{1}{T}\int_0^T S_t\,dt - K\right)_+ .$$

We propose an optimal sampling method based on an asymptotic large deviations based proxy. A point (single-date) large deviation principle for ASVM was already shown in (Jacquier et al., 2013), where the authors use their result to show that the asymptotic long-time implied volatility of a vanilla option with maturity T and strike K on a stock worth S_0 is

$$\sigma_\infty = \sqrt{2}\left(h^*\left(x\right)^{1/2} + \left(h^*\left(x\right) - x\right)^{1/2}\right)$$

where $x = T^{-1}\log(K/S_0)$ and $h^*(x) := \sup_{\theta\in\mathbb{R}}\{\theta x - h(\theta)\}$ is the convex dual of

$$h(\theta) := \lim_{t\to\infty} t^{-1}\log\mathbb{E}\left[e^{\theta\log(S_t)}\right] .$$

Since we want to price path dependent options however, a point large deviation property is insufficient for the scope of our work and we need to prove a trajectorial LDP for ASV models. Let us present the main results shown in Chapter 2.

Let $(X_t, V_t)_{t\geq 0}$ be an ASV process and let $I \subset \mathbb{R}$ be the set such that for every $u \in I$, (1.3.1b) admits a unique stable equilibrium $w(u)$. Let also $J \subset I$ be the domain of $h(u) = F(u, w(u))$.

Assumption 1. *The function h verifies the following condition.*

- *There exists $u < 0$, such that $h(u) < \infty$.*
- *$u \mapsto h(u)$ is essentially smooth.*

Assumption 2. *One of the following conditions is verified.*

1. *The interval support of F is $J = [u_-, u_+]$ and $w(u_-) = w(u_+)$.*
2. *For every $u \in \mathbb{R}$, (1.3.1b) has no unstable equilibrium.*

Theorem 1.3.1. *Let us define, for $\epsilon \in (0, 1]$, the process $X^\epsilon_t := \epsilon X_{t/\epsilon}$. If Assumptions 1 and 2 are verified, then $(X^\epsilon_t)_{0\leq t\leq T}$ satisfies a large deviations* *property on $\mathcal{F}([0,T],\mathbb{R})$ equipped with the topology of point-convergence, as $\epsilon \to 0$, with good rate function*

$$\Lambda^*(x) = \int_0^T h^*(\dot{x}_t^{ac})\,dt + \int_0^T \mathcal{H}\left(\frac{d\nu_t}{d\theta_t}\right) d\theta_t \,,$$

where

$$h^*(y) = \sup_{\theta\in J}\{\theta y - h(\theta)\}\ , \qquad \mathcal{H}(y) = \lim_{\epsilon\to 0}\epsilon\, h^*(y/\epsilon)\,,$$

$\dot{x}^{ac}$ is the derivative of the absolutely continuous part of x, ν_t is the singular component of dx_t with respect to dt and θ_t is any non-negative, finite, regular, $\mathbb{R}$-valued Borel measure, with respect to which ν_t is absolutely continuous.

We consider the class of measure changes $\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\int_0^t X_s\,d\theta_s}}{\mathbf{E}\left[e^{\int_0^t X_s\,d\theta_s}\right]}$, where θ is a finite signed measure on $[0,T]$ and wish to find the θ that minimizes

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}_\theta}\left(P(X)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right) = \mathbb{E}\left(P^2(X)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right) - \mathbb{E}^2\left(P(X)\right)\,,$$

in order to reduce the variance when pricing an option with possibly path-dependent payoff $P(X)$ using Monte-Carlo simulations. This problem is however not explicitly solvable. Denoting $H = \log P$,

$$\begin{aligned}&\lim_{\epsilon\to 0}\epsilon\log\mathbb{E}\left(P^2(X^\epsilon)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}(X^\epsilon)\right)\\ &\qquad = \sup_{x\in V_r}\left\{2H(x) - \int_0^T x_t d\theta_t - \Lambda^*(x)\right\} + \int_0^T h(\theta([t,T]))\,dt\,.\end{aligned}$$

where V_r is the set of trajectories $x:[0,t]\to\mathbb{R}$ with bounded variation. We therefore call *asymptotically optimal* a measure θ that minimises the right-hand side. A result by (Genin and Tankov, 2016) shows that, under technical hypothesis, for H concave and continuous on its domain with respect to the topology of pointwise convergence, we have

$$\begin{aligned}\inf_{\theta\in M}\sup_{x\in V_r}&\left\{2H(x) - \int_0^T x_t d\theta_t - \Lambda^*(x)\right\} + \int_0^T h(\theta([t,T]))\,dt\\ &= 2\inf_{\theta\in M}\left\{\hat{H}(\theta) + \int_0^T h(\theta([t,T]))\,dt\right\}\,,\end{aligned}$$

where

$$\hat{H}(\theta) = \sup_{x\in V_r}\left\{H(x) - \int_0^T x_t\,d\theta_t\right\}\,.$$

Furthermore, if θ^* minimises the left-hand side of the above equation, it also minimises the right-hand side. We can therefore find the asymptotically optimal measure by minimizing the right-hand side. We consider, in particular, the case of a (discretized) Asian option with payoff

$$P(X) = \left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{X_{t_j}}\right)_+ , \qquad t_j = \frac{Tj}{n} ,$$

in the Heston model

$$\begin{aligned} dX_t &= -\frac{1}{2} V_t \, dt + \sqrt{V_t} \, dW_t^1 , & X_0 = 0 \\ dV_t &= \lambda_t \left(\mu_t - V_t\right) dt + \zeta \sqrt{V_t} \, dW_t^2 , & V_0 = V_0 \\ d\left\langle W^1, W^2 \right\rangle_t &= \rho \, dt , \end{aligned} \tag{1.3.2}$$

where W is 2-dimensional correlated Brownian motion.

Proposition 1.3.2. *Consider the class* $\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\int_0^t X_s \, d\theta_s}}{\mathbf{E}\left[e^{\int_0^t X_s \, d\theta_s}\right]}$ *of measure changes. The asymptotically optimal measure in the discretized Asian option case is the* θ *supported on* $\{t_1, ..., t_n\}$ *that minimizes*

$$\log\left(\frac{K}{1-\Theta_1}\right) - \sum_{m=1}^{n} (\Theta_m - \Theta_{m+1}) \log\left(\frac{-(\Theta_m - \Theta_{m+1})}{1-\Theta_1} \frac{nK}{S_0}\right) + \frac{T}{n}\sum_{j=1}^{n} h\left(\Theta_j\right) , \tag{1.3.3}$$

where $\Theta_j = \theta([t_j, T])$*. Furthermore, under* $\mathbb{P}_\theta$*, the dynamics of the* $\mathbb{P}$*-Heston process* (1.3.2) *becomes*

$$\begin{aligned} dX_t &= \left(\Theta_{\tau_t} + \zeta\rho\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) - \frac{1}{2}\right) V_t \, dt + \sqrt{V_t} \, d\tilde{W}_t^1 , & X_0 = 0 \\ dV_t &= \tilde{\lambda}_t \left(\tilde{\mu}_t - V_t\right) dt + \zeta \sqrt{V_t} \, d\tilde{W}_t^2 , & V_0 = V_0 \\ d\left\langle \tilde{W}^1, \tilde{W}^2 \right\rangle_t &= \rho \, dt , \end{aligned}$$

where $\tilde{W}$ *is 2-dimensional correlated* $\mathbb{P}_\theta$*-Brownian motion, where* Ψ *is defined iteratively as*

$$\begin{aligned} \Psi\left(s, \Theta_j, ..., \Theta_n\right) &= \psi\left(s, \Theta_j, \Psi\left(t_{j+1} - t_j, \Theta_{j+1} ..., \Theta_n\right)\right) \\ \Psi\left(s\right) &= 0 \end{aligned}$$

and where, denoting $\tau_t = \inf\{s \in \tau \ : \ s \geq t\}$*,*

$$\tilde{\lambda}_t = \lambda - \zeta\Theta_{\tau_t}\rho - \zeta^2\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) \quad \text{and} \quad \tilde{\mu}_t = \frac{\lambda\mu}{\tilde{\lambda}_t} .$$

In Section 2.6, we suggest a dichotomy algorithm to minimize (1.3.3).

Pricing derivatives on multiple assets in the Wishart ASV model

Very early, the problem of pricing derivatives on multiple stocks, such as options on baskets, was formulated. The easiest solution naturally consists in modelling directly the price of the basket. However, the obtained prices are not consistent with the prices of single asset vanilla options and, in order for them to be, the marginals of the process used to model jointly the stock prices should be similar to the single asset model used to model the individual assets. Due to the popularity of the Heston model among practitioners, the Wishart stochastic volatility model was naturally developed.

A Wishart process is a matrix-valued symmetric non-negative definite stochastic process with dynamics

$$dX_t = \left(\alpha\, a^\top a + b\, X_t + X_t\, b^\top\right)\, dt + X_t^{1/2}\, dW_t\, a + a^\top (dW_t)^\top X_t^{1/2} \,,$$

where $(W_t)_{t\leq T}$ is a matrix Brownian motion. It was invented in (Bru, 1991) to model perturbations in biological data. Being a matrix version of the CIR process

$$dX_t = \lambda\,(\nu - X_t)\, dt + \sigma\, \sqrt{X_t}\, dW_t \,,$$

which is the instantaneous variance process of the Heston model, the Wishart model was quickly used to model the instantaneous covariance matrix in the multi-asset model developed in (Gourieroux and Sufana, 2004),

$$dS_t = \mathrm{Diag}(S_t)\, X_t^{1/2}\, dZ_t \,,$$

where $(Z_t)_{t\leq T}$ is an $\mathbb{R}^n$-dimensional Brownian motion. The Wishart stochastic volatility model[3] generalizes the Heston model to the multivariate setting. Indeed, by taking a, b and X_0 diagonal, S becomes a vector of independent Heston processes. The Wishart stochastic volatility model therefore allows to model each individual price process with a distribution that is close to the one obtained with the Heston model and, at the same time, to have a rich stochastic cross-correlation structure between the price processes. It is therefore a consistent yet flexible instrument to model simultaneously multiple assets. We consider the pricing of options on baskets, i.e. with payoff

$$\left(\sum_{k=1}^{n} S_T^k - K\right)_+ \,.$$

The Wishart stochastic volatility model is affine. Therefore, option pricing is traditionally done using Fourier inversion methods (see (Da Fonseca et al.,

[3]Note that the appellation "Wishart stochastic volatility model" also refers to the single asset models

$$dS_t = S_t \mathrm{Tr}\left[X_t^{1/2}\, dZ_t\right] \,,$$

where $(Z_t)_{t\leq T}$ is a matrix Brownian motion (see (Benabid et al., 2008)), which is a more flexible version of the Heston model.

2007)). However, since these methods require to compute integrals numerically, they suffer directly from the "curse of dimensionality" and become very quickly less tractable as the dimension increases. The use of Monte-Carlo methods then becomes necessary to price derivatives depending on numerous assets. Let us summarize the main results presented in Chapter 3.

Let (Y_t, X_t) be a Wishart stochastic volatility process, i.e. a $(\mathbb{R}^n, \mathcal{M}_n)$-valued process with dynamics

$$dY_t = \left(r\mathbf{1} - \frac{1}{2}\left((a^\top X_t\, a)_{11}\,,\, \dots\,,\, (a^\top X_t\, a)_{nn}\right)^\top\right) dt + a^\top X_t^{1/2}\, dZ_t$$
$$dX_t = (\alpha I_n + b\, X_t + X_t\, b)\; dt + X_t^{1/2}\, dW_t + (dW_t)^\top X_t^{1/2}\,, \qquad X_0 = x$$

where $r \in \mathbb{R}$, $\alpha > n-1$, a is invertible, $-b$ and x are symmetric and positive definite and $(Z_t)_{t\geq 0}$ and $(W_t)_{t\geq 0}$ are $\mathbb{R}^n$ and $\mathbb{R}^{n\times n}$-dimensional independent standard Brownian motions.

Theorem 1.3.3. *For $T > 0$, the family $(Y_T^\epsilon)_{\epsilon\in(0,1]}$ defined by $Y_t^\epsilon := \epsilon\, Y_{t/\epsilon^2}$ satisfies a large deviation property, when $\epsilon \to 0$ with good rate function*

$$\Lambda^*(y) = \sup_{\theta\in\mathbb{R}^n} \langle \theta, y\rangle - \Lambda(\theta)\,,$$

where

$$\Lambda(\theta) := \begin{cases} T\left(r\,\theta^\top\mathbf{1} - \frac{\alpha}{2}\, Tr\left[b + \phi^{1/2}(\theta)\right]\right) & \text{if } \theta \in \mathcal{U} \\ \infty & \text{if } \theta \notin \mathcal{U} \end{cases}\,,$$

for

$$\phi(\theta) := b^2 + a\,\left(Diag(\theta) - \theta\theta^\top\right) a^\top\,,$$

and

$$\mathcal{U} := \left\{\theta \in \mathbb{R}^n \,:\, \phi(\theta) \in \mathcal{S}_n^+\right\}\,.$$

The next result uses the large deviations property to characterize the limit behaviour of the implied volatility of basket options when maturity tends to infinity. Recall that the implied volatility of a basket option in a specific model is defined as the value of volatility such that the option price in this model equals the Black-Scholes option price with this volatility level obtained assuming that the entire basket follows the Black-Scholes model.

Proposition 1.3.4. *Let $\sigma(T,k)$ be the implied volatility associated with an option on a basket of stocks with payoff*

$$\left(\sum_{i=1}^n \omega_i S_T^i - e^k\right)_+\,.$$

Denote $x^* = \partial_\theta \Lambda(0)$ and $\tilde{x}_j^* = [\partial_\theta \Lambda(e_j)]_j$ for $j = 1, \ldots, n$ and let the constants $\beta^* = \max_j x_j^*$, $\hat{\beta}^* = \min_j \tilde{x}_j^*$ and $\tilde{\beta}^* = \max_j \tilde{x}_j^*$. Then using the scaling $y = k/T$, if $y \notin (\hat{\beta}^*, \tilde{\beta}^*)$, the limiting implied volatility is

$$\lim_{T\to\infty} \sigma(T, yT) = \sqrt{2}\left(\xi\sqrt{L(y)+y} + \eta\sqrt{L(y)}\right) ,$$

where

$$L(y) = \begin{cases} -y - \inf_{\lambda\in\mathbb{R}^n:\lambda_i\le 0, i=1,\ldots,n}\{\Lambda(\lambda) - y\langle\lambda, \mathbf{1}\rangle\} , & \text{if } y \le \beta^* , \\ -\max_{i,j=1,\ldots,n} \inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i + e_j)\} , & \text{if } y \ge \tilde{\beta}^* , \\ -y - \max_{i=1,\ldots,n} \inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i)\} , & \text{if } \beta^* < y < \hat{\beta}^* \end{cases}$$

and

$$(\xi, \eta) = \begin{cases} (-1, 1) , & \text{if } y \le \beta^* , \\ (1, -1) , & \text{if } y \ge \tilde{\beta}^* , \\ (1, 1) , & \text{if } \beta^* < y < \hat{\beta}^* . \end{cases}$$

In addition, if $y \in (\hat{\beta}^*, \tilde{\beta}^*)$,

$$\sigma(T, yT) = \sqrt{2y} + N^{-1}(C_\infty(y))\, T^{-1/2} + \mathcal{O}(T^{-1/2}) ,$$

as $T \to \infty$, where $C_\infty(y) = \sum_{i=1}^n \omega_i \mathbb{1}_{\tilde{x}_i^* > y}$ and N is the Gaussian distribution function.

We now discuss the variance reduction. We consider the class of measure changes $\frac{d\mathbb{P}_\theta}{d\mathbb{P}} := \frac{e^{\theta^\top Y_T}}{\mathbf{E}\left[e^{\theta^\top Y_T}\right]}$ and wish to find the parameter $\theta \in \mathbb{R}^n$ that minimizes

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}_\theta}\left(P(Y_T)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right) = \mathbb{E}\left(P^2(Y_T)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right) - \mathbb{E}(P(Y_T))^2$$

in order to minimize the variance when pricing an European option with payoff $P(Y_T)$ using Monte-Carlo simulations. Since this problem is analytically unsolvable, we use the fact that

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left(P^2(Y_T^\epsilon)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}(Y_T^\epsilon)\right) = \sup_{y\in\mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\} + \Lambda(\theta) ,$$

where $H = \log P$, to find an asymptotic proxy of the minimization problem. We say that θ is *asymptotically optimal* if it minimizes the right-hand side. The next result allows to compute the optimal measure without knowing Λ^*.

Theorem 1.3.5. *Let H be a concave upper semi-continuous function. Then*

$$\inf_{\theta\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\} + \Lambda(\theta) = 2\inf_{\theta\in\mathbb{R}^n}\left\{\hat{H}(\theta) + \Lambda(\theta)\right\} ,$$

where

$$\hat{H}(\theta) = \sup_{y\in\mathbb{R}^n}\left\{H(y) - \theta^\top y\right\} .$$

Furthermore, if θ^* minimizes the right-hand side, it also minimizes the left-hand side.

Proposition 1.3.6. *Under* $\frac{d\mathbb{P}_\theta}{d\mathbb{P}} := \frac{e^{\theta^\top Y_T}}{\mathbf{E}\left[e^{\theta^\top Y_T}\right]}$*, the process* (Y_t, X_t) *has dynamics*

$$\begin{aligned} dY_t &= \left(r\mathbf{1} - \frac{1}{2}\left((a^\top X_t\, a)_{11}\,,\,\dots,\,(a^\top X_t\, a)_{nn}\right)^\top + a^\top X_t\, a\, \theta\right) dt + a^\top X_t^{1/2}\, dZ_t^\theta \\ dX_t &= (\alpha I_n + (b + 2\,\gamma_\theta(T-t))X_t + X_t(b + 2\,\gamma_\theta(T-t)))\; dt \\ &\qquad\qquad\qquad + X_t^{1/2}\, dW_t^\theta + (dW_t^\theta)^\top X_t^{1/2}\,, \qquad X_0 = x\,, \end{aligned}$$

where $\left(Z_t^\theta\right)_{t\geq 0}$ *and* $\left(W_t^\theta\right)_{t\geq 0}$ *are* $\mathbb{R}^n$ *and* $\mathbb{R}^{n\times n}$*-dimensional independent standard* $\mathbb{P}_\theta$*-Brownian motions and* $\gamma_\theta(t)$ *is the solution of explicitly solvable matrix Riccati equations. Let*

$$P(Y_T) = \left(K - \sum_{j=1}^n \omega_j\, e^{Y_T^j}\right)_+\,,$$

where $\omega_j > 0$*, be the payoff of a put option on a basket of stocks. The asymptotically optimal* $\theta \in \mathbb{R}^n$ *is the one that minimizes* $\hat{H}(\theta) + \Lambda(\theta)$*, where*

$$\hat{H}(\theta) = \left(1 - \sum_{j=1}^n \theta_j\right)\left(\log(K) - \log\left(1 - \sum_{j=1}^n \theta_j\right)_+\right) - \sum_{j=1}^n \theta_j \log(-\theta_j/\omega_j)\,.$$

1.3.2 Options on realized variance and density expansion (Chapter 4)

We consider the problem of pricing options on realized variance when volatility is modelled as a diffusion process with general drift and constant diffusion coefficient. Define for every $\epsilon \in (0,1]$, the joint process $(Y_t^\epsilon, Z_t^\epsilon)_{t\in[0,T]}$ of the integrated variance and the instantaneous volatility

$$\begin{aligned} dY_t^\epsilon &= g(Z_t^\epsilon)\, dt \\ dZ_t^\epsilon &= \epsilon^2 b(Z_t^\epsilon)\, dt + \epsilon\, c\, dW_t\,, \end{aligned} \tag{1.3.4}$$

where $(W_t)_{t\leq T}$ is standard Brownian motion, with initial value $Y_0^\epsilon = 0$ and $Z_0^\epsilon = z_0 > 0$ and where

$$g(z) = \frac{z^2\, e^{-\frac{1}{R+1-|z|}}\mathbb{1}_{\{|z|<R+1\}} + (R+1)^2\, e^{-\frac{1}{|z|-R}}\mathbb{1}_{\{|z|>R\}}}{e^{-\frac{1}{R+1-|z|}}\mathbb{1}_{\{|z|<R+1\}} + e^{-\frac{1}{|z|-R}}\mathbb{1}_{\{|z|>R\}}}\,,$$

for R arbitrarily large, is a bounded version of $z \mapsto z^2$ with bounded derivatives of all orders. We calculate the asymptotic expansion of the density of Y_T^ϵ basing our approach on the results presented in (Deuschel et al., 2014a).

Theorem 1.3.7. *The process Y_t^ϵ in (1.3.4) admits a smooth density $f_{Y_T^\epsilon}$ and, for every $a \in \left(0,\, \frac{R^2 T}{2}\left(1+\frac{z_0/R}{\arccos(z_0/R)}\sqrt{1-z_0^2/R^2}\right)\right)$, the density $f_{Y_t^\epsilon}$ admits the expansion*

$$f_{Y_T^\epsilon}(a) = \epsilon^{-1} e^{-\Lambda(a)/\epsilon^2}\left(c_0(a) + o(1)\right), \qquad \text{as } \epsilon \to 0\,,$$

where

$$\Lambda(a) = \frac{z_0^2}{4\,c^2}\, r\, \frac{2\,r\,T - \sin(2\,r\,T)}{1+\cos(2\,r\,T)}\,,$$

$$c_0(a) = \frac{1}{\sqrt{2\pi\mathcal{A}(2rT)}}\,\frac{\cos^{3/2}(r\,T)}{2\,z_0\,c\,T^{3/2}}\, e^{\frac{1}{c^2}\int_{z_0}^{z_T} b(x)\,dx}\,,$$

$$\mathcal{A}(u) = \frac{u^3 + 6\,u\,\cos(u) + 3\,(u^2-2)\sin(u)}{6\,u^3}$$

and r is the unique solution of equation

$$1+\cos(2\,r\,T) - \frac{z_0^2\,T}{a}\left(1+\frac{\sin(2\,r\,T)}{2\,r\,T}\right) = 0$$

in the set

$$\mathcal{I} := \left(0,\, \frac{\pi}{2T}\right) \cup i\mathbb{R}_+ \subset \mathbb{C}\,.$$

We finally use the density expansion to obtain asymptotic expansions for the price and implied volatility of options on realized variance. The implied volatility of a realized variance option in a specific model is defined as the value of volatility such that the option price in this model equals the Black-Scholes option price with this volatility value, and initial value equal to the integral of the initial value of the instantaneous variance.

Theorem 1.3.8. *Let*

$$P^\epsilon(z_0, K, T) = \mathbb{E}\left((K - T^{-1}Y_T^\epsilon)_+\right)$$

be the price of a put option on realized variance with maturity T and strike K. Then P^ϵ admits the asymptotic expansion

$$P^\epsilon(z_0,K,T) = (K-z_0^2)_+ + e^{-\frac{\Lambda(KT)}{\epsilon^2}}\left(\frac{\epsilon^3\,c_0(KT)}{T\,(\Lambda'(KT))^2} + \mathcal{O}\left(\epsilon^3\right)\right) \qquad \text{as } \epsilon \to 0\,.$$

The BS implied volatility associated to $P^\epsilon(z_0,K,T)$ admits the expansion

$$\sigma_{BS} = \sigma_{BS,0} + \epsilon^2 \sigma_{BS,1} + \mathcal{O}\left(\epsilon^2\right)\,,$$

where

$$\sigma_{BS,0} = \frac{|\log(z_0^2/K)|}{\sqrt{2T}\,\Lambda^{1/2}(KT)}$$

and

$$\sigma_{BS,1} = \frac{|\log(z_0^2/K)|}{2^{3/2}\,T^{1/2}\,\Lambda^{3/2}(KT)}\,\log\left(\frac{4\sqrt{\pi}}{z_0\,K^{1/2}\,T^2}\,\frac{c_0(KT)}{(\Lambda'(KT))^2}\,\frac{\Lambda^{3/2}(KT)}{|\log(z_0^2/K)|}\right)\,,$$

as $\epsilon \to 0$.

1.3.3 Perturbation theory and interest rate derivatives pricing in the Levy Libor model (Chapter 5)

In this chapter, we study the pricing of interest rate derivatives in the Lévy Libor market model (LLMM) developed in (Eberlein and Özkan, 2005) by writing the LLMM as a perturbation of the standard log-normal LMM. Let $0 \le T_0 < ... < T_n$ be a tenor structure and denote by $L = (L^1, \dots, L^n)^\top$ the column vector of the forward Libor rates $L^j_t := L^{T_j}_t$. We assume that the dynamics of L is given by the SDE

$$dL_t = L_{t-}(b(t, L_t)dt + \Lambda(t)dX_t) \, ,$$

where X_t is a compensated d-dimensional Lévy process with non-zero diffusive part under the terminal measure $\mathbb{Q}^{T_n}$, $\Lambda(t)$ a deterministic $n \times d$ volatility matrix and $b(t, L_t)$ is the drift vector such that L^j_t is an $\mathbb{Q}^{T_j}$-martingale. Under this model, the price of a European derivative with payoff $g(L_{T_k})$ satisfies

$$P_t = B_t(T_n)\, \mathbb{E}^{\mathbb{Q}^{T_n}} \left[\prod_{j=k+1}^{n} (1 + \delta_j L^j_{T_k}) g(L_{T_k}) \mid \mathcal{F}_t \right] = B_t(T_n)\, u(t, L_t),$$

where u is the solution of the partial integro-differential equation (PIDE)

$$\begin{aligned} \partial_t u + \mathcal{A}_t u &= 0 \\ u(T_k, x) &= \prod_{j=k+1}^{n} (1 + \delta_j x_j) g(x) \, , \end{aligned}$$

where $\mathcal{A}_t$ is the generator of L_t. In order to approximate the solution $u(t, x)$, we define

$$dL^\alpha_t = L^\alpha_{t-}(b_\alpha(t, L^\alpha_t)dt + \Lambda(t)dX^\alpha_t) \, , \tag{1.3.5}$$

where $X^\alpha_t := \alpha X_{t/\alpha^2}$ and $b_\alpha(t, L^\alpha_t)$ is the corresponding drift such that L^α_t is a $\mathbb{Q}^{T_k}$-martingale. The scaling $X^\alpha_t = \alpha\, X_{t/\alpha^2}$ leaves the diffusive part unchanged, while the jump part converges to a Brownian motion as $\alpha \to 0$. Therefore, for $\alpha = 1$, (1.3.5) is the LLMM, whereas, for $\alpha = 0$, (1.3.5) corresponds to the standard LMM. The next result approximates the option price $u(t, x)$.

Theorem 1.3.9. *Let $u^\alpha(t, x)$ be the solution of the PIDE*

$$\partial_t u^\alpha + \mathcal{A}^\alpha_t u^\alpha = 0 \, , \qquad u^\alpha(T_k, x) = \prod_{j=k+1}^{n} (1 + \delta_j x_j) g(x) \, ,$$

where

$$\mathcal{A}^\alpha_t = \sum_{j=0}^{\infty} \alpha^j \mathcal{A}^j_t$$

is the infinitesimal generator of (1.3.5)*. Then* $u^\alpha(t,x)$ *admits an expansion of the form*

$$u^\alpha(t,x) = u_0(t,x) + \alpha\, u_1(t,x) + \alpha^2 u_2(t,x) + \mathcal{O}\big(\alpha^3\big) \ ,$$

where u_0 *is the LMM price of the derivative, and the correction terms* u_1 *and* u_2 *are the solutions of "LMM-like" PDEs whose explicit solutions are provided in* (5.4.13) *for* u_1 *and in* (5.4.18) *for* u_2*.*

We apply this result to obtain the expansions for the prices of caplets and swaptions in the LLMM and we test the performance of our approximation at pricing caplets in the model driven by a unidimensional CGMY process.

Chapter 2

Long-time trajectorial large deviations for affine stochastic volatility models and application to variance reduction for option pricing

2.1 Introduction

The aim of this paper is to develop efficient importance sampling estimators for prices of path-dependent options in affine stochastic volatility (ASV) models of asset prices. To this end, we establish pathwise large deviation results for these models, which are of independent interest.

An ASV model, studied in (Keller-Ressel, 2011) is a two-dimensional affine process (X, V) on $\mathbb{R} \times \mathbb{R}_+$ with special properties, where X models the logarithm of the stock price and V its instantaneous variance. This class includes many well studied and widely used models such as Heston stochastic volatility model (Heston, 1993), the model of Bates (Bates, 1996), Barndorff-Nielsen stochastic volatility model (Barndorff-Nielsen and Shephard, 2001) and time-changed Lévy models with independent affine time change. European options in affine stochastic volatility models may be priced by Fourier transform, but for path-dependent options explicit formulas are in general not available and Monte Carlo is often the method of choice. At the same time, Monte Carlo simulation of such processes is difficult and time-consuming: the convergence rates of discretization schemes are often low due to the irregular nature of coefficients of the corresponding stochastic differential equations. To accelerate Monte Carlo simulation, it is thus important to develop efficient variance-reduction algorithms for these models.

In this paper, we therefore develop an importance sampling algorithm for

ASV models. The importance sampling method is based on the following identity, valid for any probability measure $\mathbb{Q}$, with respect to which $\mathbb{P}$ is absolutely continuous. Let P be a deterministic function of a random trajectory S, then

$$\mathbb{E}[P(S)] = \mathbb{E}^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}P(S)\right].$$

This allows one to define the importance sampling estimator

$$\widehat{P}_N^{\mathbb{Q}} := \frac{1}{N}\sum_{j=1}^{N}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{(j)} P(S_{\mathbb{Q}}^{(j)}),$$

where $S_{\mathbb{Q}}^{(j)}$ are i.i.d. sample trajectories of S under the measure $\mathbb{Q}$. For efficient variance reduction, one needs then to find a probability measure $\mathbb{Q}$ such that S is easy to simulate under $\mathbb{Q}$ and the variance

$$\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left[P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right]$$

is considerably smaller than the original variance $\mathbb{V}\mathrm{ar}_{\mathbb{P}}\left[P(S)\right]$.

In this paper, following the work of (Genin and Tankov, 2016) in the context of Lévy processes, we define the probability $\mathbb{Q}$ using the path-dependent Esscher transform,

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\int_{[0,T]} X_t\cdot\theta(dt)}}{\mathbb{E}\left[e^{\int_{[0,T]} X_t\cdot\theta(dt)}\right]},$$

where X is the first component of the ASV model (the logarithm of stock price) and θ is a (deterministic) bounded signed measure on $[0,T]$. The optimal choice of θ should minimize the variance of the estimator under $\mathbb{P}_\theta$,

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}_\theta}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right) = \mathbb{E}_{\mathbb{P}}\left[P^2(S)\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right] - \mathbb{E}\left[P(S)\right]^2.$$

The computation of this variance is in general as difficult as the computation of the option price itself. Following (Dupuis and Wang, 2004; Glasserman et al., 1999; Guasoni and Robertson, 2008; Robertson, 2010) and more recently (Genin and Tankov, 2016), we propose to compute the variance reduction measure θ^* by minimizing the *proxy* for the variance computed using the theory of large deviations.

To this end, we establish a pathwise large deviation principle (LDP) for affine stochastic volatility models. A one dimensional LDP for X_t/t as $t\to\infty$ where X is the first component of an ASV model has been proven in (Jacquier et al., 2013). In this paper, we extend this result to the trajectorial setting, in the spirit of the pathwise LDP principles of (Léonard, 2000) but in a weaker topology.

The rest of this paper is structured as follows. In Section 2.2, we describe the model and recall certain useful properties of ASV processes. In Section 2.3, we recall some general results of large deviations theory. In Section 2.4, we prove a LDP for the trajectories of ASV processes. In Section 2.5, we develop the variance reduction method, using an asymptotically optimal change of measure obtained with the LDP shown in Section 2.4. In Section 2.6, we test the method numerically on several examples of options, some of which are path-dependent, in the Heston model with and without jumps.

2.2 Model description

In this paper, we model the price of the underlying $(S_t)_{t\geq 0}$ of an option as $S_t = S_0\, e^{X_t}$, where we model $(X_t)_{t\geq 0}$ as an affine stochastic volatility process. We recall, from (Keller-Ressel, 2011) and (Duffie et al., 2003), the definition and some properties of ASV models.

Definition 2.2.1. *An ASV model $(X_t, V_t)_{t\geq 0}$, is a stochastically continuous, time-homogeneous Markov process such that $\left(e^{X_t}\right)_{t\geq 0}$ is a martingale and*

$$\mathbb{E}\left(e^{uX_t+wV_t}\middle|X_0 = x, V_0 = v\right) = e^{\phi(t,u,w)+\psi(t,u,w)\,v+u\,x}\,, \tag{2.2.1}$$

for all $(t, u, w) \in \mathbb{R}_+\times \mathbb{C}^2$.

Proposition 2.2.2. *The functions ϕ and ψ satisfy generalized Riccati equations*

$$\partial_t\phi(t, u, w) = F(u, \psi(t, u, w))\,, \qquad \phi(0, u, w) = 0 \tag{2.2.2a}$$

$$\partial_t\psi(t, u, w) = R(u, \psi(t, u, w))\,, \qquad \psi(0, u, w) = w\,, \tag{2.2.2b}$$

where F and R have the Lévy-Khintchine forms

$$\begin{aligned} F(u, w) &= \begin{pmatrix} u & w \end{pmatrix} \cdot \frac{a}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + b \cdot \begin{pmatrix} u \\ w \end{pmatrix} \\ &\quad + \int_{D\backslash\{0\}} \left(e^{xu+yw} - 1 - w_F(x, y) \cdot \begin{pmatrix} u \\ w \end{pmatrix}\right) m(dx, dy)\,, \\ R(u, w) &= \begin{pmatrix} u & w \end{pmatrix} \cdot \frac{\alpha}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + \beta \cdot \begin{pmatrix} u \\ w \end{pmatrix} \\ &\quad + \int_{D\backslash\{0\}} \left(e^{xu+yw} - 1 - w_R(x, y) \cdot \begin{pmatrix} u \\ w \end{pmatrix}\right) \mu(dx, dy)\,, \end{aligned}$$

where $D = \mathbb{R} \times \mathbb{R}_+$,

$$w_F(x, y) = \begin{pmatrix} \frac{x}{1+x^2} \\ 0 \end{pmatrix} \qquad \text{and} \qquad w_R(x, y) = \begin{pmatrix} \frac{x}{1+x^2} \\ \frac{y}{1+y^2} \end{pmatrix}$$

and $(a, \alpha, b, \beta, m, \mu)$ satisfy the following conditions

- *a, α are positive semi-definite 2×2-matrices where $a_{12} = a_{21} = a_{22} = 0$.*
- *$b \in D$ and $\beta \in \mathbb{R}^2$.*
- *m and μ are Lévy measures on D and $\int_{D \setminus \{0\}} ((x^2 + y) \wedge 1)\, m(dx, dy) < \infty$.*

In the rest of the paper, we assume that there exists $u \in \mathbb{R}$ such that $R(u, 0) \neq 0$, for the law of $(X_t)_{t \geq 0}$ to depend on V_0. Define the function

$$\chi(u) = \partial_w R(u, w)|_{w=0} = \alpha_{12} u + \beta_2 + \int_{D \setminus \{0\}} y \left(e^{xu} - \frac{1}{1 + y^2} \right) \mu(dx, dy) \,.$$

A sufficient condition for $S_t = S_0\, e^{X_t}$ to be a martingale (Keller-Ressel, 2011, Corollary 2.7), which we assume to be satisfied in the sequel, is $F(1, 0) = R(1, 0) = 0$ and $\chi(0) + \chi(1) < \infty$.

In the following theorem, we compile several results of (Keller-Ressel, 2011) that describe the behaviour of the solution to eq. (2.2.2) as $t \to \infty$.

Theorem 2.2.3. *Assume that $\chi(0) < 0$ and $\chi(1) < 0$.*

- *There exists an interval $I \supseteq [0, 1]$, such that for each $u \in I$, eq. (2.2.2b) admits a unique stable equilibrium $w(u)$.*
- *For $u \in I$, eq. (2.2.2b) admits at most one other equilibrium $\tilde{w}(u)$, which is unstable.*
- *For $u \in \mathbb{R} \setminus I$, eq. (2.2.2b) does not have any equilibrium.*

We denote $\mathcal{B}(u)$ the basin of attraction of the stable solution $w(u)$ of eq. (2.2.2b) and $J = \{u \in I \,:\, F(u, w(u)) < \infty\}$, the domain of $u \mapsto F(u, w(u))$. We have that

- *J is an interval such that $[0, 1] \subseteq J \subseteq I$.*
- *For $u \in I$, $w \in \mathcal{B}(u)$ and $\Delta t > 0$, we have*

$$\psi\left(\frac{\Delta t}{\epsilon},\, u,\, w \right) \underset{\epsilon \to 0}{\longrightarrow} w(u) \,. \tag{2.2.3}$$

- *For $u \in J$, $w \in \mathcal{B}(u)$ and $\Delta t > 0$,*

$$\epsilon\, \phi\left(\frac{\Delta t}{\epsilon},\, u,\, w \right) \underset{\epsilon \to 0}{\longrightarrow} \Delta t\, h\,(u) \ , \tag{2.2.4}$$

  where $h(u) = F(u, w(u)) = \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{u X_{1/\epsilon}} \right]$.

- *For every $u \in I$, $0 \in \mathcal{B}(u)$.*

Definition 2.2.4. *A convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ with effective domain D_f is* essentially smooth *if*

i. *D_f° is non-empty;*

ii. *f is differentiable in D_f°;*

iii. *f is* steep, *that is, for any sequence $(u_n)_{n\in\mathbb{N}} \subset D_f^\circ$ that converges to a point in the boundary of D_f,*

$$\lim_{n\to\infty} ||\nabla f(u_n)|| = \infty \,.$$

In the rest of the paper, we shall make the following assumptions on the model.

Assumption 3. *The function h satisfies the following properties.*

1. *There exists $u < 0$, such that $h(u) < \infty$.*
2. *$u \mapsto h(u)$ is essentially smooth.*

In (Jacquier et al., 2013), a set of sufficient conditions is provided for Assumption 3 to be verified:

Proposition 2.2.5 (Corollary 8 in (Jacquier et al., 2013))**.** *Let (X, V) be an ASV model such that $u \mapsto R(u, 0)$ and $w \mapsto F(0, w)$ are not identically 0 and $\chi(0)$ and $\chi(1)$ are strictly negative. If either of the following conditions holds*

(i) *The Lévy measure μ of R has exponential moments of all orders, F is steep and $(0, 0), (1, 0) \in D_F^\circ$.*

(ii) *(X, V) is a diffusion,*

then function h is well defined, for every $u \in \mathbb{R}$ with effective domain J. Moreover h is essentially smooth and $\{0, 1\} \subset J^\circ$.

We now discuss the form of the basin of attraction of the unique stable solution of (2.2.2b).

Lemma 2.2.6. *(Keller-Ressel, 2011, Lemma 2.2.)*

(a) *F and R are proper closed convex functions on $\mathbb{R}^2$.*

(b) *F and R are analytic in the interior of their effective domain.*

(c) *Let U be a one-dimensional affine subspace of R^2. Then $F|U$ is either a strictly convex or an affine function. The same holds for $R|U$.*

(d) *If $(u, w) \in D_F$, then also $(u, \eta) \in D_F$ for all $\eta \le w$. The same holds for R.*

Lemma 2.2.7. *Let $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a convex function with either two zeros $w < \tilde{w}$, or a single zero w. In the latter case, we let $\tilde{w} = \infty$. Assume that there exists $y \in (w, \tilde{w})$ such that $f(y) < 0$. Then for every $x \in D_f$,*

$$\begin{cases} f(x) > 0\,, & \text{if } x < w \quad \text{or} \quad \tilde{w} < x\,, \\ f(x) < 0\,, & \text{if } x \in (w, \tilde{w})\,. \end{cases}$$

Proof. By convexity, for every $x \in D_f$ such that $x < w$,

$$\frac{y-w}{y-x} f(x) + \frac{w-x}{y-x} f(y) \geq f(w) = 0$$

and therefore $f(x) \geq -\frac{w-x}{y-w} f(y) > 0$. Furthermore, for every $x \in (w, y]$,

$$f(x) \leq \frac{y-x}{y-w} f(w) + \frac{x-w}{y-w} f(y) < 0\,.$$

Let $s = sup\{x \in D_f \ : \ f(x) < 0\}$. If f is continuous in s, then $\tilde{w} = s$ and for every $x > \tilde{w}$ in D_f, $f(x) \geq -\frac{\tilde{w}-x}{y-\tilde{w}} f(y) > 0$. If f is discontinuous in s however, then by convexity, $f(x) = +\infty$ for $x > s$. □

Proposition 2.2.8. *Let $u \in I$ and consider $w(u)$ the stable equilibrium of (2.2.2b). Then the basin of attraction of $w(u)$ is $\mathcal{B}(u) = (-\infty, \tilde{w}(u)) \cap D_{R(u,\cdot)}$, where $\tilde{w}(u) = \infty$ when (2.2.2b) admits only one equilibrium.*

Proof. By Lemma 2.2.6, $w \mapsto R(u, w)$ is convex. Since $w(u)$ is a stable equilibrium, the hypotheses of Lemma 2.2.7 are verified. Therefore, $R(u, w) > 0$ for every $w < w(u)$, whereas $R(u, w) < 0$ for every $w \in D_{R(u,\cdot)}$ such that $w(u) < w < \tilde{w}(u)$. This implies that the solution of

$$\partial_t \psi(t, u, w) = R(u, \psi(t, u, w))\,, \qquad \psi(0, u, w) = w \tag{2.2.5}$$

converges to $w(u)$ for every $w \in (-\infty, \tilde{w}(u)) \cap D_{R(u,\cdot)}$, whereas, if $w > \tilde{w}$, the solution of (2.2.5) diverges to ∞. □

2.3 Large deviations theory

In this section, we recall some useful classical results of the large deviations theory. We refer the reader to (Dembo and Zeitouni, 1998) for the proofs and for a broader overview of the theory.

Theorem 2.3.1 (Gärtner-Ellis)**.** *Let $(X^\epsilon)_{\epsilon \in]0,1]}$ be a family of random vectors in $\mathbb{R}^n$ with associated measure μ_ϵ. Assume that for each $\lambda \in \mathbb{R}^n$,*

$$\Lambda(\lambda) := \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\frac{\langle \lambda, X^\epsilon \rangle}{\epsilon}} \right]$$

as an extended real number. Assume also that 0 belongs to the interior of $D_\Lambda := \{\theta \in \mathbb{R}^n : \Lambda(\theta) < \infty\}$. *Denoting*

$$\Lambda^*(x) = \sup_{\theta \in \mathbb{R}^n} \langle \theta, x \rangle - \Lambda(\theta) \,,$$

the following hold:

(a) For any closed set F,

$$\limsup_{\epsilon \to 0} \epsilon \, \log \mu_\epsilon(F) \leq - \inf_{x \in F} \Lambda^*(x) \,.$$

(b) For any open set G,

$$\liminf_{\epsilon \to 0} \epsilon \, \log \mu_\epsilon(G) \geq - \inf_{x \in G \cap \mathcal{F}} \Lambda^*(x) \,,$$

where $\mathcal{F}$ *is the set of exposed points of* Λ^*, *whose exposing hyperplane belongs to the interior of* D_Λ.

(c) If Λ *is an essentially smooth, lower semi-continuous function, then* μ_ϵ *satisfies a LDP with good rate function* Λ^*.

Definition 2.3.2. *A partially ordered set* $(\mathcal{P}, \leq)$ *is called right-filtering if for every* $i, j \in \mathcal{P}$, *there exists* $k \in \mathcal{P}$ *such that* $i \leq k$ *and* $j \leq k$.

Definition 2.3.3. *A projective system* $(\mathcal{Y}_j, p_{ij})_{i \leq j \in \mathcal{P}}$ *on a partially ordered right-filtering set* $(\mathcal{P}, \leq)$ *is a family of Hausdorff topological spaces* $(\mathcal{Y}_j)_{j \in \mathcal{P}}$ *and continuous maps* $p_{ij} : \mathcal{Y}_j \to \mathcal{Y}_i$ *such that* $p_{ik} = p_{ij} \circ p_{jk}$ *whenever* $i \leq j \leq k$.

Definition 2.3.4. *Let* $(\mathcal{Y}_j, p_{ij})_{i \leq j \in \mathcal{P}}$ *be a projective system on a partially ordered right-filtering set* $(\mathcal{P}, \leq)$. *The projective limit of* $(\mathcal{Y}_j, p_{ij})_{i \leq j \in \mathcal{P}}$, *denoted* $\mathcal{X} = \varprojlim \mathcal{Y}_j$, *is the subset of topological spaces* $\mathcal{Y} = \prod_{j \in \mathcal{P}} \mathcal{Y}_j$, *consisting of all the elements* $x = (y_j)_{j \in \mathcal{P}}$ *for which* $y_i = p_{ij}(y_j)$ *whenever* $i \leq j$, *equipped with the topology induced by* $\mathcal{Y}$. *The projective limit of closed subsets* $F_j \subseteq \mathcal{Y}_j$ *are defined in the same way and denoted* $F = \varprojlim F_j$.

Remark 2.3.5. *The canonical projections of* $\mathcal{X}$, *i.e. the restrictions* $p_j : \mathcal{X} \to \mathcal{Y}_j$ *of the coordinate maps from* $\mathcal{X}$ *to* $\mathcal{Y}_j$, *are continuous.*

Theorem 2.3.6 (Dawson-Gärtner). *Let* $(\mathcal{Y}_j, p_{ij})_{i \leq j \in \mathcal{P}}$ *be a projective system on a partially ordered right-filtering set* $(\mathcal{P}, \leq)$ *and let* (μ_ϵ) *be a family of probabilities on* $\mathcal{X} = \varprojlim \mathcal{Y}_j$, *such that for any* $j \in \mathcal{P}$, *the Borel probability* $\mu_\epsilon \circ p_j^{-1}$ *on* $\mathcal{Y}_j$ *satisfies the LDP with the good rate function* Λ_j. *Then* μ_ϵ *satisfies the LDP with good rate function*

$$\Lambda(x) = \sup_{j \in \mathcal{P}} \Lambda_j(p_j(x)) \,.$$

Theorem 2.3.7 (Varadhan's Lemma, version of (Guasoni and Robertson, 2008)). *Let $(X^\epsilon)_{\epsilon\in]0,1]}$ be a family of $\mathcal{X}$-valued random variables, whose laws μ_ϵ satisfy a LDP with rate function Λ. If $\varphi : \mathcal{X} \to \mathbb{R} \cup \{-\infty\}$ is a continuous function which satisfies*

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{\gamma\,\varphi(X^\epsilon)}{\epsilon}\right)\right] < \infty$$

for some $\gamma > 1$, then

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{\varphi(X^\epsilon)}{\epsilon}\right)\right] = \sup_{x\in\mathcal{X}}\{\varphi(x) - \Lambda(x)\}\,.$$

2.4 Trajectorial large deviations for affine stochastic volatility model

In this section, we prove a trajectorial LDP for (X_t) when the time horizon is large. Define, for $\epsilon \in (0,1]$ and $0 \le t \le T$, the scaling $X_t^\epsilon = \epsilon X_{t/\epsilon}$. We proceed by proving first a LDP for X_t^ϵ in finite dimension, that we extend, in a second step to the whole trajectory of $(X_t^\epsilon)_{0\le t\le T}$.

2.4.1 Finite-dimensional LDP

Let $\tau = \{0 < t_1 < ... < t_n = t\}$, by convention $t_0 = 0$, and define

$$\Lambda_{\epsilon,\tau}(\theta) = \log \mathbb{E}\left[e^{\sum_{k=1}^n \theta_k X_{t_k}^\epsilon}\right]\,,$$

for $\theta \in \mathbb{R}^n$. We start by formulating our main technical assumption.

Assumption 4. *One of the following conditions is verified.*

1. *The interval support of F is $J = [u_-, u_+]$ and $w(u_-) = w(u_+)$.*
2. *For every $u \in \mathbb{R}$, $\tilde{w}(\cdot) = \infty$, i.e, the generalized Riccati equations have only one (stable) equilibrium.*

The following Lemma gives an intuition on Assumption 4.

Lemma 2.4.1. *For every $u_1, u_2 \in I$, $\tilde{w}(u_1) \ge w(u_2)$.*

Proof. If Assumption 4(2) holds, then the result is obvious. Assume then that it is Assumption 4(1), that holds. Since $u \mapsto w(u)$ is convex and $u \mapsto \tilde{w}(u)$ is concave (Keller-Ressel, 2011, Lemma 3.3), then for every $u_1, u_2 \in I$,

$$\tilde{w}(u_1) \ge \frac{u_+ - u_1}{u_+ - u_-}\tilde{w}(u_-) + \frac{u_1 - u_-}{u_+ - u_-}\tilde{w}(u_+) \ge w(u_-)\,,$$

while

$$w(u_2) \le \frac{u_+ - u_2}{u_+ - u_-}w(u_-) + \frac{u_2 - u_-}{u_+ - u_-}w(u_+) = w(u_-)\,.$$

Therefore $\tilde{w}(u_1) \ge w(u_2)$ for every $u_1, u_2 \in I$. □

As a first step to apply Theorem 2.3.1, we prove the following result.

Theorem 2.4.2. *Let $\theta \in \mathbb{R}^n$. If Assumption 4 holds, then*

$$\Lambda_\tau(\theta) := \lim_{\epsilon\to 0} \epsilon\, \Lambda_{\epsilon,\tau}(\theta/\epsilon) = \begin{cases} \sum_{j=1}^n (t_j - t_{j-1})\, h\,(\Theta_j) & \text{if } \Theta_j \in J\,,\ \forall j \\ \infty & \text{otherwise} \end{cases},$$

where $\Theta_j := \sum_{k=j}^n \theta_k$.

Proof. Since Assumption 4 holds, then, by Lemma 2.4.1, $w(\Theta_{j+1}) \in \mathcal{B}(\Theta_j)$ for every j. Assume first that $\Theta_j \in J$ for every j. Using the Markov property and eq. (2.2.1), we obtain

$$\begin{aligned}
\Lambda_\tau(\theta) &= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^n \theta_j X_{t_j/\epsilon}}\right]\right) \\
&= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{n-1} \theta_j X_{t_j/\epsilon}}\, \mathbb{E}\left(e^{\Theta_n X_{t_n/\epsilon}} \middle| X_{t_{n-1}/\epsilon}, V_{t_{n-1}/\epsilon}\right)\right]\right) \\
&= \lim_{\epsilon\to 0} \epsilon\, \phi\left(\frac{t_n - t_{n-1}}{\epsilon},\, \Theta_n,\, 0\right) \\
&\qquad + \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{n-2} \theta_j X_{t_j/\epsilon} + \Theta_{n-1} X_{t_{n-1}/\epsilon} + \psi\left(\frac{t_n - t_{n-1}}{\epsilon}, \Theta_n, 0\right) V_{t_{n-1}/\epsilon}}\right]\right)\,.
\end{aligned}$$

Since $\Theta_n \in J$ and $0 \in \mathcal{B}(\Theta_n)$, eqs. (2.2.3) and (2.2.4) apply and

$$\begin{aligned}
\Lambda_\tau(\theta) &= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{n-2} \theta_j X_{t_j/\epsilon} + \Theta_{n-1} X_{t_{n-1}/\epsilon} + \psi\left(\frac{t_n - t_{n-1}}{\epsilon}, \Theta_n, 0\right) V_{t_{n-1}/\epsilon}}\right]\right) \\
&\qquad + (t_n - t_{n-1})\, h(\Theta_n)\,.
\end{aligned}$$

Using the fact that $\Theta_j \in J$ and $w(\Theta_{j+1}) \in \mathcal{B}(\Theta_j)$ for every j, we can iterating the procedure to obtain

$$\begin{aligned}
\Lambda_\tau(\theta) &= \sum_{j=1}^n (t_j - t_{j-1})\, h\,(\Theta_j) + \lim_{\epsilon\to 0} \epsilon\, \psi\left(\frac{t_1 - t_0}{\epsilon},\, \Theta_1,\, w\,(\Theta_2)\right) V_0 + \epsilon \sum_{k=1}^n \theta_k X_0 \\
&= \sum_{j=1}^n (t_j - t_{j-1})\, h\,(\Theta_j)\,. \qquad (2.4.1)
\end{aligned}$$

Assume now that there exists k such that $\Theta_k \notin J$. Without loss of generality, we take the largest such k. Following the same procedure, we find

$$\begin{aligned}
\Lambda_\tau(\theta) &= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{k-2} \theta_j X_{t_j/\epsilon} + \Theta_{k-1} X_{t_{k-1}/\epsilon} + \psi\left(\frac{t_k - t_{k-1}}{\epsilon}, \Theta_k, w(\Theta_{k+1})\right) V_{t_{k-1}/\epsilon}}\right]\right) \\
&\qquad + \epsilon\, \phi\left(\frac{t_k - t_{k-1}}{\epsilon}, \Theta_k, w(\Theta_{k+1})\right) + \sum_{j=k+1}^n (t_j - t_{j-1})\, h(\Theta_j)\,.
\end{aligned}$$

Noting that $\phi(\cdot, u, w)$ explodes in finite time for $u \notin J$ then finishes the proof. $\square$

We now proceed to the finite-dimensional large deviations result.

Theorem 2.4.3. *Let $(X_t^\epsilon)_{t\geq 0,\,\epsilon\in(0,1]}$ and $\tau = \{t_1, ..., t_n\}$ as previously. Assuming that Assumption 4 holds, then $(X_{t_1}^\epsilon, ..., X_{t_n}^\epsilon)$ satisfies a LDP on $\mathbb{R}^n$ with good rate function*

$$\Lambda_\tau^*(x) = \sup_{\Theta\in J^n} \left\{ \sum_{j=1}^n \Theta_j(x_j - x_{j-1}) - \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\Theta_j\right) \right\} ,$$

where $\Theta_j = \sum_{k=j}^n \theta_k$.

Proof. By Assumption 3(1), there exists $u \in J$ such that $u < 0$, which implies that $[u, 1] \subset J$ and therefore 0 is in the interior of $D_{\Lambda_\tau} = J^n$. Theorem 2.4.2 implies that the limit

$$\Lambda_\tau(\theta) = \lim_{\epsilon\to 0} \epsilon\, \Lambda_{\epsilon,\tau}(\theta/\epsilon) = \begin{cases} \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\Theta_j\right) & \text{if } \Theta_j \in J\ ,\ \forall j \\ \infty & \text{otherwise} \end{cases} ,$$

where $\Theta_j := \sum_{k=j}^n \theta_k$, exists as an extended real number. Since, by Assumption 3(2), h is essentially smooth and lower semi-continuous, then so is Λ_τ. Theorem 2.3.1 then applies and $(X_{t_1}^\epsilon, ..., X_{t_n}^\epsilon)$ satisfies a LDP, on $\mathbb{R}^n$, with good rate function

$$\Lambda_\tau^*(x) = \sup_{\theta\in\mathbb{R}^n} \left\{ \theta^\top x - \Lambda_\tau(\theta) \right\} .$$

Furthermore,

$$\begin{aligned} \Lambda_\tau^*(x) &= \sup_{\theta\in\mathbb{R}^n} \left\{ \theta^\top x - \Lambda_\tau(\theta) \right\} \\ &= \sup_{\Theta\in J^n} \left\{ \sum_{j=1}^n \sum_{k=j}^n \theta_k(x_j - x_{j-1}) - \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\Theta_j\right) \right\} \\ &= \sup_{\Theta\in J^n} \left\{ \sum_{j=1}^n \Theta_j(x_j - x_{j-1}) - \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\Theta_j\right) \right\} , \end{aligned}$$

which finishes the proof. □

2.4.2 Infinite-dimensional LDP

Extension of the LDP

We now extend the LDP to the whole trajectory of $(X_t^\epsilon)_{0\leq t\leq T}$ on $\mathcal{F}([0, T], \mathbb{R}) := \{x : [0, T] \to \mathbb{R},\ x_0 = 0\}$, the set of all functions from $[0, T]$ to $\mathbb{R}$ that vanish at 0, by proving the following general lemma.

Lemma 2.4.4. *Let $(\mathcal{P}, \leq)$ be the partially ordered right-filtering set*

$$\mathcal{P} = \bigcup_{n=1}^{\infty} \{(t_1, ..., t_n) \, , \, 0 \leq t_1 \leq ... \leq t_n \leq T\}$$

ordered by inclusion. We consider, on $(\mathcal{P}, \leq)$, the projective system $(\mathcal{Y}_j, p_{ij})_{i \leq j \in \mathcal{P}}$ defined by $\mathcal{Y}_j = \mathbb{R}^{\#j}$ and $p_{ij} : \mathcal{Y}_j \to \mathcal{Y}_i$ the natural projection on shared times. Assume that for any $j = \{t_1, ..., t_n\}$, the finite-dimensional process $(X^\epsilon_{t_1}, ..., X^\epsilon_{t_n})$ satisfies a large deviation property with good rate function Λ_j. Then the family $(X^\epsilon_t)_{0 \leq t \leq T}$ satisfies a large deviation property on $\mathcal{X} = \mathcal{F}([0, T], \, \mathbb{R})$ equipped with the topology of pointwise convergence, with good rate function

$$\Lambda(x) = \sup_{j \in \mathcal{P}} \Lambda_j(p_j(x)) \, ,$$

where $p_\tau(x) = (x_{t_1}, ..., x_{t_n})$ is the canonical projection from $\mathcal{X}$ to $\mathcal{Y}_\tau$.

Proof. Let μ^ϵ be the probability measure generated by $(X^\epsilon_t)_{0 \leq t \leq T}$ on $\mathcal{X}$. Then, by hypothesis, for any $j \in \mathcal{P}$, $\mu^\epsilon \circ p_j^{-1}$ satisfies a LDP with good rate function Λ_τ. The result then follows from Theorem 2.3.6. □

Theorem 2.4.5. *Assume that Assumption 4 holds, then $(X^\epsilon_t)_{0 \leq t \leq T}$ satisfies a LDP on $\mathcal{F}([0, T], \mathbb{R})$ equipped with the topology of point-convergence, as $\epsilon \to 0$, with good rate function*

$$\Lambda^*(x) = \sup_{\tau} \Lambda^*_\tau(x) \, ,$$

where the supremum is taken over the discrete ordered subsets of the form $\tau = \{t_1, ..., t_n\} \subset [0, T]$.

Proof. The result is a direct application of Lemma 2.4.4. □

Calculation of the rate function

We finally calculate the rate function of Theorem 2.4.5.

Theorem 2.4.6. *The rate function of Theorem 2.4.5 is*

$$\Lambda^*(x) = \int_0^T h^*(\dot{x}^{ac}_t) \, dt + \int_0^T \mathcal{H}\left(\frac{d\nu_t}{d\theta_t}\right) \, d\theta_t \, ,$$

where

$$h^*(y) = \sup_{\theta \in J} \{\theta y - h(\theta)\} \, , \qquad \mathcal{H}(y) = \lim_{\epsilon \to 0} \epsilon \, h^*(y/\epsilon) \, ,$$

$\dot{x}^{ac}$ is the derivative of the absolutely continuous part of x, ν_t is the singular component of dx_t with respect to dt and θ_t is any non-negative, finite, regular, $\mathbb{R}$-valued Borel measure, with respect to which ν_t is absolutely continuous.

Proof. By identifying $(\Theta_1, ..., \Theta_n)$ with $(\theta_{t_1}, ..., \theta_{t_n})$, we find for every $x \in \mathcal{F}([0,T], \mathbb{R})$,

$$\begin{aligned}
\sup_\tau \Lambda^*_\tau(x) &= \sup_\tau \sup_{\Theta \in J^{\#\tau}} \sum_{j=1}^{\#\tau} \Theta_j (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\Theta_j) \\
&= \sup_{\theta \in \mathcal{F}([0,T],\mathbb{R})} \sup_\tau \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) \\
&= \sup_{\theta \in C([0,T],J)} \sup_\tau \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) \, .
\end{aligned}$$

Note that the supremum can be taken indifferently on $\mathcal{F}([0,T],J)$ or on $C([0,T],J)$ because the objective function depends on θ only on a finite set. Since we have assumed that there exists $u < 0$ in J, then if x has infinite variation, we immediately find that $\Lambda^*(x) = \infty$. Assume therefore that x has finite variation. We wish to show that

$$\begin{aligned}
&\sup_{\theta \in C([0,T],J)} \sup_\tau \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) \\
&\qquad = \sup_{\theta \in C([0,T],J)} \int_0^T \theta_t dx_t - \int_0^T h(\theta_t) dt \, .
\end{aligned}$$

Notice that

$$\begin{aligned}
&\sup_{\theta \in C([0,T],J)} \sup_\tau \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) \\
\geq &\sup_{\theta \in C([0,T],J)} \limsup_\tau \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) \\
= &\sup_{\theta \in C([0,T],J)} \int_0^T \theta_t dx_t - \int_0^T h(\theta_t) dt \, .
\end{aligned}$$

To prove the other inequality, we use the following construction. Fix τ and let $\theta \in C([0,T],J)$. Let also $\epsilon > 0$ such that $\epsilon < \min(t_j - t_{j-1})$ and define $\theta^{\epsilon,\tau}$ as

$$\theta^{\epsilon,\tau}_t = \begin{cases} \theta_{t_{j-1}} + \frac{t - t_{j-1}}{\epsilon} (\theta_{t_j} - \theta_{t_{j-1}}) & \text{if } t \in [t_{j-1},\, t_{j-1} + \epsilon] \, , \\ \theta_{t_j} & \text{if } t \in [t_{j-1} + \epsilon,\, t_j] \, . \end{cases}$$

Then

$$\left| \sum_{j=1}^{\#\tau} \theta_{t_j} (x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j}) - \int_0^T \theta^{\epsilon,\tau}_t dx_t + \int_0^T h(\theta^{\epsilon,\tau}_t) dt \right|$$

$$= \left| \sum_{j=1}^{\#\tau} (\theta_{t_j} - \theta_{t_{j-1}}) \int_{t_{j-1}}^{t_{j-1}+\epsilon} \left(1 - \frac{t - t_{j-1}}{\epsilon}\right) dx_t + \int_{t_{j-1}}^{t_{j-1}+\epsilon} h(\theta_t^{\epsilon,\tau}) - h(\theta_{t_j}) dt \right|$$
$$\leq \sum_{j=1}^{\#\tau} |\theta_{t_j} - \theta_{t_{j-1}}| \left| \int_{t_{j-1}}^{t_{j-1}+\epsilon} \left(1 - \frac{t - t_{j-1}}{\epsilon}\right) dx_t \right|$$
$$+ 2\epsilon \max\left\{ |h(\theta)| \,:\, \theta \in [\theta_{t_{j-1}}, \theta_{t_j}] \right\}$$
$$\leq \sum_{j=1}^{\#\tau} |\theta_{t_j} - \theta_{t_{j-1}}| \, \mu_x(]0,\epsilon]) + 2\epsilon \max\left\{ |h(\theta)| \,:\, \theta \in [\theta_{t_{j-1}}, \theta_{t_j}] \right\} \underset{\epsilon\to 0}{\to} 0\,,$$

where μ_x is the measure associated with x. Hence

$$\sup_{\theta \in C([0,T],J)} \sup_{\tau} \sum_{j=1}^{\#\tau} \theta_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1}) h(\theta_{t_j})$$
$$\leq \sup_{\theta \in C([0,T],J)} \int_0^T \theta_t dx_t - \int_0^T h(\theta_t) dt$$

and

$$\Lambda^*(x) = \sup_{\theta \in C([0,T],J)} \int_0^T \theta_t dx_t - \int_0^T h(\theta_t)\, dt\,.$$

We will now use (Rockafellar, 1971, Thm. 5.) to obtain the result. Since x has finite variations, the measure dx_t is regular. Using the notations of (Rockafellar, 1971), in our case the multifunction D is the constant multifunction $t \mapsto D(t) = J$. Therefore D is fully lower semi-continuous. Furthermore, since $[0,1] \subset J$, the interior of $D(t)$ is non-empty. The set $[0,T]$ is compact with no non-empty open sets of measure 0 and for every u in the interior of J, and $V \in [0,T]$ open,

$$\int_V |h(u)|\, dt \leq T|h(u)| < \infty\,.$$

(Rockafellar, 1971, Thm. 5.) then implies that

$$\sup_{\theta \in C([0,T],J)} \int_0^T \theta_t\, dx_t - \int_0^T h(\theta_t)\, dt = \int_0^T h^*(\dot{x}_t^{ac})\, dt + \int_0^T \mathcal{H}\left(\frac{d\nu_t}{d\theta_t}\right) d\theta_t\,,$$

where

$$h^*(y) = \lim_{\epsilon\to 0} \sup_{\theta \in J} \{\theta y - h(\theta)\}\,, \qquad \mathcal{H}(y) = \lim_{\epsilon \to 0} \epsilon\, h^*(y/\epsilon)\,,$$

$\dot{x}^{ac}$ is the derivative of the absolutely continuous part of x, ν_t is the singular component of dx_t with respect to dt and θ_t is any non-negative, finite, regular, $\mathbb{R}$-valued Borel measure, with respect to which ν_t is absolutely continuous. □

Remark 2.4.7. *In particular, the proof of theorem 2.4.6 shows that, if x does not belong to V_r, the set of trajectories $x : [0,t] \to \mathbb{R}$ with bounded variation, then $\Lambda^*(x) = \infty$.*

2.5 Variance reduction

Denote $P(S)$ the payoff of an option on $(S_t)_{0\le t\le T}$. The price of an option is generally calculated as the expectation $\mathbb{E}(P(S))$ under a certain risk-neutral measure $\mathbb{P}$. For any equivalent measure $\mathbb{Q}$, the price of the derivative can be written

$$\mathbb{E}(P(S)) = \mathbb{E}^{\mathbb{Q}}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right) .$$

The variance of $P(S)$ is

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}}\left(P(S)\right) = \mathbb{E}\left(P^2(S)\right) - \mathbb{E}^2\left(P(S)\right) ,$$

whereas the variance of

$$\begin{aligned}\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right) &= \mathbb{E}^{\mathbb{Q}}\left(P^2(S)\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^2\right) - \left(\mathbb{E}^{\mathbb{Q}}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right)\right)^2\\ &= \mathbb{E}\left(P^2(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right) - \mathbb{E}^2\left(P(S)\right) .\end{aligned}$$

We can therefore choose $\mathbb{Q}$ in order to reduce the variance of the random variable, whose expectation gives the price of the derivative.

A flexible class of measure changes introduced in (Genin and Tankov, 2016) is given by path dependent Esscher transform, that is the class $\mathbb{P}_\theta$ such that

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\int_0^T X_t\,d\theta_t}}{\mathbb{E}\left[e^{\int_0^T X_t\,d\theta_t}\right]} ,$$

where θ belong to M, the set of signed measures on $[0, T]$. Denoting $H(X) = \log P\left(S_0\, e^X\right)$, the optimization problem writes

$$\inf_{\theta\in M} \mathbb{E}\left[\exp\left(2H(X) - \int_0^T X_t\,d\theta_t + \mathcal{G}_1(\theta)\right)\right] , \tag{2.5.1}$$

where

$$\mathcal{G}_\epsilon(\theta) := \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X^\epsilon_t\,d\theta_t}\right] .$$

The optimization problem (2.5.1) cannot be solved explicitly. We therefore choose to solve the problem asymptotically using the two following lemmas. Denote $\bar{M}$ the set of measures $\theta \in M$ with support on a finite set of points. We first give a lemma that characterizes the behaviour of $\mathcal{G}_\epsilon(\theta)$ as $\epsilon \to 0$, for $\theta \in \bar{M}$ as this will be sufficient for the cases that we will consider in Section 2.6 (see Prop. 2.5.5).

Lemma 2.5.1. *If Assumption 4 holds, then for any measure $\theta \in \bar{M}$, such that for every $t \in [0, T]$, $\theta([t, T]) \in J$,*

$$\lim_{\epsilon\to 0} \mathcal{G}_\epsilon(\theta) = \int_0^T h(\theta([t, T]))\,dt .$$

Proof. Denote $\tau = \{t_1, ..., t_n\}$, the support of θ. We then obtain

$$\begin{aligned}\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X_t^\epsilon\, d\theta_t}\right] &= \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\sum_{j=1}^n X_{t_j}^\epsilon\, \theta\left((t_{j-1}, t_j\,]\right)}\right] \\ &= \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\theta\big((t_{j-1}, t_n\,]\big)\right) \\ &= \int_0^T h(\theta([t,T]))\, dt\end{aligned}$$

by applying Theorem 2.4.2 to $\theta = \big(\theta\big((t_0, t_1\,]\big), ..., \theta\big((t_{n-1}, t_n\,]\big)\big)$. □

Next, we give a result that characterizes the behaviour of the variance minimization problem 2.5.1 where X has been replaced by X^ϵ as $\epsilon \to 0$.

Lemma 2.5.2. *Let $\theta \in \bar{M}$ such that $-\theta([t,T]) \in J^\circ$ for every $t \in [0,T]$. Assume that the assumptions of Theorem 2.4.3 hold. Assume furthermore that $H : \mathcal{F}([0,T], \mathbb{R}) \to \mathbb{R}$ is bounded from above by a constant C and continuous on D the set of functions $x \in V_r$, such that $H(x) > -\infty$, with respect with to the pointwise convergence topology. Then*

$$\begin{aligned}&\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, d\theta_t + \mathcal{G}_\epsilon(\theta)}{\epsilon}\right)\right] \\ &\qquad = \sup_{x\in D}\left\{2H(x) - \int_0^T x_t d\theta_t - \Lambda^*(x)\right\} + \int_0^T h(\theta([t,T]))\, dt\,.\end{aligned}$$

Proof. First note that, by Lemma 2.5.1,

$$\begin{aligned}&\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, d\theta_t + \mathcal{G}_\epsilon(\theta)}{\epsilon}\right)\right] \\ &\qquad = \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, d\theta_t}{\epsilon}\right)\right] + \int_0^T h(\theta([t,T]))\, dt\,.\end{aligned}$$

We therefore just need to prove that

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, d\theta_t}{\epsilon}\right)\right] = \sup_{x\in D}\left\{2H(x) - \int_0^T x_t d\theta_t - \Lambda^*(x)\right\}.$$

Denote $\varphi : \mathcal{F}([0,T], \mathbb{R}) \to \mathbb{R}$ the function $\varphi(x) = 2H(x) - \int_0^T x_t\, d\theta_t$. Since H is assumed to be continuous and θ has support on τ, φ is continuous. Let us show the integrability condition of Theorem 2.3.7. For every $\gamma > 0$

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{\gamma\,\varphi(X^\epsilon)}{\epsilon}\right)\right]$$

$$= \limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2\gamma H(X^\epsilon) - \gamma \int_0^T X_t^\epsilon \, d\theta_t}{\epsilon}\right)\right]$$
$$\leq 2\gamma C + \limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X_t^\epsilon \, d(-\gamma\theta)_t}\right] .$$

Since $-\theta([t,T]) \in J^\circ$ for every $t \in [0,T]$, there exists $\gamma > 1$ such that $-\gamma\theta([t,T])$ remains in J for every t. Therefore Lemma 2.5.1 applies and

$$\limsup_{\epsilon\to 0} \epsilon \, \log \mathbb{E}\left[\exp\left(\frac{\gamma\,\varphi(X^\epsilon)}{\epsilon}\right)\right] \leq 2\gamma C + \int_0^T h(-\gamma\,\theta([t,T]))\, dt < \infty .$$

Theorem 2.3.7 then applies and yields the result. □

Definition 2.5.3. *Let $\theta \in M$. We say that θ is asymptotically optimal if it minimises*

$$\sup_{x\in V_r} \left\{2H(x) - \int_0^T x_t \, d\theta_t - \Lambda^*(x)\right\} + \int_0^T h(\theta([t,T]))\, dt .$$

In general, Λ^* is not easy to calculate explicitly. To solve this problem, we cite the following theorem of (Genin and Tankov, 2016).

Theorem 2.5.4. *Let H be concave and assume that the set $\{x \in V_r : H(x) > -\infty\}$ is non-empty and contains a constant element. Assume furthermore that H is continuous on this set with respect to the topology of pointwise convergence, that h is lower semi-continuous with open and bounded effective domain and that there exists a $\lambda > 0$ such that h is complex-analytic on $\{z \in \mathbb{C} : |Im(z)| < \lambda\}$. Then*

$$\inf_{\theta\in M} \sup_{x\in V_r} \left\{2H(x) - \int_0^T x_t d\theta_t - \Lambda^*(x)\right\} + \int_0^T h(\theta([t,T]))\, dt$$
$$= 2 \inf_{\theta\in M} \left\{\hat{H}(\theta) + \int_0^T h(\theta([t,T]))\, dt\right\} ,$$

where

$$\hat{H}(\theta) = \sup_{x\in V_r} \left\{H(x) - \int_0^T x_t \, d\theta_t\right\} .$$

Furthermore, if θ^* minimises the left-hand side of the above equation, it also minimises the right-hand side.

We finally give a result for the case where H depends on x only through $x_{t_1}, ..., x_{t_n}$.

Proposition 2.5.5. *Let $\tau = \{t_1, ..., t_n\}$ and let $H : \mathcal{F}([0,t], \mathbb{R}) \to \mathbb{R} \cup \{-\infty\}$ be a log-payoff depending on x only through x_τ. Then for every $\theta \in M$ such that $\theta(\tau) \neq \theta([0,T])$, $\hat{H}(\theta) = \infty$.*

Proof. Assume that $\theta \in M$ is such that $\theta(\tau) \neq \theta([0,T])$. Then there exists a set $A \subset [0,T]\backslash\tau$, such that $\theta(A) \neq 0$. Fix $\bar{x} \in D$. By definition, $H(\bar{x}) > -\infty$. Then

$$H(\hat{x} + \alpha\, \mathbb{1}_A) - \int_0^T \bar{x}_t + \alpha\, \mathbb{1}_A d\theta_t = H(\hat{x}) - \int_0^T \bar{x}_t d\theta_t - \alpha\, \theta(A)\,.$$

By letting α tend to $\operatorname{sgn}(\theta)\,\infty$, one can therefore increase indefinitely $H(x) - \int_0^T x_t d\theta_t$. Therefore, $\hat{H}(\theta) = \infty$. □

2.6 Numerical examples

In this section, we apply the variance reduction method to several examples. We first show a result for options on the average value of the underlying over a finite set of points.

Proposition 2.6.1. *Let $\tau = \{t_1, ..., t_n\}$ and consider an option with log-payoff*

$$H(x) = \log\left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{x_{t_j}}\right)_+ .$$

Then for any $\theta \in \bar{M}$ with support on $\theta = \{t_1, ..., t_n\}$,

$$\hat{H}(\theta) = \log\left(\frac{K}{1 - \sum_{l=1}^{n}\theta_l}\right) - \sum_{m=1}^{n}\theta_m \log\left(\frac{-\theta_m\, n\, K/S_0}{1 - \sum_{l=1}^{n}\theta_l}\right) \tag{2.6.1}$$

where we use the abuse of notation $\theta_j = \theta(\{t_j\})$.

Proof. In this case,

$$H(x) - \int_0^T x_t d\theta_t = \log\left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{x_{t_j}}\right)_+ - \sum_{j=1}^{n}\theta_j x_{t_j}\,.$$

When the option is out or at the money, the log-payoff is $-\infty$. Assume that x is such that $H(x) > -\infty$ and differentiate with respect to x_{t_j}. We obtain

$$0 = \partial_{x_{t_j}}\left\{\log\left(K - \frac{S_0}{n}\sum_{l=1}^{n} e^{x_{t_l}}\right) - \sum_{l=1}^{n} x_{t_l}\theta_l\right\} = \frac{-\frac{S_0}{n}\, e^{x_{t_j}}}{K - \frac{S_0}{n}\sum_{l=1}^{n} e^{x_{t_l}}} - \theta_j\,.$$

Therefore the x that maximises $H(x) - \int_0^t x_s d\theta_s$ satisfies

$$\frac{e^{x_{t_j}}}{\theta_j} = -n\,\frac{K}{S_0} + \sum_{l=1}^{n} e^{x_{t_l}} = -n\,\frac{K}{S_0} + \frac{e^{x_{t_j}}}{\theta_j}\sum_{l=1}^{n}\theta_l\,,$$

for every j. Therefore

$$x_{t_j} = \log\left(\frac{-\theta_j\, n\, K/S_0}{1 - \sum_{l=1}^{n}\theta_l}\right)\,.$$

Inserting x_{t_j} in the value of $H(x) - \int_0^T x_t\, d\theta_t$, we obtain the result. □

2.6.1 European and Asian put options in the Heston model

Consider the Heston model (Heston, 1993)

$$\begin{aligned} dX_t &= -\frac{V_t}{2}\,dt + \sqrt{V_t}\,dW_t^1\,, & X_0 = 0 \\ dV_t &= \lambda(\mu - V_t)\,dt + \zeta\sqrt{V_t}\,dW_t^2\,, & V_0 > 0 \\ d\big\langle W^1, W^2\big\rangle_t &= \rho\,dt\,, \end{aligned} \tag{2.6.2}$$

where W^1, W^2 are standard $\mathbb{P}$-Brownian motions. The Laplace transform of (X_t, V_t) is

$$\mathbb{E}\left(e^{uX_t + wV_t}\right) = e^{\phi(t,u,w)+\psi(t,u,w)V_0+uX_0}\,,$$

where ϕ, ψ satisfy the Riccati equations

$$\begin{aligned} \partial_t\phi(t,u,w) &= F(u,\psi(t,u,w)) & \phi(0,u,w) = 0 \\ \partial_t\psi(t,u,w) &= R(u,\psi(t,u,w)) & \psi(0,u,w) = w \end{aligned} \tag{2.6.3}$$

for $F(u,w) = \lambda\mu w$ and

$$R(u,w) = \frac{\zeta^2}{2}\,w^2 + \zeta\rho\,uw - \lambda w + \frac{1}{2}(u^2 - u)\,.$$

A standard calculation shows that the solution of the Riccati equations (2.6.3) is

$$\begin{aligned} \psi(t,u,w) &= \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \frac{\gamma}{\zeta^2}\,\frac{\tanh\left(\frac{\gamma}{2}\,t\right) + \eta}{1 + \eta\,\tanh\left(\frac{\gamma}{2}\,t\right)} \\ \phi(t,u,w) &= \mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right)t - 2\mu\,\frac{\lambda}{\zeta^2}\,\log\left(\cosh\left(\frac{\gamma}{2}\,t\right) + \eta\,\sinh\left(\frac{\gamma}{2}\,t\right)\right)\,, \end{aligned} \tag{2.6.4}$$

where $\gamma = \gamma(u) = \zeta\sqrt{\left(\frac{\lambda}{\zeta} - \rho u\right)^2 + \frac{1}{4} - \left(u - \frac{1}{2}\right)^2}$ and $\eta = \eta(u,w) = \frac{\lambda - \zeta\rho u - \zeta^2 w}{\gamma(u)}$. Furthermore, for the Heston model, the function h is given by

$$h(u) = \mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \mu\,\frac{\lambda}{\zeta^2}\,\gamma(u)\,. \tag{2.6.5}$$

Remark 2.6.2. *The log-Laplace transform h of the Heston model converges to the log-Laplace transform of an NIG process (Barndorff-Nielsen, 1997), which is complex-analytic on a strip around the real axis, thus allowing to apply Theorem 2.5.4.*

The following proposition describes the effect of the time dependent Esscher transform on the dynamics of the Heston model.

Proposition 2.6.3. *Let $\tau = \{t_1, ..., t_n\}$ and $\mathbb{P}_\theta$ the measure given by*

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\sum_{j=1}^n \theta_j X_{t_j}}}{\mathbb{E}\left[e^{\sum_{j=1}^n \theta_j X_{t_j}}\right]} .$$

Under $\mathbb{P}_\theta$, the dynamics of the $\mathbb{P}$-Heston process (X_t, V_t) becomes

$$\begin{aligned} dX_t &= \left(\Theta_{\tau_t} + \zeta\rho\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) - \frac{1}{2}\right) V_t\,dt + \sqrt{V_t}\,d\tilde{W}_t^1\,, \quad X_0 = 0 \\ dV_t &= \tilde{\lambda}_t\left(\tilde{\mu}_t - V_t\right)dt + \zeta\sqrt{V_t}\,d\tilde{W}_t^2\,, \qquad V_0 = V_0 \\ d\left\langle \tilde{W}^1, \tilde{W}^2 \right\rangle_t &= \rho\,dt\,, \end{aligned} \tag{2.6.6}$$

where $\tilde{W}$ is 2-dimensional correlated $\mathbb{P}_\theta$-Brownian motion, $\Theta_j = \sum_{m=j}^n \theta_m$, where Ψ is defined iteratively as

$$\begin{aligned} \Psi\left(s, \Theta_j, ..., \Theta_n\right) &= \psi\left(s, \Theta_j, \Psi\left(t_{j+1} - t_j, \Theta_{j+1} ..., \Theta_n\right)\right) \\ \Psi\left(s\right) &= 0 \end{aligned}$$

and where, denoting $\tau_t = \inf\{s \in \tau \,:\, s \geq t\}$,

$$\tilde{\lambda}_t = \lambda - \zeta\Theta_{\tau_t}\rho - \zeta^2\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) \quad \textit{and} \quad \tilde{\mu}_t = \frac{\lambda\mu}{\tilde{\lambda}_t} .$$

Proof. Denote

$$D(t, X_t, V_t) = \left.\frac{d\mathbb{P}_\theta}{d\mathbb{P}}\right|_{\mathcal{F}_t} .$$

Then

$$\begin{aligned} D(t, X_t, V_t) &= \frac{e^{\sum_{j=1}^{\tau_t - 1} \theta_j X_{t_j}}}{\mathbb{E}\left[e^{\sum_{j=1}^n \theta_j X_{t_j}}\right]}\,\mathbb{E}\left[e^{\sum_{j=\tau_t}^n \theta_j X_{t_j}} \,\middle|\, \mathcal{F}_t\right] \\ &= \frac{e^{\sum_{j=1}^{\tau_t - 1} \theta_j X_{t_j} + \Phi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right)}}{e^{\Phi(t_1, \Theta_1, ..., \Theta_n) + \Psi(t_1, \Theta_1, ..., \Theta_n)\,V_0 + \Theta_1 X_0}}\; e^{\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) V_t + \Theta_{\tau_t} X_t} , \end{aligned}$$

where Φ is defined iteratively as

$$\begin{aligned} \Phi\left(s, \Theta_j, ..., \Theta_n\right) &= \phi\left(s, \Theta_j, \Psi\left(t_{j+1} - t_j, \Theta_{j+1}, ..., \Theta_n\right)\right) \\ &\quad + \Phi\left(t_{j+1} - t_j, \Theta_{j+1}, ..., \Theta_n\right) \\ \Phi\left(s\right) &= 0 . \end{aligned}$$

The dynamics of $D(t, X_t, V_t)$ can then be expressed using Itō's Lemma as

$$dD(t, X_t, V_t) = D(t, X_t, V_t)\left(\Theta_{\tau_t} dX_t + \Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) dV_t\right) + ...\,dt$$

$$= D(t, X_t, V_t)\sqrt{V_t}\left(\Theta_{\tau_t} dW_t^1 + \zeta\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) dW_t^2\right) .$$

By Girsanov's theorem,

$$d\begin{pmatrix} \tilde{W}_t^1 \\ \tilde{W}_t^2 \end{pmatrix} = d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} - \sqrt{V_t}\begin{pmatrix} \Theta_{\tau_t} + \zeta\rho\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) \\ \Theta_{\tau_t}\rho + \zeta\,\Psi\left(\tau_t - t, \Theta_{\tau_t}, ..., \Theta_n\right) \end{pmatrix} dt$$

is a 2-dimensional Brownian motion under the measure P_θ. Replacing W in eq. (2.6.2) by $\tilde{W}$ gives the result. □

Remark 2.6.4. *Prop. 2.6.3 shows that the time-dependent Esscher transform changes a classical Heston process into a Heston process with time-inhomogeneous drift.*

Remark 2.6.5. *Note that Assumption 4 is verified by the Heston model only when $\rho = 0$. Indeed, $J = [u_-, u_+]$, where*

$$u_\pm = \frac{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\rho\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\rho\right)^2 + \frac{\lambda^2}{\zeta^2}(1-\rho^2)}}{(1-\rho^2)} ,$$

while

$$w(u_-) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_-\right) \qquad \text{and} \qquad w(u_+) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_+\right) .$$

However, since the actual variance reduction problem is itself unsolvable, our goal is to find a good candidate measure that we can test numerically. The fact that we do not have the full theory to justify it is therefore not problematic.

Numerical results for European put options

In this case, by Prop. 2.6.1 with $n = 1$ and $t_1 = T$, θ has support on $\{T\}$. Using the abuse of notation $\theta := \theta(\{T\})$, we have

$$\begin{aligned} &\hat{H}(\theta) + \int_0^T h(\theta([t,T]))\,dt \\ = \;& \log\left(\frac{K}{1-\theta}\right) - \theta\log\left(\frac{-\theta\,K/S_0}{1-\theta}\right) + T\,\mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho\,\theta - \frac{\gamma(\theta)}{\zeta}\right) . \end{aligned} \tag{2.6.7}$$

In order to obtain θ, we therefore differentiate (2.6.7) with respect to θ and equate the derivative to 0 by dichotomy .

We simulate $N = 10000$ trajectories of the Heston model with parameters $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$ and initial values $V_0 = 0.04$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (2.6.2) and $\mathbb{P}_\theta$, eq. (2.6.6) with $n = 1$ and $t_1 = T$, using a standard Euler scheme with 200 discretization steps. For the $\mathbb{P}$-realisations

$X^{(i)}$, we calculate the European put price as $\frac{1}{N}\sum_{j=1}^{N}\left(K - S_0\, e^{X_T^{(i)}}\right)_+$ and for the $\mathbb{P}_\theta$-realisations $X^{(i,\theta)}$, as

$$\frac{e^{\phi(T,\theta,0)+\psi(T,\theta,0)\,V_0}}{N}\sum_{j=1}^{N} e^{-\theta\, X_T^{(i,\theta)}}\left(K - S_0 e^{X_T^{(i,\theta)}}\right)_+ . \tag{2.6.8}$$

Each time, we compute the $\mathbb{P}_\theta$-standard deviation, the variance ratio and the adjusted variance ratio, i.e. the variance ratio divided by the ratio of simulation time. The latter measures the actual efficiency of the method, given the fact that simulating under the measure change takes in general slightly more time.

In Table 2.1, we fix the strike to the value $K = 1$ and let the maturity T vary from 0.25 to 3, whereas in Tables 2.2 and 2.3, we fix maturity to $T = 1$ and to $T = 3$, while we let the strike K vary between 0.25 and 1.75. We calculate each time the price, the standard error, the variance ratio adjusted and not adjusted by the ratio of simulation time.

T	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0395	$3.72 \cdot 10^{-4}$	2.46	2.14	20.2
0.5	0.0550	$4.54 \cdot 10^{-4}$	3.12	2.83	19.9
1	0.0780	$5.59 \cdot 10^{-4}$	3.92	3.66	19.5
2	0.111	$7.20 \cdot 10^{-4}$	4.21	3.89	19.7
3	0.134	$8.48 \cdot 10^{-4}$	4.19	3.79	19.8

Table 2.1: The variance ratio as function of the maturity for at-the-money European put options.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.5	0.00014	$7.65 \cdot 10^{-6}$	26.6	24.5	18.4
0.75	0.00794	$1.34 \cdot 10^{-4}$	6.53	5.91	18.7
1	0.0773	$5.60 \cdot 10^{-4}$	3.96	3.65	18.5
1.25	0.261	$8.62 \cdot 10^{-4}$	4.20	3.78	18.9
1.5	0.502	$7.92 \cdot 10^{-4}$	5.84	5.36	18.6
1.75	0.749	$6.84 \cdot 10^{-4}$	8.45	7.29	19.7

Table 2.2: The variance ratio as function of the strike for the European put option with maturity $T = 1$.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	$7.1 \cdot 10^{-5}$	$1.84 \cdot 10^{-5}$	92.0	70.9	23.1
0.5	0.00418	$6.05 \cdot 10^{-5}$	16.1	16.0	20.0
0.75	0.0369	$3.43 \cdot 10^{-4}$	6.67	6.00	20.4
1	0.133	$8.51 \cdot 10^{-4}$	4.24	4.15	20.2
1.25	0.300	$1.34 \cdot 10^{-3}$	3.61	3.13	21.3
1.5	0.517	$1.60 \cdot 10^{-3}$	3.47	3.30	19.9
1.75	0.755	$1.64 \cdot 10^{-3}$	3.89	3.53	19.9

Table 2.3: The variance ratio as function of the strike for the European put option with maturity $T = 3$.

In all the cases, we can see that the variance ratio becomes very interesting when the option gets deeply out of the money and less significant, yet still very interesting, when the option is at or in the money. This corresponds to the natural behaviour of variance reduction techniques that involve measure changes, as the measure change is going to increase the probability of choosing a trajectory that is eventually going to enter the money. Note that the simulation time is only slightly larger when simulating with the measure change, while the time required for the optimization procedure is negligible compared with the simulation time. In Figure 2.6.1, we fix the maturity to $T = 1.5$ and plot the empirical variance of the estimator (2.6.8) as a function of θ. Our method provides $\theta = -0.457$ as asymptotically optimal measure change. We can therefore see that the asymptotically optimal θ is very close to the optimal one.

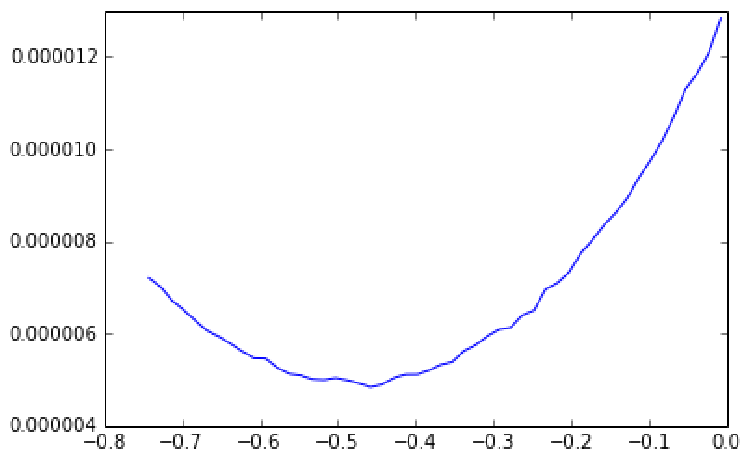

Figure 2.6.1: The variance of the Monte-Carlo estimator as a function of θ.

Numerical results for Asian put options

We now consider the case of a (discretized) Asian put option. Here, the log-payoff is

$$H(X) = \log\left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{X_{t_j}}\right)_+ ,$$

where $t_j = \frac{j}{n}T$. By Prop. 2.5.5, the support of θ is $\{t_1, ..., t_n\}$ and we can denote $\theta_j = \theta(\{t_j\})$. Using Prop. 2.6.1 and eq. (2.6.5), the function that we need to minimize

$$\log\left(\frac{K}{1-\sum_{l=1}^{n}\theta_l}\right) - \sum_{m=1}^{n}\theta_m \log\left(\frac{-\theta_m\, n\, K/S_0}{1-\sum_{l=1}^{n}\theta_l}\right) + \frac{T}{n}\sum_{j=1}^{n} h\left(\sum_{l=j}^{n}\theta_l\right)$$

or, alternatively, denoting $\Theta_j = \sum_{l=j}^{n}\theta_l$,

$$\log\left(\frac{K}{1-\Theta_1}\right) - \sum_{m=1}^{n}(\Theta_m - \Theta_{m+1})\log\left(\frac{-(\Theta_m - \Theta_{m+1})\, nK/S_0}{1-\Theta_1}\right) + \frac{T}{n}\sum_{j=1}^{n} h\left(\Theta_j\right).$$

By differentiating with respect to Θ_j, we obtain, for $j = 2, ..., n$,

$$\begin{aligned} 0 &= \partial_{\Theta_j}\left\{\hat{H}(\theta) + \frac{T}{n}\sum_{m=1}^{n} h\left(\Theta_m\right)\right\} \\ &= \frac{T\, h'\left(\Theta_j\right)}{n} - \log\left[-(\Theta_j - \Theta_{j+1})\right] + \log\left[-(\Theta_{j-1} - \Theta_j)\right] , \end{aligned} \tag{2.6.9}$$

while, for $j = 1$, we have

$$\begin{aligned} 0 &= \partial_{\Theta_1}\left\{\hat{H}(\theta) + \frac{T}{n}\sum_{m=1}^{n} h\left(\Theta_m\right)\right\} \\ &= \log\left(1-\Theta_1\right) - \log(n\,K/S_0) + \frac{T}{n} h'\left(\Theta_1\right) - \log\left[-(\Theta_1 - \Theta_2)\right] . \end{aligned} \tag{2.6.10}$$

Finally, taking the exponential in eqs. (2.6.9) and (2.6.10), we obtain

$$\begin{aligned} \Theta_2 - \Theta_1 &= (1 \quad - \Theta_1 \quad)\, e^{\frac{T}{n}h'(\Theta_1)} \cdot \frac{S_0}{n\,K} \\ \Theta_3 - \Theta_2 &= (\Theta_2 \quad - \Theta_1 \quad)\, e^{\frac{T}{n}h'(\Theta_2)} \\ \vdots \quad &= \quad \vdots \\ \Theta_n - \Theta_{n-1} &= (\Theta_{n-1} - \Theta_{n-2})\, e^{\frac{T}{n}h'(\Theta_{n-1})} \\ -\Theta_n &= (\Theta_n \quad - \Theta_{n-1})\, e^{\frac{T}{n}h'(\Theta_n)} . \end{aligned}$$

Finally, define $\mathcal{T}$ the real function that associates to Θ_n

$$\mathcal{T}(\Theta_n) = (1-\Theta_1)\mathrm{e}^{\frac{T}{n}h'(\Theta_1)} \cdot \frac{S_0}{n\,K} - \Theta_2 - \Theta_1 ,$$

where $\Theta_{n-1} = \Theta_n + \Theta_n \, e^{-\frac{T}{n} h'(\Theta_n)}$ and iteratively,

$$\Theta_{j-2} = \Theta_{j-1} - (\Theta_j - \Theta_{j-1}) \, e^{-\frac{T}{n} h'(\Theta_{j-1})} \,, \qquad j = n, ..., 3 \,.$$

Equating $\mathcal{T}$ to 0 by dichotomy then gives the asymptotically optimal measure.

Again, we simulate $N = 10000$ trajectories of the Heston model with parameters $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$ and initial values $V_0 = 0.04$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (2.6.2) and $\mathbb{P}_\theta$, eq. (2.6.6) with $n = 200$ and $t_j = \frac{j}{n} T$, using a standard Euler scheme with 200 discretization steps. For the $\mathbb{P}$-realisations $X^{(i)}$, we calculate the Asian put price as

$$\frac{1}{N} \sum_{j=1}^{N} \left(K - \frac{S_0}{n} \sum_{j=1}^{n} e^{X^{(i)}_{t_j}} \right)_+ \tag{2.6.11}$$

and for the $\mathbb{P}_\theta$-realisations $X^{(i,\theta)}$, as

$$\frac{e^{\Phi(t_1,\Theta_1,...,\Theta_n)+\Psi(t_1,\Theta_1,...,\Theta_n)\, V_0}}{N} \sum_{j=1}^{N} e^{-\sum_{j=1}^{n} \theta_j \, X^{(i,\theta)}_{t_j}} \left(K - \frac{S_0}{n} \sum_{j=1}^{n} e^{X^{(i)}_{t_j}} \right)_+ . \tag{2.6.12}$$

Again, each time, we compute the $\mathbb{P}_\theta$-standard deviation and the adjusted and non-adjusted variance ratios. In Table 2.4, we fix maturity to $T = 1.5$ and let the strike K vary between 0.6 and 1.3.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.6	$3.466 \cdot 10^{-5}$	$4.13 \cdot 10^{-6}$	16.9	14.6	19.9
0.7	0.000562	$2.60 \cdot 10^{-5}$	5.77	4.77	21.1
0.8	0.00414	$9.64 \cdot 10^{-5}$	4.36	3.77	20.1
0.9	0.0185	0.00024	3.48	3.09	20.6
1	0.0558	0.00043	3.49	3.07	20.1
1.1	0.120	0.00057	3.69	3.20	20.1
1.2	0.206	0.00062	4.27	3.80	19.7
1.3	0.301	0.00059	5.30	4.41	21.0

Table 2.4: The variance ratio as function of the strike for the Asian put option. $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$, $S_0 = 1$, $V_0 = 0.04$, $T = 1.5$, $N = 10000$, 200 discretization steps.

The conclusion is the same as for the European put. Indeed, the variance ratio explodes when the option moves away from the money. Due to the time-dependence of the measure change, the adjusted variance ratio is consistently around 13% below its non-adjusted version. The adjusted variance ratio remains however very interesting, with values above 3 around the money.

2.6.2 European put on the Heston model with negative exponential jumps

We now consider the Heston model with negative exponential jumps

$$\begin{aligned} dX_t &= \left(\delta - \frac{V_t}{2}\right) dt + \sqrt{V_t}\, dW_t^1 + dJ_t\,, \quad && X_0 = 0 \\ dV_t &= \lambda(\mu - V_t)\, dt + \zeta \sqrt{V_t}\, dW_t^2\,, && V_0 = V_0 \\ d\langle W^1, W^2\rangle_t &= \rho\, dt\,, \end{aligned} \tag{2.6.13}$$

where W^1, W^2 are standard $\mathbb{P}$-Brownian motions and $(J_t)_{t\geq 0}$ is an independent compound Poisson process with constant jump rate r and jump distribution Neg-Exp(α), i.e. the Lévy measure of $(J_t)_{t\geq 0}$ is $\nu(dx) = r\,\alpha e^{\alpha x}\mathbb{1}_{\{x<0\}}dx$. The martingale condition on $S = S_0\, e^X$ imposes $\delta = \frac{r}{\alpha+1}$. The Laplace transform of (X_t, V_t) is

$$\mathbb{E}\left(e^{uX_t + wV_t}\right) = e^{\phi(t,u,w)+\psi(t,u,w)V_0+uX_0}\,,$$

where ϕ, ψ satisfy the Riccati equations

$$\begin{aligned} \partial_t\phi(t,u,w) &= F(u, \psi(t,u,w)) \qquad && \phi(0,u,w) = 0 \\ \partial_t\psi(t,u,w) &= R(u, \psi(t,u,w)) && \psi(0,u,w) = w \end{aligned} \tag{2.6.14}$$

for $F(u,w) = \lambda\mu\, w + \tilde{\kappa}(u)$, where $\tilde{\kappa}(u) = \frac{ru(u-1)}{(\alpha+1)(\alpha+u)}$ and

$$R(u,w) = \frac{\zeta^2}{2}\, w^2 + \zeta\rho\, uw - \lambda w + \frac{1}{2}(u^2 - u)\,.$$

Again, a standard calculation shows that the solution of the Generalized Riccati equations (2.6.14) is

$$\begin{aligned} \psi(t,u,w) &= \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \frac{\gamma}{\zeta^2}\, \frac{\tanh\left(\frac{\gamma}{2}\, t\right) + \eta}{1 + \eta\, \tanh\left(\frac{\gamma}{2}\, t\right)} \\ \phi(t,u,w) &= \mu\, \frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) t - 2\mu\frac{\lambda}{\zeta^2}\, \log\Big(\cosh\left(\frac{\gamma}{2} t\right) + \eta\, \sinh\left(\frac{\gamma}{2} t\right)\Big) + t\tilde{\kappa}(u)\,, \end{aligned} \tag{2.6.15}$$

where $\gamma = \gamma(u) = \zeta\sqrt{\left(\frac{\lambda}{\zeta} - \rho u\right)^2 + \frac{1}{4} - \left(u - \frac{1}{2}\right)^2}$ and $\eta = \eta(u,w) = \frac{\lambda - \zeta\rho u - \zeta^2 w}{\gamma(u)}$. Furthermore, for the Heston model, the function h is given by

$$h(u) = \mu\, \frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \mu\, \frac{\lambda}{\zeta^2}\, \gamma(u) + \tilde{\kappa}(u)\,. \tag{2.6.16}$$

Let us now study the effect of the Esscher transform on the dynamics of the Heston model with jumps.

Proposition 2.6.6. *Let $\mathbb{P}_\theta$ be the measure given by*

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\theta X_T}}{\mathbb{E}\left[e^{\theta X_T}\right]} .$$

Under $\mathbb{P}_\theta$, the dynamics of the $\mathbb{P}$-Heston process (X_t, V_t) becomes

$$\begin{aligned} dX_t &= \delta dt + \left(\theta + \zeta\rho\,\psi\,(T-t,\theta,0) - \frac{1}{2}\right) V_t\,dt + \sqrt{V_t}\,d\tilde{W}^1_t + dJ_t\,, & X_0 = 0 \\ dV_t &= \tilde{\lambda}_t\,(\tilde{\mu}_t - V_t)\,dt + \zeta\sqrt{V_t}\,d\tilde{W}^2_t\,, & V_0 = V_0 \\ d\left\langle \tilde{W}^1, \tilde{W}^2 \right\rangle_t &= \rho\,dt\,, \end{aligned} \tag{2.6.17}$$

where $\tilde{W}$ is 2-dimensional correlated $\mathbb{P}_\theta$-Brownian motion, ϕ and ψ are given in (2.6.15),

$$\tilde{\lambda}_t = \lambda - \zeta\theta\rho - \zeta^2\,\psi\,(T-t,\theta,0) \quad \text{and} \quad \tilde{\mu}_t = \frac{\lambda\mu}{\tilde{\lambda}_t}$$

and $(J_t)_{t\geq 0}$ is a compound Poisson process with jump rate $\frac{r\alpha}{\alpha+\theta}$ and jump distribution Neg-Exp$(\alpha+\theta)$ under $\mathbb{P}_\theta$.

Proof. Denote

$$D(t, X_t, V_t) = \left.\frac{d\mathbb{P}_\theta}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{e^{\phi(T-t,\theta,0)}}{e^{\phi(T,\theta,0)+\psi(T,\theta,0)\,V_0}}\,e^{\psi(T-t,\theta,0)\,V_t+\theta\,X_t} .$$

The dynamics of $D(t, X_t, V_t)$ can then be expressed using Itō's Lemma as

$$\begin{aligned} dD(t, X_t, V_t) &= D(t, X_t, V_t)\,(\theta dX_t + \psi\,(T-t,\theta,0)\,dV_t) + \ldots dt \\ &= D(t, X_t, V_t)\Big[\sqrt{V_t}\,\big(\theta dW^1_t + \zeta\,\psi\,(T-t,\theta,0)\,dW^2_t\big) + \theta\,(\delta dt + dJ_t)\Big] \end{aligned}$$

and Girsanov's theorem then shows that

$$d\begin{pmatrix} \tilde{W}^1_t \\ \tilde{W}^2_t \end{pmatrix} = d\begin{pmatrix} W^1_t \\ W^2_t \end{pmatrix} - \sqrt{V_t}\begin{pmatrix} \theta + \zeta\rho\,\psi\,(T-t,\theta,0) \\ \theta\rho + \zeta\,\psi\,(T-t,\theta,0) \end{pmatrix} dt$$

is a 2-dimensional Brownian motion under the measure P_θ. Replacing W in eq. (2.6.2) by $\tilde{W}$ gives eq. (2.6.17). In order to finish the proof, it remains to show that the jump process $(J_t)_{t\geq 0}$ has the desired distribution under $\mathbb{P}_\theta$. Let us calculate the $\mathbb{P}_\theta$-Laplace transform of J_t.

$$\begin{aligned} \mathbb{E}^{\mathbb{P}_\theta}\left[e^{uJ_t}\right] &= \frac{\mathbb{E}\left[e^{uJ_t}\,\mathbb{E}\left[e^{\theta X_T}\,\middle|\,\mathcal{F}_t\right]\right]}{\mathbb{E}\left[e^{\theta X_T}\right]} \\ &= \frac{e^{\phi(T-t,\theta,0)}}{\mathbb{E}\left[e^{\theta X_T}\right]}\mathbb{E}\left[e^{uJ_t+\psi(T-t,\theta,0)\,V_t+\theta X_t}\right] . \end{aligned}$$

By independence of the jumps,

$$\mathbb{E}\left[e^{uJ_t+\psi(T-t,\theta,0)\,V_t+\theta X_t}\right] = e^{\theta\delta\,t}\,\mathbb{E}\left[e^{(u+\theta)J_t}\right]\,\mathbb{E}\left[e^{\psi(T-t,\theta,0)\,V_t+\theta(X_t-\delta\,t-J_t)}\right].$$

But $\mathbb{E}\left[e^{(u+\theta)J_t}\right] = e^{-rt\frac{u+\theta}{u+\theta+\alpha}}$. Furthermore, $(X_t-\delta\,t-J_t,\,V_t)_{t\geq 0}$ is a standard Heston process without jump. Therefore comparing (2.6.4) and (2.6.15), we find that

$$\mathbb{E}\left[e^{\psi(T-t,\theta,0)\,V_t+\theta(X_t-\delta\,t-J_t)}\right] = e^{\phi(t,\theta,\psi(T-t,\theta,0))-t\frac{r\theta(\theta-1)}{(\alpha+1)(\alpha+\theta)}+\psi(t,\theta,\psi(T-t,\theta,0))\,V_0}.$$

Using the fact that $\psi(t,\theta,\psi\,(T-t,\theta,0)) = \psi\,(T,\theta,0)$ and

$$\phi\,(T-t,\theta,0)+\phi(t,\theta,\psi\,(T-t,\theta,0)) = \phi\,(T,\theta,0)$$

(see eq. (2.1) in (Keller-Ressel, 2011)), we finally obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{P}_\theta}\left[e^{uJ_t}\right] &= e^{\theta\delta\,t-rt\frac{u+\theta}{u+\theta+\alpha}-t\frac{r\theta(\theta-1)}{(\alpha+1)(\alpha+\theta)}}\\ &= e^{\theta\frac{r}{\alpha+1}\,t-rt\frac{u+\theta}{u+\theta+\alpha}-t\frac{r\theta(\theta-1)}{(\alpha+1)(\alpha+\theta)}} = e^{-\frac{r\alpha}{\alpha+\theta}\,t\,\frac{u}{u+(\alpha+\theta)}},\end{aligned}$$

which is indeed the Laplace transform of a compound Poisson process with jump rate $\frac{r\alpha}{\alpha+\theta}$ and Neg-Exp$(\alpha+\theta)$-distributed jumps. □

Numerical results for the European put option

Similarly to the case of the Heston model without jump, denoting $\theta=\theta(\{T\})$, we have

$$\begin{aligned}&\hat{H}(\theta)+\int_0^T h(\theta([t,T]))\,dt\\ &= \log\left(\frac{K}{1-\theta}\right)-\theta\log\left(\frac{-\theta\,K/S_0}{1-\theta}\right)+T\,\mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta}-\rho\theta-\frac{\gamma(\theta)}{\zeta}\right)+T\,\tilde{\kappa}(\theta)\end{aligned} \tag{2.6.18}$$

and we obtain the asymptotically optimal θ by differentiate (2.6.18) with respect to θ and equating the derivative to 0 by dichotomy .

We simulate $N = 10000$ trajectories of the Heston model with jumps with parameters $\lambda = 1.1$, $\mu = 0.7$, $\zeta = 0.3$, $\rho = -0.5$, $r = 2$, $\alpha = 3$ and initial values $V_0 = 1.3$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (2.6.13) and $\mathbb{P}_\theta$, eq. (2.6.17) using a standard Euler scheme with 200 discretization step. For the $\mathbb{P}$-realisations $X^{(i)}$, we calculate the standard Monte-Carlo estimator of the European put price and for the $\mathbb{P}_\theta$-realisations $X^{(i,\theta)}$, we use (2.6.8) where ϕ and ψ are given in (2.6.15) and compute the same statistics as in the previous examples. In Table 2.5, we fix the strike to the value $K = 1$ and let the maturity T vary from 0.25 to 3, whereas in Tables 2.6 and 2.7, we fix maturity to $T = 1$ and to $T = 3$, while we let the strike K vary between 0.25 and 1.75.

T	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0945	$9.96 \cdot 10^{-4}$	3.28	3.00	23.6
0.5	0.147	$1.28 \cdot 10^{-3}$	3.20	2.99	24.5
1	0.215	$1.61 \cdot 10^{-3}$	2.95	2.77	24.7
2	0.309	$2.04 \cdot 10^{-3}$	2.61	2.43	24.7
3	0.374	$2.30 \cdot 10^{-3}$	2.40	2.20	25.0

Table 2.5: The variance ratio as function of the maturity for the European put option on the Heston model with jumps.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.00606	$7.83 \cdot 10^{-5}$	11.6	10.4	25.8
0.5	0.0377	$4.03 \cdot 10^{-4}$	5.42	5.28	24.7
0.75	0.105	$9.44 \cdot 10^{-4}$	3.76	3.19	27.3
1	0.215	$1.61 \cdot 10^{-3}$	2.93	2.89	26.1
1.25	0.369	$2.26 \cdot 10^{-3}$	2.65	2.46	25.4
1.5	0.550	$2.80 \cdot 10^{-3}$	2.43	2.24	24.9
1.75	0.766	$3.05 \cdot 10^{-3}$	2.57	2.44	24.6

Table 2.6: The variance ratio as function of the strike for the European put option with maturity $T = 1$ in the Heston model with jumps.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0280	$2.69 \cdot 10^{-4}$	5.19	4.99	24.8
0.5	0.108	$8.60 \cdot 10^{-4}$	3.32	3.05	25.1
0.75	0.226	$1.58 \cdot 10^{-3}$	2.68	2.56	26.3
1	0.374	$2.31 \cdot 10^{-3}$	2.39	2.20	27.0
1.25	0.545	$3.01 \cdot 10^{-3}$	2.20	2.19	25.2
1.5	0.730	$3.66 \cdot 10^{-3}$	2.09	1.94	24.6
1.75	0.932	$4.27 \cdot 10^{-3}$	1.97	1.83	24.8

Table 2.7: The variance ratio as function of the strike for the European put option with maturity $T = 3$ in the Heston model with jumps.

When adding negative jumps to the Heston model, one can see that the variance ratio diminishes. When the options are out of the money however it is still sufficiently important to make it interesting to use in applications. In Figure 2.6.2, we fix the maturity to $T = 1.5$ and plot again the empirical variance of the estimator (2.6.8) as a function of θ for the Heston model with jumps. The method provides $\theta = -0.312$ as asymptotically optimal measure change which is, as in the continuous case very close to the optimal one.

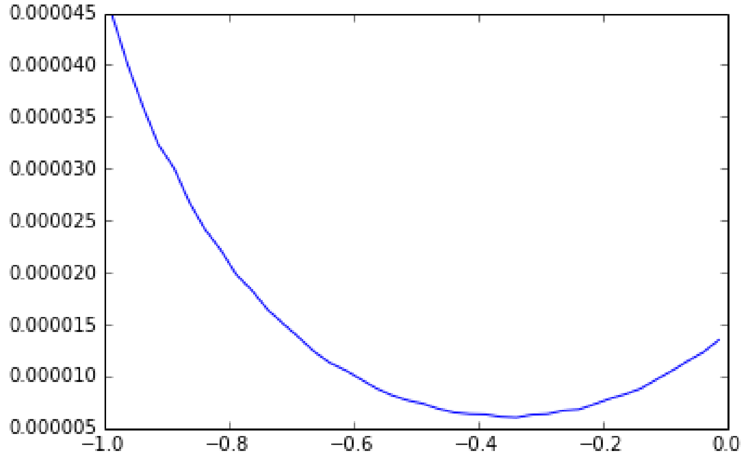

Figure 2.6.2: The variance of the Monte-Carlo estimator as a function of θ for the Heston model with jumps.

Chapter 3

Long-time large deviations for the multi-asset Wishart stochastic volatility model and option pricing

The content of this chapter is based on a paper written in collaboration with Aurélien Alfonsi.[1]

3.1 Introduction

The Heston stochastic volatility model (Heston, 1993) is one of the most popular models in quantitative finance. The Wishart stochastic volatility model is its natural extension to a basket of assets, since it coincides with the Heston model in dimension 1 and keeps the affine structure. This model, proposed in (Gourieroux and Sufana, 2004), assumes that under the risk-neutral probability, the vector of n asset prices is modelled as a diffusion process

$$dS_t = \mathrm{Diag}(S_t)\left(r\mathbf{1}\,dt + \tilde{X}_t^{1/2}\,d\tilde{Z}_t\right), \tag{3.1.1}$$

where the $n \times n$ volatility matrix $(\tilde{X}_t)$ is modelled by a Wishart process with dynamics

$$d\tilde{X}_t = \left(\alpha\, a^\top a + \tilde{b}\tilde{X}_t + \tilde{X}_t\tilde{b}^\top\right) dt + \tilde{X}_t^{1/2}\,d\tilde{W}_t\, a + a^\top (d\tilde{W}_t)^\top \tilde{X}_t^{1/2}, \tag{3.1.2}$$

where $\tilde{Z}$ and $\tilde{W}$ are independent standard Brownian motions of dimensions n and $n \times n$, and $\mathrm{Diag}(S_t)$ is the diagonal matrix whose diagonal elements are

[1] Université Paris-Est, Cermics (ENPC), INRIA, F-77455 Marne-la-Vallée, France.

given by the vector $S_t \in \mathbb{R}^n$. Then, this model has been extended by (Da Fonseca et al., 2007) to include a constant correlation between $\tilde{W}$ and $\tilde{Z}$ in a way to preserve the affine structure. The matrix process (3.1.2) has been introduced by (Bru, 1991) to model the perturbation of experimental biological data. As shown by (Bru, 1991) and (Cuchiero et al., 2011) in a more general framework, for $\alpha \geq n+1$ (resp. $\alpha \geq n-1$), the SDE (3.1.2) has a unique strong (resp. weak) solution. Furthermore, since $\tilde{X}_t$ is positive semi-definite (Bru, 1991, Prop. 4), Wishart processes turn out to be very suitable processes to model covariance matrices. This led several authors to use them in stochastic volatility models, such as (Da Fonseca et al., 2008) and (Benabid et al., 2008) for single asset models and the Wishart stochastic volatility model for multiple assets models. By using the affine property, the Laplace transform of the latter model is given by (Da Fonseca et al., 2007).

$$\mathbb{E}\left(e^{\theta^\top \log(S_t)}\right) = \exp\left(\beta_\theta(t) + \mathrm{Tr}\left[\gamma_\theta(t)\,\tilde{X}_0\right] + \delta_\theta^\top(t)\log(S_t)\right)\,, \tag{3.1.3}$$

where $\beta_\theta, \gamma_\theta$ and δ_θ satisfy the matrix Riccati equations

$$\begin{aligned}
&\partial_t\beta_\theta(t) = r\,\delta_\theta^\top(t)\,\mathbf{1} + \alpha\,\mathrm{Tr}\left[\gamma_\theta(t)\right]\\
&\partial_t\gamma_\theta(t) = \tilde{b}^\top\gamma_\theta(t) + \gamma_\theta(t)\,\tilde{b} + 2\gamma_\theta(t)\,a^\top a\,\gamma_\theta(t) - \frac{1}{2}\left(\mathrm{Diag}(\delta_\theta(t)) - \delta_\theta(t)\delta_\theta^\top(t)\right)\\
&\partial_t\delta_\theta(t) = 0\,,
\end{aligned}$$

with initial conditions $\beta_\theta(0) = 0$, $\gamma_\theta(0) = 0$ and $\delta_\theta(0) = \theta$. Since the Riccati equations can be solved explicitly, the Laplace transform can be calculated explicitly by the mean of exponential and inversion of matrices.

In the last decade, many works have studied asymptotics of the option prices under the Heston model and extensions through the volatility smile function. In particular, (Forde and Jacquier, 2011a) and (Jacquier et al., 2013) have obtained long-time asymptotics by proving a large deviations principle. The goal of the present paper is to extend these results to the Wishart stochastic volatility model (3.1.1) and (3.1.2). Even though the Laplace transform (3.1.3) is given by an explicit formula, it is not easy to calculate long-time asymptotics because of the multi-dimensional setting. Nonetheless, under some assumptions on the coefficients, we can get a simpler formula for the Laplace transform and then prove a large deviations principle. Then, we obtain asymptotics for the smile when the maturity goes to infinity.

Beyond its theoretical interest, this large deviations principle enable us to develop a generic variance reduction method for pricing derivatives. First, let us note that since the Laplace tranform is known explicitly, Fourier inversion methods can be used, as explained in (Da Fonseca et al., 2007). However, Fourier inversion methods are less competitive than in dimension 1 since they require to approximate an integral on $\mathbb{R}^n$. When, for complexity reasons, Fourier methods are not an option, the use of a large number of Monte-Carlo simulations is necessary. In (Ahdida and Alfonsi, 2013), it is given an exact simulation method for Wishart processes and a second order scheme for the Gourieroux and Sufana model (3.1.1) and (3.1.2). Thus, it is possible to sample efficiently such processes, and it is relevant to develop variance reduction techniques to reduce computational costs. Following previous works of (Guasoni and Robertson, 2008), (Robertson, 2010), (Genin and Tankov, 2016) and (Grbac et al., 2018), we develop an importance sampling method based on an asymptotically optimal Esscher transform, using large deviations theory.

In this paper, we denote $\mathcal{M}_n$ the set of real squared $n \times n$ matrices, $\mathcal{S}_n \subset \mathcal{M}_n$ the set of symmetric matrices and $\mathcal{S}_n^+$, (resp. $\mathcal{S}_n^{+,*}$), the sets of symmetric and positive semi-definite (resp.) positive definite. The paper is structured as follows.

In Section 3.2, we describe the model, make certain assumptions on the parameters and give some properties of the model. In Section 3.3, we prove that the asset log-price vector satisfies large deviations principle when maturity goes to infinity. In Section 3.4, we calculate the asymptotic put basket implied volatility, following the approach of (Jacquier et al., 2013). In Section 3.5, we develop the variance reduction method using Varadhan's lemma. Finally, in Section 3.6, we test numerically the results of Sections 3.4 and 3.5.

3.2 The Wishart stochastic volatility model

Let $(S_t)_{t\geq 0}$ be a n-dimensional vector stochastic process with dynamics

$$dS_t = \mathrm{Diag}(S_t)\left(r\mathbf{1}\,dt + a^\top X_t^{1/2}\,dZ_t\right), \qquad S_0^i > 0,\ i = 1,\dots,n, \tag{3.2.1}$$

where $\mathbf{1} = (1,...,1)^\top$, $\mathrm{Diag}(S_t)_{ij} = \mathbb{1}_{\{i=j\}}S_t^i$, Z_t is n-dimensional standard Brownian motion and the stochastic volatility matrix X is a Wishart process with dynamics

$$dX_t = (\alpha I_n + bX_t + X_t b)\,dt + X_t^{1/2}\,dW_t + (dW_t)^\top X_t^{1/2}, \qquad X_0 = x\,. \tag{3.2.2}$$

with $\alpha > n-1$, $a \in \mathcal{M}_n$ invertible, $-b, x \in \mathcal{S}_n^{+,*}$ and W is a $n \times n$ matrix standard Brownian motion independent of Z. Note again that $X_t \in \mathcal{S}_n^+$ (Bru, 1991, Prop. 4). Let us also assume that a is such that $a^\top a \in \mathcal{S}_n^{+,*}$.

Remark 3.2.1. *The model (S, X) defined in (3.2.1) and (3.2.2) is a (quite large) subclass of the one defined in (3.1.1) and (3.1.2). Indeed, defining $\tilde{X}_t := a^\top X_t\, a$, we have $a^\top X_t^{1/2}\,dZ_t = \tilde{X}_t^{1/2}\,d\tilde{Z}_t$, where $\tilde{Z}_t$ is another n-dimensional standard Brownian motion and*

$$d\tilde{X}_t = \left(\alpha\, a^\top a + \tilde{b}\tilde{X}_t + \tilde{X}_t\tilde{b}^\top\right)\,dt + \tilde{X}_t^{1/2}\,d\tilde{W}_t\, a + a^\top (d\tilde{W}_t)^\top \tilde{X}_t^{1/2}, \quad \tilde{X}_0 = a^\top x\, a,$$

where $\tilde{b} = a^\top b\,(a^\top)^{-1}$ and $\tilde{W}_t$ is another $n \times n$-Brownian motion.

Remark 3.2.2. *In dimension one, the model defined by eqs.* (3.2.1) *and* (3.2.2) *corresponds to the famous Heston model (Heston, 1993) and b being negative definite yields the mean reversion property of the stochastic volatility process.*

Defining the log-price $Y_t^k := \log(S_t^k)$, $k = 1, ..., n$, a simple application of Itō's lemma gives

$$dY_t = \left(r\mathbf{1} - \frac{1}{2} \left((a^\top X_t\, a)_{11} \,,\, \dots \,,\, (a^\top X_t\, a)_{nn} \right)^\top \right) dt + a^\top X_t^{1/2}\, dZ_t \,. \qquad (3.2.3)$$

We are interested in the Laplace transform of Y_t. In order to calculate it, we first cite the following proposition.

Proposition 3.2.3. (Alfonsi et al., 2016, Prop. 5.1.)**.** *Let $\alpha \geq n-1$, $x \in \mathcal{S}_n^+$, $b \in \mathcal{S}_n$ and X with dynamics* (3.2.2)*. Let $v, w \in \mathcal{S}_n$ be such that*

$$\exists m \in \mathcal{S}_n, \quad \frac{v}{2} - mb - bm - 2m^2 \in \mathcal{S}_n^+ \quad \text{and} \quad \frac{w}{2} + m \in \mathcal{S}_n^+ \,.$$

If $R_t := \int_0^t X_s\, ds$, then we have for $t \geq 0$

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\frac{1}{2}\, Tr\,[wX_t] - \frac{1}{2}\, Tr\,[vR_t] \right) \right] \\ &\quad = \frac{\exp\left(-\frac{\alpha}{2}\, Tr\,[b]\, t\right)}{\det\,[V_{v,w}(t)]^{\alpha/2}} \exp\left(-\frac{1}{2}\, Tr\left[\left(V'_{v,w}(t) V^{-1}_{v,w}(t) + b \right) x \right] \right) \,, \end{aligned}$$

with

$$V_{v,w}(t) = \left(\sum_{k=0}^{\infty} t^{2k+1} \frac{\tilde{v}^k}{(2k+1)!} \right) \tilde{w} + \sum_{k=0}^{\infty} t^{2k} \frac{\tilde{v}^k}{(2k)!}, \qquad \tilde{v} = v + b^2 \quad \text{and} \quad \tilde{w} = w - b.$$

If besides, $\tilde{v} \in \mathcal{S}_n^{+,*}$, then

$$V_{v,w}(t) = \tilde{v}^{-1/2} \sinh\left(\tilde{v}^{1/2} t \right) \tilde{w} + \cosh\left(\tilde{v}^{1/2} t \right)$$

and

$$V'_{v,w}(t) = \cosh\left(\tilde{v}^{1/2} t \right) \tilde{w} + \sinh\left(\tilde{v}^{1/2} t \right) \tilde{v}^{1/2} \,.$$

Proposition 3.2.4. *Let $\phi : \mathbb{R}^n \to \mathcal{S}_n$ be the function defined by*

$$\phi(\theta) := b^2 + a\, \left(Diag(\theta) - \theta\theta^\top \right) a^\top \in \mathcal{S}_n \,, \qquad (3.2.4)$$

Let $\mathcal{U} \subset \mathbb{R}^n$, be the set defined by

$$\mathcal{U} := \left\{ \theta \in \mathbb{R}^n \;:\; \phi(\theta) \in \mathcal{S}_n^+ \right\}.$$

Then, for all $\theta \in \mathcal{U}$, the Laplace transform of Y_t is

$$\mathbb{E}\left(e^{\theta^\top Y_t} \right) = \frac{e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\, t - \frac{\alpha}{2}\, Tr[b] t - \frac{1}{2}\, Tr\left[\left(b + \phi^{1/2}(\theta) \right) x - \exp\left(-t\, \phi^{1/2}(\theta) \right) \left(b + \phi^{1/2}(\theta) \right) V^{-1}(t)\, x \right]}}{\det\,[V(t)]^{\alpha/2}} \,,$$

where

$$V(t) = \cosh\left(t\, \phi^{1/2}(\theta) \right) - \phi^{-1/2}(\theta)\, \sinh\left(t\, \phi^{1/2}(\theta) \right)\, b\,.$$

Proof. By conditioning on the trajectory of X, we have

$$\mathbb{E}\left(e^{\theta^\top Y_t}\right) = \mathbb{E}\left(\mathbb{E}\left(e^{\theta^\top Y_t}\,\middle|\,(X_s)_{s\le t}\right)\right) ,$$

where

$$\begin{aligned}\mathbb{E}\left(e^{\theta^\top Y_t}\,\middle|\,(X_s)_{s\le t}\right) &= e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\,t - \frac{1}{2}\int_0^t \theta^\top\left((a^\top X_s\, a)_{11}\,,\dots,(a^\top X_s\, a)_{nn}\right)^\top - \theta^\top a^\top X_s\, a\,\theta\, ds}\\ &= e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\,t - \frac{1}{2}\int_0^t \mathrm{Tr}\left[\mathrm{Diag}(\theta)\, a^\top X_s\, a\right] - \mathrm{Tr}\left[\theta^\top a^\top X_s\, a\,\theta\right] ds}\\ &= e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\,t - \frac{1}{2}\mathrm{Tr}\left[a\left(\mathrm{Diag}(\theta)-\theta\theta^\top\right)a^\top R_t\right]} .\end{aligned}$$

Let $m = -b/2$. Then $m \in \mathcal{S}_n^+$ and

$$\frac{a\,\left(\mathrm{Diag}(\theta)-\theta\theta^\top\right)a^\top}{2} - mb - bm - 2m^2 = \frac{\phi(\theta)}{2} \in \mathcal{S}_n^+ .$$

Therefore, by Proposition 3.2.3,

$$\begin{aligned}\mathbb{E}\left(e^{\theta^\top Y_t}\right) &= e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\,t}\,\mathbb{E}\left(e^{-\frac{1}{2}\mathrm{Tr}\left[a\left(\mathrm{Diag}(\theta)-\theta\theta^\top\right)a^\top R_t\right]}\right)\\ &= e^{\theta^\top Y_0 + r\theta^\top \mathbf{1}\,t}\,\frac{\exp\left(-\frac{\alpha}{2}\mathrm{Tr}\left[b\right] t\right)}{\det\left[V(t)\right]^{\alpha/2}}\,\exp\left(-\frac{1}{2}\mathrm{Tr}\left[\left(V'(t)V^{-1}(t)+b\right)x\right]\right)\end{aligned} \tag{3.2.5}$$

where

$$\begin{cases} V(t) &= \cosh\left(t\,\phi^{1/2}(\theta)\right) - \phi^{-1/2}(\theta)\,\sinh\left(t\,\phi^{1/2}(\theta)\right)\,b\,,\\ V'(t) &= \sinh\left(t\,\phi^{1/2}(\theta)\right)\,\phi^{1/2}(\theta) - \cosh\left(t\,\phi^{1/2}(\theta)\right)\,b\,.\end{cases}$$

Since $\phi(\theta) \in \mathcal{S}_n^+$, we can write $\phi(\theta) = PDP^\top$, where D is diagonal, P is orthonormal and $\hat{b} = -P^\top b\,P \in \mathcal{S}_n^{+,*}$.

$$\begin{cases} V(t) &= P\left(\cosh\left(t\,D^{1/2}\right) + \sinh\left(t\,D^{1/2}\right)D^{-1/2}\,\hat{b}\right)P^\top\,,\\ V'(t) &= P\left(\sinh\left(t\,D^{1/2}\right)D^{1/2} + \cosh\left(t\,D^{1/2}\right)\,\hat{b}\right)P^\top\\ &= \phi^{1/2}(\theta)V(t) - \exp\left(-t\,\phi^{1/2}(\theta)\right)\left(b+\phi^{1/2}(\theta)\right)\,.\end{cases}$$

Replacing V' by the latter expression finishes the proof. □

Remark 3.2.5. *Note that, when $\phi(\theta) \in \mathcal{S}_n^+\backslash\mathcal{S}_n^{+,*}$, $\phi^{1/2}(\theta)$ is not invertible. The notation $\phi^{-1/2}(\theta)\,\sinh\left(t\,\phi^{1/2}(\theta)\right)$ is therefore abusive and is to be interpreted as the finite limit*

$$\lim_{\mathcal{S}_n^{+,*}\ni\phi\to\phi(\theta)} \phi^{-1/2}\,\sinh\left(t\,\phi^{1/2}\right) = \sum_{k=0}^{\infty}\frac{\phi(\theta)^k t^{2k+1}}{(2k+1)!} .$$

Remark 3.2.6. *The set $\mathcal{U}$ is bounded. Indeed, let $\theta = \lambda\bar{\theta}$, with $\lambda > 0$ and $\|\bar{\theta}\| = 1$. Then, letting $u = (a^\top)^{-1}\bar{\theta}$, we have*

$$u^\top\phi(\theta)u = \|b(a^\top)^{-1}\bar{\theta}\|^2 + \lambda\bar{\theta}^\top \mathit{Diag}(\bar{\theta})\bar{\theta} - \lambda^2 \le \|b(a^\top)^{-1}\|^2 + \lambda - \lambda^2$$

It follows that $\mathcal{U}$ is contained, e.g., in the set $\|\theta\| \le \lambda^*$ with

$$\lambda^* = \max\{2, \|b(a^\top)^{-1}\bar{\theta}\|\sqrt{2}\}.$$

3.3 Long-time large deviations for the Wishart volatility model

In this section, we prove that the Wishart stochastic volatility model satisfies a large deviation principle when time tends to infinity.

3.3.1 Reminder of large deviations theory

Let us recall some standard definitions and results of large deviations theory. For a wider overview of large deviations theory, we refer the reader to (Dembo and Zeitouni, 1998). We consider a family $(X^\epsilon)_{\epsilon\in(0,1]}$ of random variables on a measurable space $(\mathcal{X},\mathcal{B})$, where $\mathcal{X}$ is a topological space.

Definition 3.3.1 (Rate function). *A rate function Λ^* is a lower semi continuous mapping $\Lambda^* : \mathcal{X} \to [0,\infty]$. A good rate function is a rate function such that, for every $a \in [0,\infty]$, $\{x : \Lambda^*(x) \le a\}$ is compact.*

Definition 3.3.2 (Large deviation principle). *$(X^\epsilon)_{(0,1]}$ satisfies a large deviation principle with rate function Λ^* if, for every $A \in \mathcal{B}$, denoting $\overset{\circ}{A}$ and $\bar{A}$ the interior and the closure of A,*

$$-\inf_{x\in\overset{\circ}{A}} \Lambda^*(x) \le \liminf_{\epsilon\to 0} \epsilon \log \mathbb{P}(X^\epsilon \in A) \le \limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(X^\epsilon \in A) \le -\inf_{x\in\bar{A}} \Lambda^*(x)\,.$$

Definition 3.3.3. *Let $f : \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ be a convex function with domain $\mathcal{D} := \{x \in \mathbb{R}^n : f(x) < \infty\}$. f is called **essentially smooth** if f is differentiable on $\overset{\circ}{\mathcal{D}} \neq \emptyset$ and for every $x \in \bar{\mathcal{D}}\backslash\overset{\circ}{\mathcal{D}}$, $\lim_{y\to x} ||\nabla f(y)|| = +\infty$.*

Theorem 3.3.4 (Gärtner-Ellis). *Let $(X^\epsilon)_{\epsilon\in(0,1]}$ be a family of random vectors in $\mathbb{R}^n$. Assume that for each $\lambda \in \mathbb{R}^n$,*

$$\Lambda(\lambda) := \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{\langle\lambda, X^\epsilon\rangle}{\epsilon}}\right] \tag{3.3.1}$$

exists as an extended real number. Assume also that 0 belongs to the interior of $D_\Lambda := \{\lambda \in \mathbb{R}^n : \Lambda(\lambda) < \infty\}$. Denoting

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}^n} \langle\lambda, x\rangle - \Lambda(\lambda)\,,$$

the Fenchel-Legendre transform of Λ, the following hold.

(a) For any closed set F,

$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(X^\epsilon \in F) \le -\inf_{x\in F} \Lambda^*(x)\,.$$

(b) For any open set G,

$$\liminf_{\epsilon \to 0} \epsilon \log \mathbb{P}(X^\epsilon \in G) \geq - \inf_{x \in G \cap \mathcal{F}} \Lambda^*(x) \,,$$

where $\mathcal{F}$ is the set of exposed points of Λ^*, whose exposing hyperplane belongs to the interior of D_Λ.

(c) If Λ is an essentially smooth, lower semi-continuous function, then the family $(X^\epsilon)_{(0,1]}$ satisfies a large deviations principle with good rate function Λ^*.

Remark 3.3.5. *The function Λ of (3.3.1) is a convex function. Indeed, let $\lambda, \mu \in \mathbb{R}^n$ and $u \in (0,1)$. A direct application of Hölder's inequality yields*

$$\mathbb{E}\left[e^{\frac{\langle u\lambda + (1-u)\mu, X^\epsilon \rangle}{\epsilon}}\right] = \mathbb{E}\left[e^{\frac{\langle u\lambda, X^\epsilon \rangle}{\epsilon}} e^{\frac{\langle (1-u)\mu, X^\epsilon \rangle}{\epsilon}}\right] \leq \left(\mathbb{E}\left[e^{\frac{\langle \lambda, X^\epsilon \rangle}{\epsilon}}\right]\right)^u \left(\mathbb{E}\left[e^{\frac{\langle \mu, X^\epsilon \rangle}{\epsilon}}\right]\right)^{1-u}.$$

Applying the logarithm then proves that $\lambda \mapsto \log \mathbb{E}\left[e^{\frac{\langle \lambda, X^\epsilon \rangle}{\epsilon}}\right]$ and therefore Λ are convex.

Theorem 3.3.6 (Varadhan's Lemma, extension of (Guasoni and Robertson, 2008))**.** *Let $(\mathcal{X}, \mathcal{B})$ be a metric space with its Borel σ-field. Let $(X^\epsilon)_{\epsilon \in]0,1]}$ be a family of $\mathcal{X}$-valued random variables that satisfies a large deviations principle with rate function Λ^*. If $\varphi : \mathcal{X} \to \mathbb{R} \cup \{-\infty\}$ is a continuous function which satisfies*

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{\gamma\, \varphi(X^\epsilon)}{\epsilon}\right)\right] < \infty$$

for some $\gamma > 1$, then, for any $A \in \mathcal{B}$,

$$\begin{aligned} \sup_{x \in A^\circ} \{\varphi(x) - \Lambda^*(x)\} &\leq \liminf_{\epsilon \to 0} \epsilon \log \int_{A^\circ} \exp\left(\frac{\varphi(z)}{\epsilon}\right) d\mu_\epsilon(z) \\ &\leq \limsup_{\epsilon \to 0} \epsilon \log \int_{\bar{A}} \exp\left(\frac{\varphi(z)}{\epsilon}\right) d\mu_\epsilon(z) = \sup_{x \in \bar{A}} \{\varphi(x) - \Lambda^*(x)\} \,, \end{aligned}$$

where μ^ϵ denotes the law of X^ϵ

3.3.2 Long-time behaviour of the Laplace transform of the log-price

Let $T > 0$ and define the transformation $Y_T^\epsilon := \epsilon Y_{T/\epsilon}$, which corresponds to the long-time behaviour of Y_T. We are interested in the function

$$\theta \mapsto \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_T^\epsilon}\right] \,.$$

We first give the following lemma.

Lemma 3.3.7. *Let $A, B \in \mathcal{M}_n$ such that $A + tB$ is invertible for all $t \geq t_0$. Then, $(A + tB)^{-1}tB$ is bounded for all sufficiently large t.*

Proof. Since $A + t_0B$ is invertible, for all $t \geq t_0$,

$$(A+tB)^{-1}tB = \left\{I + (t-t_0)\tilde{B}\right\}^{-1}(t-t_0)\tilde{B}\frac{t}{t-t_0},$$

where $\tilde{B} = (A + t_0B)^{-1}B$. Now, the fact that $A + tB$ is invertible for $t \geq t_0$ means that the eigenvalues λ_i of $\tilde{B}$ satisfy $\lambda_i > 0$ or $\Im\lambda_i \neq 0$ for all i. This implies $\det[I + (t-t_0)\tilde{B}] \underset{t\to+\infty}{\sim} ct^n$ for some $c \neq 0$, and since the adjugate matrix of $I + (t-t_0)\tilde{B}$ has coefficients of order $\mathcal{O}(t^{n-1})$, we get that $\left\{I+(t-t_0)\tilde{B}\right\}^{-1}$ is bounded for $t \geq t_0$. Therefore, $\left\{I+(t-t_0)\tilde{B}\right\}^{-1}(t-t_0)\tilde{B} = I - \left\{I + (t-t_0)\tilde{B}\right\}^{-1}$ is bounded, and $(A + tB)^{-1}tB$ as well, whenever t is sufficiently large. □

We now characterise the asymptotic behaviour of the Laplace transform of Y_t^ϵ.

Proposition 3.3.8. *Define*

$$\Lambda(\theta) := \begin{cases} T\left(r\,\theta^\top\mathbf{1} - \frac{\alpha}{2}\,Tr\left[b + \phi^{1/2}(\theta)\right]\right) & \text{if } \theta\in\mathcal{U} \\ \infty & \text{if } \theta\notin\mathcal{U} \end{cases}. \tag{3.3.2}$$

For every $\theta \in \mathcal{U}$,

$$\lim_{\epsilon\to0}\epsilon\log\mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_T^\epsilon}\right] = \Lambda(\theta)\,.$$

Proof. Let $\theta \in \mathcal{U}$. By Proposition 3.2.4,

$$\begin{aligned}\epsilon\log\mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_T^\epsilon}\right] &= \epsilon\log\mathbb{E}\left[e^{\theta^\top Y_{T/\epsilon}}\right] \\ &= \epsilon\left(\theta^\top Y_0 - \frac{1}{2}\mathrm{Tr}\left[\left(b+\phi^{1/2}(\theta)\right)x\right]\right) \\ &\quad + \frac{1}{2}\epsilon\mathrm{Tr}\left[\exp\left(-T/\epsilon\,\phi^{1/2}(\theta)\right)\left(b+\phi^{1/2}(\theta)\right)V^{-1}(T/\epsilon)\,x\right] \\ &\quad + T\,r\theta^\top\mathbf{1} - \frac{T\,\alpha}{2}\mathrm{Tr}\left[b\right] - \frac{\alpha}{2}\epsilon\log\det\left[V(T/\epsilon)\right]\,.\end{aligned} \tag{3.3.3}$$

Write $\phi(\theta) = PDP^\top$, where D is diagonal, P is orthonormal and let $\hat{b} = -P^\top b\,P \in \mathcal{S}_n^{+,*}$. Then

$$V(t) = P\left(\cosh\left(t\,D^{1/2}\right) + \sinh\left(t\,D^{1/2}\right)D^{-1/2}\,\hat{b}\right)P^\top\,,$$

Let $\mathcal{E}$ and $\tilde{\mathcal{E}}$ be $n \times n$ square matrices with $\mathcal{E}_{ij} = \mathbb{1}_{\{i=j, D_{ii}=0\}}$ and $\tilde{\mathcal{E}}_{ij} = D_{ii}^{-1/2} \mathbb{1}_{\{i=j, D_{ii} \neq 0\}}$. We then have

$$\cosh\left(t\, D^{1/2}\right) = \frac{e^{tD^{1/2}}}{2}\left(I_n + e^{-2tD^{1/2}}\right) = \frac{e^{tD^{1/2}}}{2}\left(I_n + \mathcal{E} + o\left(t^{-1}\right)\right)$$

and

$$\begin{aligned}\sinh\left(tD^{1/2}\right) D^{-1/2} &= \frac{e^{tD^{1/2}}}{2} D^{-1/2}\left(I_n - e^{-2tD^{1/2}}\right)\\ &= \frac{e^{tD^{1/2}}}{2}\left(\tilde{\mathcal{E}} + 2t\mathcal{E} + o\left(t^{-1}\right)\right).\end{aligned}$$

Therefore,

$$\begin{aligned}V(t) &= \frac{1}{2} P e^{tD^{1/2}}\left((I_n + \mathcal{E}) + (2t\mathcal{E} + \tilde{\mathcal{E}})\,\hat{b} + o\left(t^{-1}\right)\right) P^\top\\ &= -\frac{1}{2} P\left(I_n + \mathcal{E}\right) e^{tD^{1/2}}\left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}}) + o\left(t^{-1}\right)\right) P^\top b\end{aligned} \tag{3.3.4}$$

and

$$V^{-1}(t) = -2\, b^{-1} P\left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}}) + o\left(t^{-1}\right)\right)^{-1} e^{-tD^{1/2}}\left(I_n - \frac{1}{2}\mathcal{E}\right) P^\top$$

where the invertibility of $\left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}}) + o\left(t^{-1}\right)\right)$ is guaranteed for every $t \geq 0$ by the existence of the Laplace transform. Since $\hat{b}^{-1} \in \mathcal{S}_n^{+,*}$ and $(t\,\mathcal{E} + \tilde{\mathcal{E}}) \in \mathcal{S}_n^{+}$, $\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}}) \in \mathcal{S}_n^{+,*}$ and is therefore invertible. Hence

$$\left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}}) + o\left(t^{-1}\right)\right) = \left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}})\right)\left(I_n + o\left(t^{-1}\right)\right)$$

and

$$V^{-1}(t) = -2\, b^{-1} P\left(I_n + o\left(t^{-1}\right)\right)\left(\hat{b}^{-1} + (t\,\mathcal{E} + \tilde{\mathcal{E}})\right)^{-1} e^{-tD^{1/2}}\left(I_n - \frac{1}{2}\mathcal{E}\right) P^\top.$$

But

$$\begin{aligned}&\left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1} e^{-tD^{1/2}}\\ &\quad= \left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1}\left(\mathcal{E} + (I_n - \mathcal{E})\right) e^{-tD^{1/2}}\\ &\quad= t^{-1}\left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1} t\mathcal{E} + \left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1}\left(I_n - \mathcal{E}\right) e^{-tD^{1/2}},\end{aligned}$$

where $\left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1} t\mathcal{E}$ is bounded by Lemma 3.3.7. Therefore,

$$\left(\hat{b}^{-1} + \left(t\,\mathcal{E} + \tilde{\mathcal{E}}\right)\right)^{-1} e^{-tD^{1/2}} \to 0$$

and $V^{-1}(t) \to 0$ as $t \to \infty$. Using (3.3.4), we find

$$\begin{aligned}
&\epsilon \log \det \left[V(T/\epsilon)\right] \\
&\quad = T\,\mathrm{Tr}\left[D^{1/2}\right] + \epsilon \log \det \left[\frac{1}{2}\left(I_n + \mathcal{E}\right)\left(I_n + (\epsilon^{-1}T\mathcal{E} + \tilde{\mathcal{E}})\hat{b} + o\left(\epsilon\right)\right)\right] \\
&\quad = T\,\mathrm{Tr}\left[\phi^{1/2}(\theta)\right] + \epsilon \log \det \left[\epsilon^{-1}T\mathcal{E}\hat{b} + \frac{1}{2}\left(I_n + \mathcal{E}\right)\left(I_n + \tilde{\mathcal{E}}\hat{b}\right) + o\left(\epsilon\right)\right] \\
&\quad = T\,\mathrm{Tr}\left[\phi^{1/2}(\theta)\right] - n\epsilon \log(\epsilon) + \epsilon \log \det \left[T\mathcal{E}\hat{b} + \frac{\epsilon}{2}\left(I_n + \mathcal{E} + \tilde{\mathcal{E}}\hat{b}\right) + o\left(\epsilon^2\right)\right].
\end{aligned}$$

We have

$$\det \left[T\mathcal{E}\hat{b} + \frac{\epsilon}{2}\left(I_n + \mathcal{E} + \tilde{\mathcal{E}}\hat{b}\right) + o\left(\epsilon^2\right)\right] \underset{\epsilon \to 0}{\sim} \det \left[T\mathcal{E}\hat{b} + \frac{\epsilon}{2}\left(I_n + \mathcal{E} + \tilde{\mathcal{E}}\hat{b}\right)\right],$$

since the latter determinant is a non-zero polynomial of ϵ (for $\epsilon = 2T$ the determinant is clearly positive). Thus, by passing to the limit,

$$\lim_{\epsilon \to 0} \epsilon \log \det \left[V(T/\epsilon)\right] = T\,\mathrm{Tr}\left[\phi^{1/2}(\theta)\right].$$

Furthermore, since $\phi \in \mathcal{S}_n^+$, $\exp\left(-\frac{T}{\epsilon}\,\phi^{1/2}(\theta)\right)$ is bounded. Therefore,

$$\mathrm{Tr}\left[\exp\left(-\frac{T}{\epsilon}\,\phi^{1/2}(\theta)\right)\left(b + \phi^{1/2}(\theta)\right)V^{-1}(T/\epsilon)\,x\right] \underset{\epsilon \to 0}{\longrightarrow} 0\,.$$

Finally, passing to the limit in (3.3.3) finishes the proof. □

The next proposition proves the essential smoothness of Λ.

Proposition 3.3.9. *The function $\theta \mapsto \Lambda(\theta)$ defined in (3.3.2) is essentially smooth.*

Proof. The function Λ defined in (3.3.2) is a lower semi-continuous proper convex function with domain $\mathcal{U}$. Furthermore, since for every $\theta \in \mathring{\mathcal{U}}$, $\phi(\theta) \in \mathcal{S}_n^{+,*}$, Λ is of class C^1 on $\mathring{\mathcal{U}}$. Only remains to prove that $||\nabla_\theta \Lambda(\theta)|| \to \infty$ when θ goes to the boundary of $\mathcal{U}$. Let $\theta \in \mathring{\mathcal{U}}$. By Proposition 3.3.8

$$\Lambda(\theta) = T\left(r\,\theta^\top \mathbf{1} - \frac{\alpha}{2}\,\mathrm{Tr}\left[b + \phi^{1/2}(\theta)\right]\right).$$

Then for every $j \in \{1, ..., n\}$,

$$\partial_{\theta_j}\Lambda(\theta) = T\left(r - \frac{\alpha}{2}\,\mathrm{Tr}\left[\partial_{\theta_j}\left[\phi^{1/2}\right](\theta)\right]\right),$$

where $\partial_{\theta_j}\left[\phi^{1/2}\right](\theta)$ satisfies

$$\partial_{\theta_j}\phi(\theta) = \partial_{\theta_j}\left[\phi^{1/2}(\theta)\phi^{1/2}(\theta)\right] = \phi^{1/2}(\theta)\,\partial_{\theta_j}\left[\phi^{1/2}\right](\theta) + \partial_{\theta_j}\left[\phi^{1/2}\right](\theta)\,\phi^{1/2}(\theta).$$

Multiplying this equation by $\phi^{-1/2}(\theta)$ and using the cyclic property of the trace, we get

$$\operatorname{Tr}\left[\partial_{\theta_j}\left[\phi^{1/2}\right](\theta)\right] = \frac{1}{2}\operatorname{Tr}\left[\phi^{-1/2}(\theta)\,\partial_{\theta_j}\phi(\theta)\right].$$

and therefore

$$\partial_{\theta_j}\Lambda(\theta) = T\left(r - \frac{\alpha}{2}\operatorname{Tr}\left[\partial_{\theta_j}\left[\phi^{1/2}\right](\theta)\right]\right) = T\left(r - \frac{\alpha}{4}\operatorname{Tr}\left[\phi^{-1/2}(\theta)\,\partial_{\theta_j}\phi(\theta)\right]\right), \tag{3.3.5}$$

where

$$\partial_{\theta_j}\phi(\theta) = a\big(e_j e_j^\top - \theta e_j^\top - e_j\theta^\top\big)\,a^\top .$$

We write $\phi(\theta) = PDP^\top$ with $D \in \mathcal{S}_n^{+,*}$ diagonal and denote $w = a^\top P$, which is invertible since P is orthonormal and $a^\top a \in \mathcal{S}_n^{+,*}$. Then

$$\begin{aligned}
\operatorname{Tr}\left[\phi^{-1/2}(\theta)\,\partial_{\theta_j}\phi(\theta)\right] &= \operatorname{Tr}\left[D^{-1/2}P^\top\partial_{\theta_j}\phi(\theta)\,P\right] \\
&= \operatorname{Tr}\left[D^{-1/2}\,w^\top\big(e_j e_j^\top - \theta e_j^\top - e_j\theta^\top\big)\,w\right] \\
&= \operatorname{Tr}\left[D^{-1/2}\,w^\top\big(e_j e_j^\top - 2e_j\theta^\top\big)\,w\right] \\
&= \sum_{i=1}^n D_{ii}^{-1/2}(w_{ji}^2 - 2w_{ji}\,(\theta^\top w e_i)).
\end{aligned}$$

Now, we observe that

$$\begin{aligned}
D_{ii} = P_i^\top\phi(\theta)\,P_i &= ||b\,P_i||^2 + e_i^\top w^\top\left(\operatorname{Diag}(\theta) - \theta\theta^\top\right)\,we_i \\
&= ||b\,P_i||^2 + \sum_{j=1}^n \theta_j w_{ji}^2 - (\theta^\top\,we_i)^2 \\
&= ||b\,P_i||^2 + (\theta^\top\,we_i)^2 + \sum_{j=1}^n \theta_j(w_{ji}^2 - 2w_{ji}\,(\theta^\top we_i)).
\end{aligned}$$

Therefore, we get by the triangular inequality

$$\begin{aligned}
\sum_{j=1}^n |\theta_j|\left|\operatorname{Tr}\left[\phi^{-1/2}(\theta)\,\partial_{\theta_j}\phi(\theta)\right]\right| &\geq \left|\sum_{j=1}^n \theta_j\sum_{i=1}^n D_{ii}^{-1/2}(w_{ji}^2 - 2w_{ji}\,(\theta^\top we_i))\right| \\
&= \left|\sum_{i=1}^n D_{ii}^{1/2} - D_{ii}^{-1/2}(||b\,P_i||^2 + (\theta^\top\,we_i)^2)\right|.
\end{aligned}$$

Then, if $\theta \to \bar\theta$ with $\bar\theta \in \mathcal{U}\backslash\overset{\circ}{\mathcal{U}}$, there exists i such that $D_{ii} \to 0$ and therefore $\sum_{i=1}^n D_{ii}^{1/2} - D_{ii}^{-1/2}(||b\,P_i||^2 + (\theta^\top\,we_i)^2) \to -\infty$ since $||b\,P_i||^2 + (\theta^\top\,we_i)^2 \geq \underline{\lambda}(-b)^2 > 0$, where $\underline{\lambda}(-b)$ is the smallest eigenvalue of $-b \in \mathcal{S}_n^{+,*}$. Therefore, $\left|\operatorname{Tr}\left[\phi^{-1/2}(\theta)\,\partial_{\theta_j}\phi(\theta)\right]\right| \to +\infty$ for some j, which implies then $|\partial_{\theta_j}\Lambda(\theta)| \to +\infty$. Thus, $||\nabla_\theta\Lambda(\theta)|| \to \infty$ and Λ is therefore essentially smooth. $\square$

Remark 3.3.10. *In fact, Prop. 3.3.8 holds, not only for $\theta \in \mathcal{U}$, but for every $\theta \in \mathbb{R}^n$. Indeed, by Remark 3.3.5,*

$$\theta \mapsto \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_t^\epsilon}\right]$$

is a convex function and by Prop. 3.3.9, Λ admits infinite derivative on $\mathcal{U} \backslash \mathring{\mathcal{U}}$. Therefore, for every $\theta \in \mathbb{R}^n \backslash \mathcal{U}$,

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_t^\epsilon}\right] = \Lambda(\theta) = \infty\,.$$

3.3.3 Long-time large deviation principle for the log-price process

We now state the large deviation principle for the family $(Y_T^\epsilon)_{\epsilon \in (0,1]}$, when $\epsilon \to 0$.

Theorem 3.3.11. *The family $(Y_T^\epsilon)_{\epsilon \in (0,1]}$ satisfies a large deviation principle, when $\epsilon \to 0$ with good rate function*

$$\Lambda^*(y) = \sup_{\lambda \in \mathbb{R}^n} \langle \lambda, y \rangle - \Lambda(\lambda)\,.$$

Proof. First note that $\phi(0) = b^2 \in \mathcal{S}_n^{+,*}$. But since

$$\theta \mapsto \phi(\theta) := b^2 + a\,\left(\mathrm{Diag}(\theta) - \theta\theta^\top\right) a^\top$$

is a continuous function, there exists a neighbourhood $B(0,\delta)$ of 0 such that $\phi(\theta) \in \mathcal{S}_n^{+,*}$ for every $\theta \in B(0,\delta)$, hence $0 \in \mathring{\mathcal{U}}$. Furthermore, Proposition 3.3.8 together with the argument in Remark 3.3.10 prove that

$$\Lambda(\theta) = \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\epsilon^{-1}\theta^\top Y_T^\epsilon}\right]\,,$$

where Λ is defined in (3.3.2). Finally, Proposition 3.3.9 yields the essential smoothness of Λ. Therefore, by the Gärtner-Ellis Theorem 3.3.4, $(Y_T^\epsilon)_{\epsilon \in (0,1]}$ satisfies a large deviation properties, when $\epsilon \to 0$ with good rate function Λ^*. □

3.4 Asymptotic implied volatility of basket options

In this section, to simplify the formulas and without loss of generality, we assume that $Y_0^j = 0$ for $j = 1, \dots, n$ and $r = 0$ so that $(e^{Y_t^j})_{t \geq 0}$ is a martingale with initial value 1 (this follows from Proposition 3.2.4). We are interested in the limiting behaviour far from maturity of basket option prices and the corresponding implied volatilities in the Wishart model. The basket call option price with log strike k and time to maturity T is defined by

$$C(T,k) = \mathbb{E}\left[\left(\sum_{i=1}^{n}\omega_i S_T^i - e^k\right)_+\right],$$

and the corresponding put option price is defined by

$$P(T,k) = \mathbb{E}\left[\left(e^k - \sum_{i=1}^{n}\omega_i S_T^i\right)_+\right]$$

where $\omega \in (\mathbb{R}_+^*)^n$ with $\sum_{i=1}^n \omega_i = 1$.

The implied volatility of basket options is defined by comparing their price to the corresponding option price in the Black-Scholes model $\frac{dS_t}{S_t} = \sigma dW_t$:

$$C^{BS}(T,k,\sigma) = N(d_1) - e^k N(d_2), \quad d_{12} = \frac{-k \pm \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}},$$

where N is the standard normal distribution function. The implied volatility for log strike k and time to maturity T is then defined as the unique value $\sigma(T,k)$ such that

$$C^{BS}(T,k,\sigma(T,k)) = C(T,k).$$

It can be equivalently defined using the put option price.

It is well known that in most models, for fixed log strike k, the implied volatility converges to a constant value independent from k as $T \to \infty$ (Tehranchi, 2009). To obtain a non-trivial limiting smile, we therefore follow (Jacquier et al., 2013) and use a renormalized log strike $k(T) = yT$. We are interested in computing the limiting implied volatility

$$\sigma_\infty(y) = \lim_{T\to\infty} \sigma(T, yT).$$

3.4.1 Asymptotic price for the Wishart model

Introduce the renormalized log-price process in the stochastic volatility Wishart model: $\tilde{Y}_T^j = T^{-1}Y_T^j$, $j = 1,\dots,n$. Note that to simplify notation, in this section we avoid using an extra parameter ϵ and simply consider the asymptotics when $T \to \infty$. For this reason, the asymptotic Laplace exponent $\Lambda(\theta)$ will be given by equation (3.3.2) with $T = 1$ and $r = 0$.

Denote the basket log price by $\mathcal{B}_T := \log \sum_{j=1}^n \omega_j e^{Y_T^j}$, and the corresponding renormalized price by $\tilde{\mathcal{B}}_T := T^{-1}\log\sum_{j=1}^n \omega_j e^{Y_T^j}$. We first show some LDP-like bounds for this quantity. In the following lemma and below, we will use that $\Lambda(0) = \Lambda(e_j) = 0$, which gives in particular that $\Lambda^*(x) \geq 0$ and

$\Lambda^*(x) - x_j \geq 0$ for all $x \in \mathbb{R}^d$. Thus, we let $x^* = \Lambda'(0)$ and $\tilde{x}_j^* = \Lambda_j'(e_j)$ for $j = 1, \dots, n$ and introduce three constants: $\beta^* = \max_j x_j^*$, $\hat{\beta}^* = \min_j \tilde{x}_j^*$ and $\tilde{\beta}^* = \max_j \tilde{x}_j^*$. It is easy to see from (3.3.5) that $x_j^* = -\tilde{x}_j^* < 0$ since $\phi(0) = \phi(e_j) = b^2$ is positive definite and a is invertible. We get $\beta^* < 0 < \hat{\beta}^* \leq \tilde{\beta}^*$.

Lemma 3.4.1. *The following estimates hold for $\tilde{\mathcal{B}}_T$.*

1. *If $\beta < \beta^*$ then*

$$\lim_{T\to\infty} T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T \in (-\infty,\, \beta]\right) = -\inf_{x\in(-\infty,\beta]^n} \Lambda^*(x) = \inf_{\lambda\in\mathbb{R}^n, \lambda_i \leq 0, i=1,\dots,n} \{\Lambda(\lambda) - \beta\langle\lambda, \mathbf{1}\rangle\} < 0\,, \tag{3.4.1}$$

otherwise

$$\lim_{T\to\infty} T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T \in (-\infty,\, \beta]\right) = 0.$$

2. *If $\beta \geq \beta^*$ then*

$$\lim_{T\to\infty} T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T \in (\beta, \infty)\right) = -\inf_{x\notin(-\infty,\beta]^n} \Lambda^*(x) \tag{3.4.2}$$

$$= \max_{i=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda\beta + \Lambda(\lambda e_i)\}, \tag{3.4.3}$$

otherwise

$$\lim_{T\to\infty} T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T \in (\beta, \infty)\right) = 0.$$

In addition if $\beta \geq \beta^*$ and $\beta \neq \tilde{x}_i^*$ for all i, then

$$\lim_{T\to\infty} T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T \in (\beta, \infty)\right) < -\beta.$$

3. *Let $j \in \{1, \dots, n\}$. Then,*

$$\lim_{T\to\infty} T^{-1} \log \mathbb{E}\left[e^{Y_T^j} \mathbb{1}_{\tilde{\mathcal{B}}_T\in(-\infty,\beta]}\right] = -\inf_{x\in(-\infty,\beta]^n} \Lambda^*(x) - x_j = \beta + \inf_{\lambda^j\leq 1, \lambda^i\leq 0, i\neq j} \{\Lambda(\lambda) - \beta\langle\lambda, \mathbf{1}\rangle\}. \tag{3.4.4}$$

In addition, if $\tilde{x}_j^* > \beta$ then

$$\lim_{T\to\infty} T^{-1} \log \mathbb{E}\left[e^{Y_T^j} \mathbb{1}_{\tilde{\mathcal{B}}_T\in(-\infty,\beta]}\right] < 0.$$

4. *Let $j \in \{1, \dots, n\}$ and assume $\beta > \tilde{x}_j^*$. Then,*

$$\lim_{T\to\infty} T^{-1} \log \mathbb{E}\left[e^{Y_T^j} \mathbb{1}_{\tilde{\mathcal{B}}_T \in (\beta,\infty)}\right] = - \inf_{x \notin (-\infty,\beta]^n} \Lambda^*(x) - x_j$$
$$= \max_{i=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda\beta + \Lambda(\lambda e_i + e_j)\} < 0. \tag{3.4.5}$$

Proof. 1. Since

$$\omega_{\min} e^{\max_j Y_T^j} \le \sum_{j=1}^n \omega_j e^{Y_T^j} \le n\omega_{\max} e^{\max_j Y_T^j}$$

with $(\omega_{\min}, \omega_{\max}) := (\min_{j=1,\dots,n} \omega_j, \max_{j=1,\dots,n} \omega_j)$, we have for every $T > 0$ and $\beta \in \mathbb{R}$,

$$\left(\tilde{Y}_T \in (-\infty, \beta - T^{-1}\log(n\,\omega_{\max}))^n\right) \subset (\tilde{\mathcal{B}}_T < \beta)$$
$$\subset \left(\tilde{Y}_T \in (-\infty, \beta - T^{-1}\log\omega_{\min})^n\right).$$

Therefore, we get for every $\delta > 0$ and T sufficiently large,

$$\mathbb{P}\left(\tilde{Y}_T \in (-\infty, \beta - \delta)^n\right) \le \mathbb{P}(\tilde{\mathcal{B}}_T < \beta) \le \mathbb{P}\left(\tilde{Y}_T \in (-\infty, \beta + \delta)^n\right).$$

Passing to the lim sup and lim inf, we get:

$$\liminf_{T\to\infty} T^{-1}\log\mathbb{P}\left(\tilde{Y}_T \in (-\infty, \beta - \delta)^n\right) \le \liminf_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta)$$
$$\le \limsup_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta) \le \limsup_{T\to\infty} T^{-1}\log\mathbb{P}\left(\tilde{Y}_T \in (-\infty, \beta + \delta)^n\right).$$

Using the large deviations principle for $\tilde{Y}_T$ (Theorem 3.3.11) further yields:

$$-\inf_{x\in(-\infty,\beta-\delta)^n} \Lambda^*(x) \le \liminf_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta)$$
$$\le \limsup_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta) \le -\inf_{x\in(-\infty,\beta+\delta]^n} \Lambda^*(x),$$

and making δ tend to zero, we see that

$$-\inf_{x\in(-\infty,\beta)^n} \Lambda^*(x) \le \liminf_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta)$$
$$\le \limsup_{T\to\infty} T^{-1}\log\mathbb{P}(\tilde{\mathcal{B}}_T < \beta) \le -\inf_{x\in(-\infty,\beta]^n} \Lambda^*(x).$$

The fact that the domain of Λ is bounded (Remark 3.2.6) implies that Λ^* is locally bounded from above and therefore continuous. The first equality of (3.4.1) then follows by continuity of Λ^*. The second equality then follows from the definition of Λ^* and the minimax theorem (see, e.g., Corollary 37.3.2 in (Rockafellar, 1970)) which can be applied because the domain of Λ is bounded (cf. Remark 3.2.6). Finally, the inequality follows from the fact that the function $f(\lambda) = \Lambda(\lambda) - \beta\langle\lambda, \mathbf{1}\rangle$ satisfies $f(0) = 0$ and $f'(0) = x^* - \beta\mathbf{1}$. Under the condition $\beta < \beta^*$ at least one component of the derivative is strictly positive, and hence the minimum of f over the set $\{\lambda_i \le 0, i = 1, \dots, n\}$ is strictly negative.

2. The first equality in (3.4.3) follows similarly to the previous item. If $\beta < \beta^*$ then $x^* \notin (-\infty, \beta]^n$ and the infimum equals 0. Otherwise by convexity of Λ^* the infimum is attained on the boundary of this set. Therefore, we can write:

$$
\begin{aligned}
-\inf_{x \notin (-\infty,\beta]^n} \Lambda^*(x) &= \max_{i=1,\dots,n} \sup_{x\in\mathbb{R}^n : x_i = \beta} \{-\Lambda^*(x)\} \\
&= \max_{i=1,\dots,n} \sup_{x\in\mathbb{R}^n : x_i = \beta} \inf_{\lambda\in\mathbb{R}^n} \{-\langle\lambda, x\rangle + \Lambda(\lambda)\} \\
&= \max_{i=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda\beta + \Lambda(\lambda e_i)\},
\end{aligned}
$$

since the inf and sup may once again be interchanged in virtue of the minimax theorem and then the supremum on $x \in \mathbb{R}^n$ such that $x_i = \beta$ is clearly $+\infty$ when there is $j \neq i$ such that $\lambda_j \neq 0$. Consider the function $f_i : \mathbb{R} \to \mathbb{R}$, $f_i(\lambda) = -\lambda\beta + \Lambda(\lambda e_i)$. Since $f_i(1) = -\beta$ and $f_i'(1) = -\beta + \tilde{x}_i^*$, it follows that

$$
\beta + \max_{i=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda\beta + \Lambda(\lambda e_i)\} < 0.
$$

when $\beta \neq \tilde{x}_i^*$ for all i.

3. For the first identity in (3.4.4), remark that, similarly to the first part, for T sufficiently large, all $\delta > 0$ and $\beta \in \mathbb{R}$ we have,

$$
\mathbb{E}[e^{Y_T^j} \mathbb{1}_{\{\tilde{Y}_T \in (-\infty, \beta-\delta]^n\}}] \ge \mathbb{E}[e^{Y_T^j} \mathbb{1}_{\{\tilde{\mathcal{B}}_T \le \beta\}}] \ge \mathbb{E}[e^{Y_T^j} \mathbb{1}_{\{\tilde{Y}_T \in (-\infty, \beta+\delta]^n\}}] \,,
$$

We can apply Theorem 3.3.6 with the function $H : x \mapsto x_j$ since $\Lambda(e_j) = 0$ and $\Lambda(\gamma e_j) < \infty$ for $\gamma > 1$ small enough. When δ goes to zero, we get

$$
\begin{aligned}
\sup_{x\in(-\infty,\beta)^n} \{x_j - \Lambda^*(x)\} &\le \liminf_{T\to\infty} T^{-1} \log \mathbb{E}[e^{Y_T^j} \mathbb{1}_{\{\tilde{\mathcal{B}}_T \le \beta\}}] \\
&\le \limsup_{T\to\infty} T^{-1} \log \mathbb{E}[e^{Y_T^j} \mathbb{1}_{\{\tilde{\mathcal{B}}_T \le \beta\}}] \le \sup_{x\in(-\infty,\beta]^n} \{x_j - \Lambda^*(x)\}.
\end{aligned}
$$

By continuity of Λ^*, the lower and the upper bounds are equal. Since $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}^n} \langle\lambda + e_j, x\rangle - \Lambda(\lambda + e_j)$, we get

$$
\sup_{x\in(-\infty,\beta]^n} \{x_j - \Lambda^*(x)\} = \sup_{x\in(-\infty,\beta]^n} \inf_{\lambda\in\mathbb{R}^n} \Lambda(\lambda + e_j) - \langle\lambda, x\rangle \,.
$$

The second identity in (3.4.4) then follows from the minimax theorem as above. Finally, to show the inequality, remark that

$$\inf_{\lambda^j \le 1, \lambda^i \le 0, i \ne j} \{\Lambda(\lambda) - \beta\langle\lambda, \mathbf{1}\rangle\} \le \inf_{\lambda \le 1} f_j(\lambda)$$

and $f_j'(1) = \tilde{x}_j^* - \beta > 0$.

4. The first identity in (3.4.5) follows as in item (3). We have $\Lambda^*(x) - x_j \ge 0$ and $\Lambda^*(\Lambda'(e_j)) = \Lambda_j'(e_j) = \tilde{x}_j^*$ since e_j is a critical point of $\lambda \mapsto \langle\lambda, \Lambda'(e_j)\rangle - \Lambda(\lambda)$. Since $\beta > \tilde{x}_j^*$ and $\Lambda'(e_j) \notin (-\infty, \beta]^n$, the supremum is attained as in item (2) on the boundary:

$$\begin{aligned}\sup_{x \in \mathbb{R}^n} x_j - \Lambda^*(x) &= \max_{i=1,\dots,n} \sup_{x \in \mathbb{R}^n : x_i = \beta} x_j - \Lambda^*(x) \\ &= \max_{i=1,\dots,n} \sup_{x \in \mathbb{R}^n : x_i = \beta} \inf_{\lambda \in \mathbb{R}^n} \Lambda(\lambda + e_j) - \langle\lambda, x\rangle .\end{aligned}$$

The second identity in (3.4.5) holds true in virtue of the minimax theorem as above, like in item (2). To prove the negativity, we consider the functions $g_i(\lambda) = -\lambda\beta + \Lambda(\lambda e_i + e_j)$. We have that $g_i(0) = 0$ and $g_i'(0) = -\beta + \Lambda_i'(e_j)$. We have $g_j'(0) = -\beta + \tilde{x}_j^* < 0$. If $g_i'(0) \ne 0$ for all i, the result is clear. Otherwise, we can find $\tilde{\beta} \in (\tilde{x}_j^*, \beta)$ such that $\tilde{\beta} \ne \Lambda_i'(e_j)$ for all i, and since $e^{Y_T^j}\mathbb{1}_{\tilde{\mathcal{B}}_T \in (\beta,\infty)} \le e^{Y_T^j}\mathbb{1}_{\tilde{\mathcal{B}}_T \in (\tilde{\beta},\infty)}$, we get the claim.

□

The following theorem characterizes the asymptotic behaviour of basket call prices in the Wishart model. There are different asymptotic regimes to consider, depending on the position of y with respect to these constants.

Theorem 3.4.2. *Assume that $y \ne \tilde{x}_i^*$ for all i. Then, as $T \to \infty$, the call option price in the Wishart model satisfies*

$$\lim_{T\to\infty} \mathbb{E}\left[(e^{\mathcal{B}_T} - e^{yT})_+\right] = \sum_{i=1}^n \omega_i \mathbb{1}_{\tilde{x}_i^* > y}. \tag{3.4.6}$$

In addition, if $y < \beta^*$ then

$$\begin{aligned}\lim_{T\to\infty} T^{-1}\log\mathbb{E}\left[(e^{yT} - e^{\mathcal{B}_T})_+\right] &= \lim_{T\to\infty} T^{-1}\log\left\{e^{yT} - 1 + \mathbb{E}\left[(e^{\mathcal{B}_T} - e^{yT})_+\right]\right\} \\ &= y - \inf_{z \in (-\infty, y]^n} \Lambda^*(z) < y; \end{aligned} \tag{3.4.7}$$

if $y > \tilde{\beta}^*$, then

$$\lim_{T\to\infty} T^{-1}\log\mathbb{E}\left[(e^{\mathcal{B}_T} - e^{yT})_+\right] = \max_{i,j=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda y + \Lambda(\lambda e_i + e_j)\} < 0. \tag{3.4.8}$$

and if $y \in (\beta^*, \hat{\beta}^*)$, *then*

$$\begin{aligned}\lim_{T\to\infty} T^{-1} \log\left(1 - \mathbb{E}[(e^{\mathcal{B}_T} - e^{yT})_+]\right) \\ = y + \max_{i=1,\dots,n} \inf_{\lambda\in\mathbb{R}} \{-\lambda y + \Lambda(\lambda e_i)\}\} < \min(0, y)\,.\end{aligned} \tag{3.4.9}$$

Proof.
Proof of (3.4.6). We remark that

$$\mathbb{E}\left[(e^{\mathcal{B}_T} - e^{yT})_+\right] = \mathbb{E}\left[e^{\mathcal{B}_T} \mathbb{1}_{\tilde{\mathcal{B}}_T > y}\right] - e^{yT}\mathbb{P}\left[\tilde{\mathcal{B}}_T > y\right] \tag{3.4.10}$$

and consider the two terms separately. If $y < 0$, the second term clearly converges to zero. Assume then that $y \geq 0$. Since $\beta^* \leq 0$, by Lemma 3.4.1 part 2,

$$\lim_{T\to\infty} T^{-1} \log e^{yT}\mathbb{P}\left(\tilde{\mathcal{B}}_T > y\right) < 0$$

This proves that the second term in (3.4.10) converges to zero. We now focus on the first term, which satisfies

$$\mathbb{E}\left[e^{\mathcal{B}_T} \mathbb{1}_{\tilde{\mathcal{B}}_T > y}\right] = \sum_{i=1}^{n} \omega_i \mathbb{E}\left[e^{Y_T^i} \mathbb{1}_{\tilde{\mathcal{B}}_T > y}\right].$$

Fix some $i \in \{1, \dots, n\}$. Then, by Lemma 3.4.1 parts 3 and 4, if $y > \tilde{x}_i^*$ then

$$\lim_{T\to\infty} \mathbb{E}\left[e^{Y_T^i} \mathbb{1}_{\tilde{\mathcal{B}}_T > y}\right] = 0,$$

and if $y < \tilde{x}_i^*$ then

$$\lim_{T\to\infty} \mathbb{E}\left[e^{Y_T^i} \mathbb{1}_{\tilde{\mathcal{B}}_T \leq y}\right] = 0.$$

Combining these estimates for different i, the proof of (3.4.6) is complete.
Proof of (3.4.7) The equality

$$e^{y\,T}(1 - e^{-\delta T})\mathbb{1}_{\{\tilde{\mathcal{B}}_T < y - \delta\}} \leq \left(e^{y\,T} - e^{\mathcal{B}_T}\right)_+ \leq e^{y\,T}\mathbb{1}_{\{\tilde{\mathcal{B}}_T < y\}}$$

holds for every $\delta > 0$ and $T > 0$. Then by successively taking the expectation, the logarithm and multiplying by T^{-1}, we find

$$\begin{aligned}y + T^{-1}\log(1 - e^{-\delta T}) + T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T < y - \delta\right) \\ \leq T^{-1} \log \mathbb{E}\left[(e^{y\,T} - e^{\mathcal{B}_T})_+\right] \leq y + T^{-1} \log \mathbb{P}\left(\tilde{\mathcal{B}}_T < y\right)\,.\end{aligned}$$

Passing to the limit $T \to \infty$ and using Lemma 3.4.1 part 1, the proof is complete.
Proof of (3.4.8). We use the inequality

$$e^{\mathcal{B}_T}(1 - e^{-\delta T})\mathbb{1}_{\{y < \tilde{\mathcal{B}}_T - \delta\}} \leq \left(e^{\mathcal{B}_T} - e^{y\,T}\right)_+ \leq e^{\mathcal{B}_T}\mathbb{1}_{\{y < \tilde{\mathcal{B}}_T\}}.$$

Consider for instance the upper bound. Taking the expectation and the logarithm, we obtain $\log \mathbb{E}[e^{\mathcal{B}_T}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y\}}] = \log \sum_{j=1}^n \omega_j \mathbb{E}\left[e^{Y_T^j}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y\}}\right]$ and thus

$$
\begin{aligned}
T^{-1}\log \mathbb{E}[e^{\mathcal{B}_T}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y\}}] &\leq \max_{j=1,\dots,n} T^{-1}\log \mathbb{E}\left[e^{Y_T^j}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y\}}\right],\\
T^{-1}\log \mathbb{E}[e^{\mathcal{B}_T}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y+\delta\}}] &\geq \max_{j=1,\dots,n} T^{-1}\log \mathbb{E}\left[e^{Y_T^j}\mathbb{1}_{\{\tilde{\mathcal{B}}_T>y+\delta\}}\right] + \log(\omega_j)/T.
\end{aligned}
$$

The result then follows from Lemma 3.4.1, part 4.

Proof of (3.4.9). We use the following identity.

$$
\begin{aligned}
1 - \mathbb{E}[(e^{\mathcal{B}_T} - e^{yT})_+] &= \mathbb{E}[e^{\mathcal{B}_T} - (e^{\mathcal{B}_T} - e^{yT})_+]\\
&= e^{yT}\mathbb{P}[\tilde{\mathcal{B}}_T > y] + \mathbb{E}[e^{\mathcal{B}_T}\mathbb{1}_{\tilde{\mathcal{B}}_T\leq y}].
\end{aligned}
$$

By Lemma 3.4.1, part 2,

$$
\lim_{T\to\infty} T^{-1}\log e^{yT}\mathbb{P}[\tilde{\mathcal{B}}_T > y] = y + \max_{i=1,\dots,n}\inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i)\} < 0.
$$

Consider the function $f_i : \mathbb{R} \to \mathbb{R}$, $f_i(\lambda) = -\lambda y + \Lambda(\lambda e_i)$. Since $f_i(0) = 0$ and $f_i'(0) = -y + x_i^* < 0$, it follows that also

$$
y + \max_{i=1,\dots,n}\inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i)\} < y.
$$

On the other hand, by Lemma 3.4.1, part 3,

$$
\begin{aligned}
\lim_{T\to\infty} T^{-1}\log \mathbb{E}[e^{\mathcal{B}_T}\mathbb{1}_{\tilde{\mathcal{B}}_T\leq y}] &= y + \max_{j=1,\dots,n}\inf_{\lambda^j\leq 1,\lambda^i\leq 0, i\neq j}\{\Lambda(\lambda) - y\langle\lambda, \mathbf{1}\rangle\}\\
&\leq y + \max_{j=1,\dots,n}\inf_{\lambda\leq 1} f_j(\lambda).
\end{aligned}
$$

Since, for $y \in (\beta^*, \hat{\beta}^*)$, $f_j'(0) < 0$ and $f_j'(1) > 0$, the infimum is attained on the interval $(0, 1)$, and the contribution of this term is less than the one of the first term. The properties of the logarithm allow to conclude the proof. □

3.4.2 Implied volatility asymptotics

In the Black-Scholes model with volatility σ, we have (see, e.g. (Forde and Jacquier, 2011a), Corollary 2.12)

$$
\begin{aligned}
\lim_{T\to\infty} T^{-1}\log(C^{BS}(T, yT, \sigma) + e^{yT} - 1) &= -\frac{1}{2}\left(\frac{\sigma}{2} - \frac{y}{\sigma}\right)^2, \quad y \leq -\frac{\sigma^2}{2}\\
\lim_{T\to\infty} T^{-1}\log C^{BS}(T, yT, \sigma) &= -\frac{1}{2}\left(\frac{\sigma}{2} - \frac{y}{\sigma}\right)^2, \quad y \geq \frac{\sigma^2}{2}\\
\lim_{T\to\infty} T^{-1}\log\left(1 - C^{BS}(T, yT, \sigma)\right) &= -\frac{1}{2}\left(\frac{\sigma}{2} - \frac{y}{\sigma}\right)^2, \quad -\frac{\sigma^2}{2} < y < \frac{\sigma^2}{2}.
\end{aligned}
$$

Under the Wishart model, for the basket option, we can write:

$$\lim_{T\to\infty} T^{-1}\log \mathbb{E}\left[(e^{yT}-e^{\mathcal{B}_T})_+\right] = -L(y), \quad y \le \beta^* \tag{3.4.11}$$
$$\lim_{T\to\infty} T^{-1}\log \mathbb{E}\left[(e^{\mathcal{B}_T}-e^{yT})_+\right] = -L(y), \quad y \ge \tilde{\beta}^*$$
$$\lim_{T\to\infty} T^{-1}\log \left(1-\mathbb{E}\left[(e^{\mathcal{B}_T}-e^{yT})_+\right]\right) = -L(y), \quad \beta^* < y < \hat{\beta}^*,$$

where

$$L(y) = -y - \inf_{\lambda\in\mathbb{R}^n:\lambda_i\le 0, i=1,\dots,n}\{\Lambda(\lambda) - y\langle\lambda,\mathbf{1}\rangle\}, \quad y \le \beta^*$$
$$L(y) = -\max_{i,j=1,\dots,n}\inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i + e_j)\}, \quad y \ge \tilde{\beta}^*$$
$$L(y) = -y - \max_{i=1,\dots,n}\inf_{\lambda\in\mathbb{R}}\{-\lambda y + \Lambda(\lambda e_i)\}, \quad \beta^* < y < \hat{\beta}^*.$$

We deduce (see (Jacquier et al., 2013) for details) that the limiting implied volatility of a basket option in the Wishart model is given by

$$\sigma_\infty(y) = \sqrt{2}\left(\xi\sqrt{L(y)+y} + \eta\sqrt{L(y)}\right), \tag{3.4.12}$$

where ξ and η are constants with $\xi^2 = \eta^2 = 1$, which must be chosen to satisfy the conditions

$$y \le -\frac{\sigma_\infty^2(y)}{2} \qquad \text{if} \quad y \le \beta^*$$
$$y \ge \frac{\sigma_\infty^2(y)}{2} \qquad \text{if} \quad y \ge \tilde{\beta}^*$$
$$-\frac{\sigma_\infty^2(y)}{2} < y < \frac{\sigma_\infty^2(y)}{2} \qquad \text{if} \quad \beta^* < y < \tilde{\beta}^*.$$

First of all remark that by taking $\lambda = 0$ and $\lambda = e_i$ it follows that $L(y) \ge y$ and $L(y) \ge 0$, so that the expressions under the square root sign are positive. It is easy to see that for $y \le \beta^*$, these conditions imply $\xi = -1$ and $\eta = 1$ since $b^* < 0$ and $-y \le L(y)$, and for $y \ge \tilde{\beta}^*$ one has $\xi = 1$ and $\eta = -1$. For $\beta^* < y < \hat{\beta}^*$, we still have $|y| \le \max(L(y), L(y)+y)$ and to satisfy the conditions in this case and $\sigma_\infty(y) > 0$, one must take $\xi = \eta = 1$.

The case when $\hat{\beta}^* < y < \tilde{\beta}^*$ requires a specific treatment. It is characterized by the following proposition.

Proposition 3.4.3. *Let $\hat{\beta}^* < y < \tilde{\beta}^*$. Then, $\sigma_\infty(y) = \sqrt{2y}$ and*

$$\sigma(T, yT) = \sqrt{2y} + N^{-1}(C_\infty(y))T^{-1/2} + o\left(T^{-1/2}\right)$$

as $T \to \infty$, where $C_\infty(y) = \sum_{i=1}^n \omega_i \mathbb{1}_{\tilde{x}_i^* > y}$.

Proof. We follow the arguments of the proof of Theorem 3.3 in (Jacquier and Keller-Ressel, 2018) with some minor changes. The Black-Scholes call option price satisfies

$$C^{BS}(T, yT, \sigma) = N\left(\frac{-y + \frac{\sigma^2}{2}}{\sigma}\sqrt{T}\right) - e^{yT} N\left(\frac{-y - \frac{\sigma^2}{2}}{\sigma}\sqrt{T}\right).$$

We have by definition of the implied volatility and equation (3.4.6),

$$C^{BS}(T, yT, \sigma(t, yT)) = C(T, yT) \underset{T\to+\infty}{\to} C_\infty(y).$$

Since $y > \hat{\beta}^* > 0$, as $T \to \infty$, we get necessarily $\frac{y+\frac{\sigma(T,yT)^2}{2}}{\sigma(T,yT)}\sqrt{T} \to +\infty$. Using the classical bound on the Mills ratio $N(-x) \leq x^{-1}\phi(x)$ for $x > 0$, where ϕ is the standard Gaussian density, we have

$$e^{yT} N\left(\frac{-y - \frac{\sigma(T,yT)^2}{2}}{\sigma(T, yT)}\sqrt{T}\right) \leq \phi\left(\frac{y - \frac{\sigma(T,yT)^2}{2}}{\sigma(T, yT)}\sqrt{T}\right) \frac{\sigma(T, yT)}{\left(y + \frac{\sigma(T,yT)^2}{2}\right)\sqrt{T}} \to 0$$

as $T \to \infty$. Therefore,

$$\frac{-y + \frac{\sigma(T,yT)^2}{2}}{\sigma(T, yT)} = N^{-1}(C_\infty(y))T^{-1/2} + o\left(T^{-1/2}\right). \tag{3.4.13}$$

Consider now the function $f(z) = -\frac{y}{z} + \frac{z}{2}$. Its inverse which is positive in the neighbourhood of zero is given by

$$f^{-1}(x) = x + \sqrt{x^2 + 2y}$$

Applying f^{-1} to both sides of (3.4.13) and neglecting terms of order $o\left(T^{-\frac{1}{2}}\right)$, the proof is complete. □

3.5 Variance reduction

3.5.1 The general variance reduction problem

Denote $P(S_T)$ the payoff of a European option on $(S_T^1, ..., S_T^n)$. The price of an option is generally calculated as the expectation $\mathbb{E}(P(S_T))$ under a certain risk-neutral measure $\mathbb{P}$. When the number of assets n is low, this expectation may be evaluated by Fourier inversion, however, when the dimension is large, as in the case of index options, Monte Carlo is the method of choice. The standard Monte Carlo estimator of $\mathbb{E}(P(S_T))$ with N samples is given by

$$\widehat{P}_N = \frac{1}{N}\sum_{j=1}^{N} P(S_T^{(j)}),$$

where $S_T^{(j)}$ are i.i.d. samples of S_T under the measure $\mathbb{P}$. The variance of the standard Monte Carlo estimator is given by

$$\mathbb{V}\mathrm{ar}[\widehat{P}_N] = \frac{1}{N}\mathbb{V}\mathrm{ar}[P(S_T)],$$

and is often too high for real-time applications. To decrease the computational time, various variance reduction methods have been proposed, the most popular being importance sampling.

The importance sampling method is based on the following identity, valid for any probability measure $\mathbb{Q}$, with respect to which $\mathbb{P}$ is absolutely continuous.

$$\mathbb{E}^{\mathbb{P}}[P(S_T)] = \mathbb{E}^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}P(S_T)\right].$$

This allows one to define the importance sampling estimator

$$\widehat{P}_N^{\mathbb{Q}} := \frac{1}{N}\sum_{j=1}^{N}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{(j)} P(S_T^{(j),\mathbb{Q}}),$$

where $S_T^{(j),\mathbb{Q}}$ are i.i.d. samples of S_T under the measure $\mathbb{Q}$. For efficient variance reduction, one needs then to find a probability measure $\mathbb{Q}$ such that S_T is easy to simulate under $\mathbb{Q}$ and the variance

$$\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left[P(S_T)\frac{d\mathbb{P}}{d\mathbb{Q}}\right] = \mathbb{E}^{\mathbb{P}}\left[P(S_T)^2\frac{d\mathbb{P}}{d\mathbb{Q}}\right] - \mathbb{E}^{\mathbb{P}}[P(S_T)]^2$$

is considerably smaller than the original variance $\mathbb{V}\mathrm{ar}_{\mathbb{P}}[P(S)]$.

In this paper we consider the class of measure changes $\{\mathbb{P}_\theta : \theta \in \mathbb{R}^n\}$, where

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\theta^\top Y_T}}{\mathbb{E}\left[e^{\theta^\top Y_T}\right]}.$$

To find the optimal variance reduction parameter θ^*, we therefore need to minimize the variance of the estimator under $\mathbb{Q}$, or, equivalently, the expectation

$$\mathbb{E}^{\mathbb{P}}\left[P(S_T)^2\frac{d\mathbb{P}}{d\mathbb{P}_\theta}\right].$$

Denoting $H(Y_T) := \log P\left(e^{Y_T}\right)$, the optimization problem writes

$$\inf_{\theta\in\mathbb{R}^n}\mathbb{E}\left[\exp\left(2H(Y_T) - \theta^\top Y_T + \mathcal{G}_1(\theta)\right)\right], \tag{3.5.1}$$

where

$$\mathcal{G}_\epsilon(\theta) := \epsilon\log\mathbb{E}\left[e^{\frac{\theta^\top Y_T^\epsilon}{\epsilon}}\right].$$

3.5.2 Asymptotic variance reduction

Since we cannot solve (3.5.1) explicitly, we instead choose to minimize the asymptotic proxy of Proposition 3.5.1, based on Theorem 3.3.6.

Proposition 3.5.1. *Let $H : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ be a continuous function and $\theta \in \mathbb{R}^n$ be such that there exists $\gamma > 1$ with*

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[\exp\left\{\gamma \frac{2H(Y_T^\epsilon) - \theta^\top Y_T^\epsilon}{\epsilon}\right\}\right] < \infty\,. \tag{3.5.2}$$

Then

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[\exp\left\{\frac{2H(Y_T^\epsilon) - \theta^\top Y_T^\epsilon + \mathcal{G}_\epsilon(\theta)}{\epsilon}\right\}\right] = \sup_{y \in \mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\} + \Lambda(\theta)\,.$$

Proof. By Theorem 3.3.6,

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[\exp\left\{\frac{2H(Y_T^\epsilon) - \theta^\top Y_T^\epsilon}{\epsilon}\right\}\right] = \sup_{y \in \mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\}\,. \tag{3.5.3}$$

Furthermore, by Proposition 3.3.8,

$$\epsilon \log \mathbb{E}\left[\exp\left\{\frac{\mathcal{G}_\epsilon(\theta)}{\epsilon}\right\}\right] = \mathcal{G}_\epsilon(\theta) \underset{\epsilon \to 0}{\longrightarrow} \Lambda(\theta)\,. \tag{3.5.4}$$

Multiplying (3.5.3) and (3.5.4) finishes the proof. □

Remark 3.5.2. *In particular, if H is continuous and bounded from above and θ is such that $\phi(-\theta) \in \mathcal{S}_n^{+,*}$, condition (3.5.2) is met.*

Definition 3.5.3. *A parameter $\theta^* \in \mathbb{R}^n$ is called asymptotically optimal if it achieves the infimum in the minimisation problem*

$$\inf_{\theta \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\} + \Lambda(\theta)\,. \tag{3.5.5}$$

Theorem 3.5.4. *Let H be a concave upper semi-continuous function. Then*

$$\inf_{\theta \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^n}\left\{2H(y) - \theta^\top y - \Lambda^*(y)\right\} + \Lambda(\theta) = 2 \inf_{\theta \in \mathbb{R}^n}\left\{\hat{H}(\theta) + \Lambda(\theta)\right\}\,,$$

where

$$\hat{H}(\theta) = \sup_{y \in \mathbb{R}^n}\left\{H(y) - \theta^\top y\right\}\,.$$

Furthermore, if θ^* minimizes the right-hand side, it also minimizes the left-hand side.

Proof. We follow the idea of the proof of (Genin and Tankov, 2016, Theorem 8), with some major simplifications due to the present finite-dimensional setting. By definition of Λ^*,

$$\begin{aligned}
&\inf_{\theta\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\Lambda^*(y)+\Lambda(\theta)\right\}\\
&\quad=\inf_{\theta\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\sup_{\lambda\in\mathbb{R}^n}\left\{\lambda^\top y-\Lambda(\lambda)\right\}+\Lambda(\theta)\right\}\\
&\qquad=\inf_{\theta\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\inf_{\lambda\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\lambda^\top y+\Lambda(\lambda)+\Lambda(\theta)\right\}\ .
\end{aligned}$$

The function

$$(y,\lambda)\mapsto 2H(y)-\theta^\top y-\lambda^\top y+\Lambda(\lambda)+\Lambda(\theta)$$

is concave-convex on $\mathbb{R}^n\times\mathcal{U}$ where $\mathcal{U}$ is bounded by Remark 3.2.6 and both $\mathbb{R}^n$ and $\mathcal{U}$ are convex. Therefore, by the minimax Theorem for concave-convex functions (see, e.g., Corollary 37.3.2 in (Rockafellar, 1970)),

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^n}\inf_{\lambda\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\lambda^\top y+\Lambda(\lambda)+\Lambda(\theta)\right\}\\
&\qquad=\inf_{\lambda\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\lambda^\top y+\Lambda(\lambda)+\Lambda(\theta)\right\}\ .
\end{aligned}$$

This allows us to rewrite

$$\begin{aligned}
&\inf_{\theta\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\Lambda^*(y)+\Lambda(\theta)\right\}\\
&\quad=\inf_{\theta\in\mathbb{R}^n}\inf_{\lambda\in\mathbb{R}^n}\sup_{y\in\mathbb{R}^n}\left\{2H(y)-\theta^\top y-\lambda^\top y+\Lambda(\lambda)+\Lambda(\theta)\right\}\\
&\qquad=2\inf_{\theta\in\mathbb{R}^n}\inf_{\lambda\in\mathbb{R}^n}\left\{\hat{H}\left(\frac{\theta+\lambda}{2}\right)+\frac{\Lambda(\lambda)+\Lambda(\theta)}{2}\right\}=2\inf_{\theta\in\mathbb{R}^n}\left\{\hat{H}(\theta)+\Lambda(\theta)\right\}\ ,
\end{aligned}$$

where the last equality is justified by the fact that, by convexity,

$$\frac{\Lambda(\lambda)+\Lambda(\theta)}{2}\geq\Lambda\left(\frac{\lambda+\theta}{2}\right)$$

with equality if $\lambda=\theta$. □

Remark 3.5.5. *Similarly to (Genin and Tankov, 2016, Definition 6) and to the discussion in Section 4 of (Robertson, 2010), it can be shown that the asymptotically optimal θ in Theorem 3.5.4 reaches the asymptotic lower bound of the variance on the log-scale over all equivalent measure changes.*

Let $\mathbb{Q}\sim\mathbb{P}$ be an equivalent measure change. Then by Jensen's inequality

$$\lim_{\epsilon\to 0}\epsilon\log\mathbb{E}^{\mathbb{Q}}\left(e^{\frac{2H(Y_T^\epsilon)}{\epsilon}}\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^2\right)\geq 2\lim_{\epsilon\to 0}\epsilon\log\mathbb{E}^{\mathbb{Q}}\left(e^{\frac{H(Y_T^\epsilon)}{\epsilon}}\frac{d\mathbb{P}}{d\mathbb{Q}}\right)$$

$$= 2 \lim_{\epsilon \to 0} \epsilon \log \mathbb{E} \left(e^{\frac{H(Y_T^\epsilon)}{\epsilon}} \right) .$$

By Theorem 3.3.6, the right-hand side is equal to

$$\begin{aligned} 2 \sup_{y \in \mathbb{R}^n} \{H(y) - \Lambda^*(y)\} &= 2 \sup_{y \in \mathbb{R}^n} \inf_{\theta \in \mathbb{R}^n} \left\{H(y) - \theta^\top y + \Lambda(\theta)\right\} \\ &= 2 \inf_{\theta \in \mathbb{R}^n} \left\{ \sup_{y \in \mathbb{R}^n} \left\{H(y) - \theta^\top y\right\} + \Lambda(\theta) \right\} , \end{aligned}$$

where the second equality is obtained by the minimax theorem for concave-convex functions (Rockafellar, 1970), already used in the proof of Theorem 3.5.4. But by the same Theorem 3.5.4, this bound is reached when θ is asymptotically optimal.

3.6 Numerical results

3.6.1 Long-time implied volatility

Let us now fix the parameters of the model to the values

$$b = - \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 0.7 \end{pmatrix} , \qquad a = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.3 \end{pmatrix}$$

and $\alpha = 1.5$, with initial values $S_0 = \mathbb{1}$ and $x = I_2$ and consider the problem of pricing a basket put option with log-payoff

$$H(Y_T) = \log \left(K - \sum_{j=1}^{n} \omega_j \, e^{Y_T^j} \right)_+$$

and weights $\omega_i = \frac{1}{2}$ for $i = 1, 2$.

Figure 3.6.1 shows the implied volatility smile for such an option, for $T = \frac{1}{3}$, computed by Monte Carlo over 100'000 trajectories, together with the 95% confidence interval. To sample the paths of the process, we use the exact simulation of the Wishart process described in (Ahdida and Alfonsi, 2013), Algorithm 3. Thus, we obtain the values of X_{t_i} on the regular time grid $t_i = i\Delta t$, with $i \in \mathbb{N}$ and $\Delta t > 0$. Then, for the stock, we use a trapezoidal rule since it gives a second-order weak convergence (see Section 4.3 in (Ahdida and Alfonsi, 2013) for details):

$$Y_{t_{i+1}} = Y_{t_i} - \frac{1}{2} \operatorname{diag} \left[a^\top \frac{X_{t_i} + X_{t_{i+1}}}{2} a \right] \Delta t + \operatorname{Chol} \left(a^\top \frac{X_{t_i} + X_{t_{i+1}}}{2} a \right) (Z_{t_{i+1}} - Z_{t_i}),$$

where Z is a Brownian motion sampled independently from X and $\operatorname{Chol}(M)$ is the Cholesky decomposition of a definite positive matrix M.

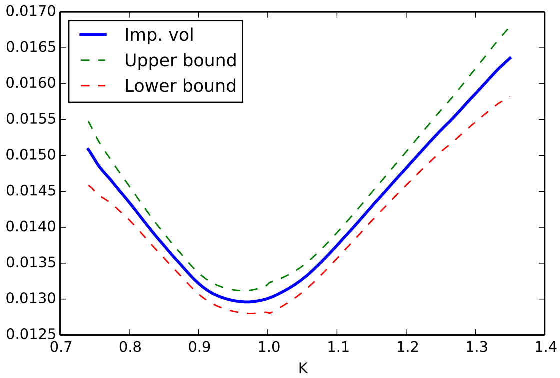

Figure 3.6.1: Basket implied volatility smile in the two-dimensional Wishart model. The upper and lower bounds correspond to the 95% confidence interval.

We next analyse the convergence of the renormalized implied volatility smile to the long-maturity limit described in Section 3.4.2. Figure 3.6.2, shows the renormalized smiles for different maturities together with the limiting smile. These smiles were computed by Monte Carlo with 100'000 trajectories and a discretization time step $\Delta t = 0.1$. We see that the convergence indeed appears to take place but it is quite slow: even for 50-year maturity using the limit as the approximation for the smile would lead to $10 - 15\%$ errors.

3.6.2 Variance reduction

We wish to test numerically the variance reduction method to price basket put options. In order to do so, we first identify the law of the Wishart process under the measure $\mathbb{P}_\theta$ and then calculate the asymptotically optimal measure change to finally test the method through Euler Monte-Carlo simulations.

Change of measure

In order to simulate from the model under $\mathbb{P}_\theta$, we need the following result.

Proposition 3.6.1. *Let $\theta \in \mathbb{R}^n$ be such that $\mathbb{E}[e^{\theta^\top Y_T}] < \infty$ and consider the change of measure $\frac{d\mathbb{P}_\theta}{d\mathbb{P}} = \frac{e^{\theta^\top Y_T}}{\mathbf{E}\left[e^{\theta^\top Y_T}\right]}$. Under $\mathbb{P}_\theta$, the process (Y_t, X_t) has dynamics*

$$dY_t = \left(r\mathbf{1} - \frac{1}{2}\left((a^\top X_t\, a)_{11}\,,\,\dots\,,\,(a^\top X_t\, a)_{nn}\right)^\top + a^\top X_t\, a\,\theta\right)\, dt + a^\top X_t^{1/2}\, dZ_t^\theta$$

and

$$dX_t = (\alpha I_n + (b + 2\,\gamma_\theta(T-t))X_t + X_t(b + 2\,\gamma_\theta(T-t)))\; dt$$

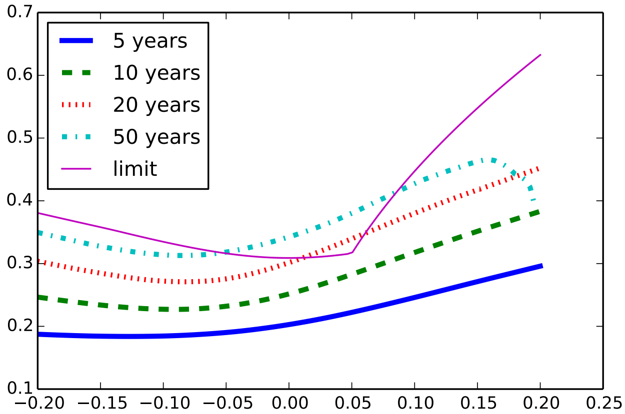

Figure 3.6.2: Convergence of the renormalized implied volatility smile to the theoretical limit in the Wishart model.

$$+ X_t^{1/2}\, dW_t^\theta + (dW_t^\theta)^\top X_t^{1/2}\,, \qquad X_0 = x\,,$$

where $\gamma_\theta(t) = -\frac{1}{2}\,(V'(t,\theta)\,V^{-1}(t,\theta) + b)$, $V(t,\theta) = V(t)$ *is given in Proposition 3.2.4 and* $\left(Z_t^\theta\right)_{t\geq 0}$ *and* $\left(W_t^\theta\right)_{t\geq 0}$ *are* $\mathbb{R}^n$ *and* $\mathbb{R}^{n\times n}$*-dimensional independent standard* $\mathbb{P}_\theta$*-Brownian motions.*

Proof. By Equation 3.2.5, the Radon-Nikodym density satisfies

$$\begin{aligned}
\zeta_t := \left.\frac{d\mathbb{P}_\theta}{d\mathbb{P}}\right|_{\mathcal{F}_t} &= \frac{\mathbb{E}\left[e^{\theta^\top Y_T}\middle|\mathcal{F}_t\right]}{\mathbb{E}\left[e^{\theta^\top Y_T}\right]} \\
&= \frac{e^{\frac{\alpha}{2}\mathrm{Tr}[b]t - \theta^\top Y_0 - r\theta^\top \mathbf{1} t - \mathrm{Tr}[\gamma_\theta(T)\,x]}}{\det[V(T,\theta)]^{-\alpha/2}\det[V(T-t,\theta)]^{\alpha/2}}\, e^{\theta^\top Y_t + \mathrm{Tr}[\gamma_\theta(T-t)X_t]}\,.
\end{aligned}$$

By Itō formula, the martingale property of ζ_t, Equations (3.2.2) and (3.2.3), and the properties of the trace, the dynamics of ζ_t is

$$\begin{aligned}
d\zeta_t = \zeta_t \Big(\theta^\top a^\top X_t^{1/2}\, dZ_t &+ \mathrm{Tr}\left[\gamma_\theta(T-t)\, X_t^{1/2}\, dW_t\right] \\
&+ \mathrm{Tr}\left[\gamma_\theta(T-t)\,(dW_t)^\top\, X_t^{1/2}\right]\Big) \\
= \zeta_t \left(\theta^\top a^\top X_t^{1/2}\, dZ_t + 2\,\mathrm{Tr}\left[\left(X_t^{1/2}\,\gamma_\theta(T-t)\right)^\top dW_t\right]\right)&\,.
\end{aligned}$$

Therefore, by Girsanov's theorem,

$$Z_t^\theta := Z_t - \int_0^t X_s^{1/2} a\,\theta\, ds$$

and

$$W_t^\theta := W_t - 2\int_0^t X_s^{1/2}\gamma_\theta(T-s)\, ds$$

are n-dimensional and $n \times n$-dimensional standard $\mathbb{P}_\theta$-Brownian motions. Replacing dZ_t and dW_t in (3.2.2) and (3.2.3) by their $\mathbb{P}_\theta$ versions finishes the proof. $\square$

We note that X is no longer a Wishart process under the probability $\mathbb{P}_\theta$, since the dynamics has time-dependent coefficients. To sample paths on the time interval $[t_i, t_{i+1}]$, we use the exact scheme for the Wishart process with the coefficient $b + 2\gamma_\theta(T - (t_i + t_{i+1})/2)$ instead of b. As explained in (Alfonsi, 2015, Section 3.3.4) on the case of the CIR process with time-dependent coefficient, this leads to a second order scheme for the weak error. Then, we can approximate Y in the same way as under $\mathbb{P}$:

$$\begin{aligned} Y_{t_{i+1}} = Y_{t_i} + \left[r\mathbf{1} - \frac{1}{2}\mathrm{diag}\left[a^\top \frac{X_{t_i} + X_{t_{i+1}}}{2} a \right] + a^\top \frac{X_{t_i} + X_{t_{i+1}}}{2} a\theta \right] \Delta t \\ + \mathrm{Chol}\left(a^\top \frac{X_{t_i} + X_{t_{i+1}}}{2} a \right) (Z_{t_{i+1}} - Z_{t_i}), \end{aligned}$$

where Z is a Brownian motion sampled independently from X. This gives a second order scheme for (X, Y).

Optimal variance reduction parameter for the European basket put option

In this section, we compute the asymptotically optimal measure to price basket put options with log-payoff $H(Y_T) = \log(K - \omega^\top e^{Y_T})_+$, for some $\omega \in (\mathbb{R}_+^*)^n$. It is shown in (Genin and Tankov, 2016, Section 4) that the function H is concave and that its convex conjugate is given by

$$\hat{H}(\theta) = \begin{cases} +\infty & \theta_k \geq 0 \text{ for some } k\,, \\ -\left(1 - \sum_k \theta_k\right) \log \frac{1 - \sum_k \theta_k}{K} - \sum_k \theta_k \log(-\theta_k/\omega_k) & \text{otherwise.} \end{cases}$$

To compute the asymptotically optimal measure change parameter θ^* using Theorem 3.5.4 we then minimize $\hat{H}(\theta) + \Lambda(\theta)$ with a numerical convex optimization algorithm.

Numerical simulations

Let us now fix the parameters of the model to the values

$$b = -\begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.5 \end{pmatrix}, \qquad a = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.12 \end{pmatrix}$$

and $\alpha = 4.5$, with initial values $S_0 = \mathbb{1}$ and $x = I_n$ and consider the problem of pricing a basket put option with log-payoff

$$H(Y_T) = \log\left(K - \sum_{j=1}^n \omega_j \, e^{Y_T^j} \right)_+$$

and weights $\omega = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$. For a wide variety of maturities T and strikes K, listed in Table 3.1, we simulate 100'000 trajectories, with the discretization step described above, with step size $\Delta = \frac{1}{40}$, under both measures $\mathbb{P}$ and $\mathbb{P}_\theta$ for the asymptotically optimal θ. The results are presented in Table 3.1.

Maturity, [Year]	Strike	Price	Std. dev.	Var. ratio	Time, [Sec]
0.50	0.7	$2.18 \cdot 10^{-7}$	$3.37 \cdot 10^{-8}$	119	202
0.50	0.8	$3.29 \cdot 10^{-5}$	$9.5 \cdot 10^{-7}$	22.5	167
0.50	0.9	$1.78 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	5.28	169
0.50	1.0	0.02620	$6.85 \cdot 10^{-5}$	3.15	167
0.50	1.1	0.10306	$9.86 \cdot 10^{-5}$	3.96	167
0.50	1.2	0.20027	$8.29 \cdot 10^{-5}$	6.68	167
0.50	1.3	0.30005	$6.41 \cdot 10^{-5}$	11.3	180
0.50	1.4	0.39999	$5.32 \cdot 10^{-5}$	16.5	168
0.25	1.0	0.01730	$5.17 \cdot 10^{-5}$	2.42	92
1.00	1.0	0.04115	$9.51 \cdot 10^{-5}$	3.76	319
2.00	1.0	0.06423	$1.39 \cdot 10^{-4}$	3.86	618
3.00	1.0	0.08319	$1.78 \cdot 10^{-4}$	3.63	934
5.00	1.0	0.11579	$2.46 \cdot 10^{-4}$	3.22	1522

Table 3.1: The variance ratio as function of the maturity and the strike for the basket put option on the Wishart stochastic volatility model.

The variance ratio is the ratio of the variance under the original measure $\mathbb{P}$ to that under the asymptotically optimal measure $\mathbb{P}_\theta$. As expected, the performance of the importance sampling algorithm is best for options far from the money, when the exercise is a rare event, but even for at the money options the variance reduction factor is significant, of the order of 3–4. The computational overhead for using the variance reduction algorithm is small: it does not exceed 20% for a small number of trajectories decreases with the number of trajectories because some precomputation steps are performed only once.

Acknowledgements. This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

Chapter 4

An asymptotic approach for the pricing of options on realized variance

I would like to thank Archil Gulisashvili[1] *for the numerous insights he shared with me during his visits to Paris Diderot, and which helped me to write this chapter.*

4.1 Introduction

The trading of variance and volatility started in the 90s with variance and volatility swaps and increased drastically in the years 2000, probably as a response of the markets to an increasing volatility risk. Gradually, a wider range of more complex financial products based on variance, such as options on realized variance appeared on the markets. An option on realized variance with strike K is a non-linear derivative with payoff

$$\left(\frac{1}{T}\int_0^T Z_t^2\,dt - K\right)_+, \tag{4.1.1}$$

where Z_t is the spot volatility process and where $\frac{1}{T}\int_0^T Z_t^2\,dt$ is the continuous version of realized variance. On the markets, only the discrete version of realized variance is traded, but the continuous version is a commonly accepted approximation when Z_t is continuous, as the first converges in probability to the second when the number of subdivisions tends to infinity (Jacod and Shiryaev, 2003, Theorem I.4.47) and as the second is the quadratic variation of the log-price process and therefore allows to solve elegantly modelling, pricing and hedging problems (see for example (Kallsen et al., 2011)). We refer

[1]Ohio University, Athens, United States of America

the reader to (Keller-Ressel and Muhle-Karbe, 2013) for a more detailed discussion on the difference between these. The authors propose an asymptotic correction term to address this issue.

The pricing of an option on realized variance is done calculating the expectation of (4.1.1). Due to the non-linearity and the path-dependence of $z \mapsto \left(\frac{1}{T}\int_0^T z_t^2\,dt - K\right)_+$ however, the calculation of the option price is a complex task that cannot be solved using straightforward methods. We model the spot volatility as a diffusion process with general drift and constant volatility and thus consider the joint process $(Y_t, Z_t)_{t\geq 0}$ with dynamics

$$\begin{aligned} dY_t &= Z_t^2\,dt & Y_0 &= 0 \\ dZ_t &= b(Z_t)\,dt + c\,dW_t & Z_0 &= z_0\,, \end{aligned}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion. Even though apparently restrictive, the generality of b allows to model Z as an Ornstein-Uhlenbeck process or as the square root of a CIR process, which are the volatility processes in the popular Stein-Stein (Stein and Stein, 1991) and Heston (Heston, 1993) models. The approach that we consider in this chapter relies on the results from (Deuschel et al., 2014a) that provide a short-time/small-noise asymptotic expansion for marginal densities of hypo-elliptic diffusions to obtain an explicit expansion of the density of Y_t in short-time.

In Section 4.2, we recall the relevant definitions and results of (Deuschel et al., 2014a). In Section 4.3, we define the process $(Y_t, Z_t)_{t\geq 0}$ and give some of its properties. In particular, we prove that the joint process of the realized variance and the volatility satisfies the hypothesis of Theorem 4.2.2 below. We then calculate, in Section 4.4, the expansion of the density of the realized variance. We finally calculate, in Section 4.5, the expansion of the price and implied volatility of put options on realized variance.

4.2 Expansion of marginal densities

In this section, we recall the main definitions and results of (Deuschel et al., 2014a). Consider a d-dimensional diffusion $(X_t^\epsilon)_{t\geq 0}$ solving the equation

$$dX_t^\epsilon = b(\epsilon, X_t^\epsilon)\,dt + \epsilon\,\sigma(X_t^\epsilon)\,dW_t\,, \qquad X_0^\epsilon = x_0^\epsilon\,, \tag{4.2.1}$$

where $(W_t)_{t\geq 0}$ is an m-dimensional Brownian motion and the functions $b : [0,1) \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma = (\sigma_1, ..., \sigma_m) : \mathbb{R}^d \to \mathrm{Lin}(\mathbb{R}^m \to \mathbb{R}^d)$ and $x_0^{\cdot} : [0,1) \to \mathbb{R}^d$ are smooth and bounded, with bounded derivatives of all orders. We assume that $b(\epsilon,\cdot) \to \sigma_0 := b(0,\cdot)$ in the sense that for every multi-index α, $\partial_x^\alpha b(\epsilon,\cdot) \to \partial_x^\alpha \sigma_0(\cdot)$ and $\partial_\epsilon b(\epsilon,\cdot) \to \partial_\epsilon b(0,\cdot)$ uniformly on compacts as $\epsilon \to 0$ and that $x_0^\epsilon = x_0 + \epsilon\,\hat{x}_0 + o(\epsilon)$ as $\epsilon \to 0$.

In order to guarantee that $Y^\epsilon_T = \Pi_l \circ X^\epsilon_T = \left(X^{\epsilon,1}_T, ..., X^{\epsilon,l}_T\right)$ admits a smooth density for every $T > 0$, we assume that the weak Hörmander condition[2] is verified at x_0, i.e. that the linear span of $\sigma_1, ..., \sigma_m$ and the Lie brackets of $\sigma_0, ..., \sigma_m$ has full rank at x_0.

Let $a \in \mathbb{R}^l$. In order to define the energy associated to a, a necessary assumption is that the set

$$\mathcal{K}_a = \{h \in H \,:\, \Pi_l \circ \phi^h_T = a\}$$

is non-empty[3]. Here, $(H, ||\cdot||_H)$ is the Cameron-Martin space of absolutely continuous trajectories starting at 0 with derivatives in $L^2([0,T], \mathbb{R}^m)$ equipped with the norm

$$||h||^2_H = \sum_{j=1}^m \int_0^T |\dot{h}^j_t|^2\, dt$$

and where ϕ^h is the solution of the controlled differential equation

$$d\phi^h_t = \sigma_0\left(\phi^h_t\right)\, dt + \sum_{j=1}^m \sigma_j\left(\phi^h_t\right)\, dh^j_t\,, \qquad \phi^h_0 = x_0\,.$$

The energy $\Lambda(a)$ associated to a, that is the minimal energy for ϕ^h_T to reach $N_a = \{(a, \cdot) \in \mathbb{R}^n\}$ in time T, is then

$$\Lambda(a) = \inf\left\{\frac{1}{2}||h||^2_H \,:\, h \in \mathcal{K}_a\right\}\,.$$

We denote $\mathcal{K}^{\min}_a \subset \mathcal{K}_a$ the set of minimizing controls. In order for $\mathcal{K}_a$ to have a manifold structure around each minimizer $h \in \mathcal{K}^{\min}_a$, we assume the invertibility, at every $h \in \mathcal{K}^{\min}_a$, of the deterministic Malliavin covariance matrix

$$C(h) = \left\langle D_h\phi^h_T, D_h\phi^h_T\right\rangle_H\,,$$

where D_h denotes the Fréchet derivative. A sufficient condition for $C(h)$ to be invertible at every $h \neq 0$ is that condition H2 in (Bismut, 1984, p. 28) is verified at x_0. Condition H2 is verified at x_0 if for every non-zero $\lambda = (\lambda_0, ..., \lambda_m) \in \mathbb{R}^{m+1}$, denoting $V = \sum_{j=0}^m \lambda_j \sigma_j$, the linear span of $\sigma_1, ..., \sigma_m$ and $[\sigma_0, V], ..., [\sigma_m, V]$ has full rank.

We define the Hamiltonian

$$\mathcal{H}(x,p) = \langle p, \sigma_0(x)\rangle + \frac{1}{2}\sum_{j=1}^m \langle p, \sigma_j(x)\rangle^2 \tag{4.2.2}$$

and give the following result.

[2]The weak Hörmander condition is a classical hypothesis for X^ϵ_T to admit a smooth density. (See (Nualart, 2006) for example.)

[3]A sufficient condition for $\mathcal{K}_a \neq \emptyset$ is the strong Hörmander condition, i.e. that $\mathrm{Lie}(\sigma_1, ..., \sigma_n)|_x$ has full rank for every $x \in \mathbb{R}^d$. (See (Jurdjevic, 1997, p. 106)

Proposition 4.2.1. *(Deuschel et al., 2014a, Prop. 2) If $h \in \mathcal{K}_a^{min}$ is a minimizing control and $C(h)$ is invertible, then there exists a unique p_0 such that*

$$\phi_t^h = \pi H_{t\leftarrow 0}(x_0, p_0)\,, \quad 0 \le t \le T\,,$$

where π denotes the projection from $\mathcal{T}\mathbb{R}^d$ to $\mathbb{R}^d$ and $H_{t\leftarrow 0}$ is the flow associated to the vector field $(\partial_p \mathcal{H}, -\partial_x \mathcal{H})$. Moreover , $(x(t), p(t)) := H_{t\leftarrow 0}(x_0, p_0)$ solve the Hamiltonian ODEs

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} \partial_p \mathcal{H}(x(t), p(t)) \\ -\partial_x \mathcal{H}(x(t), p(t)) \end{pmatrix}\,,$$

with boundary conditions $x(0) = x_0$, $x(T) = (a, \cdot)$ and $p(t) = (\cdot, 0)$. The minimizing control is recovered by $\dot{h}_t^j = \langle \sigma_j(x(t)), p(t)\rangle$.

We now give the non-degeneracy condition, which is the main hypothesis of the density expansion result (Deuschel et al., 2014a, Theorem 8). This condition generalizes the "not in the cut-locus" condition of (Ben Arous, 1988a).

Condition (ND). *We say that $\{x_0\} \times N_a$ satisfies the non-degeneracy condition (ND) if*

(i) $1 \le \#\mathcal{K}_a^{min} < \infty$,

(ii) The deterministic Malliavin covariance matrix $C(h)$ is invertible at every $h \in \mathcal{K}_a^{min}$.

(iii) x_0 is non-focal for N_a along h, for every $h \in \mathcal{K}_a^{min}$, i.e.

$$\partial_{(z,q)}|_{(z,q)=(0,0)} \pi H_{0\leftarrow T}\left(x_T + \begin{pmatrix} 0 \\ z \end{pmatrix}, \, p_T + (q\,,\, 0)\right)$$

is non-degenerate, where $(x_T, p_T) = H_{T\leftarrow 0}(x_0, p_0(h))$.

We now cite the density expansion result.

Theorem 4.2.2. *(Deuschel et al., 2014a, Theorem 8) Let X^ϵ be the solution of (4.2.1), where $b(\epsilon, \cdot)$ and x_0^ϵ converge in the previously given sense. Assume also that X^ϵ satisfies the weak Hörmander condition at x_0. Assume finally that $\{x_0\} \times N_a$ satisfies (ND). Let $a \in \mathbb{R}^l$. If $\#\mathcal{K}_a^{min} = 1$, then the energy $\Lambda(a)$ is smooth, otherwise, if $\#\mathcal{K}_a^{min} > 1$, we assume that it is. Then there exists $c_0 = c_0(x_0, a, T) > 0$ such that Y_T^ϵ admits a density $f^\epsilon(\cdot, T)$ with expansion*

$$f^\epsilon(a, T) = e^{-\frac{\Lambda(a)}{\epsilon^2}}\, e^{\frac{\max\left\{\Lambda'(a)\cdot \hat{Y}_T(h)\,:\, h\in\mathcal{K}_a^{min}\right\}}{\epsilon}}\, \epsilon^{-l}\left(c_0 + \mathcal{O}(\epsilon)\right)\,,$$

as $\epsilon \to 0$, where $\hat{Y}_t(h) = \Pi_l \hat{X}_t$ and $(\hat{X}_t)_{t\le T}$ is the solution of the controlled ODE

$$d\hat{X}_t = \left(\partial_X b(0, \phi_t^h) + \partial_X \sigma(\phi_t^h)\, \dot{h}_t\right) \hat{X}_t\, dt + \partial_\epsilon b(0, \phi_t^h)\, dt\,, \qquad \hat{X}_0 = \hat{x}_0\,.$$

4.3 Definition and properties of the integrated variance process

4.3.1 The integrated variance process

As mentioned, we define $X_t^\epsilon = (Y_t^\epsilon, Z_t^\epsilon)$ to be the joint process with dynamics

$$\begin{aligned} dY_t^\epsilon &= g(Z_t^\epsilon)\,dt & Y_0^\epsilon &= 0 \\ dZ_t^\epsilon &= \epsilon^2 b(Z_t^\epsilon)\,dt + \epsilon\, c\, dW_t & Z_0^\epsilon &= z_0\,, \end{aligned} \tag{4.3.1}$$

which corresponds to (4.2.1) where $x_0^\epsilon = (0, z_0)$,

$$b(\epsilon, (y,z)) = \begin{pmatrix} g(z) \\ \epsilon^2\, b(z) \end{pmatrix} \underset{\epsilon\to 0}{\longrightarrow} \sigma_0((y,z)) = \begin{pmatrix} g(z) \\ 0 \end{pmatrix}$$

and

$$\sigma_1(x) = \begin{pmatrix} 0 \\ c \end{pmatrix}.$$

In order for $b(\epsilon, x)$ to be in $C_b^\infty([0,1)\times\mathbb{R}^2, R^2)$, the set of bounded functions with bounded derivatives of all orders, we assume that $z \mapsto b(z) \in C_b^\infty(\mathbb{R})$. Furthermore, in order for Y_t^ϵ to correspond to realized variance, while satisfying the hypotheses, we choose

$$g(z) = \frac{z^2\, e^{-\frac{1}{R+1-|z|}}\, \mathbb{1}_{\{|z|<R+1\}} + (R+1)^2\, e^{-\frac{1}{|z|-R}}\, \mathbb{1}_{\{|z|>R\}}}{e^{-\frac{1}{R+1-|z|}}\, \mathbb{1}_{\{|z|<R+1\}} + e^{-\frac{1}{|z|-R}}\, \mathbb{1}_{\{|z|>R\}}} \tag{4.3.2}$$

for R arbitrarily large, following an idea in (Kunitomo and Takahashi, 2001). Eq. (4.3.2) is a $C_b^\infty(\mathbb{R})$ version of the $z \mapsto z^2$ function, i.e. $g(z) = z^2$ for $|z| < R$, while $g(z) = (R+1)^2$ for $|z| > R+1$.

Lemma 4.3.1. *The process X_t^ϵ satisfies the Bismut H2 condition at $x_0 = (0, z_0)$ and the deterministic Malliavin covariance matrix $c(h)$ is therefore invertible for every non-zero $h \in H$. In particular, the weak Hörmander condition is verified.*

Proof. Let $\lambda = (\lambda_0, \lambda_1) \in \mathbb{R}^2$ be a non-zero vector and

$$V(x) = \lambda_0\sigma_0(x) + \lambda_1\sigma_1(x) = \begin{pmatrix} \lambda_0\, g(z) \\ \lambda_1\, c \end{pmatrix}.$$

Then

$$[\sigma_0, V](x) = \lambda_0[\sigma_0,\sigma_0](x) + \lambda_1[\sigma_0,\sigma_1](x) = -\lambda_1 \begin{pmatrix} c\, g'(z) \\ 0 \end{pmatrix}$$

and

$$[\sigma_1, V](x) = \lambda_0[\sigma_1,\sigma_0](x) + \lambda_1[\sigma_1,\sigma_1](x) = \lambda_0 \begin{pmatrix} c\, g'(z) \\ 0 \end{pmatrix}$$

Since R is arbitrarily large, $g'(z_0) = 2z_0 \neq 0$. Therefore, since $\lambda \neq 0$, then $\sigma_1(x_0)$, $[\sigma_0, V](x_0)$ and $[\sigma_1, V](x_0)$ span $\mathbb{R}^2$. Condition H2 is then verified at x_0 and the invertibility of $C(h)$ for every non-zero $h \in H$ then follows from (Bismut, 1984, Theorem 1.10). The weak Hörmander condition is obtained by taking $\lambda = (0, 1)$. $\square$

Remark 4.3.2. *In general, the H2 condition is much stronger than the weak Hörmander condition. When $m = 1$ however, the two are equivalent.*

4.3.2 Hamiltonian equations and optimal control

In this section, we formulate and solve the Hamiltonian equations. The Hamiltonian is

$$\begin{aligned}\mathcal{H}\left(\begin{pmatrix} y \\ z \end{pmatrix}, (p\,,\,q)\right) &:= \left\langle \begin{pmatrix} p \\ q \end{pmatrix}, \sigma_0(x) \right\rangle + \frac{1}{2}\left\langle \begin{pmatrix} p \\ q \end{pmatrix}, (\sigma_1\sigma_1^\top)(x)\begin{pmatrix} p \\ q \end{pmatrix}\right\rangle \\ &= pg(z) + \frac{1}{2}\,c^2q^2\,.\end{aligned}$$

Given a certain a, provided R is taken large enough for the $(z_t)_{t\leq T}$ trajectory to stay in the ball $B(0, R)$, Prop. 4.2.1 yields the Hamiltonian equations

$$\dot{y}_t = g(z_t) = z_t^2 \tag{4.3.3}$$

$$\dot{z}_t = c^2 q_t \tag{4.3.4}$$

$$\dot{p}_t = 0 \tag{4.3.5}$$

$$\dot{q}_t = -p_t g'(z_t) = -2p_t z_t\,, \tag{4.3.6}$$

with the boundary conditions $z_0 = z_0$, $y_0 = 0$, $y_T = a$ and $q_T = 0$.

Lemma 4.3.3. *Let (y_t, z_t, p, q_t) be the solution of (4.3.3)-(4.3.6) with the associated boundary conditions. Then the minimizing control is*

$$\dot{h}_t = c\,q_t\,.$$

The energy associated to a is then

$$\Lambda(a) = pa - \frac{z_0}{2}\,q_0\,.$$

Proof. By Prop. 4.2.1, the minimizing control associated to the solution (y_t, z_t, p, q_t) of (4.3.3)-(4.3.6) with the associated boundary conditions is given by

$$\dot{h}_t = \left\langle \sigma_1\begin{pmatrix} y_t \\ z_t \end{pmatrix}, \begin{pmatrix} p_t \\ q_t \end{pmatrix}\right\rangle = c\,q_t\,,$$

where $p_t = p$ is the constant solution of (4.3.5). Following the idea of the proof of (Deuschel et al., 2014b, Lemma 7), we find

$$\begin{aligned}|\dot{h}_t|^2 = c^2\, q_t^2 = q_t\, \dot{z}_t &= \partial_t(q_t\, z_t) - \dot{q}_t z_t \\ &= \partial_t(q_t\, z_t) + 2pz_t^2 = \partial_t(q_t\, z_t) + 2p\dot{y}_t\,.\end{aligned}$$

The energy of a is therefore

$$\begin{aligned}\Lambda(a) &= \frac{1}{2}\int_0^T |\dot{h}_t|^2 dt = \frac{1}{2}\int_0^T \partial_t(q_t\, z_t) + 2p\dot{y}_t dt \\ &= \frac{1}{2}\left(q_T\, z_T - q_0\, z_0 + 2py_T - 2py_0\right) = pa - \frac{z_0}{2}\, q_0\,,\end{aligned}$$

where the last equality is obtained by using the boundary conditions. □

Below, we solve the Hamiltonian equations. In the rest of this Chapter we denote $r = r(p) = \sqrt{2p}\, c \in \mathbb{C}$, where $p = p_t$ is the constant solution of (4.3.5).

Lemma 4.3.4. *Let $s, t \in [0, T]$. Then the solution $(y_t, z_t, p_t, q_t)_{t\le T}$ of equations (4.3.4)-(4.3.6) satisfies*

$$z_t = z_s\, \cos\left(r\,(t-s)\right) + q_s\, \frac{c^2}{r}\, \sin\left(r\,(t-s)\right) \tag{4.3.7}$$

$$q_t = q_s \cos\left(r\,(t-s)\right) - z_s\, \frac{r}{c^2}\, \sin\left(r\,(t-s)\right)\,, \tag{4.3.8}$$

and

$$\begin{aligned}y_t = y_s &+ \frac{z_s^2}{2}\left(1 + \frac{\sin\left(2\,r\,(t-s)\right)}{2\,r\,(t-s)}\right)(t-s) \\ &+ z_s\, q_s\, c^2\, \frac{1-\cos\left(2\,r\,(t-s)\right)}{4\,r^2} \\ &+ q_s^2\, \frac{c^4}{2r^2}\left(1 - \frac{\sin\left(2\,r\,(t-s)\right))}{2\,r\,(t-s)}\right)(t-s)\,.\end{aligned} \tag{4.3.9}$$

Proof. Eq. (4.3.5) implies that $p_t = p$ is constant. Eqs. (4.3.4) and (4.3.6) then become a 2-dimensional linear ordinary differential equation, whose solution verifies

$$\begin{aligned}\begin{pmatrix} z_t \\ q_t \end{pmatrix} &= \begin{pmatrix} \cos\left(r\,(t-s)\right) & \frac{c}{\sqrt{2p}}\, \sin\left(r\,(t-s)\right) \\ -\frac{\sqrt{2p}}{c}\, \sin\left(r\,(t-s)\right) & \cos\left(r\,(t-s)\right) \end{pmatrix}\begin{pmatrix} z_s \\ q_s \end{pmatrix} \\ &= \begin{pmatrix} \cos\left(r\,(t-s)\right) & \frac{c^2}{r}\, \sin\left(r\,(t-s)\right) \\ -\frac{r}{c^2}\, \sin\left(r\,(t-s)\right) & \cos\left(r\,(t-s)\right) \end{pmatrix}\begin{pmatrix} z_s \\ q_s \end{pmatrix}\end{aligned}$$

thus proving (4.3.7) and (4.3.8). Then

$$y_t = y_s + \int_s^t z_u^2\, du$$

where

$$\begin{aligned}\int_s^t z_u^2\,du &= \int_s^t \left(z_s\,\cos\left(r\,(u-s)\right) + q_s\,\frac{c^2}{r}\,\sin\left(r\,(u-s)\right)\right)^2 du\\ &= \frac{z_s^2}{2}\left(1+\frac{\sin\left(2\,r\,(t-s)\right)}{2\,r\,(t-s)}\right)(t-s)\\ &\quad + z_s\,q_s\,c^2\,\frac{1-\cos\left(2\,r\,(t-s)\right)}{4\,r^2}\\ &\qquad + q_s^2\,\frac{c^4}{2\,r^2}\left(1-\frac{\sin\left(2\,r\,(t-s)\right))}{2\,r\,(t-s)}\right)(t-s)\,,\end{aligned}$$

thus proving (4.3.9). □

Theorem 4.3.5. *The energy associated with a is*

$$\Lambda(a) = \frac{r^2}{2\,c^2}\,z_0^2\,T\left(\frac{a}{z_0^2\,T}-\frac{\tan\left(r\,T\right)}{r\,T}\right)\,,$$

where r solves the equation

$$1+\cos\left(2\,r\,T\right)-\frac{z_0^2\,T}{a}\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right)=0\,. \tag{4.3.10}$$

Proof. Let (y_t, z_t, p, q_t) be the solution of (4.3.3)-(4.3.6) with the boundary conditions $z_0 = z_0$, $y_0 = 0$, $y_T = a$ and $q_T = 0$. Then by (4.3.8),

$$0 = q_T = q_0\cos\left(r\,T\right) - z_0\,\frac{r}{c^2}\,\sin\left(r\,T\right)\,.$$

Hence

$$q_0 = z_0\,\frac{r}{c^2}\,\tan\left(r\,T\right)$$

and therefore

$$\Lambda(a) = p a - \frac{z_0}{2}\,q_0 = \frac{r^2}{2\,c^2}\,z_0^2\,T\left(\frac{a}{z_0^2\,T}-\frac{\tan\left(r\,T\right)}{r\,T}\right)\,.$$

Furthermore, (4.3.7) implies that

$$z_0 = z_T\,\cos\left(r\,T\right) - q_T\,\frac{c^2}{r}\,\sin\left(r\,T\right) = z_T\,\cos\left(r\,T\right)\,.$$

Inserting this in (4.3.9) applied for $t = 0$ and $s = T$, we obtain

$$\begin{aligned}0 = y_0 &= a - \frac{z_T^2}{2}\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right)T\\ &= a - \frac{z_0^2\,T}{2\,\cos^2\left(r\,T\right)}\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right)\end{aligned}$$

$$= a - \frac{z_0^2\,T}{1+\cos\left(2\,r\,T\right)}\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right).$$

Note that $2\,r\,T = (2k+1)\pi$ for $k \in \mathbb{Z}$ is not a solution. Indeed, $1 + \cos\left((2k+1)\pi\right) = 0$, while $1 + \frac{\sin((2k+1)\pi)}{(2k+1)\pi} = 1$. We can therefore multiply by $1+\cos\left(2\,r\,T\right)$ to obtain (4.3.10). □

Theorem 4.3.5 gives the complete characterization of the energy, up to the calculation of r. Since, $p \in \mathbb{R}$, we therefore need to study the solutions of eq. (4.3.10) in $r \in \mathbb{R}_+ \cup i\mathbb{R}_+$.

Lemma 4.3.6. *If $\frac{z_0^2 T}{a} \in [1,\infty)$, (4.3.10) admits a unique solution in $i\mathbb{R}_+$, while if $\frac{z_0^2 T}{a} \in (0,1)$, (4.3.10) has no solution in $i\mathbb{R}_+$.*

Proof. Let $u = -2rTi \in \mathbb{R}_+$ and $k = \frac{z_0^2 T}{a} > 0$. Notice that 0 solves (4.3.10) if and only if $k = 1$. Let us write (4.3.10) as

$$\frac{f(u)}{u} = 0\,,$$

where

$$f(u) = \left(1+\cosh(u)\right)u - k\,\left(u+\sinh(u)\right)\,.$$

Then $r > 0$ solves (4.3.10) if and only if $u > 0$ solves $f(u) = 0$. The two first derivatives of f are

$$\begin{aligned} f'(u) &= \sinh(u)u + (1-k)(1+\cosh(u)) \\ f''(u) &= \cosh(u)u + (2-k)\sinh(u) \\ &= f(u) + (k-1)\,u + 2\,\sinh(u)\,. \end{aligned}$$

Assume first that $k > 1$. Then $f(0) = 0$ and $f'(0) = 2\,(1-k) < 0$. Therefore, there exists $u_1 > 0$ such that $f(u) < 0$ for every $u \in (0,u_1]$. Since $f(u) \to \infty$ as $u \to \infty$, there exists $u_2 > u_1$ such that $f(u) < 0$ for $u \in (0,u_2)$ and $f(u_2) = 0$. In particular, such a u_2 verifies $f'(u_2) \geq 0$. Since $f''(u) = f(u) + (k-1)\,u + 2\,\sinh(u)$, then $f''(u) > 0$ for $u \geq u_2$. Therefore, $f(u) > 0$ for $u > u_2$ and u_2 is the unique strictly positive solution of f.
Assume now that $k \in (0,1]$. Then $f(0) = 0$ and $f'(u) > 0$ for every $u > 0$. Therefore $f(u) > 0$ for every $u > 0$ and f has no strictly positive solution. □

Lemma 4.3.7. *If $\frac{z_0^2 T}{a} \in (0,1)$, (4.3.10) admits a unique solution in $\left(0, \frac{\pi}{2T}\right)$, while if $\frac{z_0^2 T}{a} \in [1,\infty)$, (4.3.10) has no solution in $\left(0, \frac{\pi}{2T}\right)$.*

Proof. Denote $u = 2rT$, $k = \frac{z_0^2 T}{a}$. Then (4.3.10) becomes

$$\frac{f(u)}{u} = 0\,,$$

where

$$f(u) = (1+\cos(u))\,u - k\,\left(u+\sin(u)\right)\,.$$

Then $r>0$ solves (4.3.10) if and only if $u>0$ solves $f(u)=0$.

Assume that $k\in(0,1)$. The derivative

$$f'(u) = -\sin(u)u + (1-k)\,(1+\cos(u))\ ,$$

in $(0,\pi)$, cancels exactly once at u_1, where u_1 is the unique solution of

$$u\tan(u/2) = 1-k$$

in $(0,\pi)$. Therefore, f is strictly monotonous on $(0,u_1)$ and (u_1,π). Since $f(0)=0$ and $f(u_1)>0$, then f is increasing on $(0,u_1)$ and does not cancel on this interval. Furthermore, $f(\pi)=-k\pi<0$ and therefore f cancels exactly once in (u_1,π).

Assume now that $k\geq 1$. Then the derivative

$$f'(u) = -\sin(u)u + (1-k)\,(1+\cos(u))\ ,$$

is always strictly negative in $(0,\pi)$. Since $f(0)=0$, then f is always strictly negative in $(0,\pi)$. Transposing from u to r finishes the proof. $\square$

Theorem 4.3.8. *The energy associated with a is*

$$\Lambda(a) = \frac{z_0^2}{4\,c^2}\,r\,\frac{2\,r\,T-\sin\left(2\,r\,T\right)}{1+\cos\left(2\,r\,T\right)}\ ,$$

where r is the unique solution of equation

$$1+\cos\left(2\,r\,T\right) - \frac{z_0^2\,T}{a}\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right) = 0 \tag{4.3.11}$$

in the set

$$\mathcal{I} := \left(0,\ \frac{\pi}{2T}\right)\cup i\mathbb{R}_+ \subset \mathbb{C}\,.$$

Proof. The combination of Lemmas 4.3.6 and 4.3.7 proves the uniqueness of the solution of (4.3.11) in $\mathcal{I}$. Due to the form of the Hamiltonian equations, Prop. 4.2.1 guaranties the uniqueness of the optimal control. Note first that if $a=z_0^2\,T$, then $h=0$ is the optimal control and $r=0$ is therefore the optimal solution. We then consider the case $a\neq z_0^2\,T$.

For $h\in H$, define (y_t^h, z_t^h) the solution of the controlled ODE

$$d\begin{pmatrix} y_t^h \\ z_t^h \end{pmatrix} = \sigma_0\left(\begin{pmatrix} y_t^h \\ z_t^h \end{pmatrix}\right)dt + \sigma_1\left(\begin{pmatrix} y_t^h \\ z_t^h \end{pmatrix}\right)dh_t = \begin{pmatrix} (z_t^h)^2\,dt \\ c\,dh_t \end{pmatrix}\ ,$$

with initial condition $(0,z_0)$. In particular, by Prop. 4.2.1, if $h\in\mathcal{K}_a^{\min}$, (y_t^h, z_t^h) solves the Hamiltonian equations. Let us start by noting that if

$h \in \mathcal{K}_a^{\min}$, then $z_t^h \geq 0$. Indeed, define $\omega \in H$ such that $\dot{\omega}_t = \dot{h}_t \mathbb{1}_{\{z_t^h > 0\}} - \dot{h}_t \mathbb{1}_{\{z_t^h < 0\}}$. We then have that

$$z_t^\omega = z_0 + c \int_0^t \dot{h}_s \mathbb{1}_{\{z_s^h > 0\}} - \dot{h}_s \mathbb{1}_{\{z_s^h < 0\}} \, ds = |z_t^h|$$

is simply z_t^h reflected at 0 and therefore $y_T^\omega = \int_0^T (z_t^\omega)^2 \, dt = a$. Furthermore $\int_0^T |\dot{w}_t|^2 \, dt = \int_0^T |\dot{h}_t|^2 \, dt$ and therefore $w \in \mathcal{K}_a^{\min}$. By uniqueness of the optimal control, $w = h$ and $z_t^h \geq 0$.

Assume that $a < z_0^2 T$. Then $\int_0^T (z_t^h)^2 \, dt < z_0^2 T$. This implies that there exists a non-trivial interval $I \subset [0, T]$ such that $\dot{z}_t^h < 0$ for $t \in I$. From eq. (4.3.4), we therefore have $q_t < 0$ for every $t \in I$. Since $z \geq 0$, eq. (4.3.6) then implies that the sign of $\dot{q}$ is constant. Since, $q_t < 0$ on I and $q_T = 0$, we have $\dot{q} \geq 0$, which implies that $p < 0$ and therefore $r \in i\,\mathbb{R}_+$.

Assume now that $a > z_0^2 T$. The same argument allows to conclude that q_t is a positive function that decreases strictly to $q_T = 0$. Combining (4.3.7) and (4.3.8), we obtain

$$q_t = \frac{z_0 \, r}{c^2 \, \cos(r\,T)} \sin(r\,(T - t)) \ .$$

Since the only values of r in $\mathbb{R}_+ \cup i\,\mathbb{R}_+$ such that

$$\frac{r \, \sin(r\,(T-t))}{\cos(r\,T)} > 0 \,,$$

for every $t < T$ are the values of r in $\left(0, \frac{\pi}{2T}\right)$, Theorem 4.3.5 implies that

$$\Lambda(a) = \frac{r^2}{2\,c^2} \, z_0^2 \, T \left(\frac{a}{z_0^2 \, T} - \frac{\tan(r\,T)}{r\,T} \right) \ ,$$

where r is the unique solution of (4.3.11) in $\mathcal{I}$. Replacing $\frac{a}{z_0^2 T}$ by

$$\frac{2\,r\,T + \sin(2\,r\,T)}{2\,r\,T\,(1 + \cos(2\,r\,T))}$$

and $\tan(r\,T)$ by

$$\frac{\sin(2\,r\,T)}{1 + \cos(2\,r\,T)}$$

finishes the proof. $\square$

Theorem 4.3.9. *$\{(0, z_0)\} \times N_a$ satisfies Condition (ND).*

Proof. As mentioned previously, the linear nature of the Hamiltonian equations implies that $\#\mathcal{K}_a^{\min} = 1$ and the H2 condition implies the invertibility of the Malliavin deterministic covariance matrix. Let us show that $(0, z_0)$ is non-focal for N_a along the optimal control h. Using Lemma 4.3.4, we find that

$$\pi H_{0\leftarrow T}\left(\left(\begin{array}{c} a \\ z_T + z \end{array}\right), (p+q\,,\,0)\right) = \left(\begin{array}{c} a - \frac{(z_T+z)^2\,T}{2}\left(1+\frac{\sin(2\,r\,T)}{2\,r\,T}\right) \\ (z_T+z)\,\cos\left(r\,T\right) \end{array}\right)$$

where $r = r(p+q) = \sqrt{2(p+q)}\,c$. Since $r' = \frac{\sqrt{2}c}{2\sqrt{p}} = \frac{c^2}{r}$, the Jacobian matrix of the projection of the backward Hamiltonian flow is

$$\begin{aligned}
&\partial_{(z,q)}|_{(z,q)=(0,0)}\pi H_{0\leftarrow T}\left(\left(\begin{array}{c} a \\ z_T + z \end{array}\right), (p+q\,,\,0)\right) \\
=&\left(\begin{array}{cc} -z_T\,T\left(1+\frac{\sin(2\,r\,T)}{2\,r\,T}\right) & -\frac{z_T^2\,T\,c^2}{2\,r^2}\left(\cos\left(2\,r\,T\right)-\frac{\sin(2\,r\,T)}{2\,r\,T}\right) \\ \cos\left(r\,T\right) & -z_T\,T^2\,c^2\,\frac{\sin(r\,T)}{r\,T} \end{array}\right).
\end{aligned}$$

The determinant of the Jacobian matrix is therefore

$$\begin{aligned}
&\det \partial_{(z,q)}|_{(z,q)=(0,0)}\pi H_{0\leftarrow T}\left(\left(\begin{array}{c} a \\ z_T + z \end{array}\right), (p+q\,,\,0)\right) \\
=&z_T^2\,c^2\,T^3\left[\left(1+\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right)\frac{\sin\left(r\,T\right)}{r\,T}\right. \\
&\qquad\left.+\frac{1}{2\,r^2\,T^2}\left(\cos\left(2\,r\,T\right)-\frac{\sin\left(2\,r\,T\right)}{2\,r\,T}\right)\cos\left(r\,T\right)\right] \\
=&z_T^2\,c^2\,T^3\left[\frac{\sin\left(r\,T\right)}{r\,T}+2\,\cos\left(r\,T\right)\,\frac{2\,r\,T-\sin\left(2\,r\,T\right)}{(2\,r\,T)^3}\right].
\end{aligned}$$

Let us show that the determinant is non-zero. Notice first that, from (4.3.7), we know that $z_T = \frac{z_0}{\cos(r\,T)} > 0$ for every $r \in \mathcal{I}$.

If $a = z_0^2\,T$, then $r = 0$, $z_T = z_0$ and

$$\det \partial_{(z,q)}|_{(z,q)=(0,0)}\pi H_{0\leftarrow T}\left(\left(\begin{array}{c} a \\ z_T + z \end{array}\right), (p+q\,,\,0)\right) = \frac{4\,z_0^2\,c^2\,T^3}{3} > 0\,.$$

If $a > z_0^2\,T$, then $r \in \left(0,\,\frac{\pi}{2\,T}\right)$. In this case, $\frac{\sin(r\,T)}{r\,T}$, $\cos\left(r\,T\right)$ and $2\,r\,T - \sin\left(2\,r\,T\right)$ are all strictly positive, hence the determinant is strictly positive. Finally, if $a < z_0^2\,T$, then $r \in i\,\mathbb{R}_+^*$. Denoting $u = -2\,i\,r\,T \in \mathbb{R}_+^*$,

$$\begin{aligned}
&\det \partial_{(z,q)}|_{(z,q)=(0,0)}\pi H_{0\leftarrow T}\left(\left(\begin{array}{c} a \\ z_T + z \end{array}\right), (p+q\,,\,0)\right) \\
&\qquad = z_T^2\,c^2\,T^3\left[\frac{2\,\sinh\left(u/2\right)}{u}+2\,\cosh\left(u/2\right)\,\frac{\sinh\left(u\right)-u}{u^3}\right],
\end{aligned}$$

where all the terms are individually strictly positive. The determinant of the Jacobian matrix of the projected backward Hamiltonian flow is therefore always strictly positive, thus proving that $(0, z_0)$ is non-focal for N_a along h. Condition (ND) is therefore verified. $\square$

4.3.3 Derivatives of the energy

We know proceed to some calculations concerning the derivatives of the energy. These results will be necessary in the rest of the chapter.

Lemma 4.3.10. *Let $a \in (0, \infty)$ and let $r = r(a)$ be the unique solution of eq. (4.3.11) in $\mathcal{I}$. Then*

$$\Lambda'(a) = \frac{r^2}{2\,c^2}\,. \tag{4.3.12}$$

As a consequence, $\Lambda'(a) \in \left(-\infty,\, \frac{\pi^2}{8\,c^2\,T^2}\right)$. Furthermore,

$$\Lambda''(a) = \frac{3}{4z_0^2\,c^2\,T^3} + \mathcal{O}\big(|r|^2\big)\ , \tag{4.3.13}$$

when $a \to z_0^2\,T$.

Proof. Consider the function $a \mapsto r(a)$. It is a 1-1 map from $(0, \infty)$ to $\mathcal{I}$ with inverse

$$r^{-1}(r) = z_0^2\,T\,\frac{1 + \frac{\sin(2\,r\,T)}{2\,r\,T}}{1 + \cos(2\,r\,T)}\,.$$

Since $1 + \cos(2\,r\,T) \neq 0$ on $\mathcal{I}$, r^{-1} is a differentiable function. Then

$$\begin{aligned}\Lambda'(a) &= \frac{r^2}{2\,c^2} + \left(\frac{r}{c^2}\,a - \frac{z_0^2}{2\,c^2}\,\tan\left(r\,T\right) - \frac{z_0^2\,T}{c^2}\,r\,\frac{1}{1+\cos\left(2\,r\,T\right)}\right)\,r' \\ &= \frac{r^2}{2\,c^2} + \frac{a\,r\,r'}{c^2\left(1+\cos(2\,r\,T)\right)}\left(1 + \cos(2\,r\,T) - \frac{z_0^2\,T}{a}\left(1 + \frac{\sin(2\,r\,T)}{2\,r\,T}\right)\right)\ ,\end{aligned}$$

where the last factor is 0 by definition of r, thus proving (4.3.12). Furthermore, $\Lambda''(a) = \frac{r\,r'}{c^2}$ where $r' = \frac{1}{(r^{-1})'(r)}$. Then, denoting $u = 2\,r\,T$, we have

$$\begin{aligned}\Lambda''(a) &= \frac{u^2}{2\,z_0^2\,c^2\,T^2\,u'}\left(1 - \frac{(u+\sin(u))(1+\cos(u) - u\sin(u))}{u\,(1+\cos(u))^2}\right)^{-1} \\ &= \frac{3}{4\,z_0^2\,c^2\,T^3} + \mathcal{O}\big(r^2\big)\ ,\end{aligned}$$

as $r \to 0$ and therefore as $a \to z_0^2\,T$. □

4.4 Asymptotic expansion of the density

In this section, we provide the asymptotic expansion of the density. The proof relies on (Deuschel et al., 2014a, Theorem 8) with a slight extension of the proof to obtain an explicit value for the c_0 coefficient.

Theorem 4.4.1. *Consider the process $X^\epsilon = (Y^\epsilon_t, Z^\epsilon_t)_{t\in[0,T]}$ defined in (4.3.1). Then, for every $a \in \left(0, \frac{R^2 T}{2}\left(1 + \frac{z_0/R}{\arccos(z_0/R)}\sqrt{1 - z_0^2/R^2}\right)\right)$ where R is the arbitrarily large constant in (4.3.2), Y^ϵ_t admits a smooth density with expansion*

$$f_{Y^\epsilon_T}(a) = \epsilon^{-1} e^{-\Lambda(a)/\epsilon^2}\left(c_0(a) + o(1)\right), \qquad \text{as } \epsilon \to 0\,,$$

where

$$c_0(a) = \frac{1}{\sqrt{2\pi \mathcal{A}(2rT)}} \frac{\cos^{3/2}(r\,T)}{2\,z_0\,c\,T^{3/2}}\, e^{\frac{1}{c^2}\int_{z_0}^{z_T} b(x)\,dx}\,,$$

and

$$\mathcal{A}(u) = \frac{u^3 + 6\,u\,\cos(u) + 3\,(u^2-2)\sin(u)}{6\,u^3}\,.$$

Proof. Let z_t be the solution of (4.3.4). Then for

$$a \in \left(0, \frac{R^2 T}{2}\left(1 + \frac{z_0/R}{\arccos(z_0/R)}\sqrt{1 - z_0^2/R^2}\right)\right)\,,$$

$|z_t| < R$ for every $t \le T$ and therefore $g(z_t) = z_t^2$. Then by Lemma 4.3.1, the weak Hörmander condition is verified at $(0, z_0)$ and by Theorem 4.3.9, $(0, z_0) \times N_a$ verifies Condition (ND). By Theorem 4.2.2, there exists $c_0 = c_0(x_0, a, T) > 0$ such that Y^ϵ_T admits a smooth density $f_{Y^\epsilon_T}(a)$ with expansion

$$f_{Y^\epsilon_T}(a) = \epsilon^{-1} e^{-\Lambda(a)/\epsilon^2}\left(c_0 + o(1)\right), \qquad \text{as } \epsilon \to 0\,,$$

as $\epsilon \to 0$. In order to calculate the c_0 coefficient we proceed, as in the proof of Theorem 4.2.2 in (Deuschel et al., 2014a), by applying the Laplace method on the Wiener space $C([0,\, T],\, \mathbb{R})$, following the methodology of (Ben Arous, 1988a).

By Fourier inversion, for any function F in $C_b^\infty(\mathbb{R})$, we have that

$$\begin{aligned}
f_{Y^\epsilon_T}(a) e^{-F(a)/\epsilon^2} &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ia\xi}\left(\int_{\mathbb{R}} e^{i\xi\theta} f_{Y^\epsilon_T}(\theta)\, e^{-F(\theta)/\epsilon^2} d\theta\right) d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{i\xi(\theta - a)}\, e^{-F(\theta)/\epsilon^2}\, f_{Y^\epsilon_T}(\theta) d\theta\, d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \mathbb{E}\left(\exp\left(\xi\,(Y^\epsilon_T - a)\right)\, e^{-F(Y^\epsilon_T)/\epsilon^2}\right) d\xi \\
&= \frac{1}{2\pi\epsilon}\int_{\mathbb{R}} \mathbb{E}\left(\exp\left(i\zeta\,\frac{Y^\epsilon_T - a}{\epsilon}\right)\, e^{-F(Y^\epsilon_T)/\epsilon^2}\right) d\zeta\,.
\end{aligned}$$

We choose $F \in C_b^\infty(\mathbb{R})$ as a smoothly bounded version of

$$\begin{aligned}
F(y) &= \lambda(y-a)^2 - \left[\Lambda'(a)(y-a) + \frac{\Lambda''(a)}{2}(y-a)^2\right] \\
&= \left(\lambda - \frac{\Lambda''(a)}{2}\right)(y-a)^2 - \Lambda'(a)\,(y-a)\,.
\end{aligned}$$

Note that F has derivatives

$$\begin{aligned} F'(y) &= (2\lambda - \Lambda''(a))\,(y-a) - \Lambda'(a) \\ F''(y) &= 2\lambda - \Lambda''(a)\,. \end{aligned}$$

The function $F(\cdot) + \Lambda(\cdot)$ then admits a non-degenerate minimum at a. Let $h \in \mathcal{K}_a^{\min}$ be the minimising control. Applying Girsanov's theorem to the martingale $-\frac{1}{\epsilon}\int_0^t \dot{h}_s\,dW_s$, we find

$$\begin{aligned} f_{Y_T^\epsilon}(a) &= \frac{1}{2\pi\epsilon}\int_{\mathbb{R}} \mathbb{E}\left(\exp\left(i\zeta\,\frac{Y_T^\epsilon - a}{\epsilon}\right)\, e^{-F(Y_T^\epsilon)/\epsilon^2}\right) d\zeta \\ &= \frac{1}{2\pi\epsilon}\int_{\mathbb{R}} \mathbb{E}\left(\exp\left(-\frac{1}{\epsilon}\int_0^T \dot{h}_s\,dW_s - \frac{\Lambda(a)}{\epsilon^2}\right)\right. \\ &\qquad\qquad \left.\cdot \exp\left(i\zeta\,\frac{\bar{Y}_T^\epsilon - a}{\epsilon}\right)\, e^{-F(\bar{Y}_T^\epsilon)/\epsilon^2}\right) d\zeta\,, \end{aligned} \tag{4.4.1}$$

where

$$\begin{aligned} d\bar{Y}_t^\epsilon &= g(Z_t^\epsilon)\,dt & Y_0^\epsilon &= 0 \\ d\bar{Z}_t^\epsilon &= \epsilon^2 b(Z_t^\epsilon)\,dt + \epsilon\, c\, dW_t + c\, dh_t & Z_0^\epsilon &= z_0\,. \end{aligned}$$

In (Deuschel et al., 2014a), the authors expand $F(\bar{Y}_T^\epsilon)$ to the first order. This leads to the presence of a $\mathcal{O}(1)$ in the definition of c_0, thus impeding its computation. In order to calculate it, let us expand Y_T^ϵ and $F(\bar{Y}_T^\epsilon)$ to the second order.

$$\bar{Y}_T^\epsilon = a + \epsilon \bar{Y}_T^1 + \frac{\epsilon^2}{2}\bar{Y}_T^2 + o(\epsilon^2) \tag{4.4.2}$$

and

$$F(\bar{Y}_T^\epsilon) = F(y_T) + \epsilon\, F'(y_T)\,\bar{Y}_T^1 + \frac{\epsilon^2}{2}\left(F''(y_T)\left(\bar{Y}_T^1\right)^2 + F'(y_T)\,\bar{Y}_T^2\right) + o(\epsilon^2)\,,$$

as $\epsilon \to 0$, where $\bar{Y}_T^j = \left(\partial_{\epsilon^j}^j \bar{Y}_t^\epsilon\right)\big|_{\epsilon=0}$ and $\bar{Z}_T^j = \left(\partial_{\epsilon^j}^j \bar{Z}_t^\epsilon\right)\big|_{\epsilon=0}$ have dynamics

$$\begin{aligned} d\bar{Y}_t^1 &= 2\, z_t\, \bar{Z}_t^1\, dt\,, & \bar{Y}_0^1 &= 0\,, \\ d\bar{Z}_t^1 &= c\, dW_t\,, & \bar{Z}_0^1 &= 0\,, \\ d\bar{Y}_t^2 &= 2\left(z_t \bar{Z}_t^2 + \left(\bar{Z}_t^1\right)^2\right) dt\,, & \bar{Y}_0^2 &= 0\,, \\ d\bar{Z}_t^2 &= 2\, b(z_t)\, dt\,, & \bar{Z}_0^2 &= 0\,. \end{aligned}$$

Since $h \in \mathcal{K}_a^{\min}$, (Ben Arous, 1988b, Lemma 1.43) proves that

$$F'(a)\,\bar{Y}_T^1 = -\int_0^T \dot{h}_s\,dW_s\,.$$

Then, since $F(y_T) = F(a) = 0$,

$$F(\bar{Y}^\epsilon_T) = -\epsilon \int_0^T \dot{h}_s \, dW_s + \frac{\epsilon^2}{2} \left(F''(a) \left(\bar{Y}^1_T\right)^2 + F'(a)\, \bar{Y}^2_T \right) + o(\epsilon^2) \,. \qquad (4.4.3)$$

Inserting (4.4.2) and (4.4.3) in (4.4.1) and setting $\lambda = \Lambda''(a)/2 > 0$, we obtain

$$f_{Y^\epsilon_T}(a) = \epsilon^{-1} \, e^{-\Lambda(a)/\epsilon^2} \left(c_0(a) + o(1) \right)$$

as $\epsilon \to 0$, where

$$c_0(a) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathbb{E} \left(e^{i\zeta \, \bar{Y}^1_T + \frac{\Lambda'(a)}{2} \, \bar{Y}^2_T} \right) d\zeta \,,$$

for

$$\bar{Y}^1_T = 2c \int_0^T z_t W_t \, dt \,, \qquad \bar{Y}^2_T = 2 \int_0^T 2 z_t \int_0^t b(z_s) \, ds \, + \, c^2 W_t^2 \, dt \,.$$

Combining (4.3.4) and (4.3.6), we obtain $\ddot{z}_t = -r^2 z_t$ with $\dot{z}_T = 0$. Then

$$\begin{aligned}
\int_0^T z_t \int_0^t b(z_s) \, ds \, dt = \int_0^T b(z_s) \int_s^T z_t \, dt \, ds = -\frac{1}{r^2} \int_0^T b(z_s) \int_s^T \ddot{z}_t \, dt \, ds \\
= \frac{1}{r^2} \int_0^T b(z_s) \, \dot{z}_s \, ds = \frac{1}{r^2} \int_{z_0}^{z_T} b(x) \, dx \,.
\end{aligned}$$

Hence

$$\bar{Y}^2_T = \frac{4}{r^2} \int_{z_0}^{z_T} b(x) \, dx + 2 \, c^2 \int_0^T W_t^2 \, dt$$

and substituting $\Lambda'(a)$ with (4.3.12),

$$c_0(a) = \frac{e^{\frac{1}{c^2} \int_{z_0}^{z_T} b(x) \, dx}}{2\pi} \int_{\mathbb{R}} \mathbb{E} \left(e^{2c \, i\zeta \int_0^T z_t W_t \, dt + \frac{r^2}{2} \int_0^T W_t^2 \, dt} \right) d\zeta \,.$$

We then have, using (Gombani and Runggaldier, 2001, Prop. 2.1)[4], that

$$\mathbb{E} \left(\exp \left\{ 2c \, i\zeta \int_0^T z_s W_s \, ds + \frac{r^2}{2} \int_0^T W_s^2 \, ds \right\} \right) = \exp\left(-A(0) \right) \,,$$

[4] The result in (Gombani and Runggaldier, 2001) concerns the price of zero-coupon bonds in exponentially quadratic term structure models. The exact same proof however holds by replacing the "HJM drift condition" by the martingale property of

$$\mathbb{E} \left[e^{-\int_0^T \left(2c \, i\zeta \, z_t W_t + \frac{r^2}{2} W_t^2 \right) dt} \middle| \mathcal{F}_s \right] \,.$$

where $A(t)$ satisfies the following Riccati equations

$$C'(t) = 2C^2(t) + \frac{r^2}{2}\,, \qquad C(T) = 0\,, \tag{4.4.4}$$

$$B'(t) = 2B(t)C(t) + 2c\,i\zeta\, z_t\,, \qquad B(T) = 0\,, \tag{4.4.5}$$

$$A'(t) = 1/2\,B^2(t) - C(t)\,, \qquad A(T) = 0\,. \tag{4.4.6}$$

Eq. (4.4.4) yields

$$C(t) = -\frac{r}{2}\tan\left(r\,(T-t)\right)\,.$$

Using variation of constants for eq. (4.4.5), we obtain

$$\begin{aligned} B(t) &= D(t)\,e^{-2\int_t^T C(s)\,ds} \\ &= -2c\,i\zeta\,e^{-2\int_t^T C(s)\,ds}\int_t^T z_s\,e^{2\int_s^T C(u)\,du}\,ds\,, \end{aligned}$$

where

$$\int_t^T C(s)\,ds = -\int_t^T \frac{r}{2}\tan\left(r\,(T-s)\right)\,ds = -\frac{1}{2}\log\left(\cos\left(r\,(T-t)\right)\right)\,.$$

Therefore, since $z_s = z_0\,\frac{\cos(r\,(T-s))}{\cos(r\,T)}$,

$$B(t) = -2c\,i\zeta\,\cos\left(r\,(T-t)\right)\int_t^T \frac{z_s}{\cos\left(r\,(T-s)\right)}\,ds = -2c\,i\zeta\,z_t\,(T-t)\,.$$

Finally, by (4.4.6)

$$\begin{aligned} A(0) &= -\frac{1}{2}\int_0^T B^2(t)\,dt + \int_0^T C(t)\,dt \\ &= \frac{2\,z_0^2\,c^2}{\cos^2(r\,T)}\,\zeta^2\int_0^T \cos^2\left(r\,(T-t)\right)\,(T-t)^2\,dt - \frac{1}{2}\log\left(\cos\left(r\,T\right)\right) \\ &= \frac{\zeta^2}{2}\,\frac{4\,z_0^2\,c^2\,T^3}{\cos^2(r\,T)}\,\mathcal{A}(2\,r\,T) - \frac{1}{2}\log\left(\cos\left(2\,r\,T\right)\right)\,, \end{aligned}$$

where

$$\mathcal{A}(u) = \frac{u^3 + 6\,u\,\cos(u) + 3\,(u^2-2)\sin(u)}{6\,u^3}\,.$$

Then

$$\mathbb{E}\left(\exp\left\{2c\,i\zeta\int_0^T z_s W_s\,ds + \frac{r^2}{2}\int_0^T W_s^2\,ds\right\}\right) = \sqrt{\cos\left(r\,T\right)}\,e^{-\frac{\zeta^2}{2}\,\frac{4\,z_0^2\,c^2\,T^3}{\cos^2(r\,T)}\,\mathcal{A}(2rT)}$$

and

$$\begin{aligned} c_0(a) &= \frac{e^{\frac{1}{c^2}\int_{z_0}^{z_T} b(x)\,dx}}{2\pi}\int_{\mathbb{R}}\mathbb{E}\left(e^{2c\,i\zeta\int_0^T z_t W_t\,dt + \frac{r^2}{2}\int_0^T W_t^2\,dt}\right)d\zeta \\ &= \frac{1}{\sqrt{2\pi\mathcal{A}(2rT)}}\,\frac{\cos^{3/2}(r\,T)}{2\,z_0\,c\,T^{3/2}}\,e^{\frac{1}{c^2}\int_{z_0}^{z_T} b(x)\,dx}\,, \end{aligned}$$

which concludes the calculation of c_0. □

4.5 Application to the pricing of options on realized variance

We now consider the pricing of options on realized variance. In this section, we provide an expansion for the price of a put option on realized variance and for its Black and Scholes implied volatility.

4.5.1 Asymptotics for the price of options on realized variance

In this section, we use Theorem 4.4.1 to calculate the asymptotic expansion of the price of a put option on realized variance. Let us first prove two asymptotic results for the expansion of integrals following the methodology of (Bleistein and Handelsman, 1975, Chapter 5, p. 180-219).

Lemma 4.5.1. *Let $a < b \in \mathbb{R}$. Let also $f \in C^2(\mathbb{R})$ and $\phi \in C^3(\mathbb{R})$, such that $f(t) = \mathcal{O}\big(e^{\lambda t}\big)$ for some λ as $t \to \infty$ and ϕ is increasing on (a,b) and $\phi'(a) \neq 0$. Then*

$$\int_a^b f(x)\, e^{-\frac{\phi(x)}{\epsilon^2}}\, dx = e^{-\frac{\phi(a)}{\epsilon^2}} \left(\frac{f(a)}{\phi'(a)}\epsilon^2 + \Big(\frac{f'(a)}{(\phi'(a))^2} - \frac{f(a)\phi''(a)}{(\phi'(a))^3}\Big)\epsilon^4 + \mathcal{O}\big(\epsilon^6\big) \right),$$

as $\epsilon \to 0$.

Proof. With a change of variable, we obtain

$$\int_a^b f(x)\, e^{-\epsilon^{-2}\phi(x)}\, dx = e^{-\epsilon^{-2}\phi(a)} \int_0^{\phi(b)-\phi(a)} G(\phi^{-1}(\phi(a)+\tau))\, e^{-\epsilon^{-2}\tau}\, d\tau\,,$$

where $G = \frac{f}{\phi'}$. Since $\phi'(a) \neq 0$,

$$\begin{aligned} G(t) &= \frac{\frac{f(a)}{\phi'(a)} + \frac{f'(a)}{\phi'(a)}(t-a) + \mathcal{O}((t-a)^2)}{1 + \frac{\phi''(a)}{\phi'(a)}(t-a) + \mathcal{O}((t-a)^2)} \\ &= \frac{f(a)}{\phi'(a)} + \left(\frac{f'(a)}{\phi'(a)} - \frac{f(a)\phi''(a)}{(\phi'(a))^2}\right)(t-a) + \mathcal{O}\big((t-a)^2\big)\,. \end{aligned}$$

Since

$$\tau := \phi(t) - \phi(a) = \phi'(a)\,(t-a) + \mathcal{O}\big((t-a)^2\big)\,,$$

then

$$G(\phi^{-1}(\phi(a)+\tau)) = \frac{f(a)}{\phi'(a)} + \left(\frac{f'(a)}{(\phi'(a))^2} - \frac{f(a)\phi''(a)}{(\phi'(a))^3}\right)\tau + \mathcal{O}\big(\tau^2\big)\,.$$

Substituting G by its development in

$$\int_0^{\phi(b)-\phi(a)} G(\phi^{-1}(\phi(a)+\tau))\, e^{-\epsilon^{-2}\tau}\, d\tau$$

and integrating using Wilson's Lemma (Bleistein and Handelsman, 1975, Chapter 4, p. 103) yields the result. □

Lemma 4.5.2 (Laplace's method)**.** *Let $a_1 < a < a_2 \in \mathbb{R}$. Let also $f \in C^3(\mathbb{R})$ and $\phi \in C^5(\mathbb{R})$, such that $f(t) = \mathcal{O}\big(e^{\lambda t}\big)$ for some $\lambda > 0$ as $t \to \infty$ and ϕ is decreasing on (a_1, a) and increasing on (a, a_2) with a global minimum at a such that $\phi''(a) > 0$. Then*

$$\int_{a_1}^{a_2} f(x)\, e^{-\frac{\phi(x)}{\epsilon^2}}\, dx = e^{-\frac{\phi(a)}{\epsilon^2}} \left(\epsilon \sqrt{\frac{2\pi}{\phi''(a)}} f(a) + \epsilon^3 \sqrt{\frac{\pi}{2}}\, \mathcal{G}(f,\phi,a) + \mathcal{O}\big(\epsilon^5\big) \right)$$

as $\epsilon \to 0$, where

$$\mathcal{G}(f,\phi,a) = \frac{f''(a)}{\sqrt{(\phi''(a))^3}} - \frac{f'(a)\phi'''(a)}{\sqrt{(\phi''(a))^5}} + \frac{5 f(a)(\phi'''(a))^2}{12\sqrt{(\phi''(a))^7}} - \frac{f(a)\phi^{(4)}(a)}{4\sqrt{(\phi''(a))^5}}\,.$$

Proof. First write

$$\begin{aligned}\int_{a_1}^{a_2} f(x)\, e^{-\epsilon^{-2}\phi(x)}\, dx &= \int_{a_1}^{a} f(x)\, e^{-\epsilon^{-2}\phi(x)}\, dx + \int_{a}^{a_2} f(x)\, e^{-\epsilon^{-2}\phi(x)}\, dx \\ &= I_1(\epsilon) + I_2(\epsilon)\,.\end{aligned}$$

With a change of variable, we obtain

$$I_2(\epsilon) = e^{-\epsilon^{-2}\phi(a)} \int_0^{\phi(a_2)-\phi(a)} G(\phi^{-1}(\phi(a)+\tau))\, e^{-\epsilon^{-2}\tau}\, d\tau$$

and

$$I_1(\epsilon) = -e^{-\epsilon^{-2}\phi(a)} \int_0^{\phi(a_1)-\phi(a)} G(\phi^{-1}(\phi(a)+\tau))\, e^{-\epsilon^{-2}\tau}\, d\tau$$

where $G = \frac{f}{\phi'}$. Since $\phi'(a) = 0$ and $\phi''(a) \neq 0$, then

$$\begin{aligned}G(t) &= \frac{f(a) + f'(a)\,(t-a) + \frac{1}{2} f''(a)\,(t-a)^2 + \mathcal{O}((t-a)^3)}{\phi''(a)(t-a) + \frac{1}{2}\phi'''(a)\,(t-a)^2 + \frac{1}{6}\phi^{(4)}(a)\,(t-a)^3 + \mathcal{O}((t-a)^4)} \\ &= \frac{f(a) + f'(a)\,(t-a) + \frac{1}{2} f''(a)\,(t-a)^2 + \mathcal{O}((t-a)^3)}{\phi''(a)(t-a)\left(1 + \frac{\phi'''(a)}{2\phi''(a)}(t-a) + \frac{\phi^{(4)}(a)}{6\phi''(a)}(t-a)^2 + \mathcal{O}((t-a)^3)\right)} \\ &= \frac{f(a)}{\phi''(a)}(t-a)^{-1} + \left(\frac{f'(a)}{\phi''(a)} - \frac{f(a)\phi'''(a)}{2\,(\phi''(a))^2}\right) + \mathcal{O}\big((t-a)^2\big) \\ &\quad + \left(\frac{f''(a)}{2\phi''(a)} - \frac{f'(a)\phi'''(a)}{2\,(\phi''(a))^2} + \frac{f(a)(\phi'''(a))^2}{4\,(\phi''(a))^3} - \frac{f(a)\phi^{(4)}(a)}{6\,(\phi''(a))^2}\right)(t-a)\end{aligned}$$

Then note that, since

$$\phi(t) = \phi(a) + \frac{\phi''(a)}{2}(t-a)^2 \left(1 + \frac{\phi'''(a)}{3\,\phi''(a)}(t-a) + \frac{\phi^{(4)}(a)}{12\,\phi''(a)}(t-a)^2 + \mathcal{O}\big((t-a)^3\big)\right),$$

denoting $\tau = \phi(t) - \phi(a)$, we have up to a $\mathcal{O}((t-a)^2)$ the following equalities,

$$\begin{aligned}
\tau^{1/2} &= \frac{\sqrt{\phi''(a)}}{\sqrt{2}}\,(t-a) + \mathcal{O}\big((t-a)^2\big)\\
\tau^{-1/2} &= \frac{\sqrt{2}}{\sqrt{\phi''(a)}}(t-a)^{-1} - \frac{\sqrt{2}\phi'''(a)}{6\sqrt{(\phi''(a))^3}}\\
&\qquad + \left(\frac{\sqrt{2}(\phi'''(a))^2}{24\sqrt{(\phi''(a))^5}} - \frac{\sqrt{2}\phi^{(4)}(a)}{24\sqrt{(\phi''(a))^3}}\right)(t-a) + \mathcal{O}\big((t-a)^2\big)
\end{aligned}$$

and hence

$$\begin{aligned}
(t-a) &= \frac{\sqrt{2}}{\sqrt{\phi''(a)}}\,\tau^{1/2}\\
(t-a)^{-1} &= \frac{\sqrt{\phi''(a)}}{\sqrt{2}}\tau^{-1/2} + \frac{\phi'''(a)}{6\,\phi''(a)} - \left(\frac{\sqrt{2}(\phi'''(a))^2}{24\sqrt{(\phi''(a))^5}} - \frac{\sqrt{2}\phi^{(4)}(a)}{24\sqrt{(\phi''(a))^3}}\right)\tau^{1/2}\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
G(\phi^{-1}(\phi(a)+\tau)) &= \frac{f(a)}{\sqrt{2\,\phi''(a)}}\,\tau^{-1/2} + \left(\frac{f'(a)}{\phi''(a)} - \frac{f(a)\phi'''(a)}{3\,(\phi''(a))^2}\right)\\
&\quad + \left(\frac{\sqrt{2}f''(a)}{2\sqrt{(\phi''(a))^3}} - \frac{\sqrt{2}f'(a)\phi'''(a)}{2\,\sqrt{(\phi''(a))^5}}\right.\\
&\qquad \left. + \frac{5\sqrt{2}f(a)(\phi'''(a))^2}{24\,\sqrt{(\phi''(a))^7}} - \frac{\sqrt{2}f(a)\phi^{(4)}(a)}{8\,\sqrt{(\phi''(a))^5}}\right)\tau^{1/2} + \mathcal{O}(\tau)\;.
\end{aligned}$$

By Watson's Lemma (Bleistein and Handelsman, 1975, Chapter 4, p. 103),

$$\begin{aligned}
I_2(\epsilon) &= e^{-\epsilon^{-2}\,\phi(a)}\int_0^{\phi(a_2)-\phi(a)} G(\phi^{-1}(\phi(a)+\tau))\,e^{-\epsilon^{-2}\,\tau}\,d\tau\\
&= e^{-\epsilon^{-2}\,\phi(a)}\left(\alpha_{-1/2}\sqrt{\pi}\,\epsilon + \alpha_0\,\epsilon^2 + \alpha_{1/2}\frac{\sqrt{\pi}}{2}\,\epsilon^3 + \mathcal{O}\big(\epsilon^4\big)\right)\;,
\end{aligned}$$

where α_k is the term of order k of $G(\phi^{-1}(\phi(a)+\tau))$. The same reasoning allows to obtain

$$I_1(\epsilon) = e^{-\epsilon^{-2}\,\phi(a)}\left(\alpha_{-1/2}\sqrt{\pi}\,\epsilon - \alpha_0\,\epsilon^2 + \alpha_{1/2}\frac{\sqrt{\pi}}{2}\,\epsilon^3 + \mathcal{O}\big(\epsilon^4\big)\right)\;.$$

Combining the two finishes the proof. □

We then proceed to the calculation of the expansion of the option price.

Theorem 4.5.3. *Let*

$$P^{\epsilon}(z_0, K, T) = \mathbb{E}\left((K - T^{-1}Y_T^{\epsilon})_+\right)$$

be the price of a put option on realized variance with maturity T and strike K. Then P^{ϵ} admits the asymptotic expansion

$$P^{\epsilon}(z_0, K, T) = (K - z_0^2)_+ + e^{-\frac{\Lambda(KT)}{\epsilon^2}}\left(\frac{\epsilon^3\, c_0(KT)}{T\,(\Lambda'(KT))^2} + o\left(\epsilon^3\right)\right) \tag{4.5.1}$$

as $\epsilon \to 0$.

Proof. First notice that, $\Lambda'(a) > 0 \Leftrightarrow a > z_0^2 T$ and $\Lambda'(a) < 0 \Leftrightarrow a < z_0^2 T$. If $K < z_0^2$, then using Lemma 4.5.1 and Theorem 4.4.1, we have

$$\begin{aligned} P^{\epsilon}(z_0, K, T) &= \frac{1}{T}\int_{-\infty}^{KT} (KT - a)\, f_{Y_T^{\epsilon}}(a)\, da \\ &= e^{-\frac{\Lambda(KT)}{\epsilon^2}}\left(\frac{\epsilon^3\, c_0(KT)}{T\,(\Lambda'(KT))^2} + o\left(\epsilon^3\right)\right)\,, \end{aligned} \tag{4.5.2}$$

whereas if $K > z_0^2$,

$$P^{\epsilon}(z_0, K, T) = K - T^{-1}\mathbb{E}(Y_T^{\epsilon}) + \mathbb{E}\left((K - T^{-1}Y_T^{\epsilon})_-\right)\,.$$

By Lemma 4.5.1 again, we have

$$\begin{aligned} \mathbb{E}\left((K - T^{-1}Y_T^{\epsilon})_-\right) &= \frac{1}{T}\int_{KT}^{\infty} (a - KT)\, f_{Y_T^{\epsilon}}(a)\, da \\ &= e^{-\frac{\Lambda(KT)}{\epsilon^2}}\left(\frac{\epsilon^3\, c_0(KT)}{T\,(\Lambda'(KT))^2} + o\left(\epsilon^3\right)\right) \end{aligned} \tag{4.5.3}$$

and, by Lemma 4.5.2,

$$\mathbb{E}(Y_T^{\epsilon}) = \int_{\mathbb{R}} a\, f_{Y_T^{\epsilon}}(a)\, da = z_0^2\, T\, \frac{\sqrt{2\pi}\, c_0(z_0^2\, T)}{|\Lambda''(z_0^2\, T)|^{1/2}} + \mathcal{O}(\epsilon)\,.$$

Since

$$c_0(z_0^2\, T) = \sqrt{\frac{3}{2\pi}}\, \frac{1}{2\, z_0\, c\, T^{3/2}} \qquad \text{and} \qquad \Lambda''(z_0^2\, T) = \frac{3}{4 z_0^2\, c^2\, T^3}\,,$$

then

$$P^{\epsilon}(z_0, K, T) = K - z_0^2 + e^{-\frac{\Lambda(KT)}{\epsilon^2}}\left(\frac{\epsilon^3\, c_0(KT)}{T\,(\Lambda'(KT))^2} + o\left(\epsilon^3\right)\right)\,. \tag{4.5.4}$$

Note that, due to the boundedness of the function g defined in (4.3.2), the process Y_T^{ϵ} is bounded uniformly in ϵ. The fact that the expansion of Theorem 4.4.1 is only uniform on compact set is not problematic to integrate the density in eqs. (4.5.2) and (4.5.3). Combining (4.5.2) and (4.5.4) proves (4.5.1). □

4.5.2 Asymptotics for the Black and Scholes implied volatility of options on realized variance

In order to calculate the asymptotic implied volatility, we start by recalling the asymptotic expansion of the price of a put option in the BS model. See (Gatheral et al., 2012) for a complete overview of the methods used in Lemma 4.5.4 and Proposition 4.5.5.

Lemma 4.5.4. *Denote Φ the Gaussian distribution function. For $x > 0$,*

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \left(\frac{1}{x} - \frac{1}{x^3} + \mathcal{O}\big(x^{-5}\big) \right) \qquad \text{as } x \to \infty\,.$$

Proof. By symmetry,

$$\sqrt{2\pi}\,\Phi(-x) = \int_x^\infty e^{-y^2/2}\,dy = \int_x^\infty \frac{-1}{y}\,(e^{-y^2/2})'\,dy\,.$$

Integrating successively by parts, we then obtain

$$\begin{aligned}
\sqrt{2\pi}\,\Phi(-x) &= \frac{1}{x}\,e^{-x^2/2} - \int_x^\infty \frac{1}{y^2}\,e^{-y^2/2}\,dy \\
&= \frac{1}{x}\,e^{-x^2/2} - \int_x^\infty \frac{-1}{y^3}\,(e^{-y^2/2})'\,dy \\
&= \frac{1}{x}\,e^{-x^2/2} - \frac{1}{x^3}\,e^{-x^2/2}\,dy + \int_x^\infty \frac{3}{y^4}\,e^{-y^2/2}\,dy\,,
\end{aligned}$$

where

$$\left| \int_x^\infty \frac{3}{y^4}\,e^{-y^2/2}\,dy \right| = \left| \int_x^\infty \frac{-3}{y^5}\,(e^{-y^2/2})'\,dy \right| \le \frac{3}{x^5}\,e^{-x^2/2}\,,$$

hence the result. □

Proposition 4.5.5. *Let $(S_t^\epsilon)_{0\le t\le T}$ with dynamics*

$$dS_t^\epsilon = \epsilon\,\sigma_{BS}\,S_t^\epsilon\,dW_t\,, \qquad S_0^\epsilon = S_0$$

and let

$$P_{BS}^\epsilon = \mathbb{E}\left((K - S_T^\epsilon)_+\right)$$

be the BS price at time 0 of a put option with strike K and maturity T. Then P_{BS}^ϵ admits the expansion

$$P_{BS}^\epsilon(S_0, K, T; \sigma_{BS}) = (K - S_0)_+ + e^{-\frac{\log(S_0/K)^2}{2\epsilon^2\sigma_{BS}^2 T}} \left(\frac{\sqrt{S_0 K}}{\sqrt{2\pi}}\, \frac{\epsilon^3\,\sigma_{BS}^3\,T^{3/2}}{|\log(S_0/K)|^2} + \mathcal{O}\big(\epsilon^5\big) \right), \tag{4.5.5}$$

as $\epsilon \to 0$.

Proof. The Black-Scholes price is given by

$$P^{\epsilon}_{BS}(S_0, K, T; \sigma_{BS}) = K\,\Phi\left(-d_-\right) - S_0\,\Phi\left(-d_+\right)\,,$$

where

$$d_{\pm} = \frac{\log(S_0/K)}{\epsilon\,\sigma_{BS}\,T^{1/2}} \pm \frac{1}{2}\,\epsilon\,\sigma_{BS}\,T^{1/2}\,.$$

If $S_0 > K$, $d_{\pm} \underset{\epsilon\to 0}{\to} \infty$, whereas if $S_0 < K$, $d_{\pm} \underset{\epsilon\to 0}{\to} -\infty$. Hence

$$P^{\epsilon}_{BS}(S_0, K, T; \sigma_{BS}) = (K - S_0)_+ - \mathrm{sgn}(K - S_0)\,\left[K\,\Phi\left(-|d_-|\right) - S_0\,\Phi\left(-|d_+|\right)\right]\,.$$

Lemma 4.5.4 indicates that

$$\begin{aligned} S_0\Phi\left(-|d_+|\right) &= \frac{S_0}{\sqrt{2\pi}}\,e^{-d_+^2/2}\left(\frac{1}{|d_+|} - \frac{1}{|d_+|^3} + \mathcal{O}\big(|d_+|^{-5}\big)\right) \\ K\Phi\left(-|d_-|\right) &= \frac{K}{\sqrt{2\pi}}\,e^{-d_-^2/2}\left(\frac{1}{|d_-|} - \frac{1}{|d_-|^3} + \mathcal{O}\big(|d_-|^{-5}\big)\right)\,, \end{aligned}$$

where

$$e^{-d_{\pm}^2/2} = e^{-\frac{\log(S_0/K)^2}{2\epsilon^2\sigma_{BS}^2 T} \mp \frac{\log(S_0/K)}{2} - \frac{1}{8}\epsilon^2\sigma_{BS}^2 T} = (S_0/K)^{\mp 1/2}\,e^{-\frac{\log(S_0/K)^2}{2\epsilon^2\sigma_{BS}^2 T}}\left(1 + \mathcal{O}\big(\epsilon^2\big)\right)\,,$$

$$\frac{1}{|d_{\pm}|} = \frac{\epsilon\,\sigma_{BS}\,T^{1/2}}{|\log(S_0/K)|} \mp \mathrm{sgn}(S_0 - K)\frac{\epsilon^3\,\sigma_{BS}^3\,T^{3/2}}{2\,|\log(S_0/K)|^2} + \mathcal{O}\big(\epsilon^5\big)$$

and

$$\frac{1}{|d_{\pm}|^3} = \frac{\epsilon^3\,\sigma_{BS}^3\,T^{3/2}}{|\log(S_0/K)|^3} + \mathcal{O}\big(\epsilon^5\big)\,.$$

Therefore,

$$\begin{aligned} &-\mathrm{sgn}(K - S_0)\left[K\Phi\left(-|d_-|\right) - S_0\Phi\left(-|d_+|\right)\right] \\ &\qquad = e^{-\frac{\log(S_0/K)^2}{2\epsilon^2\sigma_{BS}^2 T}}\left(\frac{\sqrt{S_0 K}}{\sqrt{2\pi}}\,\frac{\epsilon^3\,\sigma_{BS}^3 T^{3/2}}{|\log(S_0/K)|^2} + \mathcal{O}\big(\epsilon^5\big)\right)\,, \end{aligned}$$

which proves the result. □

With the expansion of the price in the BS model, we now proceed to the calculation of the BS implied volatility of the put option on realized variance. We define the Black-Scholes implied volatility of a realized variance option as the volatility value σ_{BS} such that

$$P_{BS}(z_0^2, K, T, \sigma_{BS}) = P(z_0, K, T).$$

Theorem 4.5.6. *The BS implied volatility of the put option on the realized variance admits the expansion* $\sigma_{BS} = \sigma_{BS,0} + \epsilon^2\sigma_{BS,1} + o\left(\epsilon^2\right)$ *where*

$$\sigma_{BS,0} = \frac{|\log(z_0^2/K)|}{\sqrt{2T}\,\Lambda^{1/2}(KT)} \tag{4.5.6}$$

and

$$\sigma_{BS,1} = \frac{|\log(z_0^2/K)|}{2^{3/2}\,T^{1/2}\,\Lambda^{3/2}(KT)}\,\log\left(\frac{4\sqrt{\pi}}{z_0\,K^{1/2}\,T^2}\frac{c_0(KT)}{(\Lambda'(KT))^2}\frac{\Lambda^{3/2}(KT)}{|\log(z_0^2/K)|}\right), \tag{4.5.7}$$

as $\epsilon \to 0$.

Proof. By replacing σ_{BS} by its expansion in (4.5.5), we obtain

$$P_{BS}^{\epsilon}(S_0,K,T;\sigma_{BS}) = (K-S_0)_+ + e^{-\frac{\log(S_0/K)^2}{2\epsilon^2\sigma_{BS,0}^2 T}}\left(\frac{\sqrt{S_0K}}{\sqrt{2\pi}}\frac{\epsilon^3\sigma_{BS,0}^3T^{3/2}}{\log(S_0/K)^2}\,e^{\frac{\log(S_0/K)^2}{\sigma_{BS,0}^3T}\sigma_{BS,1}} + \mathcal{O}\big(\epsilon^5\big)\right).$$

Equating $P^{\epsilon}(z_0,K,T)$ to $P_{BS}^{\epsilon}(z_0^2,K,T;\sigma_{BS,0}+\epsilon^2\sigma_{BS,1})$ and identifying the terms then yields (4.5.6) and (4.5.7). □

Chapter 5

Approximate option pricing in the Lévy Libor model

5.1 Introduction

The goal of this paper is to develop explicit approximations for option prices in the Lévy Libor model introduced by Eberlein and Özkan, 2005. In particular, we shall be interested in price approximations for caplets, whose pay-off is a function of only one underlying Libor rate and swaptions, which can be regarded as options on a "basket" of multiple Libor rates of different maturities.

A full-fledged model of Libor rates such as the Lévy Libor model is typically used for the purposes of pricing and risk management of exotic interest rate products. The prices and hedge ratios must be consistent with the market-quoted prices of liquid options, which means that the model must be calibrated to the available prices / implied volatilities of caplets and swaptions. To perform such a calibration efficiently, one therefore needs explicit formulas or fast numerical algorithms for caplet and swaption prices.

Computation of option prices in the Lévy Libor model to arbitrary precision is only possible via Monte Carlo. Efficient simulation algorithms suitable for pricing exotic options have been proposed in (Kohatsu-Higa and Tankov, 2010; Papapantoleon et al., 2011), however, these Monte Carlo algorithms are probably not an option for the purposes of calibration because the computation is still too slow due to the presence of both discretization and statistical error.

Eberlein and Özkan, 2005, Kluge, 2005 and (Belomestny and Schoenmakers, 2011) propose fast methods for computing caplet prices which are based on Fourier transform inversion and use the fact that the characteristic function of many parametric Lévy processes is known explicitly. Since in the Lévy Libor model, the Libor rate L^k is not a geometric Lévy process under the corresponding probability measure $\mathbb{Q}^{T_k}$, unless $k = n$ (see Remark 5.3.1

below for details), using these methods for $k < n$ requires an additional approximation (some random terms appearing in the compensator of the jump measure of L^k are approximated by their values at time $t = 0$, a method known as freezing).

In this paper we take an alternative route and develop approximate formulas for caplets and swaptions using asymptotic expansion techniques. Inspired by methods used in Černỳ et al., 2013 and Ménassé and Tankov, 2015 (see also (Benhamou et al., 2009; Benhamou et al., 2010) for related expansions "around a Black-Scholes proxy" in other models), we consider a given Lévy Libor model as a perturbation of the log-normal LMM. Starting from the driving Lévy process $(X_t)_{t\geq 0}$ of the Lévy Libor model, assumed to have zero expectation, we introduce a family of processes $X^\alpha_t = \alpha X_{t/\alpha^2}$ parameterized by $\alpha \in (0, 1]$, together with the corresponding family of Lévy Libor models. For $\alpha = 1$ one recovers the original Lévy Libor model. When $\alpha \to 0$, the family X^α converges weakly in Skorokhod topology to a Brownian motion, and the option prices in the Lévy Libor model corresponding to the process X^α converge to the prices in the log-normal LMM. The option prices in the original Lévy Libor model can then be approximated by their second-order expansions in the parameter α, around the value $\alpha = 0$. This leads to an asymptotic approximation formula for a derivative price expressed as a linear combination of the derivative price stemming from the LMM and correction terms depending on the characteristics of the driving Lévy process. The terms of this expansion are often much easier to compute than the option prices in the Lévy Libor model. In particular, we shall see the expansion for caplets is expressed in terms of the derivatives of the standard Black's formula, and the various terms of the expansion for swaptions can be approximated using one of the many swaption approximations for the log-normal LMM available in the literature.

This paper is structured as follows. In Section 5.2 we briefly review the Lévy Libor model. In Section 5.3 we show how the prices of European-style options may be expressed as solutions of partial integro-differential equations (PIDE). These PIDEs form the basis of our asymptotic method, presented in detail in Section 5.4. Finally, numerical illustrations are provided in Section 5.5.

5.2 Presentation of the model

In this section we present a slight modification of the Lévy Libor model by Eberlein and Özkan, 2005, which is a generalization, based on Lévy processes, of the Libor market model driven by a Brownian motion, introduced by Miltersen et al., 1994, Brace et al., 1997 and Miltersen et al., 1997.

Let a discrete tenor structure $0 \leq T_0 < T_1 < \ldots < T_n$ be given, and set $\delta_k := T_k - T_{k-1}$, for $k = 1, \ldots, n$. We assume that zero-coupon bonds with maturities T_k, $k = 0, \dots, n$, are traded in the market. The time-t price of a bond with maturity T_k is denoted by $B_t(T_k)$ with $B_{T_k}(T_k) = 1$.

For every tenor date T_k, $k = 1, \dots, n$, the forward Libor rate L_t^k at time $t \leq T_{k-1}$ for the accrual period $[T_{k-1}, T_k]$ is a discretely compounded interest rate defined as

$$L_t^k := \frac{1}{\delta_k} \left(\frac{B_t(T_{k-1})}{B_t(T_k)} - 1 \right). \tag{5.2.1}$$

For all $t > T_{k-1}$, we set $L_t^k := L_{T_{k-1}}^k$.

To set up the Libor model, one needs to specify the forward Libor rates L_t^k, $k = 1, \dots, n$, such that each Libor rate L^k is a martingale with respect to the corresponding forward measure $\mathbb{Q}^{T_k}$ using the bond with maturity T_k as numéraire. We recall that the forward measures are interconnected via the Libor rates themselves and hence each Libor rate depends also on some other Libor rates as we shall see below. More precisely, assuming that the forward measure $\mathbb{Q}^{T_n}$ for the most distant maturity T_n (i.e. with numéraire $B(T_n)$) is given, the link between the forward measure $\mathbb{Q}^{T_k}$ and $\mathbb{Q}^{T_n}$ is provided by

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^{T_n}}\Big|_{\mathcal{F}_t} = \frac{B_t(T_k)}{B_t(T_n)} \frac{B_0(T_n)}{B_0(T_k)} = \prod_{j=k+1}^{n} \frac{1 + \delta_j L_t^j}{1 + \delta_j L_0^j}, \tag{5.2.2}$$

for every $k = 1, \dots, n-1$. The forward measure $\mathbb{Q}^{T_n}$ is referred to as the terminal forward measure.

5.2.1 The driving process

Let us denote by $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{0 \leq t \leq T^*}, \mathbb{Q}^{T_n})$ a complete stochastic basis and let X be an $\mathbb{R}^d$-valued Lévy process $(X_t)_{0 \leq t \leq T^*}$ on this stochastic basis with Lévy measure F and diffusion matrix c. The filtration $\mathbf{F}$ is generated by X and $\mathbb{Q}^{T_n}$ is the forward measure associated with the date T_n, i.e. with the numéraire $B_t(T_n)$. The process X is assumed without loss of generality to be driftless under $\mathbb{Q}^{T_n}$.

Moreover, we assume that $\int_{|z|>1} |z| F(dz) < \infty$. This implies in addition that X is a special semimartingale and allows to choose the truncation function $h(z) = z$, for $z \in \mathbb{R}^d$. The canonical representation of X is given by

$$X_t = \sqrt{c} W_t^{T_n} + \int_0^t \int_{\mathbb{R}^d} z(\mu - \nu^{T_n})(ds, dz), \tag{5.2.3}$$

where $W^{T_n} = (W_t^{T_n})_{0 \leq t \leq T_n}$ denotes a standard d-dimensional Brownian motion with respect to the measure $\mathbb{Q}^{T_n}$, μ is the random measure of jumps of X and $\nu^{T_n}(ds, dz) = F(dz)ds$ is the $\mathbb{Q}^{T_n}$-compensator of μ.

5.2.2 The model

Denote by $L = (L^1, \dots, L^n)^\top$ the column vector of forward Libor rates. We assume that under the terminal measure $\mathbb{Q}^{T_n}$, the dynamics of L is given by the following SDE

$$dL_t = L_{t-}(b(t, L_t)dt + \Lambda(t)dX_t), \tag{5.2.4}$$

where $b(t, L_t)$ is the drift term and $\Lambda(t)$ a deterministic $n \times d$ volatility matrix. We write $\Lambda(t) = (\lambda^1(t), \dots, \lambda^n(t))^\top$, where $\lambda^k(t)$ denotes the d-dimensional volatility vector of the Libor rate L^k and assume that $\lambda^k(t) = 0$, for $t > T_{k-1}$.

One typically assumes that the jumps of X are bounded from below, i.e. $\Delta X_t > C$, for all $t \in [0, T^*]$ and for some strictly negative constant C, which is chosen such that it ensures the positivity of the Libor rates given by (5.2.4).

The drift $b(t, L_t) = (b^1(t, L_t), \dots, b^n(t, L_t))$ is determined by the no-arbitrage requirement that L^k has to be a martingale with respect to $\mathbb{Q}^{T_k}$, for every $k = 1, \dots, n$. This yields

$$\begin{aligned} b^k(t, L_t) = &- \sum_{j=k+1}^{n} \frac{\delta_j L_t^j}{1 + \delta_j L_t^j} \langle \lambda^k(t), c\, \lambda^j(t) \rangle \\ &+ \int_{\mathbb{R}^d} \langle \lambda^k(t), z \rangle \left(1 - \prod_{j=k+1}^{n} \left(1 + \frac{\delta_j L_t^j \langle \lambda^j(t), z \rangle}{1 + \delta_j L_t^j} \right) \right) F(dz). \end{aligned} \tag{5.2.5}$$

The above drift condition follows from (5.2.2) and Girsanov's theorem for semimartingales noticing that

$$\begin{aligned} dL_t^k &= L_{t-}^k (b^k(t, L_t)dt + \lambda^k(t)dX_t) \\ &= L_{t-}^k \lambda^k(t) dX_t^{T_k}, \end{aligned}$$

where

$$X_t^{T_k} = \sqrt{c} W_t^{T_k} + \int_0^t \int_{\mathbb{R}^d} z(\mu - \nu^{T_k})(ds, dz) \tag{5.2.6}$$

is a special semimartingale with a d-dimensional $\mathbb{Q}^{T_k}$-Brownian motion W^{T_k} given by

$$dW_t^{T_k} := dW_t^{T_n} - \sqrt{c} \left(\sum_{j=k+1}^{n} \frac{\delta_j L_t^j}{1 + \delta_j L_t^j} \lambda^j(t) \right) dt \tag{5.2.7}$$

and the $\mathbb{Q}^{T_k}$-compensator ν^{T_k} of μ given by

$$\begin{aligned} \nu^{T_k}(dt, dz) \;&:= \prod_{j=k+1}^{n} \left(1 + \frac{\delta_j L_{t-}^j}{1 + \delta_j L_{t-}^j} \langle \lambda^j(t), z \rangle \right) \nu^{T_n}(dt, dz) \\ &= \prod_{j=k+1}^{n} \left(1 + \frac{\delta_j L_t^j}{1 + \delta_j L_t^j} \langle \lambda^j(t), z \rangle \right) F(dz)dt \end{aligned} \tag{5.2.8}$$

$$= \quad F_t^{T_k}(dz)dt$$

with

$$F_t^{T_k}(dz) := \prod_{j=k+1}^{n}\left(1+\frac{\delta_j L_t^j}{1+\delta_j L_t^j}\langle\lambda^j(t),z\rangle\right)F(dz). \tag{5.2.9}$$

Equalities (5.2.7) and (5.2.8) and consequently also the drift condition (5.2.5), are implied by Girsanov's theorem for semimartingales applied first to the measure change from $\mathbb{Q}^{T_n}$ to $\mathbb{Q}^{T_{n-1}}$ and then proceeding backwards. We refer to Kallsen, 2006 for a version of Girsanov's theorem that can be directly applied in this case. Note that the random terms $\frac{\delta_j L_t^j}{1+\delta_j L_t^j}$ appear in the measure change due to the fact that for each $j = n, n-1, \ldots, 1$ we have

$$d(1+\delta_j L_t^j) = (1+\delta_j L_{t-}^j)\left(\frac{\delta_j L_{t-}^j}{1+\delta_j L_{t-}^j}b^j(t,L_t)dt + \frac{\delta_j L_{t-}^j}{1+\delta_j L_{t-}^j}\lambda^j(t)dX_t\right), \tag{5.2.10}$$

We point out that the predictable random terms $\frac{\delta_j L_{t-}^j}{1+\delta_j L_{t-}^j}$ can be replaced with $\frac{\delta_j L_t^j}{1+\delta_j L_t^j}$ in equalities (5.2.5), (5.2.7) and (5.2.8) due to absolute continuity of the characteristics of X.

Therefore, the vector process of Libor rates L, given in (5.2.4) with the drift (5.2.5), is a time-inhomogeneous Markov process and its infinitesimal generator under $\mathbb{Q}^{T_n}$ is given by

$$\begin{aligned}\mathcal{A}_t f(x) &= \sum_{i=1}^{n} x_i b^i(t,x)\frac{\partial f(x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} x_i x_j (\Lambda(t)c\Lambda(t)^\top)_{ij}\frac{\partial f(x)}{\partial x_i \partial x_j} \qquad (5.2.11)\\ &+ \int_{\mathbb{R}^d}\left(f(\operatorname{diag}(x)(\mathbf{1}+\Lambda(t)z)) - f(x) - \sum_{j=1}^{n} x_j(\Lambda(t)z)_j\frac{\partial f(x)}{\partial x_j}\right)F(dz),\end{aligned}$$

for a function $f \in C_0^2(\mathbb{R}^n,\mathbb{R})$ and with the function $b^i(t,x)$, for $i = 1,\ldots,n$ and $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$, given by

$$\begin{aligned}b^i(t,x) &= -\sum_{j=i+1}^{n}\frac{\delta_j x_j}{1+\delta_j x_j}\langle\lambda^i(t), c\,\lambda^j(t)\rangle\\ &+ \int_{\mathbb{R}^d}\langle\lambda^i(t),z\rangle\left(1-\prod_{j=k+1}^{n}\left(1+\frac{\delta_j x_j\langle\lambda^j(t),z\rangle}{1+\delta_j x_j}\right)\right)F(dz).\end{aligned}$$

Remark 5.2.1 (Connection to the Lévy Libor model of Eberlein and Özkan, 2005). *The dynamics of the forward Libor rate L^k, for all $k = 1,\ldots,n$, in the Lévy Libor model of Eberlein and Özkan, 2005 (compare also Eberlein and Kluge, 2007) is given as an ordinary exponential of the following form*

$$L_t^k = L_0^k \exp\left(\int_0^t \tilde{b}^k(s,L_s)ds + \int_0^t \tilde{\lambda}^k(s)d\tilde{Y}_s\right), \tag{5.2.12}$$

for some deterministic volatility vector $\tilde{\lambda}^k$ *and the drift* $\tilde{b}^k(t, L_t)$ *which has to be chosen such that the Libor rate* L^k *is a martingale under the forward measure* $\mathbb{Q}^{T_k}$*. Here* $\tilde{Y}$ *is a d-dimensional Lévy process given by*

$$\tilde{Y}_t = \sqrt{c}W_t^{T_n} + \int_0^t \int_{\mathbb{R}^d} z(\tilde{\mu} - \tilde{\nu}^{T_n})(ds, dz),$$

with the $\mathbb{Q}^{T_n}$*-characteristics* $(0, c, \tilde{F})$*, where* $\tilde{\nu}^{T_n}(ds, dz) = \tilde{F}(dz)ds$*. The Lévy measure* $\tilde{F}$ *has to satisfy the usual integrability conditions ensuring the finiteness of the exponential moments. The dynamics of* L^k *is thus given by the following SDE*

$$\begin{aligned} dL_t^k &= L_{t-}^k \left(b^k(t, L_t)dt + \sqrt{c}\tilde{\lambda}^k(t)dW_t^{T_n} + (e^{\langle \tilde{\lambda}^k(t), z\rangle} - 1)(\tilde{\mu} - \tilde{\nu}^{T_n})(dt, dz) \right) \\ &= L_{t-}^k \left(b^k(t, L_t)dt + dY_t^k \right), \end{aligned}$$

for all k, where Y^k *is a time-inhomogeneous Lévy process given by*

$$Y_t^k = \int_0^t \sqrt{c}\tilde{\lambda}^k(s)dW_s^{T_n} + \int_0^t \int_{\mathbb{R}^d} (e^{\langle \tilde{\lambda}^k(s), z\rangle} - 1)(\tilde{\mu} - \tilde{\nu}^{T_n})(ds, dz)$$

and the drift $b^k(t, L_t)$ *is given by*

$$\begin{aligned} b^k(t, L_t) &= \tilde{b}^k(t, L_t) + \frac{1}{2}\langle \tilde{\lambda}^k(t), c\tilde{\lambda}^k(t)\rangle \\ &\quad + \int_{\mathbb{R}^d} (e^{\langle \tilde{\lambda}^k(t), z\rangle} - 1 - \langle \tilde{\lambda}^k(t), z\rangle)\tilde{F}(dz). \end{aligned}$$

5.3 Option pricing via PIDEs

Below we present the pricing PIDEs related to general option payoffs and then more specifically to caplets and swaptions. We price all options under the given terminal measure $\mathbb{Q}^{T_n}$.

5.3.1 General payoff

Consider a European-type payoff with maturity T_k given by $\xi = g(L_{T_k})$, for some tenor date T_k. Its time-t price P_t is given by the following risk-neutral pricing formula

$$\begin{aligned} P_t &= B_t(T_k)\mathbb{E}^{\mathbb{Q}^{T_k}}[g(L_{T_k}) \mid \mathcal{F}_t] \\ &= B_t(T_n)\mathbb{E}^{\mathbb{Q}^{T_n}}\left[\frac{B_{T_k}(T_k)}{B_{T_k}(T_n)} g(L_{T_k}) \mid \mathcal{F}_t\right] \\ &= B_t(T_n)\mathbb{E}^{\mathbb{Q}^{T_n}}\left[\prod_{j=k+1}^{n} (1 + \delta_j L_{T_k}^j) g(L_{T_k}) \mid \mathcal{F}_t\right] \end{aligned}$$

$$= B_t(T_n)u(t, L_t),$$

where u is the solution of the following PIDE[1]

$$\begin{aligned} \partial_t u + \mathcal{A}_t u &= 0 \\ u(T_k, x) &= \tilde{g}(x) \end{aligned} \tag{5.3.1}$$

and $\tilde{g}$ denotes the transformed payoff function given by

$$\tilde{g}(x) := \tilde{g}(x_1, \ldots, x_n) = \prod_{j=k+1}^{n} (1 + \delta_j x_j) g(x_1, \ldots, x_n).$$

In what follows we shall in particular focus on two most liquid interest rate options: caps (caplets) and swaptions.

5.3.2 Caplet

Consider a caplet with strike K and payoff $\xi = \delta_k (L^k_{T_{k-1}} - K)^+$ at time T_k. Note that here the payoff is in fact a $\mathcal{F}_{T_{k-1}}$-measurable random variable and it is paid at time T_k. This is known as *payment in arrears*. There exist also other conventions for caplet payoffs, but this one is the one typically used.

The time-t price of the caplet, denoted by P^{Cpl}_t is thus given by

$$\begin{aligned} P^{Cpl}_t &= B_t(T_k)\delta_k \mathbb{E}^{\mathbb{Q}^{T_k}}[(L^k_{T_{k-1}} - K)^+ \mid \mathcal{F}_t] \\ &= B_t(T_n)\delta_k \mathbb{E}^{\mathbb{Q}^{T_n}}\left[\prod_{j=k+1}^{n} (1 + \delta_j L^j_{T_{k-1}})(L^k_{T_{k-1}} - K)^+ \mid \mathcal{F}_t\right] \\ &= B_t(T_n)\delta_k u(t, L_t) \end{aligned} \tag{5.3.2}$$

where u is the solution to

$$\begin{aligned} \partial_t u + \mathcal{A}_t u &= 0 \\ u(T_{k-1}, x) &= \tilde{g}(x) \end{aligned} \tag{5.3.3}$$

with

$$\tilde{g}(x) := (x_k - K)^+ \prod_{j=k+1}^{n} (1 + \delta_j x_j).$$

For the second equality in (5.3.2) we have used the measure change from $\mathbb{Q}^{T_k}$ to $\mathbb{Q}^{T_n}$ given in (5.2.2).

[1] A detailed proof of this statement is out of scope of this note. Here we simply assume that Equation (5.3.1) admits a unique solution which is sufficiently regular and is of polynomial growth. The existence of such a solution may be established first by Fourier methods for the case when there is no drift and then by a fixed-point theorem in Sobolev spaces using the regularizing properties of the Lévy kernel for the general case (see (De Franco, 2012, Chapter 7) for similar arguments). Once the existence of a regular solution has been established, the expression for the option price follows by the standard Feynman-Kac formula.

Remark 5.3.1. *Noting that the payoff of the caplet depends on one single underlying forward Libor rate L^k, it is often more convenient to price it directly under the corresponding forward measure $\mathbb{Q}^{T_k}$, using the first equality in (5.3.2). Thus, one has*

$$P_t^{Cpl} = B_t(T_k)\delta_k u(t, L_t),$$

where u is the solution to

$$\begin{aligned} \partial_t u + \mathcal{A}_t^{T_k} u &= 0 \\ u(T_{k-1}, x) &= \tilde{g}(x) \end{aligned} \tag{5.3.4}$$

with $\tilde{g}(x) := (x_k - K)^+$ and where $\mathcal{A}^{T_k}$ is the generator of L under the forward measure $\mathbb{Q}^{T_k}$. In the log-normal LMM this leads directly to the Black's formula for caplet prices. However, in the Lévy Libor model the driving process X under the forward measure $\mathbb{Q}^{T_k}$ is not a Lévy process anymore since its compensator of the random measure of jumps becomes stochastic (see (5.2.9)). Therefore, passing to the forward measure in this case does not lead to a closed-form pricing formula and does not bring any particular advantage. This is why in the forthcoming section we shall work directly under the terminal measure $\mathbb{Q}^{T_n}$.

5.3.3 Swaptions

Let us consider a swaption, written on a fixed-for-floating (payer) interest rate swap with inception date T_0, payment dates $T_1, \ldots, T_n$ and nominal $N = 1$. We denote by K the swaption strike rate and assume for simplicity that the maturity T of the swaption coincides with the inception date of the underlying swap, i.e. we assume $T = T_0$. Therefore, the payoff of the swaption at maturity is given by $\big(P^{Sw}(T_0; T_0, T_n, K)\big)^+$, where $P^{Sw}(T_0; T_0, T_n, K)$ denotes the value of the swap with fixed rate K at time T_0 given by

$$\begin{aligned} P^{Sw}(T_0; T_0, T_n, K) &= \sum_{j=1}^{n} \delta_j B_{T_0}(T_j) \mathbb{E}^{\mathbb{Q}^{T_j}} \left[L^j_{T_{j-1}} - K | \mathcal{F}_{T_0} \right] \\ &= \sum_{j=1}^{n} \delta_j B_{T_0}(T_j) \left(L^j_{T_0} - K \right) \\ &= (\sum_{j=1}^{n} \delta_j B_{T_0}(T_j)) \left(R(T_0; T_0, T_n) - K \right) \end{aligned}$$

where

$$R(t; T_0, T_n) = \frac{\sum_{j=1}^{n} \delta_j B_t(T_j) L^j_t}{\sum_{j=1}^{n} \delta_j B_t(T_j)} =: \sum_{j=1}^{n} w_j L^j_t \tag{5.3.5}$$

is the swap rate i.e. the fixed rate such that the time-t price of the swap is equal to zero. Here we denote

$$w_j(t) := \frac{\delta_j B_t(T_j)}{\sum_{k=1}^{n} \delta_k B_t(T_k)} \tag{5.3.6}$$

Note that $\sum_{j=1}^{n} w_j(t) = 1$. Dividing the numerator and the denominator in (5.3.5) by $B_t(T_n)$ and using the telescopic products together with (5.2.1) we see that $w_j(t) = f_j(L_t)$ for a function f_j given by

$$f_j(x) = \frac{\delta_j \prod_{i=j+1}^{n}(1+\delta_i x_i)}{\sum_{k=1}^{n} \delta_k \prod_{i=k+1}^{n}(1+\delta_i x_i)} \tag{5.3.7}$$

for $j = 1, \ldots, n$.

Therefore, the swaption price at time $t \le T_0$ is given by

$$\begin{aligned}
&P^{Swn}(t; T_0, T_n, K) \\
&\quad = B_t(T_0)\mathbb{E}^{\mathbb{Q}^{T_0}}\left[\left(P^{Sw}(T_0; T_0, T_n, K)\right)^+ |\mathcal{F}_t\right] && (5.3.8) \\
&\quad = B_t(T_0)\mathbb{E}^{\mathbb{Q}^{T_0}}\left[\left(\sum_{j=1}^{n} \delta_j B_{T_0}(T_j)\right)(R(T_0; T_0, T_n) - K)^+ |\mathcal{F}_t\right] \\
&\quad = B_t(T_n)\mathbb{E}^{\mathbb{Q}^{T_n}}\left[\frac{\sum_{j=1}^{n} \delta_j B_{T_0}(T_j)}{B_{T_0}(T_n)}(R(T_0; T_0, T_n) - K)^+ |\mathcal{F}_t\right] \\
&\quad = B_t(T_n)u(t, L_t)
\end{aligned}$$

where u is the solution to

$$\begin{aligned}
\partial_t u + \mathcal{A}_t u &= 0 \\
u(T_0, x) &= \tilde{g}(x)
\end{aligned} \tag{5.3.9}$$

with $\tilde{g}(x) := \delta_n f_n(x)^{-1}\left(\sum_{j=1}^{n} f_j(x)x_j - K\right)^+$.

5.4 Approximate pricing

5.4.1 Approximate pricing for general payoffs under the terminal measure

Following an approach introduced by Černỳ et al., 2013, we introduce a small parameter into the model by defining the rescaled Lévy process $X_t^\alpha := \alpha X_{t/\alpha^2}$ with $\alpha \in (0,1)$. The process X^α is a martingale Lévy process under the terminal measure $\mathbb{Q}^{T_n}$ with characteristic triplet $(0, c, F_\alpha)$ with respect to the truncation function $h(z) = z$, where

$$F_\alpha(A) = \frac{1}{\alpha^2}F(\{z \in \mathbb{R}^d : z\alpha \in A\}, \quad \text{for } A \in \mathcal{B}(\mathbb{R}^d).$$

We now consider a family of Lévy Libor models driven by the processes X^α, $\alpha \in (0,1)$, and defined by

$$dL_t^\alpha = L_{t-}^\alpha (b_\alpha(t, L_t^\alpha)dt + \Lambda(t)dX_t^\alpha), \tag{5.4.1}$$

where the drift b_α is given by (5.2.5) with F replaced by F_α. Substituting the explicit form of F_α, we obtain

$$\begin{aligned}
b_\alpha^k(t, L_t) = &- \sum_{j=k+1}^{n} \frac{\delta_j L_t^j}{1+\delta_j L_t^j} \langle \lambda^k(t), c\,\lambda^j(t)\rangle \\
&+ \frac{1}{\alpha} \int_{\mathbb{R}^d} \langle \lambda^k(t), z\rangle \left(1 - \prod_{j=k+1}^{n} \left(1 + \frac{\alpha \delta_j L_t^j \langle \lambda^j(t), z\rangle}{1+\delta_j L_t^j}\right)\right) F(dz) \\
&= - \sum_{j_0=k+1}^{n} \Sigma_{kj_0}(t) \frac{\delta_{j_0} L_t^{j_0}}{1+\delta_{j_0} L_t^{j_0}} \\
&- \sum_{p=1}^{n-k-1} \alpha^p \sum_{j_0=k+1}^{n} \sum_{j_1=j_0+1}^{n} \cdots \sum_{j_p=j_{p-1}+1}^{n} M_t^{p+2}(\lambda^k, \lambda^{j_0}, \dots, \lambda^{j_p}) \prod_{l=0}^{p} \frac{\delta_{j_l} L_t^{j_l}}{1+\delta_{j_l} L_t^{j_l}} \\
&=: - \sum_{p=0}^{n-k-1} \alpha^p b_p^k(t, L_t)
\end{aligned}$$

where we define

$$\Sigma_{ij}(t) := (\Lambda(t) c \Lambda(t)^\top)_{ij} + \int_{\mathbb{R}^d} \langle \lambda^i(t), z\rangle \langle \lambda^j(t), z\rangle F(dz), \tag{5.4.2}$$

for all $i, j = 1, \dots, n$, and

$$M_t^k(\lambda^1, \dots, \lambda^k) := \int_{\mathbb{R}^d} \prod_{p=1}^{k} \langle \lambda^p(t), z\rangle F(dz) \tag{5.4.3}$$

for all $k = 1, \dots, n$. We denote the infinitesimal generator of L^α by $\mathcal{A}_t^\alpha$. For a smooth function $f : \mathbb{R}^d \to \mathbb{R}$, the infinitesimal generator $\mathcal{A}_t^\alpha f$ can be expanded in powers of α as follows:

$$\begin{aligned}
\mathcal{A}_t^\alpha f(x) = &\sum_{i=1}^{n} b_\alpha^i(t, x) x_i \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \Sigma_{ij}(t) x_i x_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \\
&+ \sum_{k=3}^{\infty} \sum_{i_1,\dots,i_k=1}^{n} \frac{\alpha^{k-2}}{k!} x_{i_1} \dots x_{i_k} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} M_t^k(\lambda^{i_1}, \dots, \lambda^{i_k}).
\end{aligned}$$

Consider now a financial product whose price is given by a generic PIDE of the form (5.3.1) with $\mathcal{A}_t$ replaced by $\mathcal{A}_t^\alpha$. Assuming sufficient regularity[2], one

[2] See (Ménassé and Tankov, 2015) for rigorous arguments in a simplified but similar setting.

may expand the solution u^α in powers of α:

$$u^\alpha(t,x) = \sum_{p=0}^{\infty} \alpha^p u_p(t,x). \tag{5.4.4}$$

Substituting the expansions for $\mathcal{A}_t^\alpha$ and b_α into this equation, and gathering terms with the same power of α, we obtain an 'open-ended' system of PIDE for the terms in the expansion of u^α.

The zero-order term u_0 satisfies

$$\partial_t u_0 + \mathcal{A}_t^0 u_0 = 0, \quad u_0(T_k, x) = \tilde{g}(x)$$

with

$$\mathcal{A}_t^0 u_0(t,x) = \sum_{i=1}^n b_0^i(t,x) x_i \frac{\partial u_0(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \Sigma_{ij}(t) x_i x_j \frac{\partial^2 u_0(t,x)}{\partial x_i \partial x_j} \tag{5.4.5}$$

$$b_0^i(t,x) = -\sum_{j=i+1}^n \Sigma_{ij}(t) \frac{\delta_j x_j}{1+\delta_j x_j}. \tag{5.4.6}$$

Hence, by the Feynman-Kac formula

$$u_0(t,x) = E^{\mathbb{Q}^{T_n}}\left[\tilde{g}(X_{T_k}^{t,x})\right] \tag{5.4.7}$$

where the process $X^{t,x} = (X^{i,t,x})_{i=1}^n$ satisfies the stochastic differential equation

$$dX_s^{i,t,x} = X_s^{i,t,x}\{b_0^i(s, X_s^{i,t,x})\, ds + \sigma_i\, dW_s\}, \quad X_t^{i,t,x} = x_i, \tag{5.4.8}$$

with W a d-dimensional standard Brownian motion with respect to $\mathbb{Q}^{T_n}$ and σ an $n \times d$-dimensional matrix such that $\sigma\sigma^\top = (\Sigma_{i,j})_{i,j=1}^n$.

To obtain an explicit approximation for the higher order terms $u_1(t,x)$ and $u_2(t,x)$ given above, we consider the following proposition.

Proposition 5.4.1. *Let Y be an n-dimensional log-normal process whose components follow the dynamics*

$$dY_t^i = Y_t^i(\mu_i(t)dt + \sigma_i(t)dW_t),$$

where μ and σ are measurable functions such that

$$\int_0^T (\|\mu(t)\| + \|\sigma(t)\|^2)dt < \infty$$

and for all $y \in \mathbb{R}^n$ and some $\varepsilon > 0$,

$$\inf_{0\le t\le T} y\sigma(t)\sigma(t)^T y^T \ge \varepsilon \|y\|^2.$$

We denote by $Y^{t,y}$ the process starting from y at time t, and by $Y^{t,y,i}$ the i-th component of this process. Let f be a bounded measurable function and define

$$v(t,y) = \mathbb{E}[f(Y_T^{t,y})].$$

Then, for all $i_1, \dots, i_m$, the process

$$Y_s^{t,y,i_1} \dots Y_s^{t,y,i_m} \frac{\partial^m v(s, Y_s^{t,y})}{\partial y_{i_1} \dots \partial y_{i_m}}, \quad s \geq t,$$

is a martingale.

The proof can be carried out by direct differentiation for smooth f together with a standard approximation argument for a general measurable f.

Furthermore, we assume the following simplification for the drift terms:

> For all $i = 1, \dots, n-1$ and $p = 1, \dots, n-k-1$, the random quantities in the terms $b_p^i(t, L_t)$ in the expansion of the drift of the Libor rates under the terminal measure are constant and equal to their value at time t, i.e. for all $j = 1, \dots, n$:
>
> $$\frac{\delta_j L_s^j}{1+\delta_j L_s^j} = \frac{\delta_j L_t^j}{1+\delta_j L_t^j}, \quad \text{for all} \quad s \geq t. \tag{5.4.9}$$
>
> This simplification is known as *freezing of the drift* and is often used for pricing in the Libor market models.

Coming back now to the first-order term u_1, we see that it is the solution of

$$\partial_t u_1 + \mathcal{A}_t^0 u_1 + \mathcal{A}_t^1 u_0 = 0, \quad u_1(T_k, x) = 0 \tag{5.4.10}$$

with

$$\begin{aligned}\mathcal{A}_t^1 u_0(t,x) = &\sum_{j=1}^n b_1^j(t,x) x_j \frac{\partial u_0(t,x)}{\partial x_j} \\ &+ \frac{1}{6} \sum_{i_1,i_2,i_3=1}^n x_{i_1} x_{i_2} x_{i_3} \frac{\partial^3 u_0(t,x)}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3}} M_t^3(\lambda^{i_1}, \lambda^{i_2}, \lambda^{i_3})\end{aligned} \tag{5.4.11}$$

and the drift term

$$b_1^j(t,x) = -\sum_{j_0=j+1}^n \sum_{j_1=j_0+1}^n M_t^3(\lambda^j, \lambda^{j_0}, \lambda^{j_1}) \frac{\delta_{j_0} x_{j_0}}{1+\delta_{j_0} x_{j_0}} \frac{\delta_{j_1} x_{j_1}}{1+\delta_{j_1} x_{j_1}}. \tag{5.4.12}$$

Moreover,

$$\mathcal{A}_t^0 u_1(t,x) = \sum_{i=1}^n b_0^i(t,x) x_i \frac{\partial u_1(t,x)}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \Sigma_{ij}(t) x_i x_j \frac{\partial^2 u_1(t,x)}{\partial x_i \partial x_j}.$$

We have

Lemma 5.4.2. *Consider the model (5.4.1). Under the simplification (5.4.9), the first-order term $u_1(t,x)$ in the expansion (5.4.4) can be approximated by*

$$
\begin{aligned}
u_1(t,x) \approx & \frac{1}{6}\sum_{i_1,i_2,i_3=1}^{n} x_{i_1}x_{i_2}x_{i_3}\frac{\partial^3 u_0(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\int_t^{T_k} M_s^3(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3})ds \\
& -\sum_{j=1}^{n}\sum_{j_0=j+1}^{n}\sum_{j_1=j_0+1}^{n}\frac{\delta_{j_0}x_{j_0}}{1+\delta_{j_0}x_{j_0}}\frac{\delta_{j_1}x_{j_1}}{1+\delta_{j_1}x_{j_1}}x_j\frac{\partial u_0(t,x)}{\partial x_j}\int_t^{T_k} M_s^3(\lambda^{j},\lambda^{j_0},\lambda^{j_1})ds \\
& =: \tilde{u}_1(t,x).
\end{aligned} \tag{5.4.13}
$$

Proof. Applying the Feynman-Kac formula to (5.4.10), we have,

$$
\begin{aligned}
u_1(t,x) = & \frac{1}{6}\int_t^{T_k} ds\sum_{i_1,i_2,i_3=1}^{n} M_s^3(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3})\,\mathbb{E}^{\mathbb{Q}^{T_n}}\left[X_s^{t,x,i_1}X_s^{t,x,i_2}X_s^{t,x,i_3}\frac{\partial^3 u_0(s,X_s^{t,x})}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\right] \\
& + \int_t^{T_k} ds\sum_{j=1}^{n}\mathbb{E}^{\mathbb{Q}^{T_n}}\left[b_1^j(s,X_s^{t,x})X_s^{t,x,j}\frac{\partial u_0(s,X_s^{t,x})}{\partial x_j}\right],
\end{aligned} \tag{5.4.14}
$$

with the process $(X_s^{t,x})$ defined by (5.4.8). Under the simplification (5.4.9), we can apply Proposition 5.4.1 to obtain (5.4.13). □

Similarly, the second-order term u_2 is the solution of

$$
\partial_t u_2 + \mathcal{A}_t^0 u_2 + \mathcal{A}_t^1 u_1 + \mathcal{A}_t^2 u_0 = 0, \quad u_2(T_k,x) = 0 \tag{5.4.15}
$$

with

$$
\begin{aligned}
\mathcal{A}_t^2 u_0(t,x) = & \sum_{j=1}^{n} b_2^j(t,x)x_j\frac{\partial u_0(t,x)}{\partial x_j} \\
& + \frac{1}{24}\sum_{i_1,i_2,i_3,i_4=1}^{n} x_{i_1}x_{i_2}x_{i_3}x_{i_4}\frac{\partial^4 u_0(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}x_{i_4}}M_t^4(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3},\lambda^{i_4})
\end{aligned} \tag{5.4.16}
$$

and the drift

$$
\begin{aligned}
b_2^j(t,x) = & -\sum_{j_0=j+1}^{n}\sum_{j_1=j_0+1}^{n}\sum_{j_2=j_1+1}^{n} M_t^4(\lambda^j,\lambda^{j_0},\lambda^{j_1},\lambda^{j_2})\frac{\delta_{j_0}x_{j_0}}{1+\delta_{j_0}x_{j_0}} \\
& \cdot\frac{\delta_{j_1}x_{j_1}}{1+\delta_{j_1}x_{j_1}}\frac{\delta_{j_2}x_{j_2}}{1+\delta_{j_2}x_{j_2}}.
\end{aligned} \tag{5.4.17}
$$

Lemma 5.4.3. *Consider the model (5.4.1). Under the simplification (5.4.9), the second-order term $u_2(t,x)$ in the expansion (5.4.4) can be approximated by*

$$
u_2(t,x) \approx \tilde{u}_2(t,x) := \tilde{E}_1 + \tilde{E}_2 + \tilde{E}_3 + \tilde{E}_4, \tag{5.4.18}
$$

with

$$\tilde{E}_1 := \frac{1}{6} \sum_{i_1,i_2,i_3=1}^{n} x_{i_1} x_{i_2} x_{i_3} \int_t^{T_k} ds M_s^3(\lambda^{i_1}, \lambda^{i_2}, \lambda^{i_3}) \cdot \left[\frac{1}{6} \sum_{i_4,i_5,i_6=1}^{n} \left(\int_s^{T_k} M_v^3(\lambda^{i_4}, \lambda^{i_5}, \lambda^{i_6}) dv \right) \frac{\partial^3 v^{i_4,i_5,i_6}(t,x)}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3}} - \sum_{j_4=1}^{n} \sum_{j_5=j_4+1}^{n} \sum_{j_6=j_5+1}^{n} \left(\int_s^{T_k} M_v^3(\lambda^{j_4}, \lambda^{j_5}, \lambda^{j_6}) dv \right) \frac{\partial^3 \bar{v}^{j_4,j_5,j_6}(t,x)}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3}} \right] \tag{5.4.19}$$

$$\tilde{E}_2 := -\sum_{j=1}^{n} \sum_{j_0=j+1}^{n} \sum_{j_1=j_0+1}^{n} \frac{\delta_{j_0} x_{j_0}}{1+\delta_{j_0} x_{j_0}} \frac{\delta_{j_1} x_{j_1}}{1+\delta_{j_1} x_{j_1}} x_j \int_t^{T_k} ds M_s(\lambda^j, \lambda^{j_0}, \lambda^{j_1}) \cdot \left[\frac{1}{6} \sum_{i_4,i_5,i_6=1}^{n} \left(\int_s^{T_k} M_v^3(\lambda^{i_4}, \lambda^{i_5}, \lambda^{i_6}) dv \right) \frac{\partial v^{i_4,i_5,i_6}(t,x)}{\partial x_j} - \sum_{j_4=1}^{n} \sum_{j_5=j_4+1}^{n} \sum_{j_6=j_5+1}^{n} \left(\int_s^{T_k} M_v^3(\lambda^{j_4}, \lambda^{j_5}, \lambda^{j_6}) dv \right) \frac{\partial^3 \bar{v}^{j_4,j_5,j_6}(t,x)}{\partial x_j} \right] \tag{5.4.20}$$

$$\tilde{E}_3 := \frac{1}{24} \sum_{i_1,i_2,i_3,i_4=1}^{n} x_{i_1} x_{i_2} x_{i_3} x_{i_4} \frac{\partial^4 u_0(t,x)}{\partial x_{i_1} \partial x_{i_2} \partial x_{i_3} \partial x_{i_4}} \int_t^{T_k} ds M_s^4(\lambda^{i_1}, \lambda^{i_2}, \lambda^{i_3}, \lambda^{i_4}) \tag{5.4.21}$$

and

$$\tilde{E}_4 := -\sum_{j=1}^{n} \sum_{j_0=j+1}^{n} \sum_{j_1=j_0+1}^{n} \sum_{j_2=j_1+1}^{n} \frac{\delta_{j_0} x_{j_0}}{1+\delta_{j_0} x_{j_0}} \frac{\delta_{j_1} x_{j_1}}{1+\delta_{j_1} x_{j_1}} \frac{\delta_{j_2} x_{j_2}}{1+\delta_{j_2} x_{j_2}} x_j \frac{\partial u_0(t,x)}{\partial x_j} \cdot \int_t^{T_k} M_s^4(\lambda^j, \lambda^{j_0}, \lambda^{j_1}, \lambda^{j_2}) ds \tag{5.4.22}$$

where we define

$$v^{i,j,l}(t,x) := x_i x_j x_l \frac{\partial^3 u_0(t,x)}{\partial x_i \partial x_j \partial x_l} \tag{5.4.23}$$

for all $i, j, l = 1, \ldots, n$ *and*

$$\bar{v}^{i,j,l}(t,x) := x_i \frac{\delta_j x_j}{1+\delta_j x_j} \frac{\delta_l x_l}{1+\delta_l x_l} \frac{\partial u_0(t,x)}{\partial x_i} \tag{5.4.24}$$

for all $i = 1, \ldots, n$, $j = i+1, \ldots, n$ *and* $l = j+1, \ldots, n$.

Proof. Once again by the Feynman-Kac formula applied to (5.4.15) we have

$$\begin{aligned}
u_2(t,x) = \frac{1}{6}\int_t^{T_k} ds & \sum_{i_1,i_2,i_3=1}^{n} M_s^3(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3}) \\
& \cdot \mathbb{E}^{\mathbb{Q}^{T_n}}\left[X_s^{t,x,i_1}X_s^{t,x,i_2}X_s^{t,x,i_3}\frac{\partial^3 u_1(s,X_s^{t,x})}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\right] \\
& + \int_t^{T_k} ds \sum_{j=1}^{n} \mathbb{E}^{\mathbb{Q}^{T_n}}\left[b_1^j(s,X_s^{t,x})X_s^{t,x,j}\frac{\partial u_1(s,X_s^{t,x})}{\partial x_j}\right] \\
& + \frac{1}{24}\int_t^{T_k} ds \sum_{i_1,i_2,i_3,i_4=1}^{n} M_s^4(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3},\lambda^{i_4}) \\
& \cdot \mathbb{E}^{\mathbb{Q}^{T_n}}\left[X_s^{t,x,i_1}X_s^{t,x,i_2}X_s^{t,x,i_3}X_s^{t,x,i_4}\frac{\partial^4 u_0(s,X_s^{t,x})}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}\partial x_{i_4}}\right] \\
& + \int_t^{T_k} ds \sum_{j=1}^{n} \mathbb{E}^{\mathbb{Q}^{T_n}}\left[b_2^j(s,X_s^{t,x})X_s^{t,x,j}\frac{\partial u_0(s,X_s^{t,x})}{\partial x_j}\right] \\
& =: E_1 + E_2 + E_3 + E_4
\end{aligned} \tag{5.4.25}$$

with the process $(X_s^{t,x})$ given by (5.4.8), $b_1^j(s,x)$ by (5.4.12) and $b_2^j(s,x)$ by (5.4.17).

In order to obtain an explicit expression for $u_2(t,x)$, we apply Proposition 5.4.1 combined with the simplification (5.4.9) for the drift terms b_1^j and b_2^j above. More precisely, the expressions for the third and the fourth expectation, which are present in the terms E_3 and E_4, follow by a straightforward application of Proposition 5.4.1 after using the simplification for b_2^j. We get

$$E_3 \approx \tilde{E}_3 \qquad \text{and} \qquad E_4 \approx \tilde{E}_4$$

with $\tilde{E}_3$ and $\tilde{E}_4$ given by (5.4.21) and (5.4.22), respectively.

To obtain explicit expressions for E_1 and E_2, firstly we insert the expression for $u_1(s,X_s^{t,x})$ as given by (5.4.14). After some straightforward calculations, based again on the application of Proposition 5.4.1 and the simplification (5.4.9) for b_1^j, which yields

$$E_1 \approx \tilde{E}_1 \qquad \text{and} \qquad E_2 \approx \tilde{E}_2$$

with $\tilde{E}_1$ and $\tilde{E}_2$ given by (5.4.19) and (5.4.20), respectively. Collecting the terms above concludes the proof. □

Summarizing, we get the following expansion for the time-t price $P^\alpha(t;g)$ of the payoff $g(L_{T_k})$ when $\alpha \to 0$.

Proposition 5.4.4. *Consider the model (5.4.1) and a European-type payoff with maturity T_k given by $\xi = g(L_{T_k})$. Assuming (5.4.9), its time-t price $P^\alpha(t;g)$ for $\alpha \to 0$ satisfies*

$$P^\alpha(t;g) = P_0(t;g) + \alpha P_1(t;g) + \alpha^2 P_2(t;g) + O(\alpha^3), \tag{5.4.26}$$

with

$$P_0(t;g) := B_t(T_n)u_0(t,L_t) =: P^{LMM}(t;g)$$

$$P_1(t;g) := B_t(T_n)u_1(t,L_t) \approx B_t(T_n)\tilde{u}_1(t,L_t)$$

$$P_2(t;g) := B_t(T_n)u_2(t,L_t) \approx B_t(T_n)\tilde{u}_2(t,L_t)$$

where $P^{LMM}(t;g)$ denotes the time-t price of the payoff $g(L_{T_k})$ in the log-normal LMM with covariance matrix Σ and the drift given by (5.4.6), $M_s^3(\lambda^{i_1}, \lambda^{i_2}, \lambda^{i_3})$ and $M_s^3(\lambda^{j}, \lambda^{j_0}, \lambda^{j_1})$ are given by (5.4.3), $u_0(t,x)$ by (5.4.7) and $\tilde{u}_1(t,x)$ and $\tilde{u}_2(t,x)$ by (5.4.13) and (5.4.18), respectively.

5.4.2 Approximate pricing of caplets

Recalling that the caplet price is given by (5.3.2), where u is the solution of the PIDE (5.3.3), we can approximate this price using the development

$$u^\alpha(t,x) = u_0(t,x) + \alpha u_1(t,x) + \alpha^2 u_2(t,x) + O(\alpha^3)$$

where the zero-order term u_0 satisfies

$$\partial_t u_0 + \mathcal{A}_t^0 u_0 = 0, \quad u_0(T_{k-1},x) = (x_k - K)^+ \prod_{j=k+1}^{n} (1+\delta_j x_j)$$

$$\text{with} \quad \mathcal{A}_t^0 u_0 = \sum_{i=1}^{n} b_0^i(t,x) x_i \frac{\partial u_0(t,x)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} \Sigma_{ij}(t) x_i x_j \frac{\partial^2 u_0(t,x)}{\partial x_i \partial x_j}$$

$$\text{and} \quad b_0^i(t,x) = -\sum_{j=i+1}^{n} \Sigma_{ij}(t) \frac{\delta_j x_j}{1+\delta_j x_j}.$$

The solution to the above PDE can be found via the Feynman-Kac formula, where the conditional expectation is computed in the log-normal LMM model with covariation matrix $(\Sigma_{ij})_{i,j=1}^n$ as in Section 5.4.1. Performing a measure change from $\mathbb{Q}^{T_n}$ to $\mathbb{Q}^{T_k}$ and denoting by $P_{BS}(V,S,K)$ the Black-Scholes price of a call option with variance V,

$$P_{BS}(V,S,K) = \mathbb{E}\left[\left(Se^{-\frac{V}{2}+\sqrt{V}Z} - K\right)^+\right], \quad Z \sim N(0,1),$$

we see that the zero-order term is given by

$$u_0(t,x) = P_{BS}(V^{Cpl}_{t,T_{k-1}}, x_k, K) \prod_{j=k+1}^{n} (1+\delta_j x_j), \tag{5.4.27}$$

where

$$V^{Cpl}_{t,T} := \int_t^T \Sigma_{kk}(s)ds. \tag{5.4.28}$$

Now, in complete analogy to the case of a general payoff, the first-order term $u_1(t,x)$ and the second-order term $u_2(t,x)$ are given by (5.4.14) and (5.4.25), respectively, with $u_0(t,x)$ as in (5.4.27). Noting that $u_0(t,x)$ depends only on $x_k, x_{k+1}, \dots, x_n$, the derivatives of $u_0(t,x)$ with respect to $x_1, \dots, x_{k-1}$ are zero and the sums in (5.4.14) and (5.4.25) in fact start from the index k. An application of Proposition 5.4.1 and simplification (5.4.9) thus yields the following proposition, which provides an approximation of the caplet price $P^{Cpl,\alpha}(t;T_{k-1},T_k,K)$ when $\alpha \to 0$.

Proposition 5.4.5. *Consider the model (5.4.1) and a caplet with strike K and maturity T_{k-1}. Assuming (5.4.9), its time-t price $P^{Cpl,\alpha}(t;T_{k-1},T_k,K)$ for $\alpha \to 0$ satisfies*

$$\begin{aligned} P^{Cpl,\alpha}(t;T_{k-1},T_k,K) &= P_0^{Cpl}(t;T_{k-1},T_k,K) + \alpha P_1^{Cpl}(t;T_{k-1},T_k,K) \\ &\quad + \alpha^2 P_2^{Cpl}(t;T_{k-1},T_k,K) + O(\alpha^3), \end{aligned} \tag{5.4.29}$$

with

$$\begin{aligned} P_0^{Cpl}(t;T_{k-1},T_k,K) &:= B_t(T_n)\delta_k u_0(t,L_t) \\ &= B_t(T_n)\delta_k P_{BS}(V^{Cpl}_{t,T_{k-1}}, L_t^k, K) \prod_{j=k+1}^{n}(1+\delta_j L_t^j) \end{aligned}$$

$$\begin{aligned} &P_1^{Cpl}(t;T_{k-1},T_k,K) \\ &:= B_t(T_n)\delta_k \left\{ \frac{1}{6} \sum_{i_1,i_2,i_3=k}^{n} L_t^{i_1} L_t^{i_2} L_t^{i_3} \frac{\partial^3 u_0(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\Big|_{x=L_t} \int_t^{T_{k-1}} M_s^3(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3})ds \right. \\ &\quad - \sum_{j=k}^{n} \sum_{j_0=j+1}^{n} \sum_{j_1=j_0+1}^{n} \frac{\delta_{j_0} L_t^{j_0}}{1+\delta_{j_0} L_t^{j_0}} \frac{\delta_{j_1} L_t^{j_1}}{1+\delta_{j_1} L_t^{j_1}} L_t^j \frac{\partial u_0(t,x)}{\partial x_j}\Big|_{x=L_t} \\ &\quad \left. \cdot \int_t^{T_{k-1}} M_s^3(\lambda^j,\lambda^{j_0},\lambda^{j_1})ds \right\} \end{aligned}$$

$P_2^{Cpl}(t;T_{k-1},T_k,K)$

$$
\begin{aligned}
&:= B_t(T_n)\delta_k \Bigg\{ \frac{1}{6} \sum_{i_1,i_2,i_3=k}^{n} L_t^{i_1} L_t^{i_2} L_t^{i_3} \int_t^{T_{k-1}} ds M_s^3(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3}) \\
&\qquad \cdot \Bigg[\frac{1}{6} \sum_{i_4,i_5,i_6=k}^{n} \left(\int_s^{T_{k-1}} M_v^3(\lambda^{i_4},\lambda^{i_5},\lambda^{i_6}) dv \right) \frac{\partial^3 v^{i_4,i_5,i_6}(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\Big|_{x=L_t} \\
&\qquad\quad - \sum_{j_4=k}^{n} \sum_{j_5=j_4+1}^{n} \sum_{j_6=j_5+1}^{n} \left(\int_s^{T_{k-1}} M_v^3(\lambda^{j_4},\lambda^{j_5},\lambda^{j_6}) dv \right) \frac{\partial^3 \bar{v}^{j_4,j_5,j_6}(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}}\Big|_{x=L_t} \Bigg] \\
&- \sum_{j=k}^{n} \sum_{j_0=j+1}^{n} \sum_{j_1=j_0+1}^{n} \frac{\delta_{j_0} L_t^{j_0}}{(1+\delta_{j_0} L_t^{j_0}} \frac{\delta_{j_1} L_t^{j_1}}{(1+\delta_{j_1} L_t^{j_1}} L_t^{j} \int_t^{T_{k-1}} ds M_s(\lambda^{j},\lambda^{j_0},\lambda^{j_1}) \\
&\qquad \cdot \Bigg[\frac{1}{6} \sum_{i_4,i_5,i_6=k}^{n} \left(\int_s^{T_{k-1}} M_v^3(\lambda^{i_4},\lambda^{i_5},\lambda^{i_6}) dv \right) \frac{\partial v^{i_4,i_5,i_6}(t,x)}{\partial x_j}\Big|_{x=L_t} \\
&\qquad\quad - \sum_{j_4=k}^{n} \sum_{j_5=j_4+1}^{n} \sum_{j_6=j_5+1}^{n} \left(\int_s^{T_{k-1}} M_v^3(\lambda^{j_4},\lambda^{j_5},\lambda^{j_6}) dv \right) \frac{\partial^3 \bar{v}^{j_4,j_5,j_6}(t,x)}{\partial x_j}\Big|_{x=L_t} \Bigg] \\
&+ \frac{1}{24} \sum_{i_1,i_2,i_3,i_4=k}^{n} L_t^{i_1} L_t^{i_2} L_t^{i_3} L_t^{i_4} \frac{\partial^4 u_0(t,x)}{\partial x_{i_1}\partial x_{i_2}\partial x_{i_3}\partial x_{i_4}}\Big|_{x=L_t} \int_t^{T_{k-1}} M_s^4(\lambda^{i_1},\lambda^{i_2},\lambda^{i_3},\lambda^{i_4}) ds \\
&- \sum_{j=k}^{n} \sum_{j_0=j+1}^{n} \sum_{j_1=j_0+1}^{n} \sum_{j_2=j_1+1}^{n} \frac{\delta_{j_0} L_t^{j_0}}{1+\delta_{j_0} L_t^{j_0}} \frac{\delta_{j_1} L_t^{j_1}}{1+\delta_{j_1} L_t^{j_1}} \frac{\delta_{j_2} L_t^{j_2}}{1+\delta_{j_2} L_t^{j_2}} L_t^{j} \frac{\partial u_0(t,x)}{\partial x_j}\Big|_{x=L_t} \\
&\qquad \cdot \int_t^{T_{k-1}} M_s^4(\lambda^{j},\lambda^{j_0},\lambda^{j_1},\lambda^{j}) ds \Bigg\}
\end{aligned}
$$

with $V_{t,T_{k-1}}^{Cpl}$ given by (5.4.28), $u_0(t,x)$ by (5.4.27), the terms $M_s^3(\cdot)$ and $M_s^4(\cdot)$ by (5.4.3) and $v^{i_4,i_5,i_6}(t,x)$ and $\bar{v}^{j_4,j_5,j_6}(t,x)$ by (5.4.23) and (5.4.24), respectively.

Remark 5.4.6. *Recalling that*

$$u_0(t,x) = P_{BS}(V_{t,T}^{Cpl}, x_k, K) \prod_{j=k+1}^{n} (1+\delta_j x_j)$$

we see that the functions v and $\bar{v}$ given by

$$v^{i,j,l}(t,x) := x_i x_j x_l \frac{\partial^3 u_0(t,x)}{\partial x_i \partial x_j \partial x_l}$$

for all $i,j,l = k,\dots,n$ and

$$\bar{v}^{i,j,l}(t,x) := x_i \frac{\delta_j x_j}{1+\delta_j x_j} \frac{\delta_l x_l}{1+\delta_l x_l} \frac{\partial u_0(t,x)}{\partial x_i}$$

for all $i = k, \dots, n$, $j = i+1, \dots, n$ and $l = j+1, \dots, n$, become in fact linear combinations of the terms which are polynomials in x multiplied by derivatives of $P_{BS}(\cdot)$ up to order three.

5.4.3 Approximate pricing of swaptions

Let us consider a swaption defined in Section 5.3.3. For swaption pricing we again use the general result under the terminal measure $\mathbb{Q}^{T_n}$ given in Proposition 5.4.4. The price of the swaption $P^{Swn}(t; T_0, T_n, K)$ then satisfies

$$\begin{aligned} P^{Swn}(t; T_0, T_n, K) &= B_t(T_n)(u_0(t, L_t) + \alpha u_1(t, L_t) + \alpha^2 u_2(t, L_t)) + O(\alpha^3) \\ &=: P_0^{Swn}(t; T_0, T_n, K) + \alpha P_1^{Swn}(t; T_0, T_n, K) \\ &\quad + \alpha^2 P_2^{Swn}(t; T_0, T_n, K) + O(\alpha^3), \end{aligned}$$

where the function u_0 satisfies the equation

$$\partial_t u_0 + \mathcal{A}_t^0 u_0 = 0, \qquad u_0(T_0, x) = \tilde{g}(x)$$

with $\tilde{g}(x) = \delta_n f_n(x)^{-1} \left(\sum_{j=1}^n f_j(x) x_j - K \right)^+$. We see that the zero-order term $P_0^{Swn}(t; T_0, T_n, K)$ corresponds to the price of the swaption in the log-normal LMM model with volatility matrix $\Sigma(t)$.

The function u_0 related to the swaption price in the log-normal LMM is of course not known in explicit form but one can use various approximations developed in the literature (Jäckel and Rebonato, 2003; Schoenmakers, 2005). To introduce the approximation of (Jäckel and Rebonato, 2003), we compute the quadratic variation of the log swap rate expressed as function of Libor rates:

$$R(t; T_0, T_n) = R(L_t^1, \dots, L_t^n) = \frac{\sum_{j=1}^n \delta_j L_t^j \prod_{k=1}^j (1 + \delta_k L_t^k)}{\sum_{j=1}^n \delta_j \prod_{k=1}^j (1 + \delta_k L_t^k)}.$$

$$\begin{aligned} \langle \log R(\cdot; T_0, T_n) \rangle_T &= \int_0^T \frac{d\langle R(\cdot; T_0, T_n) \rangle_t}{R(t; T_0, T_n)^2} = \int_0^T \sum_{i,j=1}^n \frac{\partial R(L_t)}{\partial L^i} \frac{\partial R(L_t)}{\partial L^j} \frac{d\langle L^i, L^j \rangle_t}{R(t; T_0, T_n)^2} \\ &= \int_0^T \sum_{i,j=1}^n \frac{\partial R(L_t)}{\partial L^i} \frac{\partial R(L_t)}{\partial L^j} \frac{L_t^i L_t^j \Sigma_{ij}(t) dt}{R(t; T_0, T_n)^2}. \end{aligned}$$

The approximation of (Jäckel and Rebonato, 2003) consists in replacing all stochastic processes in the above integral by their values at time 0; in other words, the swap rate becomes a log-normal random variable such that $\log R(t; T_0, T_n)$ has variance

$$V_T^{swap} = \sum_{i,j=1}^n \frac{\partial R(L_0)}{\partial L^i} \frac{\partial R(L_0)}{\partial L^j} \frac{L_0^i L_0^j}{R(0; T_0, T_n)^2} \int_0^T \Sigma_{ij}(t) dt.$$

The function $u_0(0,x)$ can then be approximated by applying the Black-Scholes formula (for simplicity, $t=0$):

$$u_0(0,x) \approx P_{BS}(V_T^{swap}, R(0;T_0,T_n), K)\,.$$

5.5 Numerical examples

In this section, we test the performance of our approximation at pricing caplets on Libor rates in the model (5.2.4), where X_t is a unidimensional CGMY process (Carr et al., 2007). The CGMY process is a pure jump process, so that $c=0$, with Lévy measure

$$F(dz) = \frac{C}{|z|^{1+Y}}\left(e^{-\lambda_- z}\mathbf{1}_{\{z<0\}} + e^{-\lambda_+ z}\mathbf{1}_{\{z>0\}}\right)\,dz\,.$$

The jumps of this process are not bounded from below but the parameters we choose ensure that the probability of having a negative Libor rate value is negligible. We choose the time grid $T_0=5$, $T_1=6$, ... $T_5=10$, the volatility parameters $\lambda_i=1$, $i=1,...,5$, the initial forward Libor rates $L_0^i=0.06$, $i=1,...,5$ and the bond price for the first maturity $B_0(T_0)=1.06^{-5}$. The CGMY model parameters are chosen according to four different cases described in the following table, which also gives the standard deviation and excess kurtosis of X_1 for each case. Case 1 corresponds to a Lévy process that is close to the Brownian motion (Y close to 2 and λ_+ and λ_- large) and Case 4 is a Lévy process that is very far from Brownian motion.

Case	C	λ_+	λ_-	Y	Volatility	Excess kurtosis
1	0.01	10	20	1.8	23.2%	0.028
2	0.1	10	20	1.2	17%	0.36
3	0.2	10	20	0.5	8.7%	3.97
4	0.2	3	5	0.2	18.9%	12.7

We first calculate the price of the ATM caplet with maturity T_1 written on the Libor rate L^1 with the zero-order, first-order and second-order approximation, using as benchmark the jump-adapted Euler scheme of Kohatsu-Higa and Tankov, 2010. The first Libor rate is chosen to maximize the nonlinear effects related to the drift of the Libor rates, since the first maturity is the farthest from the terminal date. The results are shown in Table 5.1. We see that for all four cases, the price computed by second-order approximation is within or at the boundary of the Monte Carlo confidence interval, which is itself quite narrow (computed with 10^6 trajectories).

Secondly, we evaluate the prices of caplets with strikes ranging from 3% to 9% and explore the performance of our analytic approximation for estimating the caplet implied volatility smile. The results are shown in Figure 5.5.1. We see that in Cases 1, 2 and 3, which correspond to the parameter

	Case 1	Case 2	Case 3	Case 4
Order 0	0.008684	0.006392	0.003281	0.007112
Order 1	0.008677	0.006361	0.003241	0.006799
Order 2	**0.008677**	**0.006351**	**0.003172**	**0.006556**
MC lower bound	0.008626	0.006306	0.003178	0.006493
MC upper bound	0.008712	0.006361	0.003204	0.006578

Table 5.1: Price of ATM caplet computed using the analytic approximation together with the 95% confidence bounds computed by Monte Carlo over 10^6 trajectories.

values most relevant in practice given the value of the excess kurtosis, the second order approximation reproduces the volatility smile quite well (in case 1 there is actually no smile, see the scale on the Y axis of the graph). In case 4, which corresponds to very violent jumps and pronounced smile, the qualitative shape of the smile is correctly reproduced, but the actual values are often outside the Monte Carlo interval. This means that in this extreme case the model is too far from the Gaussian LMM for our approximation to be precise. We also note that the algorithm runs in $\mathcal{O}(n^6)$, for the second order approximation, due to the number of partial derivatives that one has to calculate. The algorithm may therefore run slowly, should n become too large.

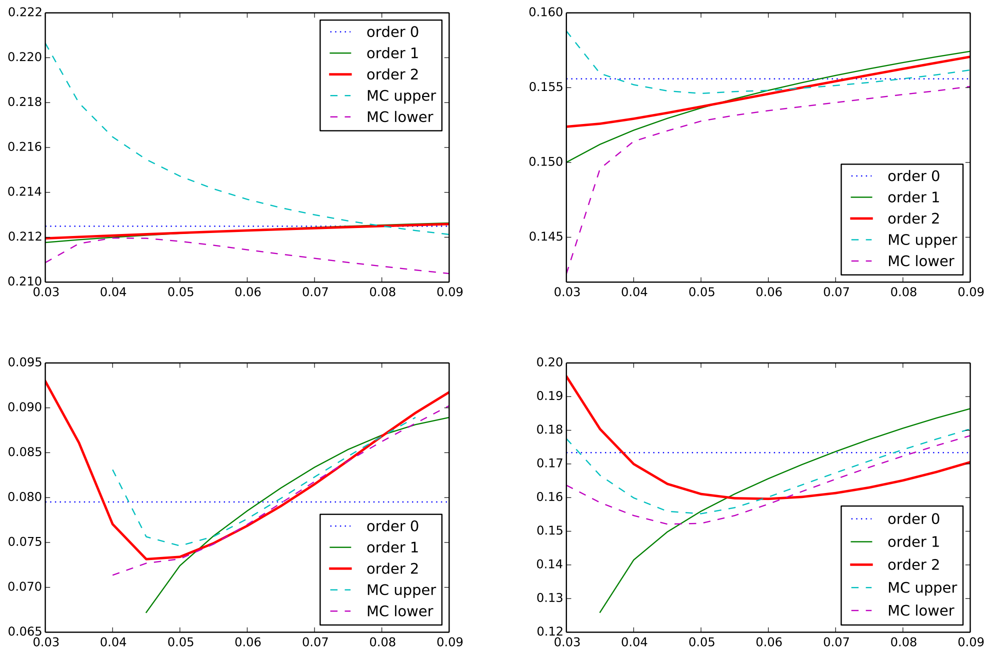

Figure 5.5.1: Implied volatilities of caplets with different strikes computed using the analytic approximation together with the Monte Carlo bound. Top graphs: Case 1 (left) and Case 2 (right). Bottom graphs: Case 3 (left) and Case 4 (right).

Bibliography

Ahdida, A. and Alfonsi, A. (2013). Exact and high-order discretization schemes for wishart processes and their affine extensions. *Annals of Applied Probabilities*, 23(3):1025–1073.

Ahn, D. H. and Gao, B. (1999). A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies*, 12(4):721–762.

Alfonsi, A. (2015). *Affine Diffusions and Related Processes: Simulation, Theory and Applications*, volume 6 of *Bocconi & Springer Series*. Springer, Cham; Bocconi University Press, Milan.

Alfonsi, A., Kebaier, A., and Rey, C. (2016). Maximum likelihood estimation for Wishart processes. *Stochastic Processes and their Applications*, 126(11):3243–3282.

Antonelli, F. and Scarlatti, S. (2009). Pricing options under stochastic volatility: a power series approach. *Finance and Stochastics*, 13(2):269–303.

Azencott, R. (1984). Densité des diffusions en temps petit: développements asymptotiques. In *Séminaire de Probabilités XVIII 1982/83*, pages 402–498. Springer.

Bachelier, L. (1900). *Théorie de la Spéculation*. Gauthier-Villars.

Barndorff-Nielsen, O. E. (1997). Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Bayer, C. and Laurence, P. (2013a). Asymptotics beats Monte-Carlo: The case of correlated local vol baskets. *Wiley, Communication on Pure and Applied Mathematics*, 67(10):1618–1657.

Bayer, C. and Laurence, P. (2013b). Asymptotics for at the money local vol basket options. *Preprint*.

Belomestny, D. and Schoenmakers, J. (2011). A jump-diffusion Libor model and its robust calibration. *Quantitative Finance*, 11(4):529–546.

Ben Arous, G. (1988a). Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. In *Annales Scientifiques de l'École Normale Supérieure*, volume 21, pages 307–331. Elsevier.

Ben Arous, G. (1988b). Methode de Laplace et de la phase stationnaire sur l'espace de Wiener. *Stochastics*, 25(3):125–153.

Benabid, A., Bensusan, H., and El Karoui, N. (2008). Wishart stochastic volatility: Asymptotic smile and numerical framework. *Preprint*.

Benhamou, E., Gobet, E., and Miri, M. (2008). Closed forms for european options in a local volatility model. *Preprint, SSRN*.

Benhamou, E., Gobet, E., and Miri, M. (2009). Smart expansion and fast calibration for jump diffusions. *Finance and Stochastics*, 13(4):563–589.

Benhamou, E., Gobet, E., and Miri, M. (2010). Time dependent Heston model. *Society for Industrial and Applied Mathematics*, 1(1):289–325.

Berestycki, H., Busca, J., and Florent, I. (2004). Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 57(10):1352–1373.

Bismut, J. (1984). *Large deviations and the Malliavin calculus*. Progress in mathematics. Birkhäuser.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Bleistein, N. and Handelsman, R. A. (1975). *Asymptotic Expansions of Integrals*. Courier Corporation.

Brace, A., Gatarek, D., and Musiela, M. (1997). The market model of interest rate dynamics. *Mathematical Finance*, 7(2):127–155.

Brigo, D. and Mercurio, F. (2001). *Interest Rate Modelling - Theory and Practice*. Springer Finance, second edition.

Bru, M. F. (1991). Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *The Journal of Business*, 75(2):305–332.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2005). Pricing options on realized variance. *Finance and Stochastics*, 9(4):453–475.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2007). Self decomposability and option pricing. *Mathematical Finance*, 17(1):31–57.

Carr, P. and Lee, R. (2008). Robust replication of volatility derivatives. In *PRMIA award for best paper in derivatives, MFA Annual Meeting 2008*. Citeseer.

Carr, P. and Lee, R. (2009). Volatility derivatives. *Annual Review of Financial Economics*, 1(1):319–339.

Carr, P. and Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73.

Carr, P. and Sun, J. (2007). A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, 10(2):87–150.

Carr, P. and Wu, L. (2003). What type of process underlies options? a simple robust test. *The Journal of Finance*, 58(6):2581–2610.

Černỳ, A., Denkl, S., and Kallsen, J. (2013). Hedging in Lévy models and the time step equivalent of jumps. *Preprint*.

Chicago Board Options Exchange, CBOE (2018). Vix white paper. *URL: https://www.cboe.com/micro/vix/vixwhite.pdf*.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*, volume 2. CRC press.

Cox, J. C. (1996). The constant elasticity of variance option pricing model. *The Journal of Portfolio Management*, 23(5):15–17.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, pages 363–384.

Cuchiero, C., Filipović, D., Mayerhofer, E., and Teichmann, J. (2011). Affine processes on positive semidefinite matrices. *The Annals of Applied Probability*, 21(2):397–463.

Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2007). Option pricing when correlations are stochastic: An analytical framework. *Review of Derivatives Research*, 10(2):151–180.

Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2008). A multifactor volatility Heston model. *Quantitative Finance*, 8(6):591–604.

Dalang, R. C., Morton, A., and Willinger, W. (1990). Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 29(2):185–201.

De Franco, C. (2012). *Two Studies in Risk Management: Portfolio Insurance under Risk Measure Constraint and Quadratic Hedge for Jump Processes*.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Springer, Application of Mathematics, second edition.

Deuschel, J. D., Friz, P. K., Jacquier, A., and Violante, S. (2014a). Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations. *Communications on Pure and Applied Mathematics*, 67(1):40–82.

Deuschel, J. D., Friz, P. K., Jacquier, A., and Violante, S. (2014b). Marginal density expansions for diffusions and stochastic volatility II: Applications. *Communications on Pure and Applied Mathematics*, 67(2):321–350.

Drimus, G. G. (2012). Options on realized variance by transform methods: a non-affine stochastic volatility model. *Quantitative Finance*, 12(11):1679–1694.

Duffie, D., Filipovic, D., and Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1).

Dupuis, P. and Wang, H. (2004). Importance sampling, large deviations, and differential games. *Stochastics: An International Journal of Probability and Stochastic Processes*, 76(6):481–508.

Eberlein, E. and Kluge, W. (2007). Calibration of Lévy term structure models. *Advances in Mathematical Finance*, pages 147–172.

Eberlein, E. and Özkan, F. (2005). The Lévy Libor model. *Finance and Stochastics*, 9(3):327–348.

Forde, M. and Jacquier, A. (2009). Small-time asymptotics for implied volatility under the Heston model. *International Journal of Theoretical and Applied Finance*, 12(06):861–876.

Forde, M. and Jacquier, A. (2011a). The large-maturity smile for the Heston model. *Finance and Stochastics*, 15(4):755–780.

Forde, M. and Jacquier, A. (2011b). Small-time asymptotics for an uncorrelated local-stochastic volatility model. *Applied Mathematical Finance*, 18(6):517–535.

Fouque, J. P., Papanicolaou, G., and Sircar, R. (2000). *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press.

Fouque, J. P., Papanicolaou, G., Sircar, R., and Solna, K. (2003). Multiscale stochastic volatility asymptotics. *Multiscale Modeling & Simulation*, 2(1):22–42.

Freidlin, M. I. and Wentzell, A. D. (2012). *Random Perturbations of Dynamical Systems*. Springer, third edition.

Frikha, N. and Kohatsu-Higa, A. (2016). A parametrix approach for asymptotic expansion of markov semigroups with applications to multidimensional diffusion processes. *Preprint*.

Gatarek, D. (2003). Libor market model with stochastic volatility. *Preprint, SSRN*.

Gatheral, J., Hsu, E. P., Laurence, P., Ouyang, C., and Wang, T. H. (2012). Asymptotics of implied volatility in local volatility models. *Mathematical Finance*, 22(4):591–620.

Genin, A. and Tankov, P. (2016). Optimal importance sampling for Lévy processes. *Preprint, arXiv: 1608.04621*.

Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999). Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance*, 9(2):117–152.

Glasserman, P. and Kou, S. G. (2003). The term structure of simple forward rates with jump risk. *Mathematical Finance*, 13(3):383–410.

Gombani, A. and Runggaldier, W. J. (2001). A filtering approach to pricing in multifactor term structure models. *International Journal of Theoretical and Applied Finance*, 4(2):303–320.

Gourieroux, C. and Sufana, R. (2004). Derivative pricing with multivariate stochastic volatility: Application to credit risk. *Les Cahiers du CREF of HEC Montréal*, (Paper No. CREF 04-09).

Grbac, Z., Krief, D., and Tankov, P. (2018). Long-time trajectorial large deviations for affine stochastic volatility models and application to variance reduction for option pricing. *Preprint*.

Grbac, Z. and Runggaldier, W. J. (2015). *Interest Rate Modeling: Post-Crisis Challenges and Approaches*. Springer.

Guasoni, P. and Robertson, S. (2008). Optimal importance sampling with explicit formulas in continuous time. *Finance and Stochastics*, 12(1):1–19.

Gulisashvili, A. and Stein, E. M. (2010). Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics and Optimization*, 61(3):287–315.

Hagan, P. and Lesniewski, A. (2008). Libor market model with SABR style stochastic volatility. *JP Morgan Chase and Ellington Management Group*, 32.

Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). Managing smile risk. *The Best of Wilmott*, 1:249–296.

Hagan, P. S. and Woodward, D. E. (1999). Equivalent Black volatilities. *Applied Mathematical Finance*, 6(3):147–157.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105.

Henry-Labordère, P. (2008). *Analysis, Geometry and Modeling in Finance: Advanced Methods in Option Pricing*. Chapman and Hall/CRC.

Henry-Labordere, P. (2017). *Model-free Hedging: A Martingale Optimal Transport Viewpoint*. Chapman and Hall/CRC.

Heston, S. L. (1993). A closed-form solutions for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Jäckel, P. and Rebonato, R. (2003). The link between caplet and swaption volatilities in a Brace-Gatarek-Musiela/Jamshidian framework: approximate solutions and empirical evidence. *Journal of Computational Finance*, 6(4):41–60.

Jacod, J. and Shiryaev, A. (2003). *Limit Theorems for Stochastic Processes*, volume 288. Springer Science & Business Media.

Jacquier, A. and Keller-Ressel, M. (2018). Implied volatility in strict local martingale models. *SIAM Journal on Financial Mathematics*, 9(1):171–189.

Jacquier, A., Keller-Ressel, M., and Mijatović, A. (2013). Large deviations and stochastic volatility with jumps: asymptotic implied volatility for affine models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 85(2):321–345.

Jacquier, A. and Lorig, M. (2015). From characteristic functions to implied volatility expansions. *Advances in Applied Probability*, 47(3):837–857.

Jacquier, A. and Mijatović, A. (2014). Large deviations for the extended Heston model: the large-time case. *Asia-Pacific Financial Markets*, 21(3):263–280.

Jourdain, B. and Lelong, J. (2009). Robust adaptive importance sampling for normal random vectors. *The Annals of Applied Probability*, 19(5):1687–1718.

Jurdjevic, V. (1997). *Geometric Control Theory*, volume 52. Cambridge university press.

Kallsen, J. (2006). A didactic note on affine stochastic volatility models. pages 343–368.

Kallsen, J., Muhle-Karbe, J., and Voß, M. (2011). Pricing options on variance in affine stochastic volatility models. *Mathematical Finance*, 21(4):627–641.

Keller-Ressel, M. (2011). Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21(1):73–98.

Keller-Ressel, M. and Muhle-Karbe, J. (2013). Asymptotic and exact pricing of options on variance. *Finance and Stochastics*, 17(1):107–133.

Kluge, W. (2005). *Time-Inhomogeneous Lévy Processes in Interest Rate and Credit Risk Models*. PhD thesis, University of Freiburg.

Kohatsu-Higa, A. and Tankov, P. (2010). Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*, 120(11):2258–2285.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

Kunitomo, N. and Takahashi, A. (1995). The asymptotic expansion approach to the valuation of interest rate contingent claims.

Kunitomo, N. and Takahashi, A. (2001). The asymptotic expansion approach to the valuation of interest rate contingent claims. *Mathematical Finance*, Volume 11(1):117–151.

Kunitomo, N. and Takahashi, A. (2003). On validity of the asymptotic expansion approach in contingent claim analysis. *The Annals of Applied Probability*, 13(3):914–952.

Léandre, R. (1987). Minoration en temps petit de la densité d'une diffusion dégénérée. *Journal of Functional Analysis*, 74(2):399–414.

Lee, R. W. (2001). Implied and local volatilities under stochastic volatility. *International Journal of Theoretical and Applied Finance*, 4(1):45–89.

Léonard, C. (2000). Large deviations for Poisson random measures and processes with independent increments. *Stochastic Processes and their Applications*, 85(1):93–121.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4):394–419.

Ménassé, C. and Tankov, P. (2015). Asymptotic indifference pricing in exponential Lévy models. *Preprint arXiv: 1502.03359.*

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of economics and management science*, pages 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144.

Miltersen, K. R., Sandmann, K., and Sondermann, D. (1994). Closed form term structure derivatives in a Heath-Jarrow-Morton model with lognormal annually compounded interest rates. *Proceedings of the Seventh Annual European Futures Research Symposium Bonn.*

Miltersen, K. R., Sandmann, K., and Sondermann, D. (1997). Closed form solutions for term structure derivatives with log-normal interest rates. *The Journal of Finance*, 52(1):409–430.

Molčanov, S. A. (1975). Diffusion processes and riemannian geometry. *Russian Mathematical Surveys*, 30(1):1–63.

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, volume 1995. Springer.

Ornstein, L. S. and Uhlenbeck, G. E. (1930). On the theory of the brownian motion. *Physical review*, 36(5):823–841.

Papanicolaou, G., Fouque, J. P., Solna, K., and Sircar, R. (2003). Singular perturbations in option pricing. *SIAM Journal on Applied Mathematics*, 63(5):1648–1665.

Papapantoleon, A., Schoenmakers, J., and Skovmand, D. (2011). Efficient and accurate log-Lévy approximations to Lévy driven Libor models. *Journal of Computational Finance*, pages 1460–1559.

Paulot, L. (2015). Asymptotic implied volatility at the second order with application to the SABR model. pages 37–69.

Piterbarg, V. (2003). A stochastic volatility forward Libor model with a term structure of volatility smiles. *Preprint, SSRN.*

Robertson, S. (2010). Sample path large deviations and optimal importance sampling for stochastic volatility models. *Stochastic Processes and their Applications*, 120(1):66–83.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.

Rockafellar, R. T. (1971). Integrals which are convex functionals. II. *Pacific Journal of Mathematics*, 39(2):439–469.

Schoenmakers, J. (2005). *Robust Libor Modelling and Pricing of Derivative Products*. Chapman & Hall/CRC, Financial Mathematics Series.

Sepp, A. (2008). Pricing options on realized variance in the Heston model with jumps in returns and volatility. *Journal of Computational Finance*, 11(4):33–70.

Shiraya, K. and Takahashi, A. (2011). Pricing average options on commodities. *Journal of Futures Markets*, 31(5):407–439.

Shiraya, K. and Takahashi, A. (2014). Pricing multiasset cross-currency options. *Journal of Futures Markets*, 34(1):1–19.

Shiraya, K. and Takahashi, A. (2016). An approximation formula for basket option prices under local stochastic volatility with jumps: An application to commodity markets. *J. Computational Applied Mathematics*, 292:230–256.

Shiraya, K. and Takahashi, A. (2017a). An asymptotic expansion for local-stochastic volatility with jump models. *Stochastics*, 89(1):65–88.

Shiraya, K. and Takahashi, A. (2017b). Pricing average and spread options under local-stochastic volatility jump-diffusion models online appendix. Technical report, Center for Advanced Research in Finance, Faculty of Economics, The University of Tokyo.

Siegmund, D. (1976). Importance sampling in the Monte Carlo study of sequential tests. *The Annals of Statistics*, 4(4):673–684.

Sircar, R. and Papanicolaou, G. C. (1999). Stochastic volatility, smile & asymptotics. *Applied Mathematical Finance*, 6:107–145.

Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *The review of financial studies*, 4(4):727–752.

Takahashi, A. (1999). An asymptotic expansion approach to pricing financial contingent claims. *Asia-Pacific Financial Markets*, 6(2):115–151.

Tehranchi, M. R. (2009). Asymptotics of implied volatility far from maturity. *Journal of Applied Probability*, 46(3):629–650.

Varadhan, S. R. S. (1967a). Diffusion processes in a small time interval. *Communications on Pure and Applied Mathematics*, 20(4):659–685.

Varadhan, S. R. S. (1967b). On the behavior of the fundamental solution of the heat equation with variable coefficients. *Communications on Pure and Applied Mathematics*, 20(2):431–455.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188.

Watanabe, S. (1987). Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *The Annals of Probability*, pages 1–39.

Widdicks, M., Duck, P. W., Andricopoulos, A. D., and Newton, D. P. (2005). The Black-Scholes equation revisited: Asymptotic expansions and singular perturbations. *Mathematical Finance*, 15(2):373–391.