IMPLIED VOLATILITY ASYMPTOTICS UNDER AFFINE STOCHASTIC VOLATILITY MODELS

by

Antoine Jacquier

Department of Mathematics
Imperial College London
London SW7 2AZ
United Kingdom

Submitted to Imperial College London
for the degree of
Doctor of Philosophy

2010

Declaration

I the undersigned hereby declare that the work presented in this thesis is my own. When material from other authors has been used, these have been duly acknowledged. This thesis has not previously been presented for this or any other PhD examinations.

Antoine Jacquier

"Start every day off with a smile and get it over with."

W.C. Fields

Abstract

This thesis is concerned with the calibration of affine stochastic volatility models with jumps. This class of models encompasses most models used in practice and captures some of the common features of market data such as jumps and heavy tail distributions of returns. Two questions arise when one wants to calibrate such a model:

(a) How to check its theoretical consistency with the relevant market characteristics?

(b) How to calibrate it rigorously to market data, in particular to the so-called implied volatility, which is a normalised measure of option prices?

These two questions form the backbone of this thesis, since they led to the following idea: instead of calibrating a model using a computer-intensive global optimisation algorithm, it should be more efficient to use a less robust—hence faster—algorithm, but with an accurate starting point. Henceforth deriving closed-form approximation formulae for the implied-volatility should provide a way to obtain such accurate initial points, thus ensuring a faster calibration.

In this thesis we propose such a calibration approach based on the time-asymptotics of affine stochastic volatility models with jumps. Mathematically since this class of models is defined via its Laplace transform, the tools we naturally use are large deviations theory as well as complex saddle-point methods. Large deviations enable us to obtain the limiting behaviour (in small or large time) of the implied volatility, and saddle-point methods are needed to obtain more accurate results on the speed of convergence. We also provide numerical evidence in order to highlight the accuracy of the closed-form approximations thus obtained, and compare them to standard pricing methods based on real calibrated data.

I dedicate this thesis to my grandparents

Acknowledgements

First and foremost I would like to thank my supervisor, Dr Aleksandar Mijatović, for guiding me through this long journey and for teaching me how to write mathematics properly. This thesis would not have been possible without Dr Claude Martini who gave me the opportunity to work as a quantitative consultant during my thesis and who constantly challenged me with new tricky questions.

I would also like to thank my coauthors, Dr M. Forde, Dr J. Gatheral and Dr R. Lee, for their valuable insights. It is a great pleasure for me to thank Professor M. Jeanblanc and all the people in Evry where I spent six months during my PhD. A special thank also goes to Professor R. Brummelhuis for his great help during the early stages of my PhD.

On the professional side I would like to express my gratitude to all the people I have worked with: everyone in Zeliade Systems for teaching me the difference between C++ and C#, Marc and Saad from AXA-IM for introducing me to the world of quantitative finance and Andry, Mitch, Karen and Jessica from Société Générale for teaching me the magic of SAS and of risk management.

I am deeply indebted to my professors who encouraged me to keep studying mathematics: Professors H. Geman, P. Heudron, J.P. Indjehagopian, J.L. Prigent and A. Roncoroni.

It is a great pleasure for me to have Professor Mark H.A. Davis and Professor Huyên Pham as examiners for my thesis and I would like to thank them for accepting this.

A warm thank you goes directly to everyone in Birkbeck College and Imperial College London for making my days flow so smoothly, particularly Bonnie, Chris, Colm, Ed, Haakon, Haluk, Michela, Nika, Pawel, Simon, Sylvie and Vimal.

As much as I would like to pretend, I have not spent days and nights buckled down in my office and I deeply thank all the Guilbs, Guilbettes and Lalalas from London, Paris, Brussels, New-York, Tokyo and Singapore for always asking the question "what's the point of this theorem?" and hence forcing me to actually think about a proper answer.

Finally I want to thank my parents, Matthieu, GG, Thomas, Alexandre and all my family for their support. This work would of course not have seen the light were it for Anne's constant encouragement and presence by my side.

Contents

Introduction

"Prediction is very difficult, especially if it's about the future." This nugget of wisdom by Niels Bohr is obviously not to be taken literally but is a serious warning against the uses and misuses of models. In particular one should not choose a model on the sole ground that past data can be fitted well. One also needs to identify what is to be predicted in the future, i.e. one needs to precisely understand the purpose of the model. In financial modelling most models are proposed as a refinement to the Black-Scholes model which—according to the general opinion—is not a good model. Depending on the data to be fitted authors and practitioners have different views on which model to select. Choosing a model amidst the plethora of literature is not an easy task, especially since each of them has its advantages and drawbacks. The underlying question, which goes slightly beyond Niels Bohr's witty quip, boils down to "What is a good model?". As we said, Black-Scholes is not a good model for many reasons. It is however an extremely useful model for possibly even more reasons, one of them being that it allows for a fast and accurate pricing of (vanilla) options. A model is by construction an approximation of the reality and the more elaborate, the closer to the reality but the less useful. The quality of a model hence arises from a subtle balance between complexity and ease of implementation. The purpose of this thesis is not to discuss at length the pros and cons of financial models but to provide some tools to make models more useful. We indeed aim at describing a calibration methodology based on approximations of models. Among the many models proposed in the literature, affine stochastic volatility models ([30] and [68]) form a wide class of tractable and realistic models for equity and foreign exchange markets, and we shall concentrate on these.

In practice (affine) stochastic volatility models are first calibrated on market data, then used for pricing. Pricing financial products is mathematically tantamount either to solving a PDE problem with boundary conditions (the final payoff of the product) or to calculating the expectation of this payoff using probabilistic tools such as Monte Carlo simulation or stochastic integration. For most models closed-form formulae are scarcely available, and accurate algorithms have been developed and used such as finite-differences [70] and quadrature [2] methods. The calibration step involves a proper selection of the data to be fitted by a model. A common practice is to calibrate the so-called implied volatility rather than option prices directly. The implied volatility is a standardised measure of option prices which makes them comparable even though the underlying assets are not the same. Since this calibration step is based on optimisation algorithms, the lack of a closed-form formula for the implied volatility makes it very time consuming. For instance, the SABR stochastic volatility model has become very popular because a closed-form approximation formula for the implied volatility was derived in [56] and hence made the model easily tractable. Likewise, perturbation methods as developed in [44] have proved to be very useful for obtaining a closed-form approximation formula of option prices. Although these methods only hold under some constraints on the parameters, they provide useful initial reference points for calibration.

In this thesis we provide closed-form approximations for the implied volatility in this large class of models, with a particular emphasis on the Heston model—the canonical continuous-path instance of affine models—since it is one of the most widely used models in financial modelling (see [49], [73]). The idea supporting these results is the following: suppose one wants to calibrate a (stochastic volatility) model to market data. This can be performed in two different ways: (i) one uses a global optimisation algorithm, which involves computing the implied volatility at each observed point until the algorithm converges; (ii) one specifies an initial set of parameters for the model and runs a local optimisation algorithm such as the least-squares method. The latter solution is the most widely used in practice since it is less computer-intensive. However its robustness heavily relies on the initial set of parameters to be specified. Simple closed-form approximations for the model make this choice robust and accurate. One first calibrates the approximation to market data—which is straightforward since this is a closed-form—then one uses this calibrated set of parameters as a starting point in the whole calibration process.

In the first part of the thesis we introduce the main tools and notations we shall need later on. In particular we recall the asymptotic methods on which our analysis is based. These methods are based on a probabilistic, complex-analytic and geometric approach to the same problem, i.e. minimising a distance (in the Riemannian sense) between two points. We also recall the definitions and main properties of affine stochastic volatility models and hint on popular extensions of these as well as calibration techniques in use.

We then move on to the application to these asymptotic methods to affine stochastic volatility models. Part II is concerned with the large-maturity asymptotics of the implied volatility smile. We first start (Chapter 4) with a pure probabilistic approach and express the large-time behaviour of the implied volatility in terms of the rate function determining the large deviations principle for the affine process. Assumptions have to be made concerning the regularity of the process for large times and we shall explain the different possible behaviours. In Chapter 5 we focus on the Heston model and use complex saddlepoint methods to provide more accurate option price and implied volatility asymptotics. We also provide numerical evidence for these closed-form approximations. In Part III we first derive a geometric characterisation of the implied volatility when the maturity is small. We are not able to state this result in full generality and only assume a certain (restrictive) form for the system of stochastic differential equations satisfied by the model. However we believe this simple case (i) sheds a light on a different approach to be pursued further and (ii) bridges the gap between the probabilistic approach based on large deviations principle and the geometric methods already derived in [57] and in [11]. In Chapter 7 we prove a large deviations principle for the extended Heston model (defined in Part I) and use it to derive the asymptotic behaviour of the implied volatility smile as the maturity tends to zero. We also give some indications on the difficulties that arise when one wants to add jumps to the models. Finally Chapter 8 parallels Part II, Chapter 5 and refines the implied volatility expansion in the Heston model when the maturity is small.

Notations

A^o	interior of a set A.
$\overline{A}$	closure of a set A.
$\mathcal{N}$	Cumulative distribution function of the standard Gaussian distribution.
$C_{\mathrm{BS}}(x, \Sigma, t)$	Black-Scholes Call option price with moneyness x, volatility Σ and maturity t.
$C(x, \Sigma, t)$	Call option price under a given model with moneyness x.
$\sigma_t(x)$	Implied volatility with maturity t and moneyness x.
$\hat{\sigma}_t(x)$	Implied volatility with maturity t and moneyness xt.
$\dot{\gamma}$	$\mathrm{d}\gamma/\mathrm{d}t$ for a curve $(\gamma(t))_{t\geq 0}$.
$\Im(z)$	imaginary part of a complex number z.
$\dfrac{f(t)}{g(t)} \sim 1$	$\lim \dfrac{f(t)}{g(t)} = 1$; the context makes it clear at which point the limit is taken.
x_+	$\max\{0, x\}$ for $x \in \mathbb{R}$.
$C^\infty([a, b])$	Space of smooth functions on $[a, b]$.

Part I

Preliminaries on asymptotic methods and affine models

Chapter 1

Review of asymptotic methods

1.1 Large deviations and the Gärtner-Ellis theorem

We provide here a brief review of the key concepts of large deviations for a family of (possibly dependent) random variables $(Z_t)_{t\geq 1}$ taking values in $\mathbb{R}$ and state the Gärtner-Ellis theorem (Theorem 1.1.1). A general reference for all the concepts in this section is [28, Section 2.3]. Let the cumulant generating function $\Lambda_t^Z(u) := \log \mathbb{E}\left(\mathrm{e}^{uZ_t}\right)$ be finite on some neighbourhood of the origin. Assume that for every real number u the following limit exists as an extended real number

$$\Lambda(u) := \lim_{t\to\infty} t^{-1}\Lambda_t^Z(ut)\,. \tag{1.1.1}$$

Let $\mathcal{D}_\Lambda := \{u \in \mathbb{R} \ :\ |\Lambda(u)| < \infty\}$ be the effective domain of Λ and assume that

$$0 \in \mathcal{D}_\Lambda^o, \tag{1.1.2}$$

where $\mathcal{D}_\Lambda^o$ denotes the interior of $\mathcal{D}_\Lambda$ (in $\mathbb{R}$). Since the function Λ_t^Z is convex by Hölder's inequality for every t, the limit Λ is also be convex and the set $\mathcal{D}_\Lambda$ is an interval. Since $\Lambda(0) = 0$, the convexity implies that for any $u \in \mathbb{R}$ we have $\Lambda(u) > -\infty$. The function $\Lambda : \mathbb{R} \to (-\infty, \infty]$ is said to be essentially smooth if (a) it is differentiable in $\mathcal{D}_\Lambda^o$ and (b) it satisfies $\lim_{n\to\infty} |\Lambda'(u_n)| = \infty$ for every sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathcal{D}_\Lambda^o$ that converges to a boundary point of $\mathcal{D}_\Lambda^o$. A cumulant generating function Λ which satisfies (b) is called steep.

The Fenchel-Legendre transform $\Lambda^* : \mathbb{R} \to \mathbb{R}_+$ of Λ is defined by the formula

$$\Lambda^*(x) := \sup\{ux - \Lambda(u) \ :\ u \in \mathbb{R}\}, \quad \text{for all } x \in \mathbb{R}, \tag{1.1.3}$$

with an effective domain $\mathcal{D}_{\Lambda^*} := \{x \in \mathbb{R} \ :\ \Lambda^*(x) < \infty\}$. Under certain assumptions Λ^* is a good rate function. By definition this means that Λ^* is lower semicontinuous (since it is a supremum of continuous functions), $\Lambda^*(\mathbb{R}) \subset [0, \infty]$ (since $\Lambda(0) = 0$) and the level sets $\{x \ :\ \Lambda^*(x) \leq y\}$ are compact for all $y \geq 0$ (see [28, Lemma 2.3.9(a)]). In general a Fenchel-Legendre transform can be discontinuous and $\mathcal{D}_{\Lambda^*}$ can be strictly contained in $\mathbb{R}$ (see [28, Section 2.3] for elementary examples of such rate functions). We say that the family of random variables $(Z_t)_{t\geq 1}$ satisfies the large deviations principle (LDP) with the good rate function Λ^* if for every Borel measurable set B in $\mathbb{R}$ the following inequalities hold

$$-\inf_{x\in B^o}\Lambda^*(x)\leq\liminf_{t\to\infty}\frac{1}{t}\log\mathbb{P}\left(Z_t\in B\right)\leq\limsup_{t\to\infty}\frac{1}{t}\log\mathbb{P}\left(Z_t\in B\right)\leq-\inf_{x\in\overline{B}}\Lambda^*(x), \tag{1.1.4}$$

where the interior B^o and the closure $\overline{B}$ of the set B are taken in the topology of $\mathbb{R}$ and $\inf\emptyset=\infty$. It is clear from definition (1.1.4) that if the family $(Z_t)_{t\geq 1}$ satisfies the LDP and Λ^* is continuous on $\overline{B}$, then $\lim_{t\to\infty}t^{-1}\log\mathbb{P}\left(Z_t\in B\right)=-\inf\left\{\Lambda^*(x)\,:\,x\in B\right\}$. An element $y\in\mathbb{R}$ is an exposed point of Λ^* if there exists $u_y\in\mathbb{R}$ such that

$$yu_y-\Lambda^*(y)>xu_y-\Lambda^*(x),\qquad\text{for all } x\in\mathbb{R}\backslash\{y\}. \tag{1.1.5}$$

Intuitively the exposed points are those at which Λ^* is strictly convex (e.g. the second derivative is continuous and strictly positive). The segments over which Λ^* is affine are not exposed. Note that (1.1.5) can only hold for $y\in\mathcal{D}_\Lambda$ and, if Λ is differentiable in $\mathcal{D}_\Lambda^o$, then u_y is the unique solution of the equation $\Lambda'(u)=y$. We now state the Gärtner-Ellis theorem, the proof of which can be found in [28, Section 2.3].

Theorem 1.1.1. *Let $(Z_t)_{t\geq 1}$ be a family of random variables and assume that the function $\Lambda:\mathbb{R}\to(-\infty,\infty]$ defined by* (1.1.1) *satisfies* (1.1.2)*. Let F be a closed set and G an open set in $\mathbb{R}$. Then the following inequalities hold*

$$\limsup_{t\to\infty}t^{-1}\log\mathbb{P}\left(Z_t\in F\right)\leq-\inf\left\{\Lambda^*(x)\,:\,x\in F\right\}, \tag{1.1.6}$$

$$\liminf_{t\to\infty}t^{-1}\log\mathbb{P}\left(Z_t\in G\right)\geq-\inf\left\{\Lambda^*(x)\,:\,x\in G\cap\mathcal{E}\right\}, \tag{1.1.7}$$

where $\mathcal{E}:=\{y\in\mathbb{R}\,:\,y$ satisfies (1.1.5) *with $u_y\in\mathcal{D}_\Lambda^o\}$. Furthermore if Λ is essentially smooth and lower semicontinuous, then the LDP holds for $(Z_t)_{t\geq 1}$ with the good rate function Λ^*.*

We have so far considered large deviations using the sole knowledge of the Laplace transform of a process. It is also possible to consider large deviations for the paths of the process itself. Let n be an integer. We shall from now on denote $\mathcal{C}_0\left([0,1]\right)$ the space of continuous functions mapping $[0,1]$ to $\mathbb{R}^n$ such that the origin is a fixed point, and equipped with the supremum norm. We also denote H the space of absolutely continuous functions with square integrable derivatives. The basic result in this direction is the following Schilder's theorem for the paths of a standard Brownian motion.

Theorem 1.1.2. *(Schilder's theorem, see [28, Theorem 5.2.3])*

Let $(W_t)_{t\in[0,1]}$ denote a standard Brownian motion on $\mathbb{R}^n$ and let $\varepsilon>0$. Define the process

$(W_t^\varepsilon)_{t\in[0,1]}$ *by* $W_t^\varepsilon := \sqrt{\varepsilon}W_t$ *for all* $t \in [0,1]$ *and let* ν_ε *denote the induced probability measure. Then the family* (ν_ε) *satisfies in* $\mathcal{C}_0([0,1])$ *a LDP with good rate function* I_W *characterised by*

$$I_W(\psi) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 \left|\dot{\psi}(t)\right|^2 \mathrm{d}t, & \text{if } \psi \in H, \\ \infty, & \text{otherwise.} \end{cases}$$

This theorem has been generalised to form the basis of the Freidlin-Wentzell theory which states a large deviations principle for a diffusion process satisfying the following SDE on $\mathbb{R}^n$

$$\mathrm{d}X_t^\varepsilon = b(X_t^\varepsilon)\,\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)\,\mathrm{d}W_t, \quad \text{for all } t \in [0,1], \quad \text{and} \quad X_0^\varepsilon = x \in \mathbb{R}^n,$$

where the function b is uniformly Lipschitz and all the coefficients of the diffusion matrix σ are also uniformly Lipschitz. Using the contraction principle (Theorem 1.1.4 below), Freidlin and Wentzell [45] proved the following.

Theorem 1.1.3. *(see [28, Theorem 5.6.7])*
If all the coefficients of σ *are bounded then the family* (X_t^ε) *satisfies a LDP in* $\mathcal{C}_0([0,1])$ *with the good rate function* I_x *defined by*

$$I_x(f) := \inf_{\left\{g\in H: f(t)=x+\int_0^t b(f(s))\mathrm{d}s+\int_0^t \sigma(f(s))\dot{g}(s)\mathrm{d}s\right\}} \frac{1}{2}\int_0^1 |\dot{g}(t)|\,\mathrm{d}t,$$

where the infimum over an empty set is taken to be infinity.

Theorem 1.1.4. *[28, Theorem 4.2.1] Contraction principle.*
Let $f : \mathbb{R} \to \mathbb{R}$ *be a continuous function. Consider a good rate function* $\mathcal{I} : \mathbb{R} \to [0,\infty]$*. For each* $y \in \mathbb{R}$ *define* $\mathcal{I}'(y) := \inf\{\mathcal{I}(x) : x \in \mathbb{R}, y = f(x)\}$*. Then* $\mathcal{I}'$ *is a good rate function on* $\mathbb{R}$ *. If* $\mathcal{I}$ *controls the LDP associated with a family of probability measures* (μ_ε) *on* $\mathbb{R}$*, then* $\mathcal{I}'$ *controls the LDP associated with a family of probability measures* $\left(\mu_\varepsilon \circ f^{-1}\right)$ *on* $\mathbb{R}$*.*

1.2 The complex analytic approach: saddlepoint methods

The large deviations approach we have presented above allowed us to obtain the exponential rate of decay of some probabilities. However it does not tell us anything about higher order expansions. Sharp large deviations theory is a probabilistic approach to tackle this issue (see [9] and [10] for instance). Saddlepoint methods in the complex plane are another powerful one in this direction. Before presenting the method, let us first briefly see how this comes into play in our analysis. Recall that we are interested in the tail probabilities $\mathbb{P}(X > a)$ for some large real number a, where X represents a random variable whose distribution F is only determined by its cumulative generating function Φ defined by $\log\mathbb{E}(\exp(zX)) = \Phi(z)$ for z in some region of the complex plane. We assume that the function Φ is analytic in some strip containing the origin. By Fourier inversion the following equalities hold:

$$\mathbb{P}(X > a) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}} \mathrm{e}^{-\varepsilon(x-a)} \mathbb{1}_{\{x>a\}} F(\mathrm{d}x) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}} \frac{\exp(-\mathtt{i}za)}{\varepsilon + \mathtt{i}z} \mathrm{e}^{\Phi(\mathtt{i}z)} \frac{\mathrm{d}z}{2\pi}.$$

We can not let ε tend to zero directly in the integral on the right-hand side since the integrand has a pole at the origin. However Cauchy's theorem ensures that for any α in the strip of analyticity of Φ we can write

$$\lim_{\varepsilon \to 0} \int_{-\mathtt{i}\infty}^{\mathtt{i}\infty} \frac{\exp(\Phi(z) - za)}{\varepsilon + z} \frac{\mathrm{d}z}{2\mathtt{i}\pi} = \lim_{\varepsilon \to 0} \int_{\alpha-\mathtt{i}\infty}^{\alpha+\mathtt{i}\infty} \frac{\exp(\Phi(z) - za)}{\varepsilon + z} \frac{\mathrm{d}z}{2\mathtt{i}\pi} = \int_{\alpha-\mathtt{i}\infty}^{\alpha+\mathtt{i}\infty} \mathrm{e}^{\Phi(z)-za} \frac{\mathrm{d}z}{2\mathtt{i}\pi z},$$

and saddlepoint theory gives us a way to choose the level α optimally. The role of this very brief preface was to introduce the link between the tail probabilities we are interested in and the analytic approach proposed by saddlepoint theory. More precisely we shall be interested here in determining the behaviour of integrals of the form

$$I := \int_{\Gamma} \psi(z) \mathrm{e}^{\lambda\phi(z)} \mathrm{d}z, \tag{1.2.1}$$

as the real parameter λ tends to infinity, where Γ is some contour in the complex plane, and ϕ and ψ are analytic functions on Γ (in our case λ shall represent the time). Define the functions $u : \Gamma \to \mathbb{R}$ and $v : \Gamma \to \mathbb{R}$ by $u(z) := \Re(\phi(z))$ and $v(z) := \Im(\phi(z))$, for all $z \in \Gamma$. The inequality

$$|I| \leq \int_a^b |\psi(z)| \mathrm{e}^{\lambda u(z)} \mathrm{d}z$$

trivially holds where a and b are the two (possibly infinite) endpoints of the contour Γ. It is clear that the main contribution of this integral as λ tends to infinity must come from the point where the function u attains its maximum, and for convenience we assume it is unique. Using Cauchy's theorem we can now deform the contour Γ such that it passes through this very point. Furthermore we deform the contour Γ such that the function u decreases as rapidly as possible along it, so that the integral I is concentrated around the value of the integrand at the point z_0. This path is called the path of steepest descent. With this in mind, we can apply Laplace's method as described in the following proposition.

Proposition 1.2.1. *(see Appendix A, Proposition 2.1 in [98])*
Let $a < b$ be two real numbers and f and g two real functions twice continuously differentiable on $[a, b]$. Assume that the function f is strictly convex on $[a, b]$ and that there exists $x_0 \in [a, b]$ such that $f'(x_0) = 0$ then the following equality holds for large t,

$$\int_a^b \mathrm{e}^{-tf(x)} g(x) \mathrm{d}x = \mathrm{e}^{-tf(x_0)} \left(\sqrt{\frac{2\pi}{tf''(x_0)}} g(x_0) + \mathcal{O}\left(t^{-1}\right) \right).$$

However in our case the complex part $\mathrm{e}^{\mathtt{i}v(z)}$ in the integrand of I in (1.2.1) adds an oscillatory term which increases as λ grows to infinity. When choosing our path of steepest descent we want to make sure to choose it in such a way that the function v is constant along it. Let us denote

$z := x + \mathtt{i}y$ and let $z_0 := x_0 + \mathtt{i}y_0$ be the solution of the equation $\nabla u|_{z_0} = 0$. From the analyticity of ϕ the Cauchy-Riemann equations lead $\partial_x u = \partial_y v$ and $\partial_y u = -\partial_x v$. Hence the point z_0 also satisfies $\nabla v|_{z_0} = 0$. However by the maximum modulus principle the two functions u and v can not have a maximum at the same time. z_0 is thus the so-called saddlepoint of ϕ and we assume it is of the first order, i.e. $\phi''(z_0) \neq 0$. From the Cauchy-Riemann equations the orthogonality property $\nabla u.\nabla v = 0$ is also satisfied and hence the new path is such that the function v is constant along it and corresponds precisely to the path where u varies the most rapidly (note that there are two paths: one along which u decreases, one along which u increases, so one must choose the correct one).
In order to formalise the discussion above let us precisely define the notions of saddlepoint and path of steepest descent.

Definition 1.2.2. (see [14]) Let $F : \mathcal{Z} \to \mathbb{C}$ be an analytic complex function on an open set $\mathcal{Z}$. A point $z_0 \in \mathcal{Z}$ such that the complex derivative $\mathrm{d}F/\mathrm{d}z$ vanishes is called a saddlepoint.

Definition 1.2.3. (see [98]) Let $z := x + \mathtt{i}y,\ x, y \in \mathbb{R}$ and $F : \mathbb{C} \to \mathbb{C}$ be an analytic complex function. The steepest descent contour $\gamma : \mathbb{R} \to \mathbb{C}$ is a map such that

(i) $\Re(F)$ has a minimum at some point $z_0 \in \gamma$ and $\Re(F''(z)) > 0$ along γ.

(ii) $\Im(F)$ is constant along γ.

These two conditions imply that $F'(z_0) = 0$.

1.3 A rough guide to differential geometry

There is a deep link between large deviations theory and differential geometry. In Theorem 1.1.3 we have seen that a large deviations principle could be stated for some diffusion processes, where the rate function corresponds to some metric. In Part III, Chapter 6 we will provide more precise results in this direction, and hence we briefly recall some notions of differential geometry. These geometric tools will allow us to precisely characterise the rate function of the large deviations principle above as a metric on a manifold. Many reference books exist on this topic and we shall mainly follow the monograph [63] by Jost. In this section M will always denote a Riemannian manifold of dimension $n \in \mathbb{N}$ and we shall use Einstein's convention, i.e. two repeated indices are summed: $v_i \frac{\partial}{\partial x^i} := \sum_{i=1}^{n} v_i \frac{\partial}{\partial x^i}$. Let us start with the definition of a smooth manifold.

Definition 1.3.1. A smooth manifold of dimension n is a locally compact Hausdorff space M equipped with an open cover $\{U_\alpha\}$ and a collection of smooth injective maps $\{x_\alpha\}$ with $x_\alpha : U_\alpha \to \mathbb{R}^n$ such that whenever $U_\alpha \cap U_\beta \neq \emptyset$, the following transition map is smooth:

$$x_\beta \circ x_\alpha^{-1} : x_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n \mapsto x_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n.$$

The collection $\{U_\alpha, x_\alpha\}$ is called a chart and x_α a map. The properties of the map transfer to the properties of the manifold, i.e. a manifold is differentiable if all its maps are. It is clear that a point $p \in U_\alpha$ is uniquely identified by its coordinates $x_\alpha(p)$, and we might use both notations. The main purpose of this rough section is to introduce the notion of a distance on a manifold. In order to do so, let us first define the derivation map and the tangent spaces.

Definition 1.3.2. Let M be smooth n-dimensional manifold and let $x_0 \in M$. A derivation D at x_0 is a linear map from $C^\infty(M)$ to $\mathbb{R}^n$ satisfying the Leibniz rule, i.e. for all functions f and g in $C^\infty(M)$ the equality $D(f.g)(x_0) = D(f)\,g(x_0) + f(x_0)\,D(g)$ holds. The tangent space of M at x_0, denoted $T_{x_0}M$, is the vector space of all derivations at x_0. If $x^1, \dots, x^n$ denote the local coordinates of x_0 on the chart U, then $\left(\frac{\partial}{\partial x^i}\right)_{i=1,\dots,n}$ forms a basis of $T_{x_0}M$.

Definition 1.3.3. The cotangent vector space of a smooth manifold M at a point $x_0 \in M$, denoted $T^*_{x_0}M$, is the dual vector space of $T_{x_0}M$, i.e. the vector space of linear forms acting on $T_{x_0}M$. The elements of $T^*_{x_0}M$ are called one-forms and we can define $(dx^i)_{i=1,\dots,n}$ as its basis, with $dx^j\left(\frac{\partial}{\partial x^i}\right) = \delta^j_i$, for all $i, j = 1, \dots, n$, where δ^j_i is the Kronecker symbol.

Definition 1.3.4. A Riemannian metric on a differentiable manifold M is given by a scalar product on each tangent space T_xM which depends smoothly on $x \in M$. A Riemannian manifold is a differentiable manifold equipped with a Riemannian metric.

In local coordinates $x := (x^1, \dots, x^n)$, a metric is a $n \times n$ positive definite symmetric matrix $(g_{ij}(x))_{1\le i,j\le n}$, such that for all $p \in M$, we can write

$$\langle v, w\rangle = g_{ij}(x(p))\, v^i w^j, \qquad \text{for all } \left(v = v^i \frac{\partial}{\partial x^i}, w = w^j \frac{\partial}{\partial x^j}\right) \in T_pM \times T_pM,$$

where $x(p)$ represents the vector of local coordinates of p. The next theorem will enable us to define a metric in the context of stochastic volatility models:

Theorem 1.3.5. *(Theorem 1.4.1 in [63])*
Every differentiable manifold may be equipped with a Riemannian metric.

From the definition of a metric we can now introduce the notion of a distance on a Riemannian manifold M.

Definition 1.3.6. Let $\gamma : [a, b] \subset \mathbb{R} \to M$ be a smooth (C^∞) curve, we can define

(i) the length of γ as $L(\gamma) := \int_a^b \left\|\frac{\mathrm{d}\gamma}{\mathrm{d}t}(t)\right\| \mathrm{d}t = \int_a^b \left(g_{ij}(x(\gamma(t)))\dot{x}^i(t)\dot{x}^j(t)\right)^{1/2} \mathrm{d}t$;

(ii) the energy of γ as $E(\gamma) := \frac{1}{2}\int_a^b \left\|\frac{\mathrm{d}\gamma}{\mathrm{d}t}(t)\right\|^2 \mathrm{d}t = \frac{1}{2}\int_a^b g_{ij}(x(\gamma(t)))\dot{x}^i(t)\dot{x}^j(t)\,\mathrm{d}t$,

where we define $\dot{x}^i(t) := \frac{\mathrm{d}}{\mathrm{d}t}x^i(\gamma(t))$. The distance d is then defined as

$$d(p, q) := \inf\Big\{L(\gamma) : \gamma : [a, b] \to M \text{ is a piecewise smooth curve such that } \gamma(a) = p, \gamma(b) = q\Big\},$$

for any points p and q in the manifold M.

Equipped with this structure, we can now define a geodesic which will be the the central object of our study.

Definition 1.3.7. A smooth curve $\gamma : [a,b] \to M$ which satisfies the system of second-order ODE

$$x(t) + \Gamma^i_{jk}(x(t))\, \dot{x}^j(t)\, \dot{x}^k(t) = 0, \quad \text{for all } i = 1, \dots n, \tag{1.3.1}$$

where $\Gamma^i_{jk} := \frac{1}{2} g^{il}(g_{jl,k} + g_{kl,j} - g_{jk,l})$ with $g^{ij} := g_{ij}^{-1}$ and $g_{jl,k} := \frac{\partial}{\partial x^k} g_{j,l}$ for all $1 \leq i,j,k \leq n$, is called a geodesic. The coefficients Γ^i_{jk} are the Christoffel symbols and the system of ODEs (1.3.1) is called the Euler-Lagrange equations.

It is easy to see that the length of a curve does not depend on its parameterisation. In particular, minimising the length of a curve is tantamount to minimising another curve parameterised by arc length (i.e. $\langle \dot{x}, \dot{x} \rangle$ is constant). Furthermore, for curves parameterised by arc length, minimising the length is equivalent to minimising the energy function, so that the connection between this definition of a geodesic and the path of shortest length is made clear. The definition of a geodesic makes complete sense as regards to the following corollary (see [63, Lemma 1.4.5]):

Corollary 1.3.8. *On any compact manifold, any two points can be connected by at least one curve of shortest length, and this curve is called a geodesic.*

We now have almost all the tools we shall need in our study. We have defined above a geodesic as the solution of a system of ODEs. In [29, Chapter 9], Do Carmo characterises it as the solution to a variational problem. We do not quote all the results of this approach but we just recall the definition of a variation of a curve as a family of neighbouring curves. The reason for introducing such a concept is that it will make it possible to characterise a geodesic curve in terms of the properties of the energy functions of its variations.

Definition 1.3.9. [29, Definition 2.1]
Let $\gamma : [0,1] \to M$ be a piecewise differentiable curve in the manifold M. A (proper) variation of γ is a continuous mapping $f : (-\varepsilon, \varepsilon) \times [0,a] \to M$ such that

(i) $f(0,t) = \gamma(t)$ for all $t \in [0,1]$;

(ii) there exists a subdivision $0 = t_0 < t_1 < \ldots < t_{n+1} = 1$ of $[0,1]$ such that the restriction of f to each $(-\varepsilon, \varepsilon) \times [t_i, t_{i+1}]$ for $i = 0, \dots, n$ is differentiable.

Chapter 2

Black-Scholes and implied volatility

Throughout this thesis we shall for convenience consider that interest rates are null so that the forward price is the same as the share price and we will denote the corresponding process $S = (S_t)_{t\geq 0}$ and its logarithm $X = (X_t)_{t\geq 0}$ defined as $X_t := \log(S_t)$ for all $t \geq 0$. The initial value S_0 will always be assumed to be a strictly positive real number. The Black-Scholes model [13] assumes that the share price process satisfies the stochastic differential equation $\mathrm{d}S_t = \Sigma S_t \,\mathrm{d}W_t$, where $(W_t)_{t\geq 0}$ is a standard Brownian motion and $\Sigma > 0$ represents the volatility of the returns. Consider a European call option at time 0 written on the share price S with strike $K > 0$, which, under the risk-neutral measure, pays at maturity $t \geq 0$ the amount $\mathbb{E}\left(S_t - K\right)_+$. We shall always consider here that the risk-neutral measure, i.e. the measure under which the share price is a true martingale, is given a priori. For convenience we shall—unless otherwise stated—consider the log-moneyness $x := \log(K/S_0) \in \mathbb{R}$ and express the quantities of interest accordingly. The Black-Scholes formula tells us that the price $C_{\mathrm{BS}}(x, \Sigma, t)$ of such an option is worth

$$C_{\mathrm{BS}}(x, \Sigma, t) = S_0 \mathcal{N}\left(-\frac{x}{\Sigma\sqrt{t}} + \frac{\Sigma\sqrt{t}}{2}\right) - K\mathcal{N}\left(-\frac{x}{\Sigma\sqrt{t}} - \frac{\Sigma\sqrt{t}}{2}\right),$$

where $\mathcal{N}$ represents the cumulative distribution function of the standard Gaussian distribution. Let $C_{\mathrm{obs}}(x, t)$ represent an observed—or given—price of a European call option with the same maturity and log-moneyness. The implied volatility is defined as the unique value $\Sigma > 0$ such that the equality

$$C_{\mathrm{obs}}(x, t) = C_{\mathrm{BS}}(x, \Sigma, t)$$

holds. It is easy to see that for any $x \in \mathbb{R}$ and $t > 0$, the function $\Sigma \mapsto C_{\mathrm{BS}}(x, \Sigma, t)$ is strictly increasing and hence the implied volatility is well defined. It is also clear that the implied volatility varies in terms of the maturity t and of the log-moneyness x. We shall hence refer to the implied volatility smile at maturity t when regarding the implied volatility as a function of x for some fixed t. Similarly, the term structure of the implied volatility will refer to the implied volatility as a function of the maturity t for some fixed x. Finally the implied volatility surface denotes the three-dimensional graph of the implied volatility as a function of both the maturity and the log-moneyness. We shall henceforth use the notation $\sigma_t(x)$ to denote the implied volatility corresponding to a European call option maturing at time t with log-moneyness x. Since we have just introduced the Gaussian cumulative distribution function, we might as well state the following trivial lemma which we shall use many times all along this thesis.

Lemma 2.0.10. *Let $n : z \in \mathbb{R} \mapsto (2\pi)^{-1/2} \exp\left(-z^2/2\right)$, and $\mathcal{N}$ denote respectively the density and the cumulative distribution function of the Gaussian distribution, then the following bounds hold for all $x > 0$:*

$$(x + 1/x)^{-1} n(x) \leq 1 - \mathcal{N}(x) \leq n(x)/x.$$

Chapter 3

Affine stochastic volatility models

3.1 Affine stochastic volatility models with jumps

In [30], Duffie, Filipović and Schachermayer introduced affine processes—a general class of realistic tractable processes—and laid the mathematical grounds to study them. Within this class, two-dimensional affine stochastic volatility processes (with jumps) encompass some of the most used models in practice, in particular the Heston [59], the Bates [5] and the Barndorff-Nielsen & Shephard [4] models. More specifically, let $(S_t)_{t\geq 0}$ represent the share price process, denote $X_t := \log(S_t)$ and let $(V_t)_{t\geq 0}$ be another process starting at $V_0 = v > 0$. We do not make specific assumptions on the filtration and it will be sufficient to consider the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by the pair $(X_t, V_t)_{t\geq 0}$. As in [68], we make the following assumptions:

A1. The pair $(X_t, V_t)_{t\geq 0}$ is a stochastically continuous, time-homogeneous Markov process.

A2. There exist two functions ϕ and ψ such that the logarithmic characteristic function Φ_t satisfies

$$\Phi_t(u, w) := \log \mathbb{E}\left(\exp\left(u\left(X_t - x_0\right) + wV_t\right)\Big| x_0, v\right) = \phi\left(t, u, w\right) + \psi\left(t, u, w\right) v, \tag{3.1.1}$$

for all $t, u, w \in \mathbb{R}_+ \times \mathbb{C}^2$, where the expectation exists.

A3. The (discounted) share price process $(S_t)_{t\geq 0}$ is a martingale.

Let us now define the functions $F : \mathbb{C}^2 \to \mathbb{C}$ and $R : \mathbb{C}^2 \to \mathbb{C}$ by $F(u, w) := \partial_t\phi(t, u, w)|_{t=0^+}$ and $R(u, w) := \partial_t\psi(t, u, w)|_{t=0^+}$, whenever $\phi\left(t, u, w\right)$ and $\psi\left(t, u, w\right)$ are defined. In [68], Keller-Ressel proved the existence (and domains) of F and R and showed that they satisfy a system of Riccati equations. Furthermore from the general theory of affine processes [30], we know that they must be of Lévy-Khintchine form:

$$F(u, w) = \left\langle \frac{a}{2}\left(u, w\right)' + b, \left(u, w\right)'\right\rangle + \int_{D\setminus\{0\}} \left(\mathrm{e}^{xu+yw} - 1 - \left\langle \omega_F(x, y), (u, w)'\right\rangle\right) m\left(\mathrm{d}x, \mathrm{d}y\right), \tag{3.1.2}$$

$$R(u,w) = \left\langle \frac{\alpha}{2}(u,w)' + \beta, (u,w)' \right\rangle + \int_{D\setminus\{0\}} \left(\mathrm{e}^{xu+yw} - 1 - \left\langle \omega_R(x,y), (u,w)' \right\rangle \right) \mu(\mathrm{d}x, \mathrm{d}y), \quad (3.1.3)$$

where $D := \mathbb{R}\times\mathbb{R}_+$, $\omega_F(x,y) := \left(x/\left(1+x^2\right), 0\right)'$ and $\omega_R(x,y) := \left(x/\left(1+x^2\right), y/\left(1+y^2\right)\right)'$ are truncation functions, a and α are two positive semi-definite matrices with $a_{12} = a_{21} = a_{22} = 0$, $b \in D$, $\beta \in \mathbb{R}^2$, m and μ are two Lévy measures on D and $\int_{D\setminus\{0\}} \left((x^2+y)\wedge 1\right) m(\mathrm{d}x, \mathrm{d}y) < \infty$. Let us make one more assumption:

A4. The function $u \mapsto R(u,w)$ is not constant.

Following the terminology in [68], a process satisfying Assumptions A1-A4 is called an affine stochastic volatility model and is fully characterised by the affine form of its characteristic function (3.1.1). For instance, take $F(u,w) := \kappa\theta w + \lambda(u)$ and $R(u,w) := u(u-1)/2 + \xi^2 w^2/2 - \kappa w + uw\rho\xi$, then one recovers the Heston model with jumps (see Section 3.2 below) for the dynamics of the logarithmic share price $\mathrm{d}X_t = (\delta - V_t/2)\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_t + \mathrm{d}J_t$, where the stochastic variance process follows $\mathrm{d}V_t = \kappa(\theta - V_t)\mathrm{d}t + \xi\sqrt{V_t}\,\mathrm{d}Z_t$, where W and Z are two Brownian motions with correlation ρ, J is an independent pure-jump Lévy process and the function λ is the compensated cumulant generating function of the jump part: $\lambda(u) = \int_{\mathbb{R}} \left(\mathrm{e}^{xu} - 1 - u(\mathrm{e}^x - 1)\right) m(\mathrm{d}x)$. In the Barndorff-Nielsen & Shephard case [4], consider $F(u,w) := \gamma\kappa(\rho u + w) - \gamma\kappa(\rho)$ and $R(u,w) := u(u-1)/2 - \gamma w$, then one obtains the corresponding dynamics for the logarithmic share price $\mathrm{d}X_t = (\delta - V_t/2)\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t + \rho\mathrm{d}J_{\gamma t}$ with stochastic variance dynamics $\mathrm{d}V_t = -\gamma V_t + \mathrm{d}J_{\gamma t}$, where the process $(J_t)_{t\geq 0}$ is a Lévy subordinator.

In this thesis we shall place great emphasis on the continuous-path version of these affine stochastic volatility models. Assume that the two Lévy measures satisfy $\mu \equiv 0$ and $m \equiv 0$, then the process $(X_t)_{t\geq 0}$ has continuous sample paths and it is easy to see that it satisfies the SDE

$$\begin{aligned}
\mathrm{d}X_t &= -\tfrac{1}{2}(a + V_t)\,\mathrm{d}t + \rho\sqrt{V_t}\,\mathrm{d}W_t + \sqrt{a + (1-\rho^2)V_t}\,\mathrm{d}Z_t, \qquad && X_0 = x \in \mathbb{R}, \\
\mathrm{d}V_t &= (b + \beta V_t)\,\mathrm{d}t + \sqrt{\alpha V_t}\,\mathrm{d}W_t, && V_0 = v \in (0,\infty),
\end{aligned} \quad (3.1.4)$$

where the admissible parameter values are given by

$$a, b \geq 0, \quad \alpha > 0, \quad \beta \in \mathbb{R} \quad \text{and} \quad \rho \in [-1,1], \quad (3.1.5)$$

and where the two Brownian motions $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are independent. The process $(V_t)_{t\geq 0}$ is a square-root diffusion process and the Yamada-Watanabe conditions [67, Section 5.2 C] ensure that a non-negative strong solution exists. It is clear that the process $\left(\mathrm{e}^{X_t}\right)_{t\geq 0}$ is a local martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, and [68, Theorem 2.5] implies that it is in fact a true martingale. The Heston model, defined in [59] with mean-reversion rate κ, positive long-time variance level θ, volatility of volatility σ and correlation ρ, falls in the class of models given by (3.1.4) (take $a = 0$, $b = \kappa\theta > 0$, $\beta = -\kappa < 0$, $\alpha = \sigma^2$; the correlation parameter ρ has the same role as in (3.1.4)).

Remark 3.1.1.

(i) This class is a continuous-path generalisation of the Heston model, both in terms of number of parameters (in Heston, $a = 0$) and in terms of range of parameters.

(ii) The general form of the instantaneous variance of a continuous affine stochastic volatility model is given by $a + \tilde{\alpha}V$ for some $\tilde{\alpha} > 0$. A simple scaling of the process V in (3.1.4) maps the class of models given by (3.1.4) to the general case. Without loss of generality we therefore assume $\tilde{\alpha} = 1$.

(iii) The process $U = (U_t)_{t\geq 0}$ defined by $U_t := a + V_t$ for all $t \geq 0$ follows the shifted square-root dynamics (see [75] for applications of the shifted square-root process in pricing theory).

3.2 The Heston model

The Heston model is the canonical example of a continuous affine stochastic volatility model defined by (3.1.4) and we shall lay particular emphasis on it in this thesis. We therefore recall here some of its well-known (and perhaps lesser known) properties.

3.2.1 The CIR process and Heston characteristic function

Under the risk-neutral measure the logarithmic share price process X satisfies the SDE

$$\begin{aligned} \mathrm{d}X_t &= -\tfrac{1}{2}V_t\,\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_t, && X_0 = x_0 \in \mathbb{R}, \\ \mathrm{d}V_t &= \kappa\,(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}Z_t, && V_0 = v > 0, \\ \mathrm{d}\,\langle W, Z\rangle_t &= \rho\,\mathrm{d}t, \end{aligned} \tag{3.2.1}$$

where W and Z are two standard Brownian motions, $\kappa > 0$ and $\theta > 0$ are respectively the mean-reversion speed and the long-term variance of the variance process, $\sigma > 0$ accounts for the volatility of volatility, v is assumed to be non random and $\rho \in [-1, 1]$ is the correlation parameter between the two driving Brownian motions.

Notation 3.2.1. For ease of notation we will from now on denote $S_t \sim \mathcal{H}(S_0, v, \kappa, \theta, \sigma, \rho, t)$ the price of an asset following the Heston dynamics (3.2.1), with a slight abuse of notation between the share price process S and its logarithm X.

The approach we shall develop in this thesis is fundamentally based on the knowledge of the Laplace transform of the process. We therefore briefly recall it here and we indicate a few references related to the pricing of European Vanilla options under the Heston model. In the literature several different—obviously equivalent—expressions exist for its characteristic function. The original formulation [59] is slightly different from the one we use here. The technical issue in the expression in [59] is that one has to take care of possible branch cuts in the complex plane. In [65] Lord &

Kahl have provided ways to avoid this difficulty. We adopt here the terminology of [2] since it bypasses the need for a step-by-step check of branch cuts. The characteristic function $\phi_t : \mathbb{R} \to \mathbb{C}$ under the Heston model takes the following form:

$$\phi_t(k) := \mathbb{E}\Big(\exp\Big(\mathtt{i}k(X_t - X_0)\Big)\Big) = \exp\Big(C(t,k) + D(t,k)\,v\Big), \tag{3.2.2}$$

where

$$C(t,k) := \frac{\kappa\theta}{\sigma^2}\left(\Big(\kappa - \rho\sigma\mathtt{i}k - d(k)\Big)t - 2\log\left(\frac{1 - g(k)\exp(-d(k)t)}{1 - g(k)}\right)\right), \tag{3.2.3}$$

$$D(t,k) := \frac{v}{\sigma^2}\Big(\kappa - \rho\sigma\mathtt{i}k - d(k)\Big)\frac{1 - \exp(-d(k)t)}{1 - g(k)\exp(-d(k)t)}, \tag{3.2.4}$$

$$g(k) := \frac{\kappa - \rho\sigma\mathtt{i}k - d(k)}{\kappa - \rho\sigma\mathtt{i}k + d(k)}, \qquad \text{and} \qquad d(k) := \sqrt{(\kappa - \rho\sigma\mathtt{i}k)^2 + \sigma^2(\mathtt{i}k + k^2)}, \tag{3.2.5}$$

where we take the principal branch for the complex logarithm. For the square root here either of the two roots may be chosen since the characteristic function is even in d. To switch from the characteristic function to the Laplace transform of the process, let us recall the following theorem.

Theorem 3.2.2. *(Lukacs [77, Theorem 7.1.1]). If a characteristic function ψ is regular in the neighborhood of $k = 0$, then it is also regular in a horizontal strip and can be represented in this strip by a Fourier integral. This strip is either the whole plane, or it has one or two horizontal boundary lines. The purely imaginary points on the boundary of the strip of regularity (if this strip is not the whole plane) are singular points of ψ.*

Recall that a characteristic function ψ is said to be regular if it is analytic and single-valued. The analyticity of the characteristic function ϕ_t is henceforth equivalent to the existence of a strictly positive real number ε such that the (normalised) logarithmic Laplace transform Λ_t defined by $\Lambda_t(p) := \log\mathbb{E}\left(\mathrm{e}^{p(X_t - x_0)}\right)$ for all $t \geq 0$ is finite for all $-\varepsilon < p < \varepsilon$. The strip of convergence of ϕ_t then corresponds to the effective domain of Λ_t. Theorem 3.2.2 together with the representation (3.2.2) lead to

$$\Lambda_t(p) = \begin{cases} C(t, -\mathtt{i}p) + D(t, -\mathtt{i}p)\,v, & \text{for all } t < t^*(p), \\ +\infty, & \text{for all } t \geq t^*(p), \end{cases} \tag{3.2.6}$$

where the explosion time function $t^* : p \in \mathbb{R} \mapsto \sup\{t \geq 0 : \Lambda_t(p) < \infty\}$ reads (see [3])

$$t^*(p) = \begin{cases} \infty, & \text{if } \delta(p) \geq 0 \text{ and } \zeta(p) < 0, \\ d(-\mathtt{i}p)^{-1}\log\left(\dfrac{\zeta(p) + d(-\mathtt{i}p)}{\zeta(p) - d(-\mathtt{i}p)}\right), & \text{if } \delta(p) \geq 0 \text{ and } \zeta(p) > 0, \\ \dfrac{2}{|d(-\mathtt{i}p)|}\left(\pi\mathbb{1}_{\{\zeta(p)<0\}} + \arctan\left(\dfrac{2|d(-\mathtt{i}p)|}{\zeta(p)\,\sigma^2}\right)\right), & \text{if } \delta(p) < 0, \end{cases} \tag{3.2.7}$$

where $\zeta(p) := (\rho\sigma p - \kappa)/\sigma^2$ and where we define $\delta(p) := (\kappa - \rho\sigma p)^2 + \sigma^2 p(1-p)$ for convenience so that $d(-\mathtt{i}p) = \sqrt{\delta(p)}$. We refer the interested reader to [3] for a complete characterisation of explosion times for more general continuous stochastic volatility models. In this thesis we will actually be interested in the other way around, i.e. for a given maturity t what is the strip of analyticity of the characteristic function ϕ_t? Since this will involve some more detailed calculations, we postpone its analysis.

3.2.2 Properties of the Heston model

We recall here some fundamental properties of the Heston model characterised by the system of SDEs (3.2.1) which we shall need later on. Even if stochastic models driven by pure Brownian terms like (3.2.1) are always local martingales, a size issue could prevent them from being fully qualified martingales, as shown by Jourdain [64]. In the case of the Heston model, this could happen when the stochasticity of the volatility makes the underlying get bigger, in case of strictly positive correlation. However everything turns out to be fine for the first moment as the following proposition [64] stipulates.

Proposition 3.2.3. *If $S_t \sim \mathcal{H}(S_0, v, \kappa, \theta, \sigma, \rho, t)$ then it is a true martingale.*

A clear self-contained proof of the true martingality of the Heston model has been provided recently by Del Baño Rollin [27] and Keller-Ressel [68] proved this under some conditions for general continuous affine stochastic volatility models as we mentioned before. The following lemma conveys a lot of information on the influence of the Heston parameters. If S_t is a Heston process, then the inverse process S_t^{-1} also follows a Heston dynamic up to a related change of measure. Let us define the Share measure as the measure with S_T/S_0 as the numeraire, then the following holds.

Lemma 3.2.4. *If $S_t \sim \mathcal{H}(S_0, v, \kappa, \theta, \sigma, \rho, t)$, then under the Share measure the inverse process S_t^{-1} is of the type $\mathcal{H}\left(S_0^{-1}, v, \bar{\kappa}, \bar{\theta}, \sigma, \bar{\rho}, t\right)$, where $\bar{\kappa} := \kappa - \sigma\rho$, $\bar{\theta} := \kappa\theta/\bar{\kappa}$. and $\bar{\rho} := -\rho$.*

Proof. In [26], Del Baño Rollin proves it using the characteristic function of Heston. We provide here a probabilistic proof. Consider (3.2.1) and define $U_t := S_t^{-1}$, then Itô's lemma implies $\mathrm{d}U_t = U_t\sqrt{V_t}\left(\sqrt{V_t}\,\mathrm{d}t - \mathrm{d}W_t\right)$. Define now a new probability measure $\mathbb{Q}$ defined by

$$\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} = \eta_t := \exp\left(\int_0^t \sqrt{V_s}\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \sqrt{V_s}\,\mathrm{d}s\right) = S_t/S_0.$$

Since S is a true martingale Girsanov theorem ensures that $\mathrm{d}U_t = U_t\sqrt{V_t}\,\mathrm{d}\overline{W}_t$ holds under $\mathbb{Q}$ with

$$\mathrm{d}V_t = \kappa\left(\theta - V_t\right)\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}Z_t + \eta_t^{-1}\mathrm{d}\left\langle \eta, V\right\rangle_t = \bar{\kappa}\left(\bar{\theta} - V_t\right)\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}Z_t,$$

and $\mathrm{d}\left\langle \overline{W}, Z\right\rangle_t = \bar{\rho}\,\mathrm{d}t$. This implies that the pair $(U_t, V_t)_{t\geq 0}$ satisfies under the measure $\mathbb{Q}$:

$$\begin{aligned}
\mathrm{d}U_t &= U_t\sqrt{V_t}\,\mathrm{d}\overline{W}_t \\
\mathrm{d}V_t &= \bar{\kappa}\left(\bar{\theta} - V_t\right)\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}Z_t \\
\mathrm{d}\left\langle \overline{W}, Z\right\rangle_t &= \bar{\rho}\,\mathrm{d}t.
\end{aligned}$$

□

Obviously this Heston symmetry is an involution on the Heston dynamics such that $\kappa - \rho\sigma > 0$, with a single fixed point which is the uncorrelated Heston model ($\rho = 0$). This symmetry property allows us to express the right wing of the Heston smile in terms of the left wing of the symmetrised model (and vice-versa). This is useful in practice since many authors provide results on smile asymptotics only on one side of the smile. A last consequence is that the at-the-money volatility structure (a.k.a the implied volatility term structure) is left invariant by the Heston symmetry. When $\rho = -1$ the Heston model is somehow degenerate in the following sense:

Proposition 3.2.5. *Assume $\rho = -1$. Define $S^*_{\bar{t}} := S_0 \exp(v_0 + \psi\bar{t})$ for all $\bar{t} := \sigma t \geq 0$ then the inequality $S_{\bar{t}} < S^*_{\bar{t}}$ holds almost surely for all $t \geq 0$. Furthermore the implied volatility is strictly decreasing below $S^*_{\bar{t}}$ and equal to zero above $S^*_{\bar{t}}$.*

Proof. This property is obtained by Fourier transform in [48]. We give here a more probabilistic proof, based on the system of SDEs (3.2.1), where we consider $\rho = -1$:

$$\begin{aligned} \mathrm{d}S_t &= S_t \sqrt{V_t}\, \mathrm{d}W_t, \\ \mathrm{d}V_t &= \kappa(\theta - V_t)\, \mathrm{d}t - \sigma\sqrt{V_t}\, \mathrm{d}W_t. \end{aligned}$$

Using Itô's formula, we have

$$\begin{aligned} S_t &= S_0 \exp\left(-\int_0^t \frac{V_s}{2}\mathrm{d}s + \int_0^t \sqrt{V_s}\,\mathrm{d}W_s\right) = S_0 \exp\left(-\int_0^t \frac{V_s}{2}\mathrm{d}s\right) \exp\left\{\int_0^t \left(\frac{\kappa}{\sigma}(\theta - V_s)\,\mathrm{d}s - \frac{\mathrm{d}V_s}{\sigma}\right)\right\} \\ &= S_0 \exp\left(-\int_0^t \frac{V_s}{2}\mathrm{d}s\right) \exp\left\{\frac{\kappa\theta}{\sigma}t + \frac{v}{\sigma} - \frac{V_t}{\sigma} - \frac{\kappa}{\sigma}\int_0^t V_s \mathrm{d}s\right\}, \quad \text{for all } t \geq 0. \end{aligned}$$

Since the variance process V is always positive almost surely (by the Feller condition), the inequality $S_t \leq S_0 \exp(v/\sigma + \kappa\theta t/\sigma)$ holds for all $t \geq 0$ almost surely. □

This anti-correlated Heston is sometimes named the Heston-Nandi model, by reference to its discrete-time version studied in [60]. By symmetry the case $\rho = 1$ has analogous properties.

3.3 Extensions of the Heston model

The Heston model has proved to be rather efficient to calibrate equity derivatives. However several drawbacks have been pointed out, either from a theoretical of from a practical point of view (see [49] for more details):

- The Heston model is rather accurate when used to calibrate an implied volatility surface, but turns out to be insufficient for short maturities as the observed skew is too steep.
- It does not provide an efficient way to perform a joint calibration vanilla and volatility derivatives (such as variance swaps) or vanilla and Foreign-Exchange options (see [17]).

To overcome these shortcomings, several approaches have been proposed such as (i) adding jumps either to the share price or to the variance process or both ([95], [69]), (ii) increasing the dimension of the process (see [15] and Section 3.3.1 below). We present hereafter a special type of multidimensional Heston process, for which our large deviations methodology in Chapter 4 will apply in a straightforward way.

3.3.1 The multidimensional Heston model

There are two types of multidimensional Heston in the literature, each corresponding to a special approach: Buehler [15] proposed the double-Heston as a Heston model where the long-term variance is also a Feller diffusion process. This is consistent with his variance swap curve approach, where each dimension of this model corresponds to a factor of the variance curve. However there is no closed or semi-closed form formula for the pricing of vanilla options in this framework. The other approach ([66], [25]) introduces several variance processes, independent one from another, each driven by a Brownian motion acting on a separate time scale. This approach is appealing since

· the characteristic function is available in closed-form (hence allowing a semi-closed form formula for vanilla options similar to the standard Heston model);

· the induced correlation between the variance process and the returns is stochastic, therefore allowing a better fit to the so-called stochastic skew observed on FX markets;

Several authors have proposed an extension of classical stochastic diffusion / jump / Lévy processes to take into account the so-called stochastic skew of the FX market, i.e. the fact that the slope of the implied volatility varies over time. The traditional approach consists of specifying some stochastic dynamics for the correlation between the share price and the volatility (for instance as a transformed Jacobi process, see [66]). We propose here (see also [66], [22]) a multidimensional Heston model for the logarithmic forward price process $X = (X_t)_{t\geq 0}$, in the following sense:

Definition 3.3.1. We define a Heston model of dimension $n \in \mathbb{N}$, or simply a multidimensional Heston model by a process $(X_t)_{t\geq 0}$ satisfying the following system of SDEs,

$$\begin{aligned}
\mathrm{d}X_t &= -\frac{1}{2}\sum_{i=1}^{n} V_t^i \mathrm{d}t + \sum_{i=1}^{n} \sqrt{V_t^i}\,\mathrm{d}W_t^i, && X_0 \in \mathbb{R},\\
\mathrm{d}V_t^i &= \kappa_i\left(\theta_i - V_t^i\right)\,\mathrm{d}t + \sigma_i\sqrt{V_t^i}\,\mathrm{d}Z_t^i, && V_0^i = v^i > 0, \text{ for all } i = 1, \dots n,\\
\mathrm{d}\left\langle W^i, W^j\right\rangle_t &= \rho_i \delta_{ij}\mathrm{d}t, \quad \text{and} \quad \mathrm{d}\left\langle Z^i, Z^j\right\rangle_t = \delta_{ij}\mathrm{d}t, && \text{for all } i, j = 1, \dots, n,
\end{aligned}$$

where δ_{ij} is the Kronecker delta and $\kappa^i > 0$, $\theta^i > 0$, $\sigma^i > 0$, $\left|\rho^i\right| \leq 1$ are satisfied for all $i = 1, \dots, n$.

Remark 3.3.2. In [22] Carr and Wu developed an estimation of a more general model, namely a double Lévy process for the process $(X_t)_{t\geq 0}$:

$$X_t = (r_d - r_f)\,t + \left(L^R_{T^R_t} - \xi^R T^R_t\right) + \left(L^L_{T^L_t} - \xi^L T^L_t\right), \quad \text{for all } t \geq 0,$$

where r_d, r_f denote the domestic and foreign risk-free interest rates, L_t^R, L_t^R, two Lévy processes exhibiting right and left skewness, T_t^R, T_t^L, two stochastic time changes, and ξ^R, ξ^L two functions ensuring that the processes $\left(L^R_{T_t^R} - \xi^R T_t^R\right)_{t\geq 0}$ and $\left(L^L_{T_t^L} - \xi^L T_t^L\right)_{t\geq 0}$ are both martingales. We refer to [22] for more details about the properties of this model.

Proposition 3.3.3. *The multidimensional Heston model can be written as a one-dimensional stochastic volatility model with stochastic correlation.*

Proof. Let $n \in \mathbb{N}$. Define the process $(\xi_t)_{t\geq 0}$ by $\xi_t := \sum_{i=1}^n V_t^i$ for all $t \geq 0$. We then have

$$\begin{aligned}
\mathrm{d}\xi_t = \sum_{i=1}^n \mathrm{d}V_t^i &= \sum_{i=1}^n \kappa_i \left(\theta_i - V_t^i\right) \mathrm{d}t + \sum_{i=1}^n \sigma_i \sqrt{V_t^i}\, \mathrm{d}Z_t^i = \sum_{i=1}^n \kappa_i \left(\theta_i - V_t^i\right) \mathrm{d}t + \left(\sum_{i=1}^n \sigma_i^2 V_t^i\right)^{1/2} \mathrm{d}Z_t \\
&= \sum_{i=1}^n \kappa_i \left(\theta_i - V_t^i\right) \mathrm{d}t + \sqrt{\xi_t} \left(\sum_{i=1}^n \sigma_i^2 V_t^i \xi_t^{-1}\right)^{1/2} \mathrm{d}Z_t,
\end{aligned}$$

where $\mathrm{d}Z_t := \left(\sum_{i=1}^n \sigma_i^2 V_t^i\right)^{-1/2} \sum_{i=1}^n \sigma_i \sqrt{V_t^i} \mathrm{d}Z_t^i$. The process Z is a standard Brownian motion by the Lévy characterisation theorem (see [97, Theorem 4.6.4]) since the process ξ never reaches zero almost surely by the Feller condition on every variance process V^i. We can also rewrite the SDE for the process X defined in Definition 3.3.1 as

$$\mathrm{d}X_t = -\frac{1}{2} \sum_{i=1}^n V_t^i \mathrm{d}t + \sum_{i=1}^n \sqrt{V_t^i}\, \mathrm{d}W_t^i = -\frac{\xi_t}{2} \mathrm{d}t + \sqrt{\xi_t} \mathrm{d}W_t,$$

where we define $W_t := \xi_t^{-1/2} \sum_{i=1}^n \sqrt{V_t^i}\, \mathrm{d}W_t^i$ for all $t \geq 0$. The process $(W_t)_{t\geq 0}$ is again well defined as a standard Brownian motion by the Lévy characterisation theorem. Finally the correlation between W and Z reads

$$\begin{aligned}
\mathrm{d}\left\langle W, Z\right\rangle_t &= \xi_t^{-1/2} \left(\sum_{i=1}^n \sigma_i^2 V_t^i\right)^{-1/2} \sum_{i=1}^n \sqrt{V_t^i} \mathrm{d}W_t^i \sum_{i=1}^n \sigma_i \sqrt{V_t^i} \mathrm{d}Z_t^i \\
&= \xi_t^{-3/2} \left(\sum_{i=1}^n \sigma_i^2 V_t^i\right)^{-1/2} \sum_{i=1}^n \sigma_i \rho_i V_t^i \mathrm{d}t
\end{aligned}$$

□

Note that this multidimensional Heston model is also an alternative to multiscale volatility models as each volatility can act on a different time scale. Although very little is known about asymptotic results for higher (greater than 2) dimensional stochastic volatility models, this particular extension is quite convenient since all the results based on the characteristic / Laplace function are readily applicable to this model. Indeed it is straightforward to see that since the variance processes are independent of each other, the logarithmic Laplace transform of the multidimensional process X reads

$$\Lambda_t(u) := \log \mathbb{E}\left(\mathrm{e}^{uX_t}\right) = \sum_{i=1}^n \Lambda_t^{(i)}(u), \quad \text{for all } u \in \mathcal{D}_{\Lambda_t}, \tag{3.3.1}$$

where $\Lambda_t^{(i)}$ is the logarithmic Laplace transform of a Heston process $\mathcal{H}\left(S_0, v^i, \kappa^i, \theta^i, \rho^i, \sigma^i, t\right)$, and the effective domain $\mathcal{D}_{\Lambda_t}$ is precisely equal to $\bigcap_{i=1}^{n} \mathcal{D}_{\Lambda_t^{(i)}}$.

3.3.2 Time-dependent Heston: the Asian option approach

In order to obtain a better fit to market data than the Heston model (for instance allowing a joint calibration equity and volatility derivatives), some authors have advocated time-dependent parameters. Consider a share price process $S_t \sim \mathcal{H}(S_0, v, r, \kappa_t, \theta_t, \sigma_t, \rho_t, t)$. Buehler [15] showed that in order to be able to fit a variance swap curve consistently, κ_t and the product $\rho_t\sigma_t$ have to be constant in time. In [41], the authors proposed a method to calibrate a Heston model with time-dependent parameters (all but κ) to a given variance swap curve. In terms of pricing Benhamou et al. [8] proposed an analytical approximation for European vanilla options using Malliavin calculus in a small volatility of volatility regime, which can be seen as an extension of the Lewis asymptotics to the time-dependent case, whereas Elices [35] calibrates a time-dependent Heston model extending the original Fourier-transform approach to the time-dependent case. For a share price process $(S_t)_{t\geq 0}$ which is a positive continuous martingale under the pricing measure, we can write

$$S_t = S_0 \exp\left(\beta_t - \frac{1}{2}\langle \beta, \beta\rangle_t\right), \quad \text{for all } t \geq 0,$$

where $\beta_t := \int_0^t \sqrt{V_s}\,\mathrm{d}W_s$, for some Brownian motion W and a predictable square integrable volatility process $\left(\sqrt{V_t}\right)_{t\geq 0}$ (see [15]). For the process S the fair price of a variance swap at inception is $\mathbb{E}_0\left(\int_0^t V_s \mathrm{d}s\right)$. The variance swap can be replicated with logarithmic contracts on S, which can in turn be replicated with a static portfolio of European call and put options of the same maturity t (see [79] and Appendix A of [15] for details). We follow here the approach developed in [41].

Let us consider a Heston model as above where the mean-reversion level θ_t and volatility-of-variance σ_t depend on time. We would like to be consistent with an observed or pre-specified variance swap curve and a pre-specified term structure for the second moment of the integrated variance. For all $t \geq 0$ we define $I_t := \int_0^t V_s \mathrm{d}s$ the realised variance of the process.

Theorem 3.3.4. *For the share price process $S_t \sim \mathcal{H}(S_0, v, r, \kappa_t, \theta_t, \sigma_t, \rho_t, t)$ the equalities*

$$\partial_t \mathbb{E}(I_t^2) = 2\,\mathbb{E}(I_t V_t), \tag{3.3.2}$$

$$\partial_t \mathbb{E}(I_t V_t) = \mathbb{E}\left(I_t(\kappa(\theta_t - V_t)\right) + \mathbb{E}(V_t^2), \tag{3.3.3}$$

$$\partial_t \mathbb{E}(V_t^2) = 2\mathbb{E}\left(V_t \kappa(\theta_t - V_t)\right) + \sigma_t^2 \mathbb{E}(V_t), \tag{3.3.4}$$

hold for all $t \geq 0$. Moreover, we can make this model consistent with a pre-specified variance swap curve $\mathbb{E}\left(\int_0^t V_s \mathrm{d}s\right)$, and a pre-specified term structure for the second moment of the integrated *variance* $\mathbb{E}\left(\left(\int_0^t V_s \mathrm{d}s\right)^2\right)$ *by choosing* v, θ_t *and* σ_t *as follows*

$$v = \partial_t \mathbb{E}\left(\int_0^t V_s \mathrm{d}s\right)\bigg|_{t=0}, \quad \theta_t = \frac{\partial_t \mathbb{E}(V_t)}{\kappa} + \mathbb{E}(V_t), \quad \sigma_t^2 = \frac{\partial_t \mathbb{E}\left(V_t^2\right) - 2\kappa \mathbb{E}\left(V_t(\theta_t - V_t)\right)}{\mathbb{E}(V_t)}, \tag{3.3.5}$$

if $0 < \theta_{\min} \leq \theta_t \leq \theta_{\max} < \infty$, $0 < \sigma_{\min} \leq \sigma_t \leq \sigma_{\max} < \infty$ *for all* $t \geq 0$, *where* $\mathbb{E}(V_t^2) = \partial_t \mathbb{E}(I_t V_t) - \mathbb{E}\left(I_t(\kappa(\theta_t - V_t)\right)$ *and* $\mathbb{E}(I_t V_t) = \partial_t \mathbb{E}\left(I_t^2\right)/2$.

Proof. Let us fix $t \geq 0$. We know that $\mathbb{E}(V_s)$ is finite for all $s \in [0,t]$. Therefore $\mathbb{E}\left(\int_0^t \sigma_s^2 V_s \mathrm{d}s\right)$ is finite as well and the stochastic integral $\int_0^t \sigma_s \sqrt{V_s}\, \mathrm{d}B_s$ has zero expectation. Using Fubini's theorem and the fundamental theorem of calculus we obtain the forward equation

$$\partial_t \mathbb{E}(V_t) = \kappa\Big(\theta_t - \mathbb{E}(V_t)\Big). \tag{3.3.6}$$

We can then rearrange this equation to recover θ_t in (3.3.5) (note that σ_t does not affect $\mathbb{E}(V_t)$). We now proceed along similar lines to Dufresne [31]: Itô's lemma implies the equality $\mathrm{d}\left(I_t^2\right) = 2 I_t V_t \mathrm{d}t$ so that Fubini's theorem and the Schwarz inequality lead

$$\begin{aligned}\mathbb{E}\left(I_t^2\right) = 2\mathbb{E}\left(\int_0^t I_s V_s \mathrm{d}s\right) &= 2\int_0^t \mathbb{E}(I_s V_s)\, \mathrm{d}s = 2\int_0^t \int_0^s \mathbb{E}\left(V_s V_u\right) \mathrm{d}u\, \mathrm{d}s \\ &\leq 2\int_0^t \int_0^s \left(\mathbb{E}\left(V_s^2\right)\mathbb{E}\left(V_u^2\right)\right)^{1/2} \mathrm{d}u\, \mathrm{d}s < \infty.\end{aligned}$$

The fundamental theorem of calculus implies $\partial_t \mathbb{E}\left(I_t^2\right) = 2\mathbb{E}(I_t V_t)$. Likewise we compute $\mathbb{E}(I_t V_t)$:

$$I_t V_t = \int_0^t \left(I_s\, \mathrm{d}V_s + V_s\, \mathrm{d}I_s\right) = \int_0^t I_s\Big(\kappa\left(\theta_s - V_s\right)\mathrm{d}s + \sigma_s \sqrt{V_s}\, \mathrm{d}B_s\Big) + V_s^2\, \mathrm{d}s. \tag{3.3.7}$$

By the Schwarz inequality we have $\mathbb{E}\left(I_t^2 V_t\right) \leq \left(\mathbb{E}\left(I_t^4\right)\mathbb{E}\left(V_t^2\right)\right)^{1/2}$, and Jensen's inequality leads

$$t^{-4}\mathbb{E}\left(I_t^4\right) = \mathbb{E}\left[\left(t^{-1}\int_0^t V_s\, \mathrm{d}s\right)^4\right] \leq \mathbb{E}\left(t^{-1}\int_0^t V_s^4\, \mathrm{d}s\right) = t^{-1}\int_0^t \mathbb{E}\left(V_s^4\right)\mathrm{d}s,$$

and we know that $\mathbb{E}(V_s^4) < \infty$ for all $s \in [0,t]$. Therefore the three expectations $\mathbb{E}\left(I_t^4\right)$, $\mathbb{E}\left(V_t^2\right)$ and $\mathbb{E}\left(I_t^2 V_t\right)$ are finite, which in turns implies that the expectation $\mathbb{E}\left(\int_0^t \sigma_s^2 I_s^2 V_s\, \mathrm{d}s\right)$ is finite as well. We hence have sufficient integrability conditions for the stochastic integral $\int_0^t \sigma_s I_s \sqrt{V}_s \mathrm{d}B_s$ in (3.3.7) to have zero expectation, and hence the forward equation $\partial_t \mathbb{E}(I_t V_t) = \mathbb{E}\left(I_t\left(\kappa(\theta_t - V_t)\right) + \mathbb{E}(V_t^2)\right)$ follows. Repeating the procedure again for V_t^2, we find

$$V_t^2 = v^2 + \int_0^t 2V_s\Big(\kappa(\theta_s - V_s)\, \mathrm{d}s + \sigma_s \sqrt{V_s}\, \mathrm{d}B_s\Big) + \int_0^t \sigma_s^2 V_s\, \mathrm{d}s.$$

Using the fact that the third moment of the process V is finite we finally obtain the ODE $\partial_t \mathbb{E}(V_t^2) = 2\mathbb{E}(V_t \kappa(\theta_t - V_t)) + \sigma_t^2 \mathbb{E}(V_t)$, which concludes the proof of the theorem. □

Much like Dupire's [32] forward equation for call options under a local volatility model, the equations in Theorem 3.3.4 have some theoretical appeal but are difficult to implement in practice since higher order derivatives need to be computed. In the following we address this issue by deriving Taylor series expansions for the expectation and the variance of the integrated variance. We also provide a bootstrapping method to enhance the calibration.

Proposition 3.3.5. *Let $(\mathcal{F}_u)_{u\geq 0}$ denote the filtration generated by I. For all $0 \leq s \leq t$, we have*

$$\mathbb{E}(I_t/\mathcal{F}_s) = \mathbb{E}(I_s) + \mathbb{E}(V_s)(t-s) + \frac{\kappa}{2}\left(\theta_s - \mathbb{E}(V_s)\right)(t-s)^2 + \mathcal{O}\left((t-s)^3\right), \tag{3.3.8}$$

and

$$\begin{aligned}\mathbb{V}(I_t/\mathcal{F}_s) &= \mathbb{V}(I_s) + 2\Big(\mathbb{E}(I_sV_s) - \mathbb{E}(V_s)\mathbb{E}(I_s)\Big)(t-s) \\ &\quad + \Big\{\mathbb{E}(V_s^2) - \mathbb{E}^2(V_s) - \kappa\Big(\mathbb{E}(I_sV_s) - \mathbb{E}(I_s)\mathbb{E}(V_s)\Big)\Big\}(t-s)^2 + \mathcal{O}\left((t-s)^3\right).\end{aligned} \tag{3.3.9}$$

Corollary 3.3.6. *The following expansions hold for t close enough to zero:*

$$\begin{aligned}\mathbb{E}\left(t^{-1}I_t\right) &= v - \frac{1}{2}\kappa(v-\theta_0)t + \frac{\kappa}{6}\Big(\theta_0' - \kappa(\theta_0 - v)\Big)t^2 + \mathcal{O}\left(t^3\right), \\ \mathbb{V}\left(t^{-1}I_t\right) &= \frac{\sigma_0^2}{3}vt + \left(-\kappa v\sigma_0 + \frac{\kappa}{4}\theta_0\sigma_0 + \frac{\sigma_0'}{2}v\right)\frac{\sigma_0}{3}t^2 + \mathcal{O}\left(t^3\right),\end{aligned}$$

where θ_0' denote the derivative of the map $t \mapsto \theta_t$ evaluated at $t = 0$, and likewise for σ_0'.

Proof. Let $0 \leq s \leq t$. We can then solve the ODE in (3.3.6), expand the solution as a Taylor series around $t = s$, integrate and then square it and we obtain

$$\mathbb{E}(I_t/\mathcal{F}_s)^2 = \mathbb{E}(I_s)^2 + 2\mathbb{E}(I_s)\mathbb{E}(V_s)(t-s) + \Big(\kappa\theta_s\mathbb{E}(I_s) + \mathbb{E}^2(V_s) - \kappa\mathbb{E}(V_s)\mathbb{E}(I_s)\Big)(t-s)^2 + \mathcal{O}\left((t-s)^3\right).$$

We also have

$$\begin{aligned}\mathbb{E}\left(V_t^2/\mathcal{F}_s\right) &= \mathbb{E}(V_s^2) + \Big(\mathbb{E}(V_s)\sigma_s^2 + 2\kappa\theta_s\mathbb{E}(V_s) - 2\kappa\mathbb{E}\left(V_s^2\right)\Big)(t-s) \\ &\quad + \left\{\kappa\theta_s\left(\kappa\theta_s + \frac{\sigma_s^2}{2}\right) + \left(\sigma_s\sigma_s' - \frac{3\kappa}{2}\sigma_s^2 + \kappa\theta_s' - 3\kappa^2\theta_s\right)\mathbb{E}(V_s) + 2\kappa^2\mathbb{E}(V_s^2)\right\}(t-s)^2 \\ &\quad + \mathcal{O}\left((t-s)^3\right),\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left(I_tV_t/\mathcal{F}_s\right) &= \mathbb{E}(I_sV_s) + \Big(\kappa\theta_s\mathbb{E}(I_s) - \kappa\mathbb{E}(I_sV_s) + \mathbb{E}(V_s^2)\Big)(t-s) \\ &\quad + \frac{1}{2}\Big(\kappa^2(\mathbb{E}(I_sV_s) - \theta_s\mathbb{E}(I_s)) + 3\kappa(\mathbb{E}(V_s)\theta_s - \mathbb{E}(V_s^2)) + \sigma_s^2\mathbb{E}(V_s) + \kappa\theta_s'\mathbb{E}(I_s)\Big)(t-s)^2 \\ &\quad + \mathcal{O}\left((t-s)^3\right).\end{aligned}$$

Thus we obtain

$$\begin{aligned}\mathbb{E}\left(I_t^2/\mathcal{F}_s\right) &= 2\int_0^t \mathbb{E}(I_uV_u/\mathcal{F}_s)\mathrm{d}u = 2\int_0^s \mathbb{E}(I_uV_u)\mathrm{d}u + 2\int_s^t \mathbb{E}(I_uV_u/\mathcal{F}_s)\mathrm{d}u \\ &= \mathbb{E}(I_s^2) + 2\mathbb{E}(I_sV_s)(t-s) + \Big(\kappa\theta_s\mathbb{E}(I_s) - \kappa\mathbb{E}(I_sV_s) + \mathbb{E}(V_s^2)\Big)(t-s)^2 + \mathcal{O}((t-s)^3),\end{aligned}$$

and

$$\begin{aligned}\mathbb{V}\left(\int_0^t V_u\mathrm{d}u/\mathcal{F}_s\right) &= \mathbb{E}(I_s^2) - \mathbb{E}^2(I_s) + 2\left\{\mathbb{E}(I_sV_s) - \mathbb{E}(V_s)\mathbb{E}(I_s)\right\}(t-s) \\ &\quad + \left\{\mathbb{E}(V_s^2) - \mathbb{E}^2(V_s) - \kappa\left[\mathbb{E}(I_sV_s) - \mathbb{E}(I_s)\mathbb{E}(V_s)\right]\right\}(t-s)^2 + \mathcal{O}((t-s)^3),\end{aligned}$$

which proves the proposition. Since $\mathbb{E}(I_0) = 0$ and $\mathbb{E}(V_0) = v$, the corollary follows immediately. □

The following procedure outlines how one can use Corollary 3.3.6 and Proposition 3.3.5 to calibrate a time-dependent Heston model with piecewise linear θ_t and σ_t:

(1) Choose a value for κ.

(2) Fit the variance swap rate curve $\mathbb{E}\left(I_t/t\right)$ from $t_0 = 0$ to $t_1 > 0$ with a quadratic, and the $\mathbb{V}\left(I_t/t\right)$ curve with a quadratic, whose leading order term is linear. Using Corollary 3.3.6, we can then back out v, θ_0, θ_0', σ_0 and σ_0'.

(3) Using the values for θ_0, θ_0', σ_0 and σ_0', construct affine functions for θ_t and σ_t over $[0, t_1]$.

(4) Solve (3.3.2), (3.3.3) and (3.3.4) analytically over $[0, t_1]$, so as to compute $\mathbb{E}(v_{t_1}^2), \mathbb{E}(I_{t_1} V_{t_1})$, $\mathbb{E}(I_{t_1}^2)$, and hence $\mathbb{V}(I_{t_1}^2)$.

(5) From pre-specified values for $\mathbb{E}(I_{t_2})$ and $\mathbb{V}(I_{t_2})$ with $t_2 > t_1$, back out θ_{t_1}' and σ_{t_1}' using (3.3.8) and (3.3.9) respectively with $s = t$ and $t = t_2$. We then use these derivatives to construct affine functions for θ_t and σ_t over $[t_1, t_2]$, so that θ_t and σ_t are piecewise linear (and continuous) over $[0, t_2]$.

(6) Repeat steps (1)-(5).

Part II

Large-maturity implied volatility asymptotics

Introduction

As we mentioned in the general introduction, the goal of this thesis is to develop a set of closed-form approximations for the implied volatility in order (i) to have a better understanding of models and (ii) to obtain more accurate and more stable calibration procedures. In this part we focus on the large-time asymptotics of affine stochastic volatility models. Our main result is the following large-time closed-form formula for the implied volatility $\hat{\sigma}_t(x)$ corresponding to a European option with maturity t and maturity-dependent strike $S_0 \exp(xt)$:

$$\hat{\sigma}_t^2(x) = \hat{\sigma}_\infty^2(x) + t^{-1}\frac{8\hat{\sigma}_\infty^4(x)}{4x^2 - \hat{\sigma}_\infty^4(x)} \log\left(\frac{A(x)}{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), 0)}\right) + o\left(t^{-1}\right)$$

for any $x \in \mathbb{R}$, as t tends to infinity, where the function $\hat{\sigma}_\infty^2$ is the genuine limit of the implied volatility smile. The limiting behaviour $\hat{\sigma}_\infty$ is first proved for a large class of models in Proposition 4.1.3 and the expansion above including the correction term is derived for the Heston model in Theorem 5.2.1. The functions $\hat{\sigma}_\infty^2$, A_{BS} and A are all available in closed form. For a constant strike $S_0 \exp(x)$, the above formula simplifies to

$$\sigma_t^2(x) = 8\Lambda^*(0) + 4t^{-1}\left((2u^*(0) - 1)\,x - 2\log\left(-A(0)\sqrt{2\Lambda^*(0)}\right)\right) + o\left(t^{-1}\right)$$

as t tends to infinity, where the function Λ^* is the Fenchel-Legendre transform of the limiting Laplace transform of the model under consideration, and u^* is the corresponding saddlepoint. This formula is proved for the Heston model in Theorem 5.2.3. It is well-known [86] that for a fixed strike, the implied volatility flattens as the maturity increases; this is confirmed by the formula for σ_t above, the zeroth order term of which was already known for the Heston model (see [73]). However, the maturity-dependent strike formulation $\hat{\sigma}_t$ above reveals that the implied volatility smile does not flatten but rather spreads out in a very specific way as the maturity increases. In the fixed-strike case, Lewis [73] pioneered the research on large-time asymptotics of implied volatility in stochastic volatility models by studying the first eigenvalue and eigenfunction of the generator of the underlying stochastic process. Recently Tehranchi [101] studied the large-time behaviour of the implied volatility when the share price is a non-negative local martingale and obtained analogues of the maturity-independent strike formula above in that setting.

In Chapter 4 we consider the general class of affine stochastic volatility models and use large deviations techniques to derive the asymptotic behaviour of option prices and the limit $\hat{\sigma}_\infty$ of the implied volatility as the maturity tends to infinity. We are not concerned here with the correction term for $\hat{\sigma}_t$. Section 4.1 assumes that the scaled process $(X_t/t)_{t\geq 1}$ satisfies a large deviations principle and derives the implied volatility behaviour for large maturities. In Section 4.2 we focus specifically on continuous affine stochastic volatility models. This is justified by the fact that the continuous case is extremely informative and points out precisely where the difficulties arise. In this section we prove a full large deviations principle for the Heston model and explains why this does not hold for general continuous affine stochastic volatility models for which a more refined analysis is needed. Finally, in Section 4.3 we extend these results to the multidimensional Heston model and to affine stochastic volatility models with jumps, for which more assumptions are needed. Incidentally we shall present a simple non-affine case—the Schöbel-Zhu model—for which calculations follow straightforwardly from the standard Heston case.

Motivated by these results, we push the analysis further in order to study precisely both the rate of convergence and higher-order terms. Large deviations theory acts on a logarithmic scale and a different tool is needed to carry out this program. Chapter 5 focuses on the Heston model and fully proves the expansion for $\hat{\sigma}_t$ above by means of complex saddlepoint methods and contour integration. The idea of studying the behaviour of the call price function as an inverse Fourier transform has already been applied by several authors ([19], [20], [52], [1] and [87]) in order to speed up the computation of option pricing algorithms based on inverse Fourier transforms. In the Heston model we are able to obtain the saddlepoint $u^*(x)$ in closed form, thus avoiding any numerical approximations, and we test our asymptotic formulae on calibrated data.

Chapter 4

The large-maturity smile: a large deviations approach

4.1 From large deviations to implied volatility asymptotics

In this section we consider a stochastic process $(X_t)_{t\geq 0}$ such that $\left(\mathrm{e}^{X_t}\right)_{t\geq 0}$ is a true martingale with respect to a reference filtration and we assume that the scaled process $(X_t/t)_{t\geq 1}$ satisfies a large deviations principle. From these two assumptions only we derive in Section 4.1.1 the behaviours of call and put options as the maturity tends to infinity. As a corollary, we obtain an equivalent formulation for call and put options under the Black-Scholes model. These behaviours are intermediate results which we use to determine the asymptotic behaviour of the implied volatility.

4.1.1 The rate function and the asymptotics of option prices

Before stating the main theorem that makes this relationship precise, let us define a new probability measure $\widetilde{\mathbb{P}}$—the *Share measure*—where the normalised share price process $\left(\mathrm{e}^{X_t}\right)_{t\geq 0}$ is the new numeraire, i.e. for any event A in the filtration generated by $(X_t)_{t\geq 0}$, we have $\widetilde{\mathbb{P}}(A) := \mathbb{E}\left(\exp\left(X_t\right)\mathbb{1}_A\right)$. $\widetilde{\mathbb{P}}$ is a well-defined probability measure since by assumption the process $\left(\mathrm{e}^{X_t}\right)_{t\geq 0}$ is a martingale. Let Λ defined in (1.1.1) be the limiting moment generating function of the process $(X_t/t)_{t\geq 1}$ and Λ^* defined in (1.1.3) its Fenchel-Legendre transform. It is straightforward to see that the cumulant generating functions and consequently the Fenchel-Legendre transforms of X under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are related by

$$\widetilde{\Lambda}(u) = \Lambda(u+1), \quad \text{if } (1+u)\in\mathcal{D}_\Lambda, \quad \text{and} \quad \widetilde{\Lambda}^*(x) = \Lambda^*(x) - x, \quad \text{for all } x\in\mathbb{R}, \tag{4.1.1}$$

where $\mathcal{D}_\Lambda$ is the effective domain of Λ. The first equality is obivous and the second follows from

$$\widetilde{\Lambda}^*(x) = \sup_{u\in\mathcal{D}_{\widetilde{\Lambda}}}\left\{ux - \widetilde{\Lambda}(u)\right\} = \sup_{u\in\mathcal{D}_{\widetilde{\Lambda}}}\left\{(u+1)x - \Lambda(u+1) - x\right\} = \sup_{u\in\mathcal{D}_\Lambda}\left\{ux - \Lambda(u)\right\} - x = \Lambda^*(x) - x.$$

The following theorem is fundamental and establishes the link between large deviations principles for the random process X and option price asymptotics. We shall say that a random variable satisfies a *full* large deviations principle under $\mathbb{P}$ if it satisfies a LDP as defined in Part I, Section 1.1 and if its limiting cumulant generating function Λ is essentially smooth with the origin lying in the interior of its domain. Note that this implies that Λ^* is a good continuous rate function. Furthermore since $\mathcal{D}_\Lambda$ always contains the compact interval $[0,1]$ and $\Lambda(0)=\Lambda(1)=0$, then the convexity of Λ clearly implies that Λ^* attains its minimum at $\Lambda'(0)$. Likewise a full LDP under $\widetilde{\mathbb{P}}$ implies that $\widetilde{\Lambda}^*$ attains its minimum at $\Lambda'(1)$. The full LDP statement is a rather strong assumption and we shall see in Section 4.2 how to relax part of it.

Theorem 4.1.1. *Let x be a fixed real number.*

(i) If $(X_t/t)_{t\geq 1}$ satisfies a full LDP under the measure $\mathbb{P}$ with the good rate function Λ^*, the asymptotic behaviour of a put option with strike $\exp(xt)$ *is given by the following formula*

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left[\left(\mathrm{e}^{xt}-\mathrm{e}^{X_t}\right)_+\right] = \begin{cases} x-\Lambda^*(x) & \text{if } x\leq\Lambda'(0), \\ x & \text{if } x>\Lambda'(0). \end{cases}$$

(ii) If $(X_t/t)_{t\geq 1}$ satisfies a full LDP under the measure $\widetilde{\mathbb{P}}$ with the good rate function $\widetilde{\Lambda}^*$, the asymptotic behaviour of a call option struck at e^{xt} *is given by the formula*

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right] = \begin{cases} x-\Lambda^*(x) & \text{if } x\geq\Lambda'(1), \\ 0 & \text{if } x<\Lambda'(1), \end{cases}$$

(iii) If $(X_t/t)_{t\geq 1}$ satisfies a full LDP under the measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ with the respective good rate functions Λ^* and $\widetilde{\Lambda}^*$, the asymptotic behaviour of a covered call option with payoff $\mathrm{e}^{X_t}-\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+$ *is given by the formula*

$$\lim_{t\to\infty} t^{-1}\log\left(1-\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\right) = x-\Lambda^*(x) \quad \text{if } x\in\left[\Lambda'(0),\Lambda'(1)\right].$$

Proof. The following inequalities hold for all $t\geq 1$ and $\varepsilon>0$:

$$\mathrm{e}^{xt}\left(1-\mathrm{e}^{-\varepsilon}\right)\mathbb{1}_{\{X_t/t<x-\varepsilon\}} \leq \left(\mathrm{e}^{xt}-\mathrm{e}^{X_t}\right)_+ \leq \mathrm{e}^{xt}\mathbb{1}_{\{X_t/t<x\}}.$$

Taking expectations, logarithms, dividing by t and applying the LDP for $(X_t/t)_{t\geq 1}$ give

$$x-\inf_{y<x-\varepsilon}\Lambda^*(y)\leq\liminf_{t\to\infty}\frac{1}{t}\log\mathbb{E}\left[\left(\mathrm{e}^{xt}-\mathrm{e}^{X_t}\right)_+\right]\leq\limsup_{t\to\infty}\frac{1}{t}\log\mathbb{E}\left[\left(\mathrm{e}^{xt}-\mathrm{e}^{X_t}\right)_+\right]\leq x-\inf_{y\leq x}\Lambda^*(y).$$

Statement (i) in Theorem 4.1.1 then follows from the continuity of Λ^*.

The following inequalities hold for all $t\geq 1$ and $\varepsilon>0$:

$$\mathrm{e}^{X_t}\left(1-\mathrm{e}^{-\varepsilon}\right)\mathbb{1}_{\{X_t/t>x+\varepsilon\}} \leq \left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+ \leq \mathrm{e}^{X_t}\mathbb{1}_{\{X_t/t>x\}}.$$

The same steps as above under $\widetilde{\mathbb{P}}$ lead to the inequalities

$$-\inf_{y>x+\varepsilon}\widetilde{\Lambda}^*(y)\leq\liminf_{t\to\infty}\frac{1}{t}\log\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\leq\limsup_{t\to\infty}\frac{1}{t}\log\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\leq-\inf_{y\geq x}\widetilde{\Lambda}^*(y),$$

and the continuity of $\widetilde{\Lambda}^*$ implies (ii) in Theorem 4.1.1.

Assume now that $x\in[\Lambda'(0),\Lambda'(1)]$ and note that the following holds

$$\mathrm{e}^{xt}\mathbb{1}_{\{X_t/t\geq x\}}\leq\mathrm{e}^{X_t}-\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+=\mathrm{e}^{X_t}\mathbb{1}_{\{X_t/t<x\}}+\mathrm{e}^{xt}\mathbb{1}_{\{X_t/t\geq x\}}.$$

This, together with the LDP under both measures, implies the inequalities

$$\begin{aligned}\mathrm{e}^{xt}\mathbb{P}\left[X_t/t\geq x\right] &\leq 1-\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\leq\widetilde{\mathbb{P}}\left[X_t/t<x\right]+\mathrm{e}^{xt}\mathbb{P}\left[X_t/t\geq x\right]\\ &\leq \exp\left(-t\widetilde{\Lambda}^*(x)+\varepsilon t\right)+\mathrm{e}^{xt}\exp\left(-t\Lambda^*(x)+\varepsilon t\right)\end{aligned}$$

for any $\varepsilon>0$ and t large enough since x is smaller (resp. greater) than the minimum of $\widetilde{\Lambda}^*$ (resp. Λ^*). By (4.1.1) we get

$$x+t^{-1}\log\mathbb{P}\left[X_t/t\geq x\right]\leq t^{-1}\log\left(1-\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\right)\leq x-\Lambda^*(x)+\varepsilon+t^{-1}\log 2.$$

for any $\varepsilon>0$ and all large t. Therefore we find the inequalities

$$\begin{aligned}x-\Lambda^*(x) &\leq \liminf_{t\to\infty}t^{-1}\log\left(1-\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\right)\\ &\leq \limsup_{t\to\infty}t^{-1}\log\left(1-\mathbb{E}\left[\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right]\right)\leq x-\Lambda^*(x)+\varepsilon\end{aligned}$$

for all $\varepsilon>0$. This proves the theorem. □

Our final goal is to translate these option price asymptotics into implied volatility asymptotics. Since the implied volatility is defined in terms of the Black-Scholes model, we shall need an equivalent formulation for call and put options under the Black-Scholes model as in Theorem 4.1.1. Let us consider the Black-Scholes model where the process $(X_t)_{t\geq0}$ satisfies the SDE $\mathrm{d}X_t=-\Sigma^2/2\mathrm{d}t+\Sigma\mathrm{d}W_t$, with $\Sigma>0$. The limiting logarithmic Laplace transform in the Black-Scholes model reads

$$\Lambda_{\mathrm{BS}}(u)=u\,(u-1)\,\Sigma^2/2,\quad\text{for all } u\in\mathbb{R},\tag{4.1.2}$$

and the functions $\Lambda^*_{\mathrm{BS}}:\mathbb{R}\times\mathbb{R}^*_+\to\mathbb{R}$ and $u^*_{\mathrm{BS}}:\mathbb{R}\to\mathbb{R}$ are easily characterised by

$$\Lambda^*_{\mathrm{BS}}(x,\Sigma):=\left(x+\Sigma^2/2\right)^2/\left(2\Sigma^2\right),\quad\text{for all } x\in\mathbb{R},\ \Sigma\in\mathbb{R}^*_+,\tag{4.1.3}$$

$$u^*_{\mathrm{BS}}(x):=\left(x+\Sigma^2/2\right)/\Sigma^2,\quad\text{for all } x\in\mathbb{R}.\tag{4.1.4}$$

For each $x\in\mathbb{R}$, $u^*_{\mathrm{BS}}(x)$ is the unique solution to $\Lambda'_{\mathrm{BS}}(u)=x$, i.e. corresponds to the saddlepoint of the function $u\mapsto ux-\Lambda_{\mathrm{BS}}(u)$. Furthermore we clearly have $(\Lambda^*_{\mathrm{BS}})'(x)=0$ if and only if $x=-\Sigma^2/2$ and $(\Lambda^*_{\mathrm{BS}})'(x)=1$ if and only if $x=\Sigma^2/2$. Corollary 4.1.2 below follows directly from Theorem 4.1.1 and provides the asymptotic behaviour of option prices under the Black-Scholes model.

Corollary 4.1.2. *Under the Black-Scholes model, we have the following option price asymptotics.*

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left(\mathrm{e}^{xt}-\mathrm{e}^{X_t}\right)_+ = \begin{cases} x-\Lambda^*_{\mathrm{BS}}(x) & \text{if } x \le -\Sigma^2/2, \\ x & \text{if } x > -\Sigma^2/2, \end{cases}$$

$$\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+ = \begin{cases} x-\Lambda^*_{\mathrm{BS}}(x) & \text{if } x \ge \Sigma^2/2, \\ 0 & \text{if } x < \Sigma^2/2, \end{cases}$$

$$\lim_{t\to\infty} t^{-1}\log\left(1-\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^{xt}\right)_+\right) = \begin{cases} 2x+\Sigma^2 & \text{if } x \le -3\Sigma^2/2, \\ x-\Lambda^*_{\mathrm{BS}}(x) & \text{if } x \in \left(-3\sigma^2/2, \Sigma^2/2\right], \\ 0 & \text{if } x > \Sigma^2/2. \end{cases}$$

4.1.2 Implied volatility asymptotics

In the previous section we have seen how to turn tail probabilities into asymptotics for option prices. In this section we take on the next step and prove how to translate these option prices into implied volatility asymptotics. The strikes we have considered are of the form $S_0\exp(xt)$. For strikes of this form, the corresponding implied volatility is $\sigma_t(xt)$ and for simplicity we shall use the notation $\hat{\sigma}_t(x) := \sigma_t(xt)$ throughout this thesis. Let us define the function $\hat{\sigma}^2_\infty : \mathbb{R} \to \mathbb{R}_+$ by

$$\hat{\sigma}^2_\infty(x) := 2\left(2\Lambda^*(x) - x + 2\left(\mathbb{1}_{x\in(x^*,\tilde{x}^*)} - \mathbb{1}_{x\in\mathbb{R}\setminus(x^*,\tilde{x}^*)}\right)\sqrt{\Lambda^*(x)\left(\Lambda^*(x)-x\right)}\right), \tag{4.1.5}$$

where Λ^* is the Fenchel-Legendre transform defined above, and where $x^* := \Lambda'(0)$ and $\tilde{x}^* := \Lambda'(1)$ are the unique real numbers satisfying of $\Lambda^*(x^*) = 0$ and $\Lambda^*(\tilde{x}^*) = \tilde{x}^*$ (note that the latter is equivalent to $\widetilde{\Lambda}^*(\tilde{x}^*) = 0$). From the properties of Λ^* on page 13, $\Lambda^*(x)$ and $\Lambda^*(x) - x$ are non-negative, so that $\hat{\sigma}^2_\infty(x)$ is a well defined real number for all $x \in \mathbb{R}$. The following proposition gives us the behaviour of $\hat{\sigma}_t$ as t tends to infinity.

Proposition 4.1.3. *If the random variable $(X_t/t)_{t\ge1}$ satisfies a full large deviations principle under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$, then the function $\hat{\sigma}_\infty$ is continuous on the whole real line and is the uniform limit of $\hat{\sigma}_t$ as t tends to infinity.*

Proof. We prove the formula in the proposition in the case $x > \tilde{x}^*$. We have to prove that, for all $\delta > 0$, there exists $t^*(\delta) > 0$ such that for all $t > t^*(\delta)$, the inequality $|\hat{\sigma}_t(x) - \hat{\sigma}_\infty(x)| \le \delta$ holds. Note that $\hat{\sigma}_\infty(x)$ defined in (4.1.5) satisfies the quadratic equation

$$\Lambda^*(x) - x = \Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x)) - x, \quad \text{for all } x \in \mathbb{R}, \tag{4.1.6}$$

where Λ^*_{BS} is given by (4.1.3). By (4.1.6) and Theorem 4.1.1 we know that for all $\varepsilon > 0$, there exists $t^*(\varepsilon)$ such that for all $t > t^*(\varepsilon)$ we have the lower bound

$$\exp\Big(-\left(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x)) - x + \varepsilon\right)t\Big) = \exp\Big(-\left(\Lambda^*(x) - x + \varepsilon\right)t\Big) \le \frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^{xt}\right)_+}{S_0}, \tag{4.1.7}$$

and the upper bound

$$\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^{xt}\right)_+}{S_0} \le \exp\Big(-\left(\Lambda^*(x) - x - \varepsilon\right)t\Big) = \exp\Big(-\left(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x)) - x - \varepsilon\right)t\Big). \tag{4.1.8}$$

Note that

$$\hat{\sigma}_\infty^2(x) - 2x = 4\Big((\Lambda^*(x) - x) - \sqrt{(\Lambda^*(x) - x)^2 + (\Lambda^*(x) - x)x}\Big) < 0,$$

since $\Lambda^*(x) - x > 0$. For x fixed, the function $\Sigma \mapsto \Lambda^*_{\mathrm{BS}}(x, \Sigma) - x$ defined on $\big(0, \sqrt{2x}\big)$ is continuous and strictly monotonically decreasing. Thus for any $\delta > 0$ such that $\hat{\sigma}_\infty(x) \pm \delta \in (0, \sqrt{2x})$, define

$$\varepsilon_1(\delta) := \Big(\Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x)) - \Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x) + \delta)\Big)/2 > 0,$$
$$\varepsilon_2(\delta) := \Big(\Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x) - \delta) - \Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x))\Big)/2 > 0,$$

where $\lim_{\delta \to 0} \varepsilon_1(\delta) = \lim_{\delta \to 0} \varepsilon_2(\delta) = 0$. Combining (4.1.7), (4.1.8) and Corollary 4.1.2, there exists $t^*(\delta)$ such that for all $t > t^*(\delta)$,

$$\frac{C_{\mathrm{BS}}\left(xt, t, \hat{\sigma}_\infty(x) - \delta\right)}{S_0} \leq \exp\Big(-\Big(\Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x) - \delta) - x - \varepsilon_2(\delta)\Big)\, t\Big) \leq \frac{\mathbb{E}\left(S_t - S_0 \mathrm{e}^{xt}\right)_+}{S_0},$$

and

$$\frac{\mathbb{E}\left(S_t - S_0 \mathrm{e}^{xt}\right)_+}{S_0} \leq \exp\Big(-\Big(\Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x) + \delta) - x + \varepsilon_1(\delta)\Big)\, t\Big) \leq \frac{C_{\mathrm{BS}}\left(xt, t, \hat{\sigma}_\infty(x) + \delta\right)}{S_0}.$$

Thus by the monotonicity of the Black-Scholes call option formula as a function of the volatility, we have the following bounds for the implied volatility $\hat{\sigma}_t(x)$ at maturity t

$$\hat{\sigma}_\infty(x) - \delta \leq \hat{\sigma}_t(x) \leq \hat{\sigma}_\infty(x) + \delta.$$

The cases $x < x^*$ and $x \in (x^*, \tilde{x}^*)$ are analogous. All that is left to prove is that the limit in the theorem holds uniformly on compact subsets of the complement $\mathbb{R} \backslash \{x^*, \tilde{x}^*\}$. To every point in a compact set we can associate a small interval that contains it such that the limit holds for all x in that interval and all large times t. This defines a cover of the compact set. We can hence find a finite collection of such intervals that also covers our compact set. It follows that the limit holds for any x in the compact set and all times t that are larger than the maximum of the finite number of $t^*(\delta)$ that correspond to the intervals in the finite family that covers the original set. □

We can use this result to determine the implied volatility asymptotics in the case where the strike does not depend on the maturity anymore.

Corollary 4.1.4. *The implied volatility $\sigma_t(x)$ of a European call option struck at $S_0 \exp(x)$ satisfies*

$$\lim_{t \to \infty} \sigma_t^2(x) = 8\Lambda^*(0), \quad \textit{for all } x \in \mathbb{R}.$$

In the standard Heston case, this result has already been obtained by Lewis [73]. We note in passing that the implied volatility smile is indeed flat in the large-maturity limit when the strike is independent of the maturity.

4.2 Asymptotics of continuous affine stochastic volatility models

In this section we use the results above to determine the asymptotic behaviour of the implied volatility for large maturities when the share price process belongs to the class of continuous affine stochastic volatility models defined in (3.1.4). We will show that four different cases can happen, only one of which gives rise to a full LDP. In the other cases, we provide some preliminary results and we leave the full analysis for future research.

4.2.1 The model and its effective domain

We consider here a logarithmic share price process $(X_t)_{t\geq 0}$ satisfying (3.1.4) and we assume for clarity that $X_0 = 0$. We denote its logarithmic moment generating function Λ_t defined by $\Lambda_t(u) := \log \mathbb{E}(\exp(uX_t))$ for all $u \in \mathbb{R}$ as an extended real number in $(-\infty, \infty]$, with effective domain $\mathcal{D}_t$. Note that by Hölder's inequality the function Λ_t is convex on $\mathcal{D}_t$. In order to give the structure of $\Lambda_t(u)$ explicitly we need to define

$$\chi(u) := \beta + u\rho\sqrt{\alpha}, \tag{4.2.1}$$

as well as

$$\gamma(u) := \left(\chi(u)^2 + \alpha u(1-u)\right)^{1/2} \quad \text{and} \quad f_t(u) := \cosh\left(\frac{\gamma(u)t}{2}\right) - \frac{\chi(u)}{\gamma(u)}\sinh\left(\frac{\gamma(u)t}{2}\right). \tag{4.2.2}$$

In Proposition 4.2.1 below we show how to express the cumulant generating function of X in terms of the logarithmic moment generating function of the model (3.1.4) with $a = 0$.

Proposition 4.2.1. *The logarithmic moment generating function Λ_t reads*

$$\Lambda_t(u) = \Lambda_t^H(u) + \frac{a}{2}u(u-1)t, \quad \textit{for all } t \geq 0 \textit{ and } u \in \mathcal{D}_t,$$

where Λ_t^H is the logarithmic moment generating function of the process X in (3.1.4) with $a = 0$. It follows that the effective domains of Λ_t and of Λ_t^H coincide, and the following formula holds

$$\Lambda_t^H(u) = -\frac{2b}{\alpha}\left(\frac{\chi(u)t}{2} + \log f_t(u)\right) + \frac{u(u-1)}{f_t(u)\gamma(u)}\sinh\left(\frac{\gamma(u)t}{2}\right)v, \quad \textit{for all } u \in \mathcal{D}_t. \tag{4.2.3}$$

Proof. It is well known that the logarithmic moment generating function of an affine process X given as a solution of SDE (3.1.4) is of the form

$$\Lambda_t(u) = \phi_t(u) + \psi_t(u)v \quad \text{for all } t \geq 0 \text{ and } u \in \mathcal{D}_t,$$

where the functions $\phi_t, \psi_t : \mathcal{D}_t \to \mathbb{R}$ satisfy the system of Riccati equations

$$\begin{aligned} \partial_t\phi_t(u) &= F(u, \psi_t(u)), \qquad \phi_0(u) = 0, \\ \partial_t\psi_t(u) &= R(u, \psi_t(u)), \qquad \psi_0(u) = 0, \end{aligned} \tag{4.2.4}$$

with

$$R(u,w) := \frac{1}{2}u(u-1) + \frac{\alpha}{2}w^2 + uw\rho\sqrt{\alpha} + \beta w \qquad \text{and} \qquad F(u,w) := \frac{a}{2}u(u-1) + bw$$

(see e.g. [68]). The Riccati equation equation for ψ_t can be solved in closed form,

$$\psi_t(u) = \sinh\left(\frac{\gamma(u)\,t}{2}\right)\frac{u(u-1)}{\gamma(u)\,f_t(u)},$$

where the functions γ and f_t are defined in (4.2.2). The function ϕ_t can be determined by noting that equation (4.2.4) is equivalent to $\phi_t(u) = \int_0^t F(u,\psi_s(u))\,\mathrm{d}s$. Therefore $\phi_t(u) = au(u-1)\,t/2 + b\int_0^t \psi_s(u)\,\mathrm{d}s$. The function Λ_t^H can be constructed in an analogous way on the set $\{u\in\mathbb{R} : \Lambda_t^H(u)<\infty\}$ with R and F as above and $a=0$. This concludes the proof. □

In order to analyse the effective domain $\mathcal{D}_t$ we introduce the quantities u_- and u_+ given by

$$u_- := \begin{cases} \dfrac{\sqrt{\alpha}+2\rho\beta-\sqrt{\left(\sqrt{\alpha}+2\beta\rho\right)^2+4\beta^2\left(1-\rho^2\right)}}{2\sqrt{\alpha}\left(1-\rho^2\right)}, & \text{if } |\rho|<1,\\ -\infty, & \text{if } |\rho|=1 \text{ and } \sqrt{\alpha}+2\rho\beta\le 0,\\ -\beta^2/\left(\alpha+2\rho\beta\sqrt{\alpha}\right), & \text{if } |\rho|=1 \text{ and } \sqrt{\alpha}+2\rho\beta>0, \end{cases} \tag{4.2.5}$$

and

$$u_+ := \begin{cases} \dfrac{\sqrt{\alpha}+2\rho\beta+\sqrt{\left(\sqrt{\alpha}+2\beta\rho\right)^2+4\beta^2\left(1-\rho^2\right)}}{2\sqrt{\alpha}\left(1-\rho^2\right)}, & \text{if } |\rho|<1,\\ \infty, & \text{if } |\rho|=1 \text{ and } \sqrt{\alpha}+2\rho\beta\ge 0,\\ -\beta^2/\left(\alpha+2\rho\beta\sqrt{\alpha}\right), & \text{if } |\rho|=1 \text{ and } \sqrt{\alpha}+2\rho\beta<0. \end{cases} \tag{4.2.6}$$

The inequalities $u_-\le 0$ and $u_+\ge 1$ hold for all admissible values of the parameters and in the case $|\rho|<1$ the parabola γ^2 is strictly positive on the interior of the interval $[u_-,u_+]$ between its distinct zeros. In the case $|\rho|=1$ the graph of the function γ^2 is a line and either u_- or u_+ is infinite. For notational convenience we shall understand the interval $[x,y]\subset\mathbb{R}$ as $[x,\infty)$ if $y=\infty$ and as $(-\infty,y]$ if $x=-\infty$. Using this notation note that $\gamma(u)\in\mathbb{R}$ for all $u\in[u_-,u_+]$. Proposition 4.2.2 analyses the structure of the effective domain $\mathcal{D}_t$ of the function Λ_t.

Proposition 4.2.2. *The effective domain $\mathcal{D}_t$ of the function Λ_t satisfies $[0,1]\subset\mathcal{D}_t$ for all $t\ge 0$ and any set of admissible parameter values from* (3.1.5)*. Furthermore the following statements hold.*

(i) If $\chi(0)\le 0$ we have:

(a) if $\chi(1)\le 0$ then $[u_-,u_+]\subset\mathcal{D}_t$ for any $t>0$;

(b) if $\chi(1)>0$ then for all t large enough there exists $\overline{u}(t)\in(1,u_+)$ such that

$$\lim_{t\to\infty}\overline{u}(t)=1 \quad \textit{and} \quad [u_-,\overline{u}(t))\subset\mathcal{D}_t\subset(-\infty,\overline{u}(t)).$$

(ii) If $\chi(0)>0$ we have:

(a) *if* $\chi(1) \le 0$ *then for all large* t *there exists* $\underline{u}(t) \in (u_-, 0)$ *such that*

$$\lim_{t\to\infty} \underline{u}(t) = 0 \quad \textit{and} \quad (\underline{u}(t), u_+] \subset \mathcal{D}_t \subset (\underline{u}(t), \infty);$$

(b) *if* $\chi(1) > 0$ *then for large* t *there exist* $\underline{u}(t) \in (u_-, 0)$ *and* $\overline{u}(t) \in (1, u_+)$ *such that*

$$\lim_{t\to\infty} \underline{u}(t) = 0, \quad \lim_{t\to\infty} \overline{u}(t) = 1 \quad \textit{and} \quad \mathcal{D}_t = (\underline{u}(t), \overline{u}(t)).$$

Remark 4.2.3. The following elementary facts are useful in the proof of Proposition 4.2.2.

(I) $u_- = -\infty$ and $u_+ = \infty$ if and only if the condition $|\rho| = 1$ and $\sqrt{\alpha} + 2\rho\beta = 0$ holds.

(II) The condition $\chi(1) \neq 0$ implies that $u_+ > 1$ since u_+ is the largest root of the quadratic $\gamma(u)^2$ given in (4.2.2). In particular in (i)(b) and (ii)(b) of Proposition 4.2.2 the interval $(1, u_+)$ is not empty.

(III) The condition $\chi(0) \neq 0$ implies that $u_- < 0$. In particular in (ii) we have $\chi(0) = \beta > 0$ and hence the interval $(u_-, 0)$ is not empty.

(IV) The interval $[0, 1]$ is contained in $\mathcal{D}_t$ for all $t \ge 0$ since the process $(\mathrm{e}^{X_t})_{t\ge 0}$ is a martingale.

(V) If $\chi(0) = 0$ then $u_- = 0$ and $u_+ = 1/(1-\rho^2)$ for $|\rho| < 1$ and $u_+ = \infty$ for $|\rho| = 1$.

Proof. Proposition 4.2.1 implies that it is enough to study the effective domain of the cumulant generating function Λ_t^H of the Heston model. It is clear that the function f_t, defined in (4.2.2) by

$$f_t(u) = \cosh\left(\frac{\gamma(u)\, t}{2}\right) - \frac{\chi(u)}{\gamma(u)} \sinh\left(\frac{\gamma(u)\, t}{2}\right),$$

will play a key role in in understanding the set $\mathcal{D}_t$.

Case (i): If we can prove that

$$f_t(u) > 0, \qquad \text{for all } u \in [u_-, 1], \tag{4.2.7}$$

then Proposition 4.2.1 implies that $[u_-, 1] \subset \mathcal{D}_t$ since the functions on both sides of equality (4.2.3) can be analytically extended to a neighbourhood of $[u_-, 1]$ in the complex plane and therefore coincide on the interval.

We now prove (4.2.7). It follows from the definition of γ in (4.2.2) that $|\chi(u)/\gamma(u)| \le 1$ for all $u \in [0, 1]$ and hence (4.2.7) holds on $[0, 1]$. It is easy to see that $\lim_{u \searrow u_-} \chi(u) \le 0$. Since $\chi(0) = \beta \le 0$ we have $\chi(u) \le 0$ for all $u \in [u_-, 0]$ which implies (4.2.7).

In case (i)(a) assume first that $u_+ < \infty$. Then elementary algebra shows that $\chi(u_+) \le 0$. Therefore $\chi(u) \le 0$, and hence $f_t(u) > 0$, for all $u \in [1, u_+]$. If $u_+ = \infty$ the condition $\chi(1) \le 0$ implies that $\rho = -1$ and therefore $\chi(u) < 0$ for all $u \ge 1$. Hence $f_t(u) \in (0, \infty)$ for all $u \in [1, \infty) = [1, u_+]$. Proposition 4.2.1 and the analytic continuation argument as above imply $[u_-, u_+] \subset \mathcal{D}_t$.

Recall that in case (i)(b) we have $u_+ > 1$ (see Remark 4.2.3 (II)). Let $\overline{u}(t)$ be the smallest solution of the equation $f_t(u) = 0$ in the interval $(1, u_+)$. Note that, since γ is strictly positive on the interval $(1, u_+)$, for a fixed t the equation $f_t(u) = 0$ can be rewritten as

$$t = F(u), \qquad \text{where} \quad F(u) := \frac{2}{\gamma(u)} \operatorname{arctanh}\left(\frac{\gamma(u)}{\chi(u)}\right). \tag{4.2.8}$$

The latter equation clearly has a solution in $(1, u_+)$ for large t since the continuous function F tends to ∞ as u decreases to 1 (note that $\lim_{u \searrow 1} \gamma(u)/\chi(u) = 1$). This also implies that the smallest solution $\overline{u}(t)$ decreases to one. The functions on both sides of the equality in (4.2.3) coincide on $[u_-, 1]$, are analytic on some neighbourhood of this interval in the complex plane and the right-hand side is (4.2.3) is real and finite on $[u_-, \overline{u}(t))$. The two functions must therefore also coincide on $[u_-, \overline{u}(t))$, which in particular implies $[u_-, \overline{u}(t)) \subset \mathcal{D}_t$. Formula (4.2.3) implies that $\overline{u}(t)$ is not an element of $\mathcal{D}_t$ and the convexity of Λ_t yields that $\mathcal{D}_t \cap [\overline{u}(t), \infty) = \emptyset$.

Case (ii): In case (ii)(a) the condition $\chi(1) \le 0$ implies $\rho < 0$ and hence $\chi(u) \le 0$ for all $u \in [1, u_+]$. Therefore $f_t(u) > 0$ on $[1, u_+]$ and hence $[0, u_+] \subset \mathcal{D}_t$. Let $\underline{u}(t)$ be the largest solution of the equation $f_t(u) = 0$ in the interval $(u_, 0)$. Since $\lim_{u \nearrow 0}(\gamma(u)/\chi(u)) = 1$, an analogous argument as in the proof of (i)(b) shows that $\underline{u}(t)$ is well defined and the limit in the proposition holds. The proof for the inclusions follows the same steps as in the proof of (i)(b).

In case (ii)(b) we have $\chi(0) = \beta > 0$ and $\chi(1) > 0$. Therefore the definition of γ, given in (4.2.2), implies

$$\lim_{u \nearrow 0} \frac{\gamma(u)}{\chi(u)} = 1 \qquad \text{and} \qquad \lim_{u \searrow 1} \frac{\gamma(u)}{\chi(u)} = 1$$

and hence, by (4.2.8), there exist solutions to the equation $f_t(u) = 0$ in both intervals $(u_-, 0)$ and $(1, u_+)$. Let $\underline{u}(t)$ be the largest solution in $(u_-, 0)$ and $\overline{u}(t)$ the smallest solution in $(1, u_+)$. An analogous argument to the one in the proofs of (i)(b) and (ii)(a) gives the form of $\mathcal{D}_t$. $\square$

4.2.2 LDP for continuous affine stochastic volatility models

In this section we analyse the large deviations behaviour of the family of random variables $Z_t := X_t/t$ for $t \ge 1$, where $(X_t)_{t \ge 0}$ is defined by SDE (3.1.4). Proposition 4.2.4 describes the properties of the cumulant generating function Λ defined in (1.1.1). The Fenchel-Legendre transform Λ^* is studied in Proposition 4.2.7. The following proposition follows from Propositions 4.2.1 and 4.2.2.

Proposition 4.2.4. *The limiting cumulant generating function Λ of $(X_t/t)_{t \ge 1}$ is given by*

$$\Lambda(u) = -\frac{b}{\alpha}\left(\chi(u) + \gamma(u)\right) + \frac{a}{2} u (u-1) \quad \text{for all } u \in \mathcal{D}_\Lambda,$$

with the functions χ and γ given in (4.2.1) and (4.2.2) respectively. The function Λ is infinitely differentiable on the interior $\mathcal{D}_\Lambda^o$ of its effective domain. The boundary points u_- and u_+, defined in (4.2.5) and (4.2.6), can be used to describe the effective domain $\mathcal{D}_\Lambda$ as follows.

(i) If $\chi(0) \leq 0$ we have:

(a) if $\chi(1) \leq 0$ then $\mathcal{D}_\Lambda = [u_-, u_+]$;

(b) if $\chi(1) > 0$ then $\mathcal{D}_\Lambda = [u_-, 1]$.

(ii) If $\chi(0) > 0$ we have:

(a) if $\chi(1) \leq 0$ then $\mathcal{D}_\Lambda = [0, u_+]$;

(b) if $\chi(1) > 0$ then $\mathcal{D}_\Lambda = [0, 1]$.

Remark 4.2.5. The following facts are relevant for the large deviations behaviour of the family of random variables $(X_t/t)_{t\geq 1}$.

(I) In case (i)(a) of Proposition 4.2.4 the function Λ is essentially smooth.

(II) In case (i)(b) (resp. (ii)(a)) of Proposition 4.2.4 the function Λ is steep at the left boundary u_- (resp. right boundary u_+) but not at the right (resp. left) boundary of the effective domain.

(III) In case (i)(b) (resp. (ii)(a)) of Proposition 4.2.4 the right (resp. left) boundary point of the effective domain is strictly smaller (resp. greater) than u_+ (resp. u_-). This is a consequence of Remarks 4.2.3 (II) and 4.2.3 (III).

(IV) In case (ii)(b) of Proposition 4.2.4 Λ is not steep at either of the two boundaries of its effective domain. Furthermore $\mathcal{D}_\Lambda$ is contained in the interior of the interval $[u_-, u_+]$ by Remarks 4.2.3 (II) and 4.2.3 (III).

(V) It is a consequence of the remarks in (I)–(IV) above that the limiting cumulant generating function Λ is steep at a boundary point of the effective domain if and only if this point is an element of the set $\{u_-, u_+\}$.

Note that when u_- (resp. u_+) is not in $\mathcal{D}_\Lambda$ then the function Λ is discontinuous at 0 (resp. at 1). We henceforth define the following extended real numbers

$$\Lambda_-(1) := \lim_{u \nearrow 1} \Lambda(u), \quad \Lambda_+(0) := \lim_{u \searrow 0} \Lambda(u), \quad \Lambda'_-(1) := \lim_{u \nearrow 1} \Lambda'(u), \quad \Lambda'_+(0) := \lim_{u \searrow 0} \Lambda'(u).$$

The functions Λ and Λ' are monotone on the intervals $(0, \varepsilon)$ and $(1-\varepsilon, 1)$ for small enough ε, hence all the limits exist. Note further that the limit $\Lambda'_+(0)$ (resp. $\Lambda'_-(1)$) is equal to $-\infty$ (resp. ∞) if and only if $\chi(0) = 0$ (resp. $\chi(1) = 0$).

Remark 4.2.6. At zero and one the following identities hold

$$\Lambda_+(0) = -\frac{b}{\alpha}\left(\chi(0) + |\chi(0)|\right) \quad \text{and} \quad \Lambda'_+(0) = \begin{cases} \dfrac{1}{|\chi(0)|}\left(\left(\chi(1) - \chi(0)\right)\Lambda_+(0) - \dfrac{b}{2}\right) - \dfrac{a}{2}, & \text{if } \chi(0) \neq 0, \\ -\infty, & \text{if } \chi(0) = 0, \end{cases}$$

$$\Lambda_-(1) = -\frac{b}{\alpha}\left(\chi(1) + |\chi(1)|\right) \quad \text{and} \quad \Lambda'_-(1) = \begin{cases} \dfrac{1}{|\chi(1)|}\left(\left(\chi(1) - \chi(0)\right)\Lambda_-(1) + \dfrac{b}{2}\right) + \dfrac{a}{2}, & \text{if } \chi(1) \neq 0, \\ \infty, & \text{if } \chi(1) = 0. \end{cases}$$

Note that the inequalities $\Lambda_+(0) \leq 0$ and $\Lambda_-(1) \leq 0$ hold for any admissible set of parameters.

Proposition 4.2.7. *The Fenchel-Legendre transform Λ^* of Λ can be represented as*

$$\Lambda^*(x) = \begin{cases} xu_x - \Lambda(u_x), & \text{for all } x \in \Lambda'(\mathcal{D}^o_\Lambda), \\ x - \Lambda_-(1), & \text{for all } x \in \left[\Lambda'_-(1), \infty\right) \cap \left(\mathbb{R}\backslash\Lambda'(\mathcal{D}^o_\Lambda)\right), \\ -\Lambda_+(0), & \text{for all } x \in \left(-\infty, \Lambda'_+(0)\right] \cap \left(\mathbb{R}\backslash\Lambda'(\mathcal{D}^o_\Lambda)\right), \end{cases} \tag{4.2.9}$$

where u_x is the unique solution in $\mathcal{D}^o_\Lambda$ to the equation $\Lambda'(u) = x$ for all $x \in \Lambda'(\mathcal{D}^o_\Lambda)$. Furthermore Λ^* is continuously differentiable on its effective domain $\mathcal{D}_{\Lambda^*}$ and $\mathcal{D}_{\Lambda^*} = \mathbb{R}$.

(i) The function Λ^* attains its global minimal value $-\Lambda_+(0)$ at $\Lambda'_+(0)$. If $0 \in \mathcal{D}^o_\Lambda$ then the minimum is attained at the unique point $\Lambda'_+(0) = \Lambda'(0)$ and the minimal value is $\Lambda^*(\Lambda'(0)) = \Lambda_+(0) = 0$. If $0 \notin \mathcal{D}^o_\Lambda$ the minimal value is attained at every $x \in \left(-\infty, \Lambda'_+(0)\right] \cap \left(\mathbb{R}\backslash\Lambda'(\mathcal{D}^o_\Lambda)\right)$.

(ii) The function $x \mapsto \Lambda^*(x) - x$ attains its global minimal value $-\Lambda_-(1)$ at $\Lambda'_-(1)$. If $1 \in \mathcal{D}^o_\Lambda$ then the minimum value $\Lambda_-(1) = \Lambda(1) = 0$ is attained at the unique point $\Lambda'_-(1) = \Lambda'(1)$ which is therefore the unique solution of the equation $\Lambda^*(x) = x$. If $1 \notin \mathcal{D}^o_\Lambda$ the function $x \mapsto \Lambda^*(x) - x$ attains the minimal value at every $x \in \left[\Lambda'_-(1), \infty\right) \cap \left(\mathbb{R}\backslash\Lambda'(\mathcal{D}^o_\Lambda)\right)$.

Remark 4.2.8.

(i) Since Λ is a strictly convex smooth function on $\mathcal{D}^o_\Lambda$, the first derivative Λ' is invertible on this interval and u_x is a strictly increasing, differentiable function of x on $\Lambda'(\mathcal{D}^o_\Lambda)$. Furthermore the equality $(\Lambda^*)'(x) = u_x$ holds for any $x \in \Lambda'(\mathcal{D}^o_\Lambda)$.

(ii) Proposition 4.2.4 implies the following form for the interval $\Lambda'(\mathcal{D}^o_\Lambda)$:

$$\Lambda'(\mathcal{D}^o_\Lambda) = \begin{cases} \mathbb{R} & \text{if } \chi(0) \leq 0,\ \chi(1) \leq 0, \\ \left(-\infty, \Lambda'_-(1)\right) & \text{if } \chi(0) \leq 0,\ \chi(1) > 0, \\ \left(\Lambda'_+(0), \infty\right) & \text{if } \chi(0) > 0,\ \chi(1) \leq 0, \\ \left(\Lambda'_+(0), \Lambda'_-(1)\right) & \text{if } \chi(0) > 0,\ \chi(1) > 0. \end{cases} \tag{4.2.10}$$

Hence the second case in (4.2.9) corresponds to $\chi(1) > 0$ and the third case to $\chi(0) > 0$.

(iii) If we take $a = 0$ in the model given by (3.1.4), there exists an explicit formula for the zero u_x of the equation $\Lambda'(u) = x$, when $x \in \Lambda'(\mathcal{D}_\Lambda^o)$, given by

$$u_x = \frac{1}{2(1-\rho^2)\sqrt{\alpha}}\left(2\rho\beta + \sqrt{\alpha} + \frac{p(x)\xi}{\sqrt{p(x)^2 + b^2(1-\rho^2)}}\right),$$

where

$$p(x) := b\rho + x\sqrt{\alpha}, \qquad \text{and} \qquad \xi := \sqrt{\left(2\rho\beta + \sqrt{\alpha}\right)^2 + 4\beta^2(1-\rho^2)}.$$

Note that u_x is well defined as a limit when $|\rho|$ tends to 1. Together with (4.2.9) this yields an explicit formula for the rate function Λ^*, the details of which are postponed to Section 4.2.3.

(iv) In view of Part I Section 3.3.1, all the results above hold for the multidimensional Heston model. However, even in the standard case $a = 0$ the Fenchel-Legendre transform is not available in closed-form anymore. However determining it numerically simply boils down to a root-finding exercise.

Proof. Let $u_x \in \mathcal{D}_\Lambda^o$ be the unique solution of $\Lambda'(u) = x$, which exists by Remark 4.2.8 (i). It is clear from definition (1.1.3) that, for $x \in \Lambda'(\mathcal{D}_\Lambda^o)$, the Fenchel-Legendre Λ^* takes the form given in the proposition.

Assume now that $\Lambda'_-(1)$ is finite. This is equivalent to $\chi(1) \neq 0$ which implies that for every $u \in \mathcal{D}_\Lambda^o$ we have $u < 1$. Then for any $x \in \left[\Lambda'_-(1), \infty\right) \cap (\mathbb{R}\backslash\Lambda'(\mathcal{D}_\Lambda^o))$ the inequality $\Lambda_-(1) - \Lambda(u) \leq x(1-u)$ holds by Lagrange theorem (and the fact that Λ' is strictly increasing) and (4.2.9) follows.

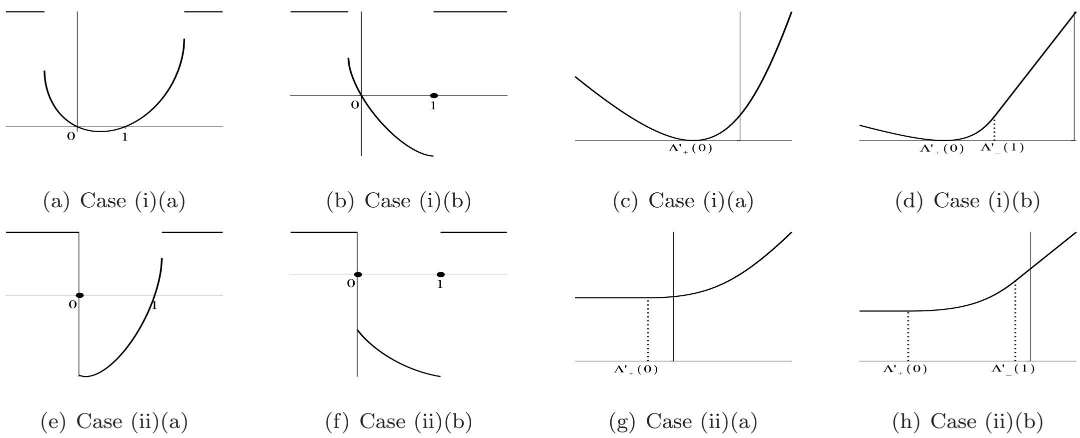
If $\Lambda'_+(0)$ is finite, then for every $u \in \mathcal{D}_\Lambda^o$ we have $u > 0$. For any $x \in \left(-\infty, \Lambda'_+(0)\right] \cap (\mathbb{R}\backslash\Lambda'(\mathcal{D}_\Lambda^o))$ the inequality $ux - \Lambda(u) \leq -\Lambda_+(0)$ holds for all $u \in \mathcal{D}_\Lambda^o$ and (4.2.9) follows.

The function Λ^* is continuously differentiable on $\mathbb{R}$ by (4.2.9) and Remark 4.2.8 (i). Note that, if $0 \in \mathcal{D}_\Lambda^o$, at the minimum we have $u_x = 0$. This implies by definition that the minimum of Λ^* is attained at $\Lambda'(0) = x$. The case $0 \notin \mathcal{D}_\Lambda^o$ follows in a similar way.

If $1 \in \mathcal{D}_\Lambda^o$, then by differentiating the formula in (4.2.9) we find that the minimum of $x \mapsto \Lambda^*(x) = x$ is attained if and only if $u_x = 1$, which is equivalent to $\Lambda'(1) = x$. If $1 \notin \mathcal{D}_\Lambda^o$, it is easy to see that the minimum is attained for all $x \geq \Lambda'_-(1)$. This concludes the proof. □

Remark 4.2.9. It is clear from Proposition 4.2.7 (and visually from Figure 4.1) that the function Λ^* is a *good* rate function only in cases (i)(a) and (i)(b). In the other two cases it is straightforward to show that its level sets are closed but not necessarily compact. However Λ^* and $\widetilde{\Lambda}^*$ are both good rate functions only in case (i)(a).

If $\chi(0) < 0$ and $\chi(1) < 0$, i.e. case (i)(a) with $u_- < 0$ and $u_+ > 0$, a full large deviations principle clearly holds under $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ since Λ and $\widetilde{\Lambda}$ are both essentially smooth on the effective domains, the interiors of which contain the origin. Therefore Theorem 4.1.1 can be applied and we



(a) Case (i)(a) (b) Case (i)(b) (c) Case (i)(a) (d) Case (i)(b)

(e) Case (ii)(a) (f) Case (ii)(b) (g) Case (ii)(a) (h) Case (ii)(b)

Figure 4.1: The four figures on the left represent the function Λ characterised in Proposition 4.2.4. The four figures on the right represent the Fenchel-Legendre Λ^* determined in Proposition 4.2.7. The dotted line on the graphs for Λ^* represent the thresholds $\Lambda'_-(1)$ and $\Lambda'_+(0)$ above or below which Λ^* becomes linear.

obtain the asymptotic implied volatility defined in (4.1.5) by Proposition 4.1.3. We provide below in Section 4.2.3 a clear and simple closed-form expression for the implied volatility in this case. In all other cases several problems may arise: first we loose steepness at one or both boundaries of the domain $\mathcal{D}_\Lambda$, which implies that the function Λ is not essentially smooth anymore. More importantly, in all these cases the origin is not in the interior of $\mathcal{D}_\Lambda$ or of $\mathcal{D}_{\widetilde{\Lambda}}$ and hence we do not have the large deviations upper and lower bounds for the tail probabilities. Removing this assumption is a work under progress and we shall not mention any result in this direction here. However a deeper analysis can be carried out in case the origin is in $\mathcal{D}_\Lambda^o$ and the function Λ is not steep at the right boundary of its effective domain, i.e. in the case $\chi(1) > 0$ and $\chi(0) < 0$. In this case we have $\mathcal{D}_\Lambda = [u_-, 1]$ with $u_- < 0$ and we are able to prove the following theorem. Note that similar results in the case where the limiting cumulant generating function is not steep have been obtained in [38] and in [104] for other processes.

Theorem 4.2.10. *Assume $\chi(0) < 0$ and $\chi(1) \geq 0$ (i.e. $\mathcal{D} = [u_-, 1]$, with $u_- < 0$). For any real number x, the following limits hold*

$$-\lim_{t\to\infty} t^{-1} \log \mathbb{P}\left(X_t/t \geq x\right) = \inf_{y \geq x} \Lambda^*(y), \tag{4.2.11}$$

$$-\lim_{t\to\infty} t^{-1} \log \mathbb{P}\left(X_t/t \leq x\right) = \inf_{y \leq x} \Lambda^*(y), \tag{4.2.12}$$

where the function Λ^* is given in Proposition 4.2.7.

Remark 4.2.11. Note that we have included the case $\chi(1) = 0$ in the statement of the theorem. In this case $u_+ = 1$ and the function Λ is clearly essentially smooth so the proof follows immediately from the Gärtner-Ellis theorem under the probability $\mathbb{P}$.

Remark 4.2.12. Since we know that the function Λ^* is continuous and convex, in view of the proof of Theorem 4.1.1 it is clear that Theorem 4.2.10 above is sufficient for (i) in Theorem 4.1.1 to hold. This implies that the implied volatility smile $\hat{\sigma}_\infty(x)$ has the form (4.1.5) for all $x \leq \Lambda'(0)$.

Lemma 4.2.13. *Assume $\chi(1) > 0$. As t tends to infinity, the following behaviour for the cumulant generating function Λ_t holds as u approaches 1 from below:*

$$t^{-1}\Lambda_t(u) = \Lambda(u) - \frac{2b}{\alpha t}\log(1-u) + t^{-1}R_t(u),$$

where $R_t : \mathcal{D}_t \to \mathbb{R}$ is an analytic function for any $t \geq 0$ which converges on any compact subsets of $\mathcal{D}_t$ as t tends to infinity.

Proof. From (4.2.2), we have $f_t(u) \sim \cosh(\gamma(u)t/2)\,|1 - \chi(u)/\gamma(u)|$ as t tends to infinity, so that

$$\log(f_t(u)) = \log\left|1 - \frac{\chi(u)}{\gamma(u)}\right| + \gamma(u)t/2 - \log(2) + o\left(\mathrm{e}^{-\gamma(u)t}\right), \quad \text{as } t \text{ tends to infinity.}$$

Note further that the condition $\chi(1) > 0$ implies $1 - \frac{\chi(u)}{\gamma(u)} = -\frac{\alpha}{2}(u-1)/\chi(1)^2 + \mathcal{O}\left((u-1)^2\right)$. Since the right boundary of the domain $\mathcal{D}_t$ tends to one from above as t tends to infinity, we can perform a Taylor expansion on the left of 1 and we obtain as t tends to infinity,

$$\log(f_t(u)) = \log(1-u) + \gamma(u)\frac{t}{2} + \widetilde{R}_t(u) + o\left(\mathrm{e}^{-\gamma(u)t}\right).$$

It is straightforward to see that the terms left in the expression (4.2.3) for Λ_t satisfy the properties we need for the function R_t as in the statement of the lemma. Combining this result with the expression (4.2.3) and Proposition 4.2.4 concludes the proof. □

Proof. (of Theorem 4.2.10)
By Remark 4.2.11 we can assume $\chi(1) > 0$ which implies that $u_+ > 1$. In the case $\mathcal{D}_\Lambda = [u_-, 1]$ with $u_- < 0$, the limiting cumulant generating function Λ is not essentially smooth anymore, since it is not steep at the right boundary 1. The proof of the theorem follows similar steps as the standard Gärtner-Ellis theorem, but some refinements are needed. For each limit we first prove an upper bound and then a lower bound. For any real numbers x and y such that $y > x$, Chebycheff inequality gives us an upper bound on the compact interval $[x, y]$, namely

$$\limsup_{t\to\infty} t^{-1}\log\mathbb{P}(X_t/t \in [x,y]) \leq -\inf_{z\in[x,y]}\Lambda^*(z). \tag{4.2.13}$$

Since the equality $\mathbb{P}(X_t/t \geq x) = \lim_{y\to\infty}\mathbb{P}(X_t/t \in [x,y])$ holds then for all $\varepsilon > 0$ there exists $t^* > 0$ such that for all $t > t^*$, (4.2.13) implies

$$t^{-1}\log\mathbb{P}(X_t/t \in [x,y]) \leq \varepsilon - \inf_{z\in[x,y]}\Lambda^*(z) \leq \varepsilon - \inf_{z\geq x}\Lambda^*(z).$$

The right-hand side does not depend on the value y, hence we can take the limit on both sides as y tends to infinity and we obtain the upper bound for $\limsup_{t\to\infty} t^{-1}\log\mathbb{P}(X_t/t \geq x)$ in (4.2.11).

We now prove the lower bound. The function Λ is strictly convex on the interval $\left(-\infty, \Lambda'_{-}(1)\right)$ so that the lower bound follows immediately from the Gärtner-Ellis theorem for all $x < \Lambda'_{-}(1)$. Let us therefore consider $x \geq \Lambda'_{-}(1)$. Since the function Λ is continuously differentiable and convex on $\mathcal{D}_\Lambda^0$, two possible cases arise: either it attains its minimum at a unique point $u_0 \in \mathcal{D}_\Lambda^o$ or it is strictly decreasing on its effective domain. The first case uses similar arguments as in the standard Gärtner-Ellis theorem. Proving the lower bound is tantamount to proving the following equality

$$\lim_{\delta\to 0}\liminf_{t\to\infty} t^{-1}\log\mathbb{P}\left(X_t/t \in (x, x+\delta)\right) \geq -\Lambda^*(x).$$

From Proposition 4.2.13 let us define the function $\widetilde{\Lambda}_t : \mathcal{D}_t \cap (-\infty, 1) \to \mathbb{R}$ by

$$\widetilde{\Lambda}_t(u) := \Lambda_t(u) - \frac{2b}{\alpha t}\log(1-u), \quad \text{for all } u \in \mathcal{D}_t \cap (-\infty, 1).$$

The properties of the function Λ and its behaviour as t tends to infinity given in Proposition 4.2.13 imply that for each $t > 0$ the function $\widetilde{\Lambda}'_t$ has a unique root $u_t \in \mathcal{D}_t \cap (-\infty, 1)$ and that this root converges to 1 as t tends to infinity. Let us further define a new measure $\mathbb{Q}_t$ by

$$\frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}\mathbb{P}_t}(z) := \exp\left(u_t z t - \Lambda_t(u_t)\right), \quad \text{for any } z \in \mathbb{R}. \tag{4.2.14}$$

For any $\delta > 0$ and $x \geq \Lambda'_{-}(1)$ we then have

$$\begin{aligned}\log\mathbb{P}_t\Big(X_t/t \in (x, x+\delta)\Big) &= \log\int_{(x,x+\delta)}\exp\Big(\Lambda_t(u_t) - u_t z t\Big)\mathrm{d}\mathbb{Q}_t(z)\\ &= \Lambda_t(u_t) - u_t x t + \log\int_{(x,x+\delta)} \mathrm{e}^{-u_t(z-x)t}\mathrm{d}\mathbb{Q}_t(z)\\ &\geq \Lambda_t(u_t) - u_t(x+\delta)t + \log\mathbb{Q}_t\Big(X_t/t \in (x, x+\delta)\Big),\end{aligned}$$

for t large enough so that $u_t > 0$, and hence

$$\begin{aligned}\lim_{\delta\to 0}\liminf_{t\to\infty} t^{-1}\log\mathbb{P}\Big(X_t/t \in (x, x+\delta)\Big) \geq{}& \liminf_{t\to\infty}\left(t^{-1}\Lambda(u_t) - u_t x\right)\\ &+ \lim_{\delta\to 0}\liminf_{t\to\infty} t^{-1}\log\mathbb{Q}_t\Big(X_t/t \in (x, x+\delta)\Big).\end{aligned} \tag{4.2.15}$$

We now have to find a lower bound for both terms on the right-hand side of this inequality. Since the function Λ_t is convex for all $t > 0$, we have $\Lambda_t(u_t) - \Lambda(u) \geq (u_t - u)\Lambda'(u)$ for all $u < 1$. From [85, Theorem 25.7] we have $\lim_{t\to\infty} t^{-1}\Lambda'_t(u) = \Lambda'(u)$ and $\liminf_{t\to\infty} t^{-1}\Lambda_t(u_t) \geq \Lambda(u) + (1-u)\Lambda'(u)$, which implies that $\liminf_{t\to\infty} t^{-1}\Lambda_t(u_t) \geq \Lambda_{-}(1)$. The fact that u_t converges to 1 as t tends to infinity leads to $\liminf_{t\to\infty}\left(t^{-1}\Lambda(u_t) - u_t x\right) \geq \Lambda^*(x)$. We are left to find a lower bound for the last term on the right-hand side of the inequality (4.2.15) as t tends to infinity and δ to zero. By Lemma 4.2.14, we know that $\mathbb{Q}_t$ converges to a probability measure $\mathbb{Q}$ as t tends to infinity, which concludes the proof for the lower bound, and hence for the limit in (4.2.11). The proof of the limit (4.2.12) is analogous. □

We conclude this section with the following technical lemma needed above. Recall that the logarithmic Laplace transform of a Gamma-distributed random variable Y with parameters $g_1, g_2 > 0$ (and we denote it $Y \sim \Gamma(g_1, g_2)$) reads $\log \mathbb{E}(\exp(uY)) = g_1 \log(g_2/(g_2+u))$, for all $u > -g_2$. We also denote $\delta(g)$ the distribution of a Dirac random variable with parameter g, $\mathcal{N}(\mu, \nu)$ a standard Gaussian with mean μ and variance ν, and $*$ stands for the convolution operator.

Lemma 4.2.14. *Under the measure $\mathbb{Q}_t$ defined in* (4.2.14)*, the random variable $(X_t/t)_t > 0$ converges weakly to the random variable Y where*

$$Y \sim \begin{cases} \delta\left(\Lambda'_-(1)\right) * \Gamma\left(\dfrac{2b}{\alpha}, -\dfrac{2b}{\alpha\Lambda'_-(1)}\right), & \text{if } \Lambda'_-(1) < 0; \\ \\ \mathcal{N}\left(\sqrt{\dfrac{2b\Lambda''_-(1)}{\alpha}}, \Lambda''_-(1)\right) * \Gamma\left(\dfrac{2b}{\alpha}, \sqrt{\dfrac{\alpha}{\Lambda''_-(1)}}\right), & \text{if } \Lambda'_-(1) = 0. \end{cases}$$

Proof. We first prove the lemma in the case $\Lambda'_-(1) < 0$. For all $t > 0$ and all ξ such that $u_t + \xi/t \in \mathcal{D}_t$, we can write

$$\log \mathbb{E}_{Q_t}\left(\exp\left(\xi \frac{X_t}{t}\right)\right) = \log \mathbb{E}_{Q_t}\left(\exp\left(\left(u_t + \frac{\xi}{t}\right) X_t - \Lambda_t(u_t)\right)\right) = \Lambda_t\left(u_t + \frac{\xi}{t}\right) - \Lambda_t(u_t).$$

From Lemma 4.2.13 and the fact that $\widetilde{\Lambda}'_t(u_t) = 0$, a Taylor expansion of Λ around 1 gives

$$\Lambda'_-(1) + (u_t - 1)\Lambda''_-(1) + \frac{2b}{\alpha t(1-u_t)} + \mathcal{O}\left((1-u_t)^2\right) = 0. \tag{4.2.16}$$

From (4.2.16) we have

$$\lim_{t\to\infty} t(1-u_t) = -\frac{2b}{\alpha\Lambda'_-(1)}, \tag{4.2.17}$$

and hence using Lemma 4.2.13,

$$\Lambda_t\left(u_t + \frac{\xi}{t}\right) - \Lambda_t(u_t) = t\left(\Lambda\left(u_t + \frac{\xi}{t}\right) - \Lambda(u_t)\right) - \frac{2b}{\alpha}\log\left(1 - \frac{t^{-1}\xi}{1-u_t}\right) + R_t\left(u_t + \frac{\xi}{t}\right) - R_t(u_t),$$

therefore the limit in (4.2.17) implies the equalities

$$t\left(\Lambda\left(u_t + \frac{\xi}{t}\right) - \Lambda(u_t)\right) = \xi\Lambda'_-(1) + o(1) \quad \text{and} \quad -\frac{2b}{\alpha}\log\left(1 - \frac{t^{-1}\xi}{1-u_t}\right) = -\frac{2b}{\alpha}\log\left(1 + \frac{\alpha\xi}{2b}\Lambda'_-(1)\right) + o(1),$$

as t tends to infinity. Since the function R_t converges on any compact subset of $\mathcal{D}_t$ we obtain

$$\lim_{t\to\infty} \mathbb{E}_{\mathbb{Q}_t}(\exp(\xi X_t/t)) = \xi\Lambda'_-(1) - \frac{2b}{\alpha}\log\left(1 + \frac{\alpha\xi}{2b}\Lambda'_-(1)\right),$$

which proves the lemma in the case $\Lambda'_-(1) < 0$.

If $\Lambda'_-(1) = 0$, then for all $t > 0$ and all ξ such that $u_t + \xi/\sqrt{t} \in \mathcal{D}_t$, we have

$$\log \mathbb{E}_{Q_t}\left(\exp\left(\xi \frac{X_t}{\sqrt{t}}\right)\right) = \log \mathbb{E}_{Q_t}\left(\exp\left(\left(u_t + \frac{\xi}{\sqrt{t}}\right) X_t - \Lambda_t(u_t)\right)\right) = \Lambda_t\left(u_t + \frac{\xi}{\sqrt{t}}\right) - \Lambda_t(u_t).$$

The expansion (4.2.16) now gives us

$$\lim_{t\to\infty} t(1-u_t)^2 = -\frac{2b}{\alpha\Lambda''_-(1)}, \tag{4.2.18}$$

and hence using Lemma 4.2.13 we again have

$$\Lambda_t\left(u_t+\frac{\xi}{\sqrt{t}}\right)-\Lambda_t\left(u_t\right)=t\left(\Lambda\left(u_t+\frac{\xi}{\sqrt{t}}\right)-\Lambda\left(u_t\right)\right)-\frac{2b}{\alpha}\log\left(1-\frac{t^{-1/2}\xi}{1-u_t}\right)+R_t\left(u_t+\frac{\xi}{\sqrt{t}}\right)-R_t\left(u_t\right),$$

therefore the limit in (4.2.18) implies

$$t\left(\Lambda\left(u_t+\frac{\xi}{\sqrt{t}}\right)-\Lambda\left(u_t\right)\right)=\frac{\xi^2}{2}\Lambda_-''\left(1\right)+\xi\sqrt{\frac{2b\Lambda_-''\left(1\right)}{\alpha}}+o\left(1\right),$$

and

$$-\frac{2b}{\alpha}\log\left(1-\frac{t^{-1/2}\xi}{1-u_t}\right)=-\frac{2b}{\alpha}\log\left(1-\xi\sqrt{\frac{\Lambda_-''\left(1\right)}{\alpha}}\right)+o\left(1\right).$$

Since the function R_t converges on any compact subset of $\mathcal{D}_t$ we obtain

$$\lim_{t\to\infty}\mathbb{E}_{\mathbb{Q}_t}\left(\exp\left(\xi X_t/t\right)\right)=\xi^2\Lambda_-''\left(1\right)+\xi\sqrt{\frac{2b\Lambda_-''\left(1\right)}{\alpha}}-\frac{2b}{\alpha}\log\left(1-\xi\sqrt{\frac{\Lambda_-''\left(1\right)}{\alpha}}\right),$$

which proves the lemma in the case $\Lambda_-'\left(1\right)=0$. □

4.2.3 From Heston to SVI

We focus here on the Heston model (3.2.1) (with $a=0$) where the functions Λ and Λ^* are available in closed-form. As mentioned on page 23 the correspondence between the Heston parameters and the affine parameters of (3.1.4) is $b=\kappa\theta>0$, $\beta=-\kappa<0$, $\alpha=\sigma^2$. In this section, we use the standard Heston parameters κ, θ, σ, ρ in order not to confuse the reader with the rest of the literature on the Heston model. Note that the condition $\chi(0)<0$ now reads $\kappa>0$ and $\chi(1)<0$ reads $\kappa-\rho\sigma>0$. The first condition is assumed a priori. The second condition is not that restrictive in practice on equity markets since the correlation ρ is usually negative. Proposition 4.2.17 below provides a clear and closed-form formula for the asymptotic implied volatility smile in the Heston model when the condition $\kappa-\rho\sigma>0$ is satisfied, and gives a partial result when this condition fails. Recall that from Proposition 4.2.7, Remark 4.2.8 and Proposition 4.2.4 we have

$$\Lambda\left(u\right)=-\frac{b}{\alpha}\left(\chi\left(u\right)+\gamma\left(u\right)\right),\quad\text{for all } u\in\mathcal{D}_\Lambda,\tag{4.2.19}$$

$$\Lambda^*\left(x\right)=u^*\left(x\right)x-\Lambda\left(u^*(x)\right),\quad\text{for all } x\in\mathbb{R},\tag{4.2.20}$$

$$u^*(x)=\frac{\sigma-2\kappa\rho+\left(\kappa\theta\rho+x\sigma\right)\eta\left(x^2\sigma^2+2x\kappa\theta\rho\sigma+\kappa^2\theta^2\right)^{-1/2}}{2\sigma\bar{\rho}^2},\quad\text{for all } x\in\mathbb{R},\tag{4.2.21}$$

where $u^*(x):=u_x$ is characterised in (4.2.8). From now on we use the notation $u^*(x)$ to emphasise that u^* is a function of x. The following trivial proposition summarises the properties of u^*.

Proposition 4.2.15. *The function $u^*:\mathbb{R}\to(u_-,u_+)$, where u_- and u_+ are defined in* (4.2.5) *and* (4.2.6)*, is strictly increasing, infinitely differentiable and satisfies the following properties*

$$u^*\left(-\theta/2\right)=0,\quad u^*\left(\bar{\theta}/2\right)=1,\quad\lim_{x\to-\infty}u^*(x)=u_-\quad\textit{and}\quad\lim_{x\to+\infty}u^*(x)=u_+,$$

as well as the equation

$$\Lambda'\left(u^*(x)\right)=x,\quad\textit{for all } x\in\mathbb{R}.\tag{4.2.22}$$

Proof. The equation (4.2.22) follows from the definition of the saddlepoint. The other properties in the proposition are a direct consequence of the explicit formula for u^* given in (4.2.21). □

Let $\bar{\theta} := \kappa\theta/(\kappa - \rho\sigma)$ and recall the following properties of Λ^*, which will help us later on.

(a) $\Lambda^{*'}(x) = u^*(x)$ for all $x \in \mathbb{R}$;

(b) $\Lambda^{*''}(x) > 0$ for all $x \in \mathbb{R}$;

(c) $x \mapsto \Lambda^*(x)$ is non-negative, has a unique minimum at $-\theta/2$ and $\Lambda^*(-\theta/2) = 0$;

(d) $x \mapsto \Lambda^*(x) - x$ is non-negative, has a unique minimum at $\bar{\theta}/2$ and $\Lambda^*\left(\bar{\theta}/2\right) = \bar{\theta}/2$.

Remark 4.2.16. Consider the change of variables $\gamma := \sqrt{\alpha + 4\beta^2 + \beta\rho\sqrt{\alpha}}/\left(2\sqrt{\alpha}\left(1-\rho^2\right)\right)$, $\eta := -\left(2\beta\rho + \sqrt{\alpha}\right)/\left(2\sqrt{\alpha}\left(1-\rho^2\right)\right)$, $\mu := -b\rho/\sqrt{\alpha}$ and $\delta := b\sqrt{1-\rho^2}/\sqrt{\alpha}$. When $a = 0$ the limiting logarithmic Laplace transform Λ corresponds exactly to the Laplace transform Λ_{NIG} of a Normal inverse Gaussian $\Lambda_{\mathrm{NIG}}(u) = \mu u + \delta\left(\sqrt{\gamma^2 - \eta^2} - \sqrt{\gamma^2 - (\eta + u)^2}\right)$. Since the proof of Proposition 4.1.3 follows entirely from the behaviour of the function Λ, this means that the Heston model has the same rate function Λ^* and the same asymptotic implied volatility $\hat{\sigma}_\infty$ as the Normal Inverse Gaussian model (see also [6], [7], [68]).

In [48], Jim Gatheral proposed the following parameterisation for the implied squared volatility

$$\hat{\sigma}^2_{\mathrm{SVI}}(x) = \frac{\omega_1}{2}\left(1 + \omega_2\rho x + \sqrt{\left(\omega_2 x + \rho\right)^2 + 1 - \rho^2}\right), \quad \text{for all } x \in \mathbb{R}, \tag{4.2.23}$$

where x represents the time-scaled log-moneyness. We prove here by an appropriate change of variables that this SVI implied volatility parameterisation and the large-time asymptotic of the Heston implied volatility proved in Proposition 4.1.3 agree algebraically, thus confirming a conjecture from [48] as well as providing a simpler expression for the asymptotic implied volatility in the Heston model. We show how this result can help in interpreting SVI parameters. Let us first give an intuitive motivation for the conjecture in [48] that the limit as the maturity tends to infinity of the Heston volatility smile should be SVI. Consider [49, Equation (5.7)] which relates the implied volatility $\sigma_{\mathrm{BS}}(k,t)$ at log-moneyness k and maturity t to the characteristic function ϕ_t of the log-stock price. We rewrite this equation in the form

$$\int_{-\infty}^{\infty} \frac{\exp(-\mathtt{i}uk)\,\mathrm{d}u}{u^2 + 1/4}\phi_T\left(u - \frac{\mathtt{i}}{2}\right) = \int_{-\infty}^{\infty} \frac{\exp(-\mathtt{i}uk)\,\mathrm{d}u}{u^2 + 1/4}\exp\left(-\frac{t}{2}\left(u^2 + \frac{1}{4}\right)\sigma^2_{\mathrm{BS}}(k,t)\right). \tag{4.2.24}$$

In the limit as the maturity t tends to infinity, the Heston characteristic function has the form $\phi_t(u - \mathtt{i}/2) \sim \exp(-\psi(u)t)$. Then as pointed out in [73, page 186], we may apply the saddlepoint method to both sides in equation (4.2.24) to obtain

$$\mathrm{e}^{-\mathtt{i}k\tilde{u}}\frac{\exp(-\psi(\tilde{u})t)}{\tilde{u}^2 + 1/4}\sqrt{\frac{2\pi}{\psi''(\tilde{u})t}} \sim 4\exp\left(-\frac{\zeta t}{8} - \frac{k^2}{2\zeta t}\right)\sqrt{\frac{2\pi}{\zeta t}}, \tag{4.2.25}$$

where ζ is short-form notation for the total Black-Scholes implied variance $\sigma_{\mathrm{BS}}^2(k,t)\,t$ and $\tilde{u}$ is such that $\psi'(\tilde{u}) = -\mathtt{i}k/t$, so that $\tilde{u}$—which is in general a function of k—is a saddlepoint, which in the Heston model at least, may be computed explicitly. Defining $k := xT$ and equating the arguments of the exponentials in equation (4.2.25), the dependence on t cancels and we obtain

$$\frac{\zeta(x)}{8} + \frac{x^2}{2\zeta(x)} = \psi(\tilde{u}(x)) + \mathtt{i}x\tilde{u}(x), \tag{4.2.26}$$

where we have reinstated explicit dependence on x for emphasis. One can verify that in the limit as the maturity t tends to infinity and with a suitable change of parameters, expression (4.2.23) exactly solves the saddlepoint condition (4.2.26):

$$\frac{\hat{\sigma}_{\mathrm{SVI}}^2(x)}{8} + \frac{x^2}{2\hat{\sigma}_{\mathrm{SVI}}^2(x)} = \psi(\tilde{u}(x)) + \mathtt{i}x\tilde{u}(x).$$

We are thus led to conjecture that the implied variance under the Heston model converges to the SVI parameterisation as the maturity T tends to infinity.

Consider now the following choice of SVI parameters in terms of the Heston parameters,

$$\omega_1 := \frac{4\kappa\theta}{\sigma^2(1-\rho^2)}\left(\sqrt{(2\kappa-\rho\sigma)^2 + \sigma^2(1-\rho^2)} - (2\kappa-\rho\sigma)\right), \quad \text{and} \quad \omega_2 := \frac{\sigma}{\kappa\theta}. \tag{4.2.27}$$

and define $\eta := \sqrt{4\kappa^2 + \sigma^2 - 4\kappa\rho\sigma}$. The following proposition makes the intuition above precise.

Proposition 4.2.17. *Let $\hat{\sigma}_\infty$ be as in (4.1.5) and consider the change of parameters (4.2.27).*

(i) If $\kappa - \rho\sigma > 0$ then $\hat{\sigma}_{\mathrm{SVI}}^2(x) = \hat{\sigma}_\infty^2(x)$ for all $x \in \mathbb{R}$.

(ii) If $\kappa - \rho\sigma \le 0$ then $\hat{\sigma}_{\mathrm{SVI}}^2(x) = \hat{\sigma}_\infty^2(x)$ for all $x \le \Lambda'(0)$.

Proof. Statement (ii) follows directly from (i), from Theorem 4.2.10 and from Remark 4.2.12. Let us introduce the notation $\Delta(x) := \left(\sigma^2x^2 + 2\kappa\theta\rho\sigma x + \kappa^2\theta^2\right)^{1/2}$. Under the change of variables (4.2.27), the SVI implied variance takes the form

$$\hat{\sigma}_{\mathrm{SVI}}^2(x) = \frac{2}{\sigma^2\bar{\rho}^2}\Big(\eta - (2\kappa-\rho\sigma)\Big)\Big(\kappa\theta + \rho\sigma x + \Delta(x)\Big), \quad \text{for all } x \in \mathbb{R}. \tag{4.2.28}$$

We now move on to simplify the expression for $\hat{\sigma}_\infty^2$ as expressed in Proposition 4.1.3. We first start by the expression for $\Lambda^*(x)$ in (4.2.20).

$$\Lambda^*(x) = \frac{A(x)\Delta(x) + B(x)\eta}{2\sigma^2\bar{\rho}^2\Delta(x)},$$

with

$$A(x) := x\sigma^2 - 2x\kappa\rho\sigma - 2\kappa^2\theta + \kappa\theta\rho\sigma, \qquad \text{and} \qquad B(x) := 2x\sigma\kappa\theta\rho + x^2\sigma^2 + \kappa^2\theta^2\rho^2 + \kappa^2\theta^2\bar{\rho}^2.$$

Note that $B(x) = \Delta^2(x)$, so that $\Lambda^*(x) = (A(x) + \Delta(x)\eta)\,/\left(2\sigma^2\bar{\rho}^2\right)$. We further have

$$2\Lambda^*(x) - x = \frac{A(x) + \Delta(x)\eta - x\sigma^2\bar{\rho}^2}{\sigma^2\bar{\rho}^2} = \frac{\Delta(x)\eta - (2\kappa-\rho\sigma)(\kappa\theta + x\rho\sigma)}{\sigma^2\bar{\rho}^2}, \tag{4.2.29}$$

where we use the factorisation $A(x) - x\sigma^2\bar{\rho}^2 = -(2\kappa - \rho\sigma)(\kappa\theta + x\rho\sigma)$. Returning back to the expression for $\hat{\sigma}_\infty$, let us define $\Psi(x) := \Lambda^*(x)(\Lambda^*(x) - x)$, so that we have

$$\Psi(x) = \left(\frac{\Delta(x)\eta}{2\sigma^2\bar{\rho}^2}\right)^2 + \alpha(x)\Delta(x) + \beta(x),$$

where

$$\alpha(x) := -\frac{\eta(2\kappa - \rho\sigma)(\kappa\theta + x\rho\sigma)}{2\sigma^4\bar{\rho}^4}, \qquad \text{and} \qquad \beta(x) := \frac{(2\kappa - \rho\sigma)^2(\kappa\theta + x\rho\sigma)^2 - x^2\sigma^4\bar{\rho}^4}{4\sigma^4\bar{\rho}^4}.$$

Using the factorisations $\Delta^2(x) = (\kappa\theta + x\rho\sigma)^2 + x^2\sigma^2\bar{\rho}^2$ and $\eta^2 = (2\kappa - \rho\sigma)^2 + \sigma^2\bar{\rho}^2$, we can write $\beta(x) = (4\sigma^4\bar{\rho}^4)^{-1}\left((2\kappa - \rho\sigma)^2\Delta^2(x) - x^2\sigma^2\bar{\rho}^2\eta^2\right)$ and some rearrangements lead to

$$\Psi(x) = \frac{1}{4\sigma^4\bar{\rho}^4}\left\{\eta(\kappa\theta + x\rho\sigma) - (2\kappa - \rho\sigma)\Delta(x)\right\}^2, \tag{4.2.30}$$

with $a(x) := 4\sigma^4\bar{\rho}^4\alpha(x)$. To complete the proof we need to take the square root of $\Psi(x)$. Note that

$$\eta(\kappa\theta + x\rho\sigma) - (2\kappa - \rho\sigma)\Delta(x) = \sqrt{\gamma(x) + \sigma^2\bar{\rho}^2(\kappa\theta + x\rho\sigma)^2} - \sqrt{\gamma(x) + x^2\sigma^2\bar{\rho}^2(2\kappa - \rho\sigma)^2},$$

where $\gamma(x) := (2\kappa - \rho\sigma)^2(\kappa\theta + x\rho\sigma)^2$. Since $\gamma(x)$ is non negative for all x, the sign of the whole expression is simply given by the sign of the difference $\psi(x) := \sigma^2\bar{\rho}^2(\kappa\theta + x\rho\sigma)^2 - x^2\sigma^2\bar{\rho}^2(2\kappa - \rho\sigma)^2$. The function ψ also reads $\psi(x) = \kappa\sigma^2\bar{\rho}^2(2x + \theta)(2x\rho\sigma + \kappa\theta - 2\kappa x)$, and hence this polynomial has exactly two real roots $-\theta/2$ and $\bar{\theta}/2$, and its second-order coefficient $-4\kappa\sigma^2\bar{\rho}^2(\kappa - \rho\sigma)$ is strictly negative whenever $\kappa - \rho\sigma > 0$. So, plugging (4.2.29) and (4.2.30) into the expression for $\hat{\sigma}_\infty$ in Proposition 4.1.3, we obtain (4.2.28) and the proposition follows. □

We now use Proposition 4.2.17 to help interpret the SVI parameters. From [48], the standard SVI parameterisation in terms of the log-strike k reads

$$\sigma^2_{\mathrm{SVI}}(k) = a + b\left\{\tilde{\rho}(k - m) + \sqrt{(k - m)^2 + \xi^2}\right\}. \tag{4.2.31}$$

Equating (4.2.31) with (4.2.23) and with the parameter choice (4.2.27), we find the following correspondence between SVI parameters and Heston parameters;

$$a = \frac{\omega_1}{2}\left(1 - \rho^2\right), \qquad b = \frac{\omega_1\omega_2}{2t}, \qquad \tilde{\rho} = \rho, \qquad m = -\frac{\rho t}{\omega_2}, \qquad \xi = \frac{t\sqrt{1 - \rho^2}}{\omega_2}. \tag{4.2.32}$$

For concreteness, imagine that we are given an SVI fit to the implied volatility smile generated from the Heston model with t very large so that we have the SVI parameters a, b, $\tilde{\rho}$, m and ξ. Our first observation is that the SVI parameter $\tilde{\rho}$ is exactly the correlation ρ between changes in the instantaneous variance ζ and changes in the underlying process, i.e we can read off the correlation directly from the orientation of the smile. In particular, the smile is symmetric when $\rho = 0$. The parameter b gives the angle between the asymptotes of the implied variance smile. We see from (4.2.31) and (4.2.32) that the angle between the asymptotes of the total variance smile $\sigma^2_{\mathrm{SVI}}(k)\,t$

is constant for large t but that the overall level increases with t. From equation (4.2.23), ω_1 is the at-the-money implied variance $\sigma^2_{\mathrm{SVI}}(0)\,t$. From (4.2.27), in the limit $\sigma \ll \kappa$, we have

$$\omega_1 = \theta\left\{1 + \frac{\rho\sigma}{\kappa} + \mathcal{O}\left(\left(\frac{\sigma}{\kappa}\right)^2\right)\right\},$$

so that the at-the-money volatility is given directly by θ when the volatility of volatility is small. In the limit $\sigma \gg \kappa$, we have

$$\omega_1 = \frac{4\kappa\theta}{\sigma(1-\rho)}\left\{1 - \frac{2\kappa}{\sigma} + \mathcal{O}\left(\left(\frac{\kappa}{\sigma}\right)^2\right)\right\},$$

showing that at-the-money volatility decreases as the volatility-of-volatility increases and as the volatility becomes less correlated with the underlying. Finally, the minimum of the variance smile is attained at $x = -2\rho/\omega_2$, providing a simple interpretation of the parameter ω_2. In particular, if $\rho = 0$, this minimum is exactly the at-the-money point. The minimum shifts to the upside $x > 0$ if $\rho < 0$ and to the downside $x < 0$ if $\rho > 0$.

4.2.4 Moment explosions and wings

In [72], Lee initiated a whole stream of research on the asymptotics of the implied volatility. He provided a rigorous explanation of the relation between moment explosions of the asset price process under no-arbitrage conditions and the slope of the implied volatility smile in the wings, i.e. for very large and very small strikes. Benaim & Friz ([6] and [7]) sharpened Lee's results and provided some more precise results for exponential (time-changed) Lévy processes. The following theorem states Lee's result for the right wing of the implied volatility smile. A similar result holds for the left side of the smile.

Theorem 4.2.18. *([72, Theorem 3.2]*
For any maturity $t > 0$, define the moment explosion $u^* := \sup\left\{u : \mathbb{E}\left(S_t^{u+1}\right) < \infty\right\}$ as well as the right slope of the implied (squared) volatility smile $\beta_R := \limsup_{x\to\infty} \sigma_t^2(x)\,t/x$. Then $\beta_R = 2 - 4\left(\sqrt{u^*(u^*-1)} - u^*\right) \in [0,2]$.

To be able to use this result in practice one henceforth needs the explosion times of the process, which are derived in [3] for some classes of stochastic volatility models. In the Heston case we are able to provide more precise results, which we state in the following corollary:

Proposition 4.2.19. *If the condition $\kappa > \rho\sigma$ is satisfied then for any $t > 0$ the equality*

$$\beta_R(t) = 2 - 4\left(\sqrt{u^*(t)\,(u^*(t)-1)} - (u^*(t)-1)\right),$$

holds, where $u^*(t) > 1$ can also be characterised as the unique positive solution of the equation $t^*(u^*(t)) = t$ where the explosion time t^* is defined in (3.2.7)*.*

The equation for $u^*(t)$ in the proposition above, though not in closed form, is straightforward to solve numerically. When t tends to infinity, $u^*(t)$ tends to the strictly positive root $u_+ \geq 1$ of $\delta(u)$ (defined just after (3.2.7)) and the formula boils down to the following: As the maturity t tends to infinity, the following closed-form formula holds:

$$\beta_R(\infty) = \frac{2}{\sigma(1-\rho)}\left(\sqrt{(2\kappa-\rho\sigma)^2 + \sigma^2(1-\rho^2)} - (2\kappa-\rho\sigma)\right). \tag{4.2.33}$$

The proofs of Proposition 4.2.19 and Formula (4.2.33) follow directly from [3, Proposition 3.1 and Corollary 6.2] as well as [72]. The left slope of the asymptotic smile is obtained easily using the symmetry property detailed in Section 3.2. Very recently Gulisashvili, in [54] and [55], and Friz et al. [46] provided sharp asymptotic estimates for the wings of the implied volatility smile, thus refining Lee's result in the Heston case. The above results were stated under the condition $\kappa - \rho\sigma > 0$. It is also interesting to note that from this limiting slope and Proposition 4.2.17 the limit (as t tends to infinity) and the derivative in the wings can be interchanged. When the condition $\kappa - \rho\sigma > 0$ is not satisfied, one can see that Proposition 4.2.19 still holds and in this case we have $\beta_R(\infty) = 2$. Indeed, as t tends to infinity, the unique root $\omega^*(t)$ of $t^*(\omega^*(t)) = t$ tends to 1. So that the result follows from Proposition 4.2.19 with $\omega^*(t) = 1$.

4.3 Beyond continuous affine stochastic volatility models

We have stated and proved above a large deviations principle for the Heston model, from which we were able to determine a closed-form expression for the implied volatility smile as the maturity tends to infinity. It is a natural step forward to try and extend this methodology to other and larger classes of models. In Remark 4.2.8 (iv) we have already mentioned a simple extension, namely the multidimensional Heston model. We also present a simple extension to a non-affine case: the Schöbel-Zhu model. The reason for presenting these two results is to highlight the fact that the Heston model is in some sense a canonical model for continuous stochastic volatility models and hence a natural one to work with at first. Finally, in Section 4.3.2 we extend our results to affine stochastic volatility models with jumps, which for instance include exponential Lévy processes.

4.3.1 A non-affine alternative: the Schöbel-Zhu model

We present here a simple alternative to the Heston model. Since the analysis is very similar than the one for the Heston model above, we shall skip some details. We highlight however how the study of the implied volatility for large maturities can be immediately derived from the one in the Heston model. As introduced in [93] the Schöbel-Zhu stochastic volatility model is an extension to non zero spot-volatility correlation of the Stein & Stein [99] model and the logarithmic spot price process $(X_t)_{t\geq 0}$ satisfies the following system of SDEs

$$\begin{aligned}
\mathrm{d}X_t &= -\tfrac{1}{2}\sigma_t^2\mathrm{d}t + \sigma_t \mathrm{d}W_t, && X_0 = x_0 \in \mathbb{R},\\
\mathrm{d}\sigma_t &= \kappa\left(\theta - \sigma_t\right)\mathrm{d}t + \xi \mathrm{d}Z_t, && \sigma_0 \in \mathbb{R}_+^*,\\
\mathrm{d}\left\langle W, Z\right\rangle_t &= \rho \mathrm{d}t,
\end{aligned}$$

where κ, θ and ξ are strictly positive real numbers, $\rho \in [-1,1]$ and $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are two standard Brownian motions. The logarithmic characteristic function of the process defined by $\Phi_t(u) := \log \mathbb{E}\left(\mathrm{e}^{\mathtt{i}u(X_t - x_0)}\right)$ reads (see [65])

$$\Phi_t(u) = A(u,t) + B_\sigma(u,t)\,\sigma_0 + B_v(u,t)\,\sigma_0^2, \quad \text{for all } u \in \mathbb{C}, t \geq 0, \tag{4.3.1}$$

where

$$\begin{aligned}
A(u,t) &:= A_\sigma(u,t) + \frac{t}{4}\left(b_u - d_u\right) - \frac{1}{2}\log\left(\frac{\gamma_u \exp\left(-d_u t\right) - 1}{\gamma_u - 1}\right),\\
B_\sigma(u,t) &:= \frac{\kappa\theta}{\xi^2}\frac{b_u - d_u}{d_u}\frac{\left(1 - \exp\left(-d_u t/2\right)\right)^2}{1 - \gamma_u \exp\left(-d_u t\right)}, \qquad B_v(u,t) := \frac{b_u - d_u}{4\xi^2}\frac{1 - \exp\left(-d_u t\right)}{1 - \gamma_u \exp\left(-d_u t\right)},\\
a_u &:= -\frac{u}{2}\left(u + \mathtt{i}\right), \quad b_u := 2\left(\kappa - \mathtt{i}\rho\xi u\right), \quad d_u := \left(b_u^2 - 8\xi^2 a_u\right)^{1/2}, \quad \gamma_u := \frac{b_u - d_u}{b_u + d_u},\\
A_\sigma(u,t) &:= \frac{\kappa^2\theta^2}{2d_u^3\xi^2}\left(b_u - d_u\right)\left(b_u\left(d_u t - 4\right) + d_u\left(d_u t - 2\right) + 4\frac{2b_u + \frac{d_u^2 - 2b_u^2}{b_u + d_u}\exp\left(-\frac{1}{2}d_u t\right)}{1 - \gamma_u \exp\left(-d_u t\right)}\mathrm{e}^{-\frac{1}{2}d_u t}\right).
\end{aligned}$$

It is clear from the form of the characteristic function that this model is not affine in the sense of Section 3.1. In the large deviations framework developed in the preceding sections we are interested in the behaviour of the function $\Lambda_{\mathrm{SZ}} : u \mapsto \lim_{t\to\infty} t^{-1}\Phi_t(-\mathtt{i}u)$, for all $u \in \mathbb{R}$ such that it is well defined as an extended real number. From the representation of the characteristic function above, it is easy to show that the only contribution comes from the function A and we have

$$\Lambda_{\mathrm{SZ}}(u) = \frac{\kappa^2\theta^2}{2\xi^2}\frac{\beta_u^2 - \delta_u^2}{\delta_u} + \frac{1}{2}\left(\beta_u - \delta_u\right),$$

where

$$\beta_u := \left(\kappa - \rho\xi u\right), \qquad \text{and} \qquad \delta_u := \left(\beta_u^2 - \xi^2 u\left(u - 1\right)\right)^{1/2}.$$

If Λ_H denote the limiting normalised cumulant generating function of the Heston model, then

$$\Lambda_{\mathrm{SZ}}(u) = \Lambda_H(u)\frac{\beta_u + \delta_u}{2\delta_u} + \frac{1}{2}\left(\beta_u - \delta_u\right), \quad \text{for all } u \in \left(u_-, u_+\right),$$

where $u_\pm := \frac{\xi - 2\kappa\rho \pm \sqrt{(2\kappa\rho - \xi)^2 + 4\kappa^2(1-\rho^2)}}{2\xi(1-\rho^2)}$ are the same bounds as in the Heston model (see (4.2.5) and (4.2.6)). Note further that the condition under which a full LDP is available is exactly the same as in the Heston case, i.e. $\kappa - \rho\xi > 0$. However, contrary to the Heston model the Fenchel-Legendre transform of the function Λ is not here available in closed-form, but can be numerically determined by a simple root-finding algorithm.

4.3.2 Affine stochastic volatility with jumps

In this section we extend the large-time asymptotic behaviour we have proved above to the class of affine stochastic volatility models with jumps defined in Section 3.1. From (3.1.1) we see that for all $t \geq 0$ the equality $\Lambda_t(u) = \Phi_t(u, 0)$ holds for all real u such that both sides are well defined. We assume that the process $(\exp(X_t))_{t\geq 0}$ is a martingale (Assumption A.3) so that the function Λ_t is well defined at least on $[0, 1]$ for all $t \geq 0$. Define now the function $\chi : \mathbb{R} \to \mathbb{R}$ by $\chi(u) := \partial_w R(u, w)|_{w=0}$, for all $u \in \mathbb{R}$. As in [68], we make the additional assumption

A5. $\chi(0) < 0$ and $\chi(1) < 0$.

In the continuous case this assumption enabled us to state a full large deviations principle for the process $(X_t/t)_{t\geq 1}$ as t tends to infinity. It is hence natural to see this condition here as well. However it might not be a sufficient condition when jumps arise. From Lemma 3.2 in [68], under Assumption A5 there exist a maximal interval $\mathcal{I} \subset \mathbb{R}$ and a unique function $w \in C(\mathcal{I}) \cap C^1(\mathcal{I}^\circ)$ such that $R(u, w(u)) = 0$ for all $u \in \mathcal{I}$ with $w(0) = w(1) = 0$. Define now the set $\mathcal{J} := \{u \in \mathcal{I} : F(u, w(u)) < \infty\}$ as well as the function $\Lambda : \mathcal{J} \to \mathbb{R}$ by $\Lambda(u) := F(u, w(u))$, then

$$\lim_{t\to\infty} t^{-1}\Lambda_t(u) = \lim_{t\to\infty} t^{-1}\phi(t, u, 0) = \Lambda(u), \quad \text{for all } u \in \mathcal{J}, \tag{4.3.2}$$

$$\lim_{t\to\infty} \psi(t, u, 0) = w(u), \quad \text{for all } u \in \mathcal{I}, \tag{4.3.3}$$

and $[0, 1] \subseteq \mathcal{J} \subseteq \mathcal{I}$. By construction, the function w is continuous and convex on its effective domain $\mathcal{I}$; by Lemma 2.2 in [68], the function F is convex as well on its effective domain, so that Λ is convex on the interval $\mathcal{J} := \{u \in \mathcal{I} : F(u, w(u)) < \infty\}$.

Remark 4.3.1. In the general case (3.1.3) implies

$$\partial_u R(u, w) = \alpha_{11}u + \alpha_{12}w + \beta + \int_D x\left(\mathrm{e}^{xu+yw} - \frac{1}{1+x^2}\right)\mu(\mathrm{d}x, \mathrm{d}y),$$
$$\partial_w R(u, w) = \alpha_{22}w + \alpha_{12}u + \beta + \int_D y\left(\mathrm{e}^{xu+yw} - \frac{1}{1+y^2}\right)\mu(\mathrm{d}x, \mathrm{d}y).$$

The function $w : \mathbb{R} \to \mathbb{R}$ is defined as the unique solution to the equation $R(u, w(u)) = 0$, and hence can implicitly represented as

$$w(u) = -\int^u \frac{\partial_p R(p, w(p))}{\partial_w R(p, w(p))}\mathrm{d}p, \quad \text{for all } u \in \mathcal{I}.$$

If we assume no variance-dependent jump in the measure μ, i.e. $\mu(\mathrm{d}x, \mathrm{d}y) = \mu(\mathrm{d}x) \otimes \delta_0(\mathrm{d}y)$ (δ_0 represents the Dirac measure with mass at the origin), then the equality $R(u, w(u)) = 0$ reads

$$\frac{1}{2}\alpha_{22}w(u)^2 + (\alpha_{12}u + \beta_2)w(u) + g(u) = 0, \tag{4.3.4}$$

where the function $g : \mathcal{I} \to \mathbb{R}$ is defined by

$$g(u) := \frac{1}{2}\alpha_{11}u^2 + \beta_1 u + \int_{\mathbb{R}^*}\left(\mathrm{e}^{ux} - 1 - \frac{ux}{1+x^2}\right)\mu(\mathrm{d}x), \quad \text{for all } u \in \mathcal{I}. \tag{4.3.5}$$

We can now solve equation (4.3.4) explicitly as

$$w(u) = -\alpha_{22}^{-1}\left(\alpha_{12}u + \beta_2 \mp \sqrt{(\alpha_{12}u + \beta_2)^2 - 2\alpha_{22}g(u)}\right), \quad \text{for all } u \in \mathcal{I}. \tag{4.3.6}$$

Indeed, we know the square root exists because the equation $R(u, w(u)) = 0$ has a unique solution for all $u \in \mathcal{I}$ with $w(0) = w(1) = 0$, by Lemma 3.2 in [68]. From (4.3.5), we have $g(0) = 0$, so that $w(u) = 0$ if and only if we take the positive root in (4.3.6) when β_2 is negative and the negative root when β_2 is positive; when $\beta_2 = 0$, $w(0) = 0$ if and only if $\alpha_{12} = 0$, so that

$$w(u) = -\alpha_{22}^{-1}\left(\alpha_{12}u + \beta_2 - \operatorname{sgn}(\beta_2)\sqrt{(\alpha_{12}u + \beta_2)^2 - 2\alpha_{22}g(u)}\right), \quad \text{for all } u \in \mathcal{I},$$

and $w(u)$ is everywhere null if $\beta_2 = 0$. Hence we can write the function Λ explicitly as

$$\Lambda(u) := F(u, w(u)) = \frac{1}{2}a_{11}u^2 + b(u, w(u))' + \int_{D\setminus\{0\}}\left(\mathrm{e}^{xu+yw(u)} - 1 - \frac{ux}{1+x^2}\right) m(\mathrm{d}x, \mathrm{d}y).$$

Let us now define the Fenchel-Legendre transform $\Lambda^* : \mathbb{R} \to \mathbb{R}_+$ of the function Λ as

$$\Lambda^*(x) := \sup_{u\in\mathcal{J}} \{ux - \Lambda(u)\}, \quad \text{for all } x \in \mathbb{R} \tag{4.3.7}$$

By the smoothness properties of Λ, we have $\Lambda^*(x) = xu^*(x) - \Lambda(u^*(x))$ for all $x \in \mathbb{R}$, where $u^*(x)$ satisfies the equation $\Lambda'(u^*(x)) = x$. As before we define the Share measure $\widetilde{\mathbb{P}}$ as the probability taking the normalised share price process $(\exp(X_t - x_0))_{t\geq 0}$ as the numeraire. The equality

$$\widetilde{\Phi}_t(u, w) := \mathbb{E}_{\widetilde{\mathbb{P}}}\left(\mathrm{e}^{u(X_t - x_0) + wV_t}\Big| x_0, v\right) = \mathbb{E}_{\mathbb{P}}\left(\mathrm{e}^{u(X_t - x_0) + wV_t}\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\Bigg| x_0, v\right) = \Phi_t(u+1, w),$$

therefore holds for all $t, u, w \in \mathbb{R}_+ \times \mathbb{C}^2$ such that the expectation exists. Let us define the set $\widetilde{\mathcal{I}} := \{u \in \mathbb{R} : (u+1) \in \mathcal{I}\}$ as well as the function $\widetilde{\Lambda}_t : \widetilde{\mathcal{I}} \to \mathbb{R}$ for all $t > 0$ by

$$\widetilde{\Lambda}_t(u) := \log \mathbb{E}_{\widetilde{\mathbb{P}}}(\exp(u(X_t - x_0))) = \widetilde{\Phi}_t(u, 0) = \Lambda_t(u+1), \quad \text{for all } t, u \in \mathbb{R}_+ \times \widetilde{\mathcal{I}}.$$

We can therefore rewrite the equalities in (4.1.1) in terms of the effective domains $\widetilde{\mathcal{I}}$ and $\widetilde{\mathcal{J}}$. The properties of Λ and Λ^* translate directly into those of the functions $\widetilde{\Lambda}$ and $\widetilde{\Lambda}^*$ where

$$\widetilde{\Lambda}(u) := \lim_{t\to\infty} t^{-1}\widetilde{\Lambda}_t(u), \text{ for all } u \in \widetilde{\mathcal{J}}, \quad \text{and} \quad \widetilde{\Lambda}^*(x) := \sup_{u\in\widetilde{\mathcal{J}}}\left\{ux - \widetilde{\Lambda}(u)\right\}, \text{ for all } x \in \mathbb{R}, \tag{4.3.8}$$

with $\widetilde{\mathcal{J}} := \{u \in \mathbb{R} : (u+1) \in \mathcal{J}\}$ and the following trivial lemma holds

Lemma 4.3.2. *The following two equalities hold*

$$\widetilde{\Lambda}(u) = \Lambda(u+1) \quad \textit{for all } u \in \widetilde{\mathcal{J}}, \qquad \textit{and} \qquad \widetilde{\Lambda}^*(x) = \Lambda^*(x) - x \quad \textit{for all } x \in \mathbb{R}.$$

Let us now make one final assumption.

Assumption 4.3.3. The function Λ is essentially smooth on its effective domain $\mathcal{D}_\Lambda$ and there exist two real numbers $\varepsilon_0, \varepsilon_1 > 0$ such that $[-\varepsilon_0, 1+\varepsilon_1] \subset \mathcal{D}_\Lambda$.

This assumption is fundamental for our analysis. Keller-Ressel [68] has shown that under Assumptions A1-A5 the limiting logarithmic Laplace transform Λ in (4.3.2) is the cumulant generating function of an infinitely divisible random variable. In Proposition 4.2.4 we have seen that the limiting Laplace transform is indeed essentially smooth on its effective domain which is larger than $[0,1]$ if and only if $\chi(0) < 0$ and $\chi(1) < 0$. Keller-Ressel also proved that the interval $[0,1]$ is contained in the effective domain of Λ but we do not know whether this domain is actually larger than this interval or not. In the continuous case, we have seen that this domain can be precisely $[0,1]$ when Assumption A5 is not satisfied. The link between the essential smoothness of Λ and the conditions on $\chi(0)$ and $\chi(1)$ is not entirely clear. In particular, as we shall see later, Assumption A5 is trivially not satisfied for exponential Lévy processes, but the Lévy exponent might or might not be essentially smooth on its effective domain. Under these assumptions, the functions Λ^* and $\widetilde{\Lambda}^*$ are both good rate functions on $\mathbb{R}$, they are strictly convex and admit a unique minimum attained respectively at x^* and $\tilde{x}^*$. By Lemma 4.3.2 and the fact that the two good rate functions Λ^* and $\widetilde{\Lambda}^*$ are non-negative everywhere, it is straightforward to see that the inequalities $x^* < 0 < \tilde{x}^*$ are satisfied, and the following theorem holds.

Theorem 4.3.4. *Under Assumptions A1-A5 and Assumption 4.3.3 the process* $\left(t^{-1}\left(X_t - x_0\right)\right)_{t\geq 1}$ *satisfies, as t tends to infinity,*

(i) a large deviations principle under $\mathbb{P}$ with the good rate function Λ^* defined in (4.3.7)*;*

(ii) a large deviations principle under $\widetilde{\mathbb{P}}$ with the good rate function $\widetilde{\Lambda}^*$ defined in (4.3.8)*.*

We can now apply the results of Section 4.1 which gives us the asymptotic behaviours of option prices (Theorem 4.1.1), and of the implied volatility (Proposition 4.1.3).

One-dimensional exponential Lévy processes

We now move on to deriving similar results for exponential Lévy processes. Let $(X_t)_{t\geq 0}$ be a Lévy process taking values in $\mathbb{R}$ with characteristic triplet (σ, η, ν). From [92], we know that for all $t \geq 0$, its logarithmic Laplace transform $\Phi_t(u) := \log \mathbb{E}\left(\exp\left(uX_t\right)\right)$ reads (for ease of notation, we drop the w dependence in formula (3.1.1))

$$\Phi_t(u) = t\phi_X(u), \quad \text{for all } u \in \mathcal{D}_X,$$

where $\mathcal{D}_X := \{u \in \mathbb{R} : \phi_X(u) < \infty\}$ represents the effective domain of ϕ_X and where the function $\phi_X : \mathcal{D}_X \to \mathbb{R}$ has the Lévy-Khintchine representation

$$\phi_X(u) := \frac{\sigma^2}{2}u^2 + \eta u + \int_{\mathbb{R}} \left(\mathrm{e}^{ux} - 1 - ux\mathbb{1}_{\{|x|\leq 1\}}\right)\nu(\mathrm{d}x), \quad \text{for all } u \in \mathcal{D}_X, \tag{4.3.9}$$

where $\sigma > 0$, $\eta \in \mathbb{R}$ and ν is a Lévy measure defined on $\mathbb{R}$ such that

$$\nu(\{0\}) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}} \left(x^2 \wedge 1\right)\nu(\mathrm{d}x) < \infty. \tag{4.3.10}$$

From Assumption A3 on page 22, we need the process $\left(\mathrm{e}^{X_t}\right)$ to be a martingale. According to [23, Proposition 3.18], this holds if and only if

$$\int_{|x|\geq 1} \mathrm{e}^x \nu(\mathrm{d}x) < \infty \qquad \text{and} \qquad \frac{\sigma^2}{2} + \eta + \int_{\mathbb{R}} \left(\mathrm{e}^x - 1 - x\mathbb{1}_{\{|x|\leq 1\}}\right) \nu(\mathrm{d}x) = 0. \tag{4.3.11}$$

From the notations for affine stochastic volatility models in (3.1.1), (3.1.2) and (3.1.3), we have $\phi_X(u)t = \phi(t,u,0)$ and $\psi(t,u,0) = 0$, so that the functions F and R simply read

$$F(u,0) = \phi_X(u) \qquad \text{and} \qquad R(u,0) = 0.$$

Note here that Condition A5 on page 60 is not satisfied since the function χ is null everywhere. This is actually a borderline case, and the analysis can still be carried out with some precautions. In our case, this means that we can not define the function w as in (4.3.3), but we can however directly work with the function F. Note that a direct application of Jensen's inequality proves that the function ϕ_X is convex on its effective domain $\mathcal{D}_X$ and we immediately have $\lim_{t\to\infty} t^{-1}\log\mathbb{E}\left(\mathrm{e}^{uX_t}\right) = \phi_X(u)$, for all $u \in \mathcal{D}_X$, and hence the function ϕ_X is exactly the limiting cumulant generating function Λ defined in (4.3.2). Now, note that from (4.3.9), the function F can be rewritten as

$$\Lambda(u) = F(u,0) = \frac{\sigma^2}{2}u + \widetilde{\eta}u + \int_{\mathbb{R}} \left(\mathrm{e}^{ux} - 1 - \frac{ux}{1+x^2}\right) \nu(\mathrm{d}x), \quad \text{for all } u \in \mathcal{D}_X.$$

where $\widetilde{\eta} := \eta + \int_{\mathbb{R}} \left(\left(1+x^2\right)^{-1} - \mathbb{1}_{\{|x|\leq 1\}}\right) x\,\nu(\mathrm{d}x)$, and it is straightforward to see that $\widetilde{\eta}$ is well-defined using the properties of the Lévy measure ν given in (4.3.10). The effective domain $\mathcal{D}_X$ precisely contains all the moments of the random variable X (by a slight abuse of notation, we denote $X = X_1$) that exist—since Lévy processes are infinitely divisible, the time-dependence does not appear in the moments of the process. Moment explosions in Lévy processes have been studied in [72], [6] and [7], so following [72], define the critical moments $u^*_\pm$ of the random variable X as

$$u^*_- := \sup\{u \in \mathbb{R}_+ : \phi_X(-u) < \infty\} \qquad \text{and} \qquad u^*_+ := \sup\{u \in \mathbb{R}_+ : \phi_X(u) < \infty\},$$

and it is clear that $\mathcal{D}^o_X = \left(u^*_-, u^*_+\right)$ (note that the boundary points of $\mathcal{D}_X$ are not necessary included). However, as we have noted in the case of affine stochastic volatility models, we do not have necessary and sufficient conditions ensuring that the function ϕ_X is essentially smooth, and we henceforth suppose that Assumption 4.3.3 holds. Under Assumptions A1-A5 and Assumption 4.3.3, it is then straightforward to see that Theorem 4.1.1 holds and hence the large-time implied volatility is given by Proposition 4.1.3. Note that if the interval the effective domain of $\mathcal{D}_X$ is of the form (u_-, u_+) with $u_- < 0$ ad $u_+ > 1$ then since Λ is continuously differentiable and convex on $\mathcal{D}^o_X$ then it is essentially smooth. However if $\mathcal{D}_X$ contains at least one of its boundary points the we can not conclude directly. For instance the logarithmic Laplace transform of the Variance-Gamma model $\mathrm{VG}(a,b,c)$ reads

$$\Lambda_{\mathrm{VG}}(u) = \left(\frac{ab}{(a-u)(b+u)}\right)^c, \qquad \text{for all } u \in \mathcal{D}_X = (-b,a),$$

which is clearly essentially smooth on $\mathcal{D}_X^o$, thus a full large deviations principle holds.

The Heston model with jumps

We have built up our methodology from the Heston model. It is then quite natural to study precisely how jumps influence this model, at least for large maturities. We consider the following dynamics for the logarithmic share price process $(X_t)_{t\geq 0}$ under the risk-neutral measure,

$$\begin{aligned} \mathrm{d}X_t &= \left(\delta - \tfrac{1}{2}V_t\right)\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_t + \mathrm{d}J_t, && X_0 = x_0 \in \mathbb{R},\\ \mathrm{d}V_t &= \kappa\left(\theta - V_t\right)\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}Z_t, && V_0 = v > 0,\\ \mathrm{d}\left\langle W, Z\right\rangle_t &= \rho \mathrm{d}t, \end{aligned}$$

where $\kappa, \sigma, \theta > 0$, $\rho \in (-1,1)$, $\kappa - \rho\sigma > 0$, and $J := (J_t)_{t\geq 0}$ is a pure-jump Lévy process independent of the Brownian motion $(W_t)_{t\geq 0}$. The parameter $\delta \in \mathbb{R}$ ensures that the process $(\exp(X_t))_{t\geq 0}$ is a martingale. The logarithmic moment generating function of the Heston model with jumps reads

$$\log \mathbb{E}\left(\mathrm{e}^{u(X_t - x_0)}\right) = \Lambda_H(u,t) + \widetilde{\Lambda}_J(u)\,t,$$

where Λ_H is the cumulant generating function of the Heston model given in Proposition 4.2.1, Λ_J the cumulant generating function of the jump process J, and $\widetilde{\Lambda}_J(u) := \Lambda_J(u) - u\Lambda_J(1)$ accounts for the compensated part of the jumps. In terms of the functions F and R, we have

$$F(u,w) = \kappa\theta w + \widetilde{\Lambda}_J(u) \qquad \text{and} \qquad R(u,w) = \frac{u}{2}(u-1) + \frac{\sigma^2}{2}w^2 - \kappa w + \rho\sigma u w.$$

Let us consider the following three examples:

(i) J is a Poisson process with intensity $\eta > 0$ and Lévy exponent $\Lambda_P(u) = -\eta\left(1 - \mathrm{e}^u\right)$.

(ii) J is a Gamma process with parameters $a > 0$, $b > 0$ and Lévy exponent $\Lambda_\Gamma(u) = -b\log\left(1 - u/a\right)$.

(iii) J is a Normal Inverse Gaussian process with parameters $(\alpha, \beta, \mu, \delta)$ and Lévy exponent $\Lambda_{\mathrm{NIG}}(u) = \mu u + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + u)^2}\right)$.

Let $\left[u_-^h, u_+^h\right]$ denote the effective domain (see (4.2.5) and (4.2.6)) of the Heston limiting cumulant generating function Λ_H^∞, and $u_\pm^J$ represent the lower and upper bounds of the effective domain of the function Λ_J. For the three examples above we have

$$\Lambda(u) := \lim_{t\to\infty} t^{-1}\Lambda_t(u) = \Lambda_H^\infty(u) + \widetilde{\Lambda}_J(u), \quad \text{for all } u \in \left[u_-^h \vee u_-^J, u_+^h \wedge u_+^J\right].$$

Since Λ_H is essentially smooth, the essential smoothness properties of the limiting cumulant generating function of the Heston model with jumps will only depend on the properties of Λ_J. From the moment generating functions above, it is easy to see that

$$u_\pm^P = \pm\infty, \quad u_-{}^\Gamma = -\infty, \quad u_+^\Gamma = a, \quad u_\pm^{\mathrm{NIG}} = -\beta \pm \alpha,$$

and that the function $\widetilde{\Lambda}_J$ is essential smoooth on its effective domain. Therefore Theorem 4.1.1 and the large-time asymptotic formula (4.1.5) for the implied volatility hold by Proposition 4.1.3. This is of course a simple example since all the functions are essentially smooth and the (necessary and sufficient) conditions for a cumulative generating function to be essentially smooth are left for future research, as we mentioned in the preceding paragraph on exponential Lévy processes.

We provide now numerical evidence on the standard Heston model (Figure 4.2) and the Heston model with Normal Inverse Gaussian jumps (Figure 4.3). We compare our asymptotic formula (4.1.5) for the large-maturity smile and the implied volatility calculated numerically using the Zeliade Quant Framework from Zeliade Systems. We consider the spot price to be worth 100 and a maturity of ten years. For the Heston dynamics, the parameters read

$$\kappa = 1.15, \quad \sigma = 0.2, \quad \theta = 0.04, \quad \rho = -0.4, \quad v = 0.04. \tag{4.3.12}$$

For the Heston model with Normal Inverse Gaussian jumps, we consider the same parameters as in (4.3.12) for the Heston dynamics, and the NIG parameters read

$$\alpha = 7.104, \quad \beta = -3.3, \quad \delta = 0.193, \quad \mu = 0.092. \tag{4.3.13}$$

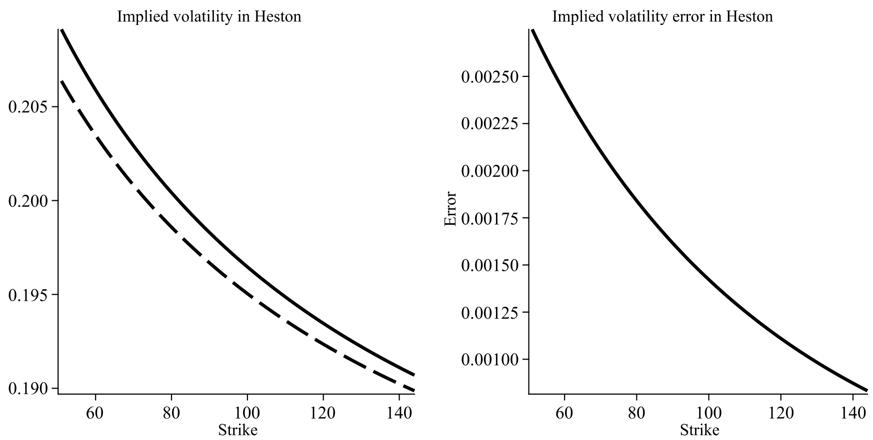


Figure 4.2: Implied volatility smiles under the Heston model without jumps using the parameters given in (4.3.12). The dash curve is computed numerically, whereas the solid one corresponds to the asymptotic formula (4.1.5).

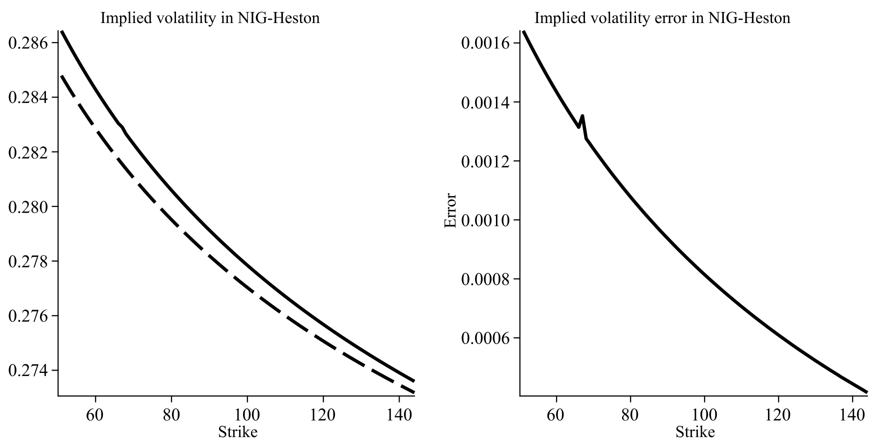


Figure 4.3: Implied volatility smiles under the Heston model with NIG-jumps using the parameters given in (4.3.13). The dash curve is computed numerically, whereas the solid curve corresponds to the asymptotic formula (4.1.5).

The Barndorff-Nielsen & Shephard model

We provide a final example, namely the Barndorff-Nielsen & Shephard model given by the following SDEs under the risk-neutral measure (see [94] and [80]),

$$\begin{aligned} \mathrm{d}X_t &= -\left(\gamma k(\rho) + \tfrac{1}{2}V_t\right)\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_t + \rho\,\mathrm{d}J_{\gamma t}, \qquad && X_0 = x_0 \in \mathbb{R}, \\ \mathrm{d}V_t &= -\gamma V_t \mathrm{d}t + \mathrm{d}J_{\gamma t}, && V_0 = v > 0, \end{aligned}$$

where $\gamma > 0$, $\rho < 0$ and $(J_t)_{t\geq 0}$ is a Lévy subordinator where the cumulant generating function of J_1 is given by $k(u) = \log \mathbb{E}\left(\mathrm{e}^{uJ_1}\right)$. The logarithmic moment generating function reads

$$\Lambda_t(u) = -ut\gamma k(\rho) - \frac{vu}{2\gamma}(1-u)\left(1-\mathrm{e}^{-\gamma t}\right) + \gamma\int_0^t k\left(\rho u + \frac{u(1-u)}{2\gamma}\left(1-\mathrm{e}^{-\gamma(t-s)}\right)\right)\mathrm{d}s,$$

for all $t \geq 0$ and $p \in \mathbb{R}$ such that it is well defined. As mentioned in [68], it is easy to see that the domain $\mathcal{D}_{\Lambda_t}$ of the function Λ_t reads $\mathcal{D}_{\Lambda_t} = (u_-(t), u_+(t))$ for all $t \geq 0$, where

$$u_\pm(t) := \frac{1}{2} - \frac{\rho\gamma}{1-\exp(-\gamma t)} \pm \sqrt{\frac{1}{4} + \frac{(2k^* - \rho)\gamma}{1-\exp(-\gamma t)} + \frac{\rho^2\gamma^2}{(1-\exp(-\gamma t))^2}},$$

and $k^* := \sup\{u > 0 : k(u) < \infty\}$. Since $\gamma > 0$, we have that $\lim_{t\to\infty} u_\pm(t) = u_\pm$, with

$$u_\pm := \frac{1}{2} - \rho\gamma \pm \sqrt{\frac{1}{4} - (2k^* - \rho)\gamma + \rho^2\gamma^2}.$$

Furthermore the two functions ϕ and ψ read

$$\phi(t,u,0) = -ut\gamma k(\rho) + \gamma\int_0^t k\left(\rho u + \frac{1}{2\gamma}\left(u-u^2\right)\left(1-\mathrm{e}^{-\gamma(t-s)}\right)\right)\mathrm{d}s,$$

and

$$\psi(t,u,0) = \frac{1}{2\gamma}\left(u^2-u\right)\left(1-\mathrm{e}^{-\gamma t}\right).$$

We can hence deduce the two functions F and R,

$$R(u,0) = \partial_t\psi(t,u,0)|_{t=0} = \frac{1}{2}\left(u^2-u\right), \quad \text{and} \quad F(u,0) = \partial_t\phi(t,u,0)|_{t=0} = \gamma k(\rho u) - u\gamma k(\rho).$$

We will present results for two types of BNS models (i) the Γ-BNS, where the subordinator is $\Gamma(a,b)$-distributed with a, $b > 0$ and hence $k_\Gamma(u) = (b-u)^{-1}au$, (ii) the IG-BNS, where the subordinator is IG(a,b)-distributed with a, $b > 0$ and hence $k_{\mathrm{IG}}(u) = ap\left(b^2-2u\right)^{-1/2}$. It is easy to see (see [80] for more details on this feature) that the jumps in the Γ-BNS model arise only a finite number of times in finite time intervals, whereas they occur infinitely often in finite time intervals when the driving Lévy process J is of inverse Gaussian type as its Lévy density $w : \mathbb{R} \to \mathbb{R}$ satisfies $\int_0^{+\infty} w(x)\mathrm{d}x = \infty$. Let us define the two auxiliary functions f_1 and f_2 by

$$f_1 := u\rho - \frac{1}{2}\left(u-u^2\right)\left(1-\mathrm{e}^{-\gamma t}\right), \quad \text{and} \quad f_2 := u\rho - \frac{1}{2}\left(u-u^2\right).$$

We then have

$$\Lambda_t^\Gamma(u) = -\frac{a\gamma\rho u}{b-\rho}t - \frac{v}{2\gamma}\left(u-u^2\right)\left(1-\mathrm{e}^{-\gamma t}\right) + \frac{a}{b-f_2}\left(f_2\gamma t + b\log\left(\frac{b-f_1}{b-u\rho}\right)\right),$$

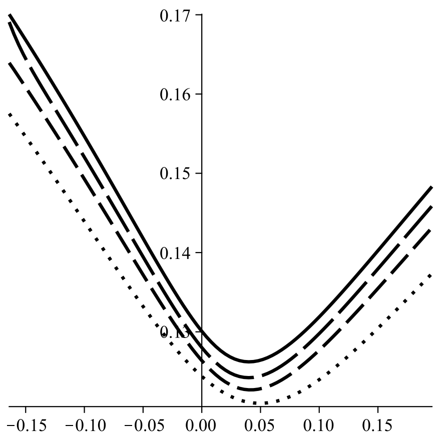


Figure 4.4: Implied volatility smiles under the Γ-BNS model with maturities 5 years (dots), 10 years (dash), 20 years (long dash). The solid line represents the large-time implied volatility (4.1.5).

and

$$\Lambda_t^{IG}(u) = -\frac{a\rho\gamma ut}{b\sqrt{1-2\rho/b^2}} - \frac{v}{2\gamma}\left(u-u^2\right)\left(1-\mathrm{e}^{-\gamma t}\right) + a\sqrt{b^2-2f_1} - a\sqrt{b^2-2\rho u}$$
$$+\frac{2af_2}{\sqrt{2f_2-b^2}}\left(\arctan\left(\sqrt{\frac{b^2-2\rho u}{2f_2-b^2}}\right) - \arctan\left(\sqrt{\frac{b^2-2f_1}{2f_2-b^2}}\right)\right).$$

From [68], we know that the two functions w and Λ read

$$w(u) = \frac{1}{2\gamma}\left(u^2-u\right) \qquad \text{and} \qquad \Lambda(u) = \gamma k\circ\zeta(u) - u\gamma k(\rho), \qquad \text{for all } u\in(u_-,u_+),$$

where $\zeta(u) := \frac{p^2}{2\gamma} + u\left(\rho - \frac{1}{2\gamma}\right)$. Let us concentrate on the Γ-BNS model. We see that the function Λ is well defined on the interval $\left(u_-^\Gamma, u_+^\Gamma\right)$, where $u_\pm^\Gamma := \frac{1}{2} - \rho\gamma \pm \sqrt{\left(\frac{1}{2}-\rho\gamma\right)^2 + 2b\gamma} \in \mathbb{R}_\pm$ are the two roots of the equation $\zeta(u) = b$. It is straightforward to check that the limiting function Λ satisfies all the necessary conditions to apply the Gärtner-Ellis in the large-time framework and hence we can define its Fenchel-Legendre transform $\Lambda^*(x) := \sup_{u\in[u_-,u_+]}\{ux - \Lambda(u)\} = u^*(x)\,x - \Lambda(u^*(x))$, for all $x\in\mathbb{R}$, where $u^*(x)$ is the unique solution to the equation $\Lambda'(u^*(x)) = x$, for any $x\in\mathbb{R}$. The same result applies to the IG-BNS model, but in this case we have

$$u_\pm^{IG} := \frac{1}{2} - \rho\gamma \pm \sqrt{\left(\frac{1}{2}-\rho\gamma\right)^2 + b^2\gamma} \in \mathbb{R}_\pm.$$

We use the following values given in [94, Section 7.3] calibrated on the S&P for the Γ-BNS model: $a = 1.4338$, $b = 11.6641$, $v = 0.0145$, $\gamma = 0.5783$, $\rho = -1.2606$.

Chapter 5

A short ride along the steepest descent

In the preceding chapter we have developed a framwork to study the limit of the implied volatility of continuous affine stochastic volatility models as the maturity tends to infinity. In particular we have obtained a completely explicit formula for this limit for the Heston model. In this chapter we wish to take the analysis in the Heston case a step further and derive higher-order terms (in t) in order to obtain more precise asymptotics and rate of convergence. The Gärtner-Ellis theorem does not allow us to do so, and we hence use complex saddlepoint methods to obtain such results. In this chapter we shall use the standard notations for the Heston model as in Section 4.2.3, and we will always assume the condition $\kappa - \rho\sigma > 0$ (i.e. Assumption A5 on page 60).

5.1 Large-time behaviour of call options

5.1.1 Large-time behaviour of call options under the Heston model

In this section, we derive the asymptotic behaviour of call option prices under the Heston dynamics (3.2.1) as the maturity t tends to infinity, both for maturity-dependent and for fixed strikes. Before diving into the details, let us introduce the function $\mathcal{I} : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$ by

$$\mathcal{I}\left(x, t; a, b\right) := \left(1 - \mathrm{e}^{xt}\right) \mathbb{1}_{\{x \leq a\}} + \mathbb{1}_{\{a < x < b\}} + \frac{1}{2}\mathbb{1}_{\{x = b\}}. \tag{5.1.1}$$

The next theorem is the main result of this chapter and its long proof is postponed to Section 5.4.

Theorem 5.1.1. *In the Heston model* (3.2.1) *with $\kappa > \rho\sigma$ the following asymptotic behaviour for the price of a call option with strike $S_0 \exp(xt)$ for all $x \in \mathbb{R}$,*

$$\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^{xt}\right)_+}{S_0} = \mathcal{I}\left(x, t; -\frac{\theta}{2}, \frac{\bar{\theta}}{2}\right) + \frac{A(x)}{\sqrt{2\pi t}} \exp\Big(- \left(\Lambda^*(x) - x\right) t\Big) \left(1 + \mathcal{O}\left(\frac{1}{t}\right)\right), \quad \text{as } t \to \infty,$$

holds where

$$A(x) := \frac{1}{\sqrt{\Lambda''(u^*(x))}} \begin{cases} \dfrac{U(u^*(x))}{u^*(x)\,(u^*(x)-1)}, & \text{if } x \in \mathbb{R} \setminus \left\{-\frac{\theta}{2}, \frac{\bar{\theta}}{2}\right\}, \\ -1 - \operatorname{sgn}(x) \left(\dfrac{1}{6} \dfrac{\Lambda'''(u^*(x))}{\Lambda''(u^*(x))} - U'(u^*(x)) \right), & \text{if } x \in \left\{-\frac{\theta}{2}, \frac{\bar{\theta}}{2}\right\}, \end{cases} \tag{5.1.2}$$

and

$$U(u) := \left(\frac{2d(-\mathtt{i}u)}{\kappa - \rho\sigma u + d(-\mathtt{i}u)} \right)^{2\kappa\theta/\sigma^2} \exp\left(\frac{v}{\kappa\theta} \Lambda(u) \right). \tag{5.1.3}$$

The function Λ *is defined in* (4.2.19)*,* u^* *in* (4.2.21)*,* Λ^* *in* (4.2.20)*,* d *in* (3.2.5) *and* $\operatorname{sgn}(x) = \mathbb{1}_{\{x>0\}}$.

Remark 5.1.2. Note that, from property (b) on page 54, the square root in the function A is a strictly positive real number.

Remark 5.1.3. Theorem 5.1.1 is similar in spirit to the saddlepoint approximation for a density of random variable X given in [16]

$$f_X(x) \approx \left(2\pi K''(u^*(x))\right)^{-1/2} \exp\left(K(u^*(x)) - xu^*(x)\right),$$

where $K(u) := \log \mathbb{E}(\exp(uX))$, and $u^*(x)$ is the unique solution to $K'(u^*(x)) = x$. Here $X := X_t - x_0$, $K(u) = \Lambda(u)t + \mathcal{O}(1)$, and we substitute x to xt so that

$$f_{X_t - x_0}(xt) \approx \frac{\exp\left(\Lambda(u^*(x))\,t - xu^*(x)\,t\right)}{\sqrt{2\pi\Lambda''(u^*(x))\,t}} = \frac{\exp\left(-\Lambda^*(x)\,t\right)}{\sqrt{2\pi\Lambda''(u^*(x))\,t}}.$$

In order to precisely compare our result to the existing literature, we prove the following lemma, which gives the asymptotic behaviour of call options for a fixed strike. This result was derived in [73, Chapter 6] and we provide a rigorous proof in Section 5.5.1.

Lemma 5.1.4. *Under the same assumptions as Theorem 5.1.1, for any* $x \in \mathbb{R}$*, we have the following behaviour for a call option with fixed strike* $S_0 \mathrm{e}^x$ *as* t *tends to infinity,*

$$\frac{1}{S_0} \mathbb{E}\left(S_t - S_0 \mathrm{e}^x\right)_+ = 1 + \frac{A(0)}{\sqrt{2\pi t}} \exp\left(x(1 - u^*(0)) - \Lambda^*(0)\,t\right) \left(1 + \mathcal{O}\left(\frac{1}{t}\right)\right).$$

5.1.2 Large-time behaviour of the the Black-Scholes call option formula

By a similar analysis, we can deduce the large-time asymptotic call price for the Black-Scholes model. This result is of fundamental importance for us as it will allow us to compute the implied volatility by comparing the Black-Scholes and the Heston call option prices. In this section, we further assume that the variance is of the form $\Sigma^2 + a_1/t$, for each maturity t, where Σ is a strictly positive constant and $a_1 > -\Sigma^2 t$. The following proposition proved in Section 5.5.2 gives the behaviour of the Black-Scholes price as the maturity tends to infinity.

Proposition 5.1.5. *With the assumptions above, for all $x \in \mathbb{R}$, we have the following asymptotic behaviour for the Black-Scholes call option formula in the large-strike, large-time case*

$$\frac{C_{\mathrm{BS}}\left(xt, t, \sqrt{\Sigma^2 + a_1/t}\right)}{S_0} = \mathcal{I}\left(x, t; -\frac{\Sigma^2}{2}, \frac{\Sigma^2}{2}\right) + \frac{A_{\mathrm{BS}}(x, \Sigma, a_1)}{\sqrt{2\pi t}} \exp\left(-\left(\Lambda^*_{\mathrm{BS}}(x, \Sigma) - x\right)t\right)\left(1 + \mathcal{O}\left(\frac{1}{t}\right)\right),$$

where the function $\mathcal{I}$ is defined in (5.1.1) *and*

$$A_{\mathrm{BS}}(x, \Sigma, a_1) := \exp\left(\frac{1}{8}a_1\left(\frac{4x^2}{\Sigma^4} - 1\right)\right)\frac{\Sigma^3}{x^2 - \Sigma^4/4}\mathbb{1}_{\{x \neq \pm\Sigma^2/2\}} + \frac{a_1/2 - 1}{\Sigma}\mathbb{1}_{\{x = \pm\Sigma^2/2\}}. \tag{5.1.4}$$

Remark 5.1.6. If we set $a_1 = 0$, we obtain the large-time expansion for a call option under the standard Black-Scholes model with volatility Σ and log-moneyness equal to xt.

As in the Heston model above, we derive here the equivalent of Proposition 5.1.5 when the strike does not depend on the maturity anymore. This lemma is immediate from the Black-Scholes formula and the approximation (5.5.2) for the Gaussian cumulative distribution function.

Lemma 5.1.7. *With the assumptions above, we have the following behaviour for the Black-Scholes call option formula in the fixed-strike, large-time case*

$$\frac{C_{\mathrm{BS}}\left(x, t, \sqrt{\Sigma^2 + a_1/t}\right)}{S_0} = 1 - \frac{2}{\Sigma\sqrt{2\pi t}}\exp\left(-\frac{\Sigma^2 t}{8} + \frac{x}{2} - \frac{a_1}{8}\right)\left(1 + \mathcal{O}\left(\frac{1}{t}\right)\right).$$

5.2 Large-time behaviour of implied volatility

The previous section dealt with large-time asymptotics for call option prices. In this section, we translate these results into asymptotics for the implied volatility. Recall that [40], [73], and [68] have already derived the leading order term for the implied volatility in the large-time, fixed-strike case. Our goal here is to obtain the leading order and the correction term in the large-time, large-strike case. Theorem 5.2.1 provides the main result, i.e. the large-time behaviour of the implied volatility in the large strike case. In the following, $\hat{\sigma}_t(x)$ will denote the implied volatility corresponding to a vanilla call option with maturity t and (maturity-dependent) strike $S_0\exp(xt)$ in the Heston model (3.2.1). Recall the function $\hat{\sigma}^2_\infty : \mathbb{R} \to \mathbb{R}_+$ from (4.1.5),

$$\hat{\sigma}^2_\infty(x) := 2\left(2\Lambda^*(x) - x + 2\left(\mathbb{1}_{x\in(-\theta/2,\bar{\theta}/2)} - \mathbb{1}_{x\in\mathbb{R}\setminus(-\theta/2,\bar{\theta}/2)}\right)\sqrt{\Lambda^*(x)^2 - \Lambda^*(x)\,x}\right), \quad \text{for all } x \in \mathbb{R} \tag{5.2.1}$$

and define the function $\hat{a}_1 : \mathbb{R} \to \mathbb{R}$ by

$$\hat{a}_1(x) := 2\begin{cases} \dfrac{\hat{\sigma}^4_\infty(x) - 1/4}{x^2}\log\left(\dfrac{A(x)}{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), 0)}\right), & x \in \mathbb{R}\setminus\left\{-\frac{\theta}{2}, \frac{\bar{\theta}}{2}\right\}, \\ 1 - \dfrac{\hat{\sigma}_\infty(x)}{\sqrt{\Lambda''(u^*(x))}}\left(1 + \mathrm{sgn}(x)\left(\dfrac{\Lambda'''(u^*(x))}{6\Lambda''(u^*(x))} - U'(u^*(x))\right)\right), & x \in \left\{-\frac{\theta}{2}, \frac{\bar{\theta}}{2}\right\}, \end{cases} \tag{5.2.2}$$

where A is defined in (5.1.2), A_{BS} in (5.1.4), U in (5.1.3), Λ in (4.2.19), u^* in (4.2.21) and Λ^* in (4.2.20). They are all completely explicit, so that the functions $\hat{\sigma}^2_\infty$ and $\hat{a}_1$ are explicit as well. From the properties of Λ^* proved on page 54, $\Lambda^*(x)$ and $\Lambda^*(x)-x$ are non-negative, so that $\hat{\sigma}^2_\infty(x)$ is a well defined real number for all $x \in \mathbb{R}$. Then the following theorem holds:

Theorem 5.2.1. *The functions $\hat{\sigma}_\infty$ and $\hat{a}_1$ are continuous on $\mathbb{R}$ and*

$$\hat{\sigma}^2_t(x) = \hat{\sigma}^2_\infty(x) + \hat{a}_1(x)/t + o\left(1/t\right) \qquad \text{for all } x \in \mathbb{R} \text{ as } t \text{ tends to } \infty.$$

Furthermore the error term $|\hat{\sigma}^2_t(x) - \hat{\sigma}^2_\infty(x) - \hat{a}_1(x)/t|t$ tends to zero as t goes to infinity uniformly in x on compact subsets of $\mathbb{R}\backslash\{-\theta/2, \bar{\theta}/2\}$.

Remark 5.2.2. The functions A and A_{BS} are not continuous at $-\theta/2$ and $\bar{\theta}/2$, (see Figure 5.1) however $\hat{a}_1$ is continuous by this theorem.

Proof. We first prove that the functions $\hat{\sigma}_\infty$ and $\hat{a}_1$ are continuous. In fact, the continuity of the function $\hat{\sigma}_\infty$ follows from properties (c) and (d) on page 54. We observe that the two functions $x \mapsto A(x)$ and $x \mapsto A_{\mathrm{BS}}\left(x, \hat{\sigma}_\infty(x), 0\right)$ have poles at $\bar{\theta}/2$ and $-\theta/2$ and their quotient is strictly positive for $x \in \mathbb{R} \setminus \left\{-\theta/2, \bar{\theta}/2\right\}$. Therefore the function $\hat{a}_1$ is continuous on this complement. Elementary calculations show that in the neighbourhood of $\bar{\theta}/2$, we have the following expansion (where $\Theta := (\bar{\theta}/\Lambda''(1))^{1/2}$)

$$\hat{\sigma}^2_\infty(x) = \bar{\theta} + 2\left(1-\Theta\right)\left(x - \bar{\theta}/2\right) + \frac{2}{\Lambda''(1)}\left(1 - \frac{1}{\Theta} + \frac{\Lambda'''(1)}{6\Lambda''(1)^2}\Theta\right)\left(x-\bar{\theta}/2\right)^2 + \mathcal{O}\left(\left|x - \bar{\theta}/2\right|^3\right).$$

This expansion can be used to obtain

$$\frac{A(x)}{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), 0)} = 1 + \frac{1}{\Lambda''(1)}\left(U'(1) - 1 - \frac{\Lambda'''(1)}{6\Lambda''(1)} + \frac{1}{\Theta}\right)(x - \bar{\theta}/2) + \mathcal{O}\left(|x-\bar{\theta}/2|^2\right),$$

which implies that $\lim_{x\to\bar{\theta}/2} \hat{a}_1(x) = \hat{a}_1\left(\bar{\theta}/2\right)$. A similar argument shows continuity of $\hat{a}_1$ at $-\theta/2$.

We now prove the formula in the theorem in the case $x > \bar{\theta}/2$. Note that $\hat{\sigma}_\infty(x)$ as defined in (5.2.1) satisfies the following quadratic equation

$$\Lambda^*(x) - x = \Lambda^*_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x)) - x, \quad \text{for all } x \in \mathbb{R}, \tag{5.2.3}$$

where Λ^* is given by (4.2.20) and Λ^*_{BS} by (4.1.3). We prove the theorem in two steps: first, we prove the convergence of the implied variance to $\hat{\sigma}^2_\infty(x)$, then we prove the first order correction term. The first step has already been proved in Proposition 4.1.3. For the second step of the proof, we show that for all $\delta > 0$, there exists $t^*(\delta)$ such that,

$$|\hat{\sigma}^2_t(x) - \hat{\sigma}^2_\infty(x) - \hat{a}_1(x)/t| \le \delta/t, \qquad \text{for all } t > t^*(\delta). \tag{5.2.4}$$

Note that by definition in (5.2.2), we have $A(x) = A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x))$ for all $x \in \mathbb{R}$, so that Theorem 5.1.1 implies that for all $\varepsilon > 0$, there exists $t^*(\varepsilon)$ such that for all $t > t^*(\varepsilon)$ we have

$$\frac{1}{S_0}\mathbb{E}\left(S_t - S_0\mathrm{e}^{xt}\right)_+ \le \frac{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x))}{\sqrt{2\pi t}}\mathrm{e}^{-(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x))-x)t}\mathrm{e}^{\varepsilon}, \tag{5.2.5}$$

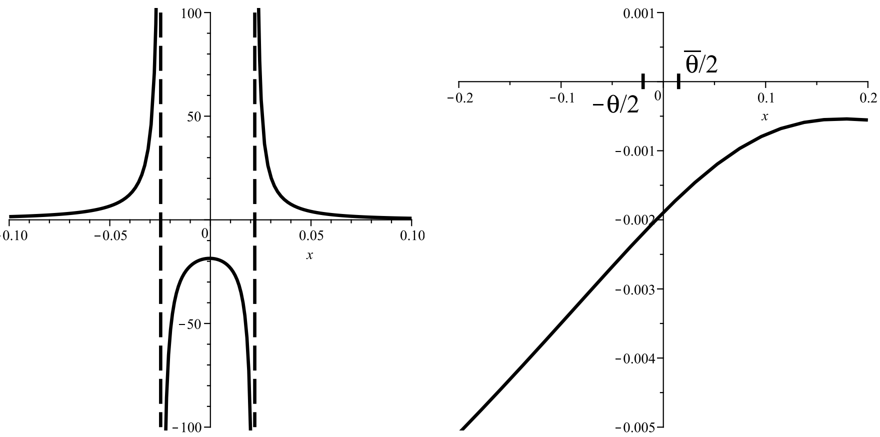


Figure 5.1: We plot the two functions A (left) and $\hat{a}_1$ (right) with the calibrated parameters on page 74. The function A_{BS} in Proposition 5.1.5 has the same shape as A. Here, the two poles of A and A_{BS} are $-\theta/2 \approx -0.025$ and $\bar{\theta}/2 \approx 0.022$.

and

$$\frac{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x))}{\sqrt{2\pi t}} \mathrm{e}^{-(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x))-x)t} \mathrm{e}^{-\varepsilon} \leq \frac{1}{S_0} \mathbb{E}\left(S_t - S_0 \mathrm{e}^{xt}\right)_+ .$$

Let $\delta > 0$ and $\varepsilon(\delta) := \left(4x^2/\hat{\sigma}^4_\infty(x) - 1\right)\delta/16 > 0$. From (5.2.5), there exists $t^*(\delta)$, such that for all $t > t^*(\delta)$,

$$\frac{\mathbb{E}\left(S_t - S_0 \mathrm{e}^{xt}\right)_+}{S_0} \leq \frac{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x) + \delta)}{\sqrt{2\pi t}} \mathrm{e}^{-(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x))-x)t} \mathrm{e}^{-\varepsilon(\delta)} \leq \frac{C_{\mathrm{BS}}\left(xt, t, \sqrt{\hat{\sigma}^2_\infty(x) + (\hat{a}_1(x) + \delta)/t}\right)}{S_0},$$

and

$$\frac{\mathbb{E}\left(S_t - S_0 \mathrm{e}^{xt}\right)_+}{S_0} \geq \frac{A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x) - \delta)}{\sqrt{2\pi t}} \mathrm{e}^{-(\Lambda^*_{\mathrm{BS}}(x,\hat{\sigma}_\infty(x))-x)t} \mathrm{e}^{\varepsilon(\delta)} \geq \frac{C_{\mathrm{BS}}\left(xt, t, \sqrt{\hat{\sigma}^2_\infty(x) + (\hat{a}_1(x) - \delta)/t}\right)}{S_0}.$$

By strict monotonicity of the Black-Scholes price as a function of the volatility, we obtain

$$\hat{\sigma}^2_\infty(x) + (\hat{a}_1(x) - \delta)/t \leq \hat{\sigma}^2_t(x) \leq \hat{\sigma}^2_\infty(x) + (\hat{a}_1(x) + \delta)/t,$$

which proves (5.2.4). An analogous argument implies the inequality in (5.2.4) for $x' \in \mathbb{R}\backslash\{-\theta/2, \bar{\theta}/2\}$ since the functions A and $x \mapsto A_{\mathrm{BS}}(x, \hat{\sigma}_\infty(x), \hat{a}_1(x))$ are continuous on this set and by (5.2.3) the identity $\mathcal{I}\left(x, t; -\theta/2, \bar{\theta}/2\right) = \mathcal{I}\left(x, t; -\hat{\sigma}^2_\infty(x)/2, \hat{\sigma}^2_\infty(x)/2\right)$ holds for all $(x,t) \in \mathbb{R} \times \mathbb{R}_+$ (the function $\mathcal{I}$ is defined in (5.1.1)). In the cases $x' \in \{-\theta/2, \bar{\theta}/2\}$ a similar argument proves (5.2.4) at the point x' but not necessarily on its neighbourhood. The uniform convergence on compact subsets of the complements of $\mathbb{R}\backslash\{x^*, \tilde{x}^*\}$ follows the same lines as in the proof of Proposition 4.1.3. $\square$

5.2.1 The large-time, fixed-strike case

This section is the translation of Lemmas 5.1.4 and 5.1.7 in terms of implied volatility asymptotics, and improves the understanding of the behaviour of the Heston implied volatility in the long term.

Let $\sigma_t(x)$ denote the implied volatility corresponding to a vanilla call option with maturity t and fixed strike $K = S_0 \exp(x)$ in the Heston model (3.2.1). Let us define the function $a_1 : \mathbb{R} \to \mathbb{R}$ by

$$a_1(x) := -8\log\left(-A(0)\sqrt{2\Lambda^*(0)}\right) + 4\left(2u^*(0) - 1\right)x, \quad \text{for all } x \in \mathbb{R}, \tag{5.2.6}$$

where A is defined in (5.1.2), Λ^* in (4.2.20) and u^* in (4.2.21). Elementary calculations show that $A(0) < 0$. From the properties of Λ^* on page 54, $a_1(x)$ is a well defined real number for all $x \in \mathbb{R}$.

Theorem 5.2.3. *With the assumptions above, we have*

$$\sigma_t^2(x) = 8\Lambda^*(0) + a_1(x)/t + o\left(1/t\right), \qquad \textit{for all } x \in \mathbb{R}, \ \textit{as } t \textit{ tends to } \infty,$$

where Λ^* *is given by* (4.2.20) *and* a_1 *by* (5.2.6).

Proof. The proof of this theorem is very similar to the proof of Theorem 5.2.1, so we do not detail it too much, in particular we only prove the upper bound, the lower bound being analogous; also we deal with both the zeroth and the first order term at the same time. By Lemma 5.1.4 and (5.2.6), we know that for all $\varepsilon > 0$, there exists a $t^*(\varepsilon)$ such that for all $t > t^*(\varepsilon)$, we have

$$\begin{aligned}\frac{1}{S_0}\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+ &\le \left(1 + (2\pi t)^{-1/2}A(0)\exp\left(x\left(1 - u^*(0)\right) - \Lambda^*(0)t\right)\right)\mathrm{e}^{\varepsilon} \\ &= \left(1 - \frac{1}{2\sqrt{\pi\Lambda^*(0)t}}\exp\left(x/2 - a_1(x)/8 - \Lambda^*(0)t\right)\right)\mathrm{e}^{\varepsilon}.\end{aligned}$$

Now, for any $\delta > 0$, by a continuity argument, we can then find $\varepsilon\left(\delta\right) > 0$ such that

$$\left(1 - \frac{1}{2\sqrt{\pi\Lambda^*(0)t}}\mathrm{e}^{x/2 - a_1(x)/8 - \Lambda^*(0)t}\right)\mathrm{e}^{\varepsilon(\delta)} = \left(1 - \frac{1}{2\sqrt{\pi\Lambda^*(0)t}}\mathrm{e}^{x/2 - (a_1(x)+\delta)/8 - \Lambda^*(0)t}\right)\mathrm{e}^{-\varepsilon(\delta)}.$$

Combining these equations, there exists $t^*(\delta) > 0$ such that, for all $t > t^*(\delta)$,

$$\frac{1}{S_0}\mathbb{E}\left(S_t - \mathrm{e}^x\right)_+ \le \frac{1}{S_0}C_{\mathrm{BS}}\left(x, t, \sqrt{8\Lambda^*(0) + (a_1(x) + \delta)/t}\right).$$

Thus, by the monotonicity of the Black-Scholes formula as a function of the volatility, we obtain $\sigma_t^2(x) \le \Lambda^*(0) + \left(a_1(x) + \delta\right)/t$. Since δ can be chosen as small as we wish, the theorem follows. □

5.3 Numerical results

We present here some numerical evidence of the validity of the volatility asymptotic formula in Theorem 5.2.1. We calibrate the Heston model on the European vanilla options on the Eurostoxx 50 on February, 15th, 2006. The option maturities range from one year to nine years and the initial spot S_0 equals 3729.79. The calibration, performed using the Zeliade Quant Framework, by Zeliade Systems, on the whole implied volatility surface gives the following parameters: $\kappa = 1.7609$, $\theta = 0.0494$, $\sigma = 0.4086$, $v = 0.0464$ and $\rho = -0.5195$. We plot the implied volatility smile of the calibrated Heston model for maturities $t = 5$ and $t = 9$ years (see Figure 5.2).

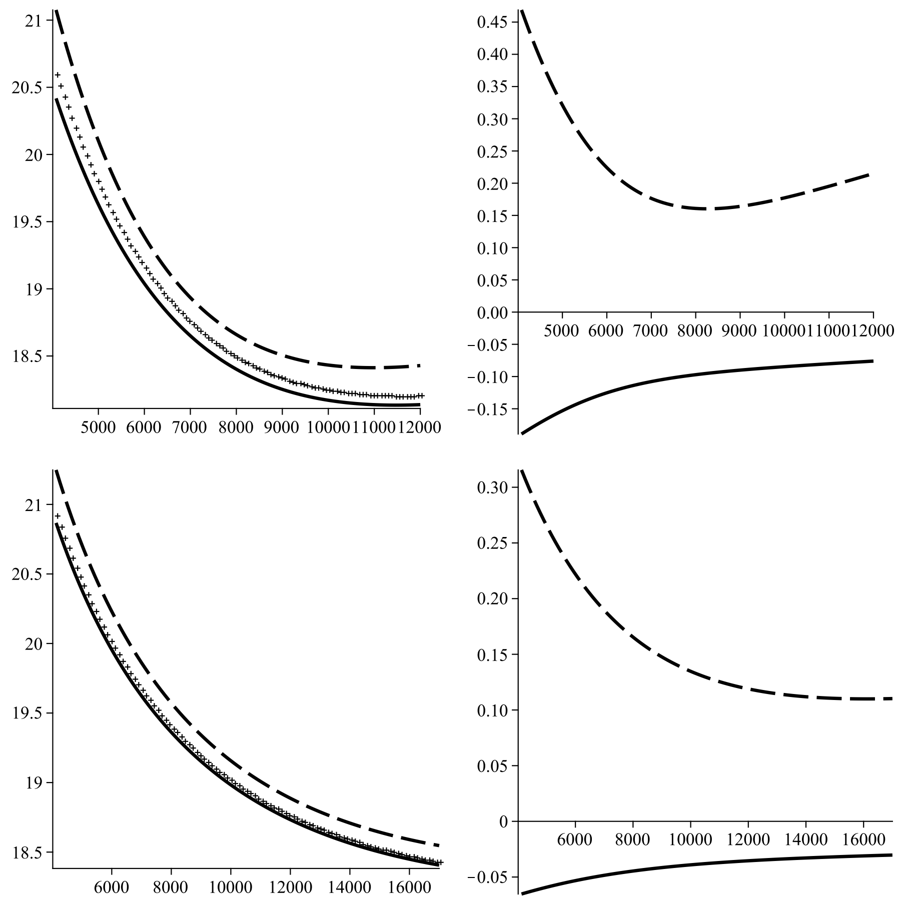


Figure 5.2: The left plots represent the leading order term $\hat{\sigma}_\infty$ (dashed) defined in (5.2.1), the asymptotic formula in Theorem 5.2.1 (solid line) and the true implied volatility (crosses) for the calibrated parameters (see page 74) as functions of the strike K. The right plots represent the errors between the true implied volatility and $\hat{\sigma}_\infty$ (dashed) and between the true implied volatility and our formula. From top to bottom, these plots correspond to the maturities $t = 5$ and 9 years. All the values are given as percentage.

5.4 Proof of Theorem 5.1.1

The proof of the theorem is divided into a series of steps: we first write the Heston call price in terms of an inverse Fourier transform of the characteristic function of the stock price. Then we prove a large-time estimate for the characteristic function (Lemma 5.4.1). The next step is to deform the contour of integration of the inverse Fourier transform through the saddlepoint of the integrand (4.2.21). Finally, studying the behaviour of the integral around this saddlepoint (Proposition 5.4.6) and bounding the remaining terms (Lemma 5.4.4) completes the proof. The special cases $x = -\theta/2$ and $x = \bar{\theta}/2$ in formula (5.1.2) are proved in Sections 5.4.3 and 5.4.4.

5.4.1 The Lee-Fourier inversion formula for call options

Using similar notation to Lee [71], set

$$A_{t,X} := \left\{\nu \in \mathbb{R} : \mathbb{E}\left(\exp\left(\nu\left(X_t - x_0\right)\right)\right) < \infty\right\}, \quad \text{for all } t \geq 0,$$

and define the characteristic function $\phi_t : \mathbb{R} \to \mathbb{C}$ of $X_t - x_0$ by

$$\phi_t(z) := \mathbb{E}\left(\exp\left(\mathtt{i}z(X_t - x_0)\right)\right), \quad \text{for all } t \geq 0.$$

From Theorem 1 in [65] and Proposition 3.1 in [3], we know that $\phi_t(z)$ can be analytically extended for any $z \in \mathbb{C}$ such that $-\Im(z) \in A_{t,X}$, and from our assumptions on the parameters in Section 5.1, $(u_-, u_+) \subseteq A_{t,X}$ for all $t \geq 0$. By Theorem 5.1 in [71], for any $\alpha \in (u_-, u_+)$, we have the following Fourier inversion formula for the price of a call option on S_t

$$\begin{aligned}\frac{1}{S_0}\mathbb{E}\left(S_t - K\right)_+ &= \phi_t(-\mathtt{i})1\!\!1_{\{0<\alpha<1\}} + \left(\phi_t(-\mathtt{i}) - \mathrm{e}^x\phi_t(0)\right)1\!\!1_{\{\alpha<0\}} + \frac{1}{2}\phi_t(-\mathtt{i})1\!\!1_{\{\alpha=1\}} \\ &+ \left(\phi_t(-\mathtt{i}) - \frac{\phi_t(0)}{2}\mathrm{e}^x\right)1_{\{\alpha=0\}} + \frac{1}{\pi}\int_{\gamma_+}\Re\left(\mathrm{e}^{-\mathtt{i}zx}\frac{\phi_t(z-\mathtt{i})}{\mathtt{i}z - z^2}\right)\mathrm{d}z,\end{aligned}$$

where $x := \log(K/S_0)$ and $\gamma_+ : \mathbb{R}_+ \to \mathbb{C}$ is a contour such that $\gamma_+(u) := u - \mathtt{i}(\alpha - 1)$. The first four terms on the right hand side are complex residues that arise when we cross the pole of $\left(\mathtt{i}z - z^2\right)^{-1}$ at $z = 0$. We now set $k = \mathtt{i} - z$, substitute x to xt, and use the fact that S_t is a true martingale for all $t \geq 0$ (see Proposition 2.5 in [3]). Since k will always denote a complex number, we shall from now on use the notation $k = k_r + \mathtt{i}k_\mathtt{i}$ for $k_r, k_\mathtt{i} \in \mathbb{R}$. Note that

$$\Re\left(\mathrm{e}^{\mathtt{i}kxt}\frac{\phi_t(-k)}{\mathtt{i}k - k^2}\right) = \mathbb{E}\left(\Re\left(\mathrm{e}^{\mathtt{i}kxt}\frac{\mathrm{e}^{-\mathtt{i}k(X_t - x_0)}}{\mathtt{i}k - k^2}\right)\right),$$

that $k_r \mapsto \Re\left(\mathrm{e}^{\mathtt{i}kxt}\frac{\mathrm{e}^{-\mathtt{i}k(X_t-x_0)}}{\mathtt{i}k-k^2}\right)$ is an even function and that $k_r \mapsto \Im\left(\mathrm{e}^{\mathtt{i}kxt}\frac{\mathrm{e}^{-\mathtt{i}k(X_t-x_0)}}{\mathtt{i}k-k^2}\right)$ is an odd function. Clearly the normalised call price $\frac{1}{S_0}\mathbb{E}(S_t - S_0\exp(xt))_+$ is real, so if we take the real part of both sides and break up the integral, we obtain

$$\begin{aligned}\frac{1}{S_0}\mathbb{E}\left(S_t - S_0\mathrm{e}^{xt}\right)_+ &= 1\!\!1_{\{0<\alpha<1\}} + \left(1 - \mathrm{e}^{xt}\right)1\!\!1_{\{\alpha<0\}} + \frac{1}{2}1\!\!1_{\{\alpha=1\}} + \left(1 - \frac{1}{2}\mathrm{e}^{xt}\right)1\!\!1_{\{\alpha=0\}} \\ &+ \frac{\exp(xt)}{2\pi}\Re\left(\left(\int_{\gamma_\alpha} + \int_{\zeta_\alpha}\right)\mathrm{e}^{\mathtt{i}kxt}\frac{\phi_t(-k)}{\mathtt{i}k - k^2}\mathrm{d}k\right),\end{aligned} \tag{5.4.1}$$

for any $R > 0$, where, for any $\alpha \in \mathbb{R}$, we define the contours

$$\gamma_\alpha : (-\infty, -R] \cup [R, +\infty) \to \mathbb{C} \text{ such that } \gamma_\alpha(u) := u + \mathtt{i}\alpha, \tag{5.4.2}$$

and

$$\zeta_\alpha : (-R, R) \to \mathbb{C} \text{ such that } \zeta_\alpha(u) := u + \mathtt{i}\alpha. \tag{5.4.3}$$

For ease of notation, we do not write explicitly the dependence of these contours on R. We will see later how to choose R. In the following lemma, we characterise the large-time asymptotic behaviour of the characteristic function ϕ_t.

Lemma 5.4.1. *For all $k \in \mathbb{C}$ such that $-k_{\mathtt{i}} \in (u_-, u_+)$, we have*

$$\phi_t(k) = \exp\Big(\Lambda(\mathtt{i}k)t\Big) U(\mathtt{i}k)\left(1 + \varepsilon(k,t)\right), \quad \text{as } t \text{ tends to } \infty,$$

$\Re(d(k)) > 0$, and $\varepsilon(k,t) = \mathcal{O}\left(\mathrm{e}^{-t\Re(d(k))}\right)$, where U is defined in (5.1.3)*, u_-, u_+ in* (4.2.5) *and* (4.2.6) *and Λ is the analytic continuation of formula* (4.2.19)*.*

If $-k_{\mathtt{i}}$ is not in (u_-, u_+), this large-time behaviour of ϕ_t still holds, but $\Re(d(k))$ might be null (for instance if $k_r = 0$) so that $\varepsilon(k,t)$ does not tend to 0 as $t \to \infty$.

Proof. From Equation (3.2.2) we have for all $k \in \mathbb{C}$ such that $-k_{\mathtt{i}} \in (u_-, u_+)$,

$$\phi_t(k) = \exp\left(\Lambda(\mathtt{i}k)t - \frac{2\kappa\theta}{\sigma^2}\log\left(\frac{1 - g(k)\mathrm{e}^{-d(k)t}}{1 - g(k)}\right)\right)\exp\left(\frac{v}{\kappa\theta}\Lambda(\mathtt{i}k)\frac{1 - \mathrm{e}^{-d(k)t}}{1 - g(k)\mathrm{e}^{-d(k)t}}\right),$$

where d is defined in (3.2.5), Λ is the analytic extension of formula (4.2.19) and the correct branch for the complex logarithm and the complex square root function is the principal branch (see also [2] and [65]) and we recall the function $g : \mathbb{C} \to \mathbb{C}$ from (3.2.5),

$$g(k) := \frac{\kappa - \mathtt{i}\rho\sigma k - d(k)}{\kappa - \mathtt{i}\rho\sigma k + d(k)}, \quad \text{for all } k \in \mathbb{C}.$$

For all $k \in \mathbb{C}$ such that $-k_{\mathtt{i}} \in (u_-, u_+)$, we have $\Re(d(k)) > 0$. Let

$$\varepsilon_1(k,t) := \left(1 - g(k)\mathrm{e}^{-d(k)t}\right)^{-2\kappa\theta/\sigma^2} \text{ and } \varepsilon_2(k,t) := \exp\left\{-\frac{2d(k)\Lambda(\mathtt{i}k)v}{\kappa\theta\left(\kappa - \rho\sigma\mathtt{i}k + d(k)\right)}\left(\mathrm{e}^{d(k)t} - g(k)\right)^{-1}\right\}.$$

Then we have

$$\phi_t(k) = \exp\Big(\Lambda(\mathtt{i}k)t\Big)U(\mathtt{i}k)\varepsilon_1(k,t)\varepsilon_2(k,t), \quad \text{for all } t \geq 0.$$

Then, as t tends to infinity we have

$$\varepsilon_1(k,t) = 1 + \frac{2\kappa\theta g}{\sigma^2}\mathrm{e}^{-d(k)t} + \mathcal{O}\left(\mathrm{e}^{-2d(k)t}\right) \text{ and } \varepsilon_2(k,t) = 1 + \frac{c}{\mathrm{e}^{d(k)t} - 1} + \mathcal{O}\left(\left(\mathrm{e}^{d(k)t} - 1\right)^{-2}\right)$$

for some constant c. Set $\varepsilon(k,t) := \varepsilon_1(k,t)\varepsilon_2(k,t) - 1$ and the lemma follows. $\square$

5.4.2 The saddlepoint and its properties

The function $\Lambda : (u_-, u_+) \to \mathbb{R}$ defined in (4.2.19) can be analytically extended and we define the function $F : \mathcal{Z} \to \mathbb{C}$ by

$$F(k) := -\mathtt{i}kx - \Lambda(-\mathtt{i}k), \quad \text{where } \mathcal{Z} := \{k \in \mathbb{C} : k_{\mathtt{i}} \in (u_-, u_+)\}. \tag{5.4.4}$$

The exponent of the integrand in (5.4.1) has the form $-F(k)t$ by Lemma 5.4.1, therefore the saddlepoint properties of F given in the following elementary lemma are fundamental. Saddlepoints in the complex plane and their properties were defined in Section 1.2.

Lemma 5.4.2. *The saddlepoints of the complex function $F : \mathcal{Z} \to \mathbb{C}$ are given by*

$$z_0^{\pm} = \mathtt{i}\frac{\sigma - 2\kappa\rho \pm (\kappa\theta\rho + x\sigma)\eta(x^2\sigma^2 + 2x\kappa\theta\rho\sigma + \kappa^2\theta^2)^{-1/2}}{2\sigma\bar{\rho}^2} \in \mathcal{Z},$$

where $\eta := \sqrt{\sigma^2 + 4\kappa^2 - 4\kappa\rho\sigma}$.

Proof. Since we are looking for a saddlepoint in $\mathcal{Z}$, we can use the representation (4.2.19) for the function Λ. The equation $F'(z) = 0$ is quadratic and hence has two purely imaginary solutions $z_0^{\pm}$ since the expression $(x^2\sigma^2 + 2x\kappa\theta\rho\sigma + \kappa^2\theta^2) = (x\sigma + \kappa\theta\rho)^2 + \kappa^2\theta^2(1-\rho^2)$ is strictly positive for any $x \in \mathbb{R}$. It is also clear from the definition of u_- and u_+ in (4.2.5) and (4.2.6) and the assumptions on the coefficients that $\Im\left(z_0^{\pm}\right) \in (u_-, u_+)$, and hence $z_0^{\pm} \in \mathcal{Z}$ are saddlepoints of F. $\square$

The next task is to choose the saddlepoint of the function F in such a way that it converges to the saddlepoint of the function $F_{\mathrm{BS}}(k) := -\mathtt{i}kx - \Lambda_{\mathrm{BS}}(-\mathtt{i}k)$ for all $k \in \mathcal{Z}$, where Λ_{BS} is given by (4.1.2), in the Black-Scholes model as both the volatility of volatility and the correlation in model (3.2.1) tend to zero. It is easy to see that the saddlepoint of F_{BS} equals $\mathtt{i}u^*_{\mathrm{BS}}(x)$ for any $x \in \mathbb{R}$, were u^*_{BS} is given by (4.1.4). We can rewrite $\Im\left(z_0^{\pm}\right)$ defined in Lemma 5.4.2 as

$$\Im\left(z_0^{\pm}\right) = \frac{1}{2\bar{\rho}^2} - \frac{\kappa\rho}{\sigma\bar{\rho}^2} \pm \frac{\kappa\theta\rho\eta(x^2\sigma^2 + 2x\kappa\theta\rho\sigma + \kappa^2\theta^2)^{-1/2}}{2\sigma\bar{\rho}^2} \pm \frac{x\sigma\eta(x^2\sigma^2 + 2x\kappa\theta\rho\sigma + \kappa^2\theta^2)^{-1/2}}{2\sigma\bar{\rho}^2}, \tag{5.4.5}$$

where $\bar{\rho} := \sqrt{1-\rho^2}$. The first term converges to $1/2$ and the last one to $\pm x/\theta$ as (ρ, σ) tends to 0. When both ρ and σ tend to 0, a Taylor expansion at first order of the third term gives

$$\frac{\kappa\theta\rho\eta(x^2\sigma^2 + 2x\kappa\theta\rho\sigma + \kappa^2\theta^2)^{-1/2}}{2\sigma\bar{\rho}^2} = \frac{\rho\eta}{2\sigma\bar{\rho}^2}.$$

Take now the positive sign in (5.4.5), then the second and third terms cancel out in the limit since η converges to 2κ as (ρ, σ) tends to zero. In that case $\lim_{(\rho,\sigma)\to 0} \Im\left(z_0^+\right) = 1/2 + x/\theta = u^*_{\mathrm{BS}}(x)$ holds for all $x \in \mathbb{R}$, where the Black-Scholes variance is equal to θ. If we take the negative sign in (5.4.5), we do not recover u^*_{BS} because the function $(\rho, \sigma) \mapsto \rho/\sigma$ has no limit as the pair (ρ, σ) tends to 0. Hence the correct saddlepoint is z_0^+ and (4.2.21) implies $\Im(z_0^+(x)) = u^*(x)$ for all $x \in \mathbb{R}$. Note that in the fixed strike case (i.e. $x = 0$), we have $u^*(0) = (\sigma + \rho\eta - 2\kappa\rho) / \left(2\sigma\bar{\rho}^2\right)$, and the corresponding saddlepoint $\mathtt{i}u^*(0)$ is the same as the one in [73, Chapter 6]. The following lemma is of fundamental importance and will be the key tool for Proposition 5.4.6.

Lemma 5.4.3. *Let $k \in \mathcal{Z}$. Then the function $k_r \mapsto \Re(-\mathtt{i}kx - \Lambda(-\mathtt{i}k))$ has a unique minimum at 0 and is strictly decreasing (resp. increasing) on $(-\infty, 0)$ (resp. $(0,\infty)$).*

Proof. Note that the statement in the lemma is equivalent to the map $k_r \mapsto -\Re(\Lambda(-\mathtt{i}(k_r + \mathtt{i}k_{\mathtt{i}})))$ having a unique minimum at 0 for any $k_{\mathtt{i}} \in (u_-, u_+)$ and being increasing (resp. decreasing) on the positive (resp. negative) half line. Let $k_{\mathtt{i}} \in (u_-, u_+)$, then

$$\Re\Big(\Lambda(-\mathtt{i}(k_r + \mathtt{i}k_{\mathtt{i}}))\Big) = \frac{\kappa\theta}{\sigma^2}\left(\kappa - \rho\sigma k_{\mathtt{i}} - \Re\left(\sqrt{w(k_r) + \mathtt{i}v(k_r)}\right)\right),$$

where

$$w(k_r) := \sigma^2\bar{\rho}^2 k_r^2 - \sigma^2\bar{\rho}^2 k_{\mathtt{i}}^2 - \sigma(2\kappa\rho - \sigma)k_{\mathtt{i}} + \kappa^2 \quad \text{and} \quad v(k_r) := \left(2\kappa\rho - \sigma + 2\sigma\bar{\rho}^2 k_{\mathtt{i}}\right)\sigma k_r.$$

From the identity and the fact that the principal value of the square-root is used, we get

$$\Re\left(\sqrt{w(k_r) + \mathtt{i}v(k_r)}\right) = \frac{1}{2}\sqrt{2w(k_r) + 2\sqrt{w(k_r)^2 + v(k_r)^2}} \tag{5.4.6}$$

is monotonically increasing in w, w^2 and v^2. First, note that $w'(k_r) = 2\sigma^2\bar{\rho}^2 k_r$, hence w is a parabola with a unique minimum at 0, so that from (5.4.6), it suffices to prove the following claim:
Claim: For every $k_{\mathtt{i}} \in (u_-, u_+)$, the function $g := w^2 + v^2$ has a unique (strictly positive) minimum attained at $k_r = 0$ and is strictly increasing (resp. decreasing) for $k_r > 0$ (resp. $k_r < 0$).
Let us write $w(k_r) = \sigma^2\bar{\rho}^2 k_r^2 + \psi(k_{\mathtt{i}})$, for all $k_r \in \mathbb{R}$, where $\psi(k_{\mathtt{i}}) := \kappa^2 - \sigma^2\bar{\rho}^2 k_{\mathtt{i}}^2 - \sigma(2\kappa\rho - \sigma)k_{\mathtt{i}}$. We have

$$g(k_r) = \sigma^4\bar{\rho}^4 k_r^4 + \left(2\bar{\rho}^2\psi(k_{\mathtt{i}}) + \left(2\kappa\rho - \sigma + 2\sigma\bar{\rho}^2 k_{\mathtt{i}}\right)^2\right)\sigma^2 k_r^2 + \psi(k_{\mathtt{i}})^2, \quad \text{for all } k_r \in \mathbb{R}. \tag{5.4.7}$$

The coefficient $\sigma^2\bar{\rho}^4$ and the constant κ^2 are strictly positive, so the claim follows if $\chi(k_{\mathtt{i}}) > 0$ for all $k_{\mathtt{i}} \in (u_-, u_+)$, where

$$\chi(k_{\mathtt{i}}) := \sigma^2\left(2\bar{\rho}^2\psi(k_{\mathtt{i}}) + \left(2\kappa\rho - \sigma + 2\sigma\bar{\rho}^2 k_{\mathtt{i}}\right)^2\right) = 2\sigma^4\bar{\rho}^4 k_{\mathtt{i}}^2 + 2\sigma^3\bar{\rho}^2(2\kappa\rho - \sigma)k_{\mathtt{i}} + \sigma^2\left(\kappa^2 + (2\kappa\rho - \sigma)^2\right).$$

The discriminant $\Delta_\chi = -4\sigma^6\bar{\rho}^4\left(2\kappa^2 + (2\kappa\rho - \sigma)^2\right)$ is strictly negative, so that χ has no real root and is always strictly positive. This proves the claim and concludes the proof of the lemma. □

The following two results complete the proof of Theorem 5.1.1, by studying the behaviour of the two integrals in (5.4.1) as the time to maturity tends to infinity. The following lemma proves that the integral along $\gamma_{u^*(x)}$ is negligible and Proposition 5.4.6 hereafter provides the asymptotic behaviour of the integral along the contour $\zeta_{u^*(x)}$.

Lemma 5.4.4. *For any $m > \Lambda^*(x)$, there exists $R(m) > 0$ such that for every $k \in \mathcal{Z}$ with $|k_r| > R(m)$, we have*

$$|\exp(\mathtt{i}kxt)\,\phi_t(-k)| \le \exp(-mt), \quad \textit{for all } t \ge 1. \tag{5.4.8}$$

Therefore

$$\left|\exp\left(xt\right)\int_{\gamma_{u^*(x)}}\exp\left(\mathtt{i}kxt\right)\frac{\phi_t(-k)}{\mathtt{i}k-k^2}\,\mathrm{d}k\right| = \mathcal{O}\left(\mathrm{e}^{-(m-x)t}\right), \tag{5.4.9}$$

where the contour $\gamma_{u^*(x)}$ *is defined in* (5.4.2).

Remark 5.4.5. (i) For every $x \in \mathbb{R}$, we have $\Lambda^*(x) > x$ by (d) on page 54 and hence $m - x > 0$. Therefore the modulus of the integral (5.4.9) tends to zero exponentially in time and in m.
(ii) Recall that $u^*(x) \in (u_-, u_+)$ by Proposition 4.2.15 and hence inequality (5.4.8) can be applied when estimating integral (5.4.9).

Proof. We only need to prove (5.4.9). Recall from Lemma 5.4.1, after some rearrangements, that

$$\phi_t(-k) = \exp\Big(\left(t + v/(\kappa\theta)\right)\Lambda(-\mathtt{i}k)\Big)\left(1 - g(-k)\right)^{2\kappa\theta/\sigma^2}\left(1 + \mathcal{O}\left(\mathrm{e}^{-d(k)t}\right)\right).$$

It follows from equations (3.2.5) and (4.2.19) that

$$\begin{aligned}
&d\left(-(k_r + \mathtt{i}k_\mathtt{i})\right) \sim \sigma\bar{\rho}|k_r|, && \text{as } |k_r| \to \infty,\\
&\Lambda\left(-\mathtt{i}(k_r + \mathtt{i}k_\mathtt{i})\right) \sim -\kappa\theta\bar{\rho}|k_r|/\sigma, && \text{as } |k_r| \to \infty,\\
&\lim_{|k_r|\to\infty} g\left(-(k_r + \mathtt{i}k_\mathtt{i})\right) = \left(\rho - \mathtt{i}\bar{\rho}\right)^2 \neq 1, && \text{since } |\rho| < 1.
\end{aligned}$$

Hence, there exists a constant $C > 0$ such that

$$\left|\mathrm{e}^{\mathtt{i}kxt}\phi_t(-k)\right| \leq C\exp\left(-\left(t + \frac{v}{\kappa\theta}\right)\frac{\kappa\theta}{\sigma}\right)|k_r|.$$

Define $R(m) := \sigma\left(m + \log(C)\right)/(\kappa\theta)$. Then, if $|k_r| > R(m)$, (5.4.9) follows. □

Proposition 5.4.6. *For any* $R > 0$ *and* $x \in \mathbb{R} \setminus \left\{-\theta/2, \bar{\theta}/2\right\}$, *we have as* $t \to \infty$,

$$\frac{\exp(xt)}{2\pi}\Re\left(\int_{\zeta_{u^*(x)}}\mathrm{e}^{\mathtt{i}kxt}\frac{\phi_t(-k)}{\mathtt{i}k-k^2}\mathrm{d}k\right) = \frac{\exp\left(-\left(\Lambda^*(x) - x\right)t\right)}{\sqrt{2\pi t}}\left(A(x) + \mathcal{O}\left(1/t\right)\right), \tag{5.4.10}$$

where A *is given in* (5.1.2), Λ^* *in* (4.2.20) *and* $\zeta_{u^*(x)}$ *in* (5.4.3).

Proof. Let $x \in \mathbb{R} \setminus \left\{-\theta/2, \bar{\theta}/2\right\}$. Applying Lemma 5.4.1 on the compact interval $[-R, R]$, we have

$$\int_{\zeta_{u^*(x)}}\mathrm{e}^{\mathtt{i}kxt}\frac{\phi_t(-k)}{\mathtt{i}k-k^2}\mathrm{d}k = \int_{\zeta_{u^*(x)}}\frac{U\left(-\mathtt{i}k\right)}{\mathtt{i}k-k^2}\mathrm{e}^{(\mathtt{i}kx+\Lambda(-\mathtt{i}k))t}\left(1 + \varepsilon\left(k,t\right)\right)\mathrm{d}k,$$

for t large enough. By Lemma 5.4.3, we know that $k_r \mapsto -\Re(\mathtt{i}(k_r + \mathtt{i}u^*(x))x + \Lambda(-\mathtt{i}(k_r + \mathtt{i}u^*(x))))$ has a unique minimum at $k_r = 0$ and the value of the function at this minimum equals $\Lambda^*(x)$ by the definition of Λ^*. The functions Λ and U are analytic along the contour of integration and thus, by [81, Theorem 7.1, Chapter 4], we have

$$\begin{aligned}
\Re\left(\int_{\zeta_{u^*(x)}}\frac{U(-\mathtt{i}k)}{\mathtt{i}k-k^2}\mathrm{e}^{(\mathtt{i}kx+\Lambda(-\mathtt{i}k))t}\mathrm{d}k\right) &= \frac{\mathrm{e}^{xt}}{\sqrt{\pi t}}\mathrm{e}^{-\Lambda^*(x)t}\left(\frac{U\left(u^*(x)\right)}{\sqrt{2\Lambda''\left(u^*(x)\right)}} + \mathcal{O}\left(1/t\right)\right)\\
&= \frac{\exp(-(\Lambda^*(x) - x)t)}{\sqrt{2\pi t}}\left(A(x) + \mathcal{O}\left(1/t\right)\right)
\end{aligned}$$

as t tends to infinity. The $\varepsilon\left(k,t\right)$ term is a higher order term which we can ignore at this level. □

Lemma 5.4.4 and Proposition 5.4.6 complete the proof of Theorem 5.1.1 for the general case. Concerning the two special cases, we first introduce a new contour, the path of steepest descent, which represents the optimal (in a sense made precise below) path of integration. Note that, the general case can also be proved using this path, but Lemma 5.4.3 simplifies the proof.

5.4.3 Construction of the path of steepest descent

The following lemma computes the path of steepest descent—see Definition 1.2.3—explicitly in the Heston case for the function F given in (5.4.4) passing through the saddlepoint $\mathtt{i}u^*(x)$.

Lemma 5.4.7. *The path of steepest descent is the map $\gamma : \mathbb{R} \to \mathbb{C}$ defined by*

$$\gamma(s) := s + \mathtt{i}k_{\mathtt{i}}(s), \quad \textit{for all } s \in \mathbb{R},$$

where

$$k_{\mathtt{i}}(s) := -\frac{\beta - (\kappa\theta\rho + x\sigma)\sqrt{\psi(s)}}{2\kappa\theta\sigma\xi\bar{\rho}^2}, \tag{5.4.11}$$

$\beta := \kappa\theta\xi(2\kappa\rho - \sigma)$, $\psi(s) := 4\sigma^2\bar{\rho}^2\xi^2 s^2 + \kappa^2\theta^2\left((2\kappa\rho - \sigma)^2 + 4\kappa^2\bar{\rho}^2\right)\xi$ *and* $\xi := (\kappa\theta\rho + x\sigma)^2 + \kappa^2\theta^2\bar{\rho}^2$.

Note that ξ is strictly positive, so that the function $k_{\mathtt{i}}$ is well defined.

Proof. By definition, the contour of steepest descent is such that the function $\Im(F \circ \gamma)$ remains constant. So we look for the map γ such that $\Im(F(\gamma(s))) = 0$, for all $s \in \mathbb{R}$ because $F(\gamma(0))$ is already real. Using the identity $\Im\left(\sqrt{x + \mathtt{i}y}\right) = 4\left(2x + 2\sqrt{x^2 + y^2}\right)^{-1/2}$, for all $x, y \in \mathbb{R}$, we find that the function $F \circ \gamma$ is real along the contour $\gamma : s \mapsto s + \mathtt{i}k_{\mathtt{i}}(s)$. Note also that this contour is orthogonal to the imaginary axis at $k_{\mathtt{i}}(0)$ (see [98, Exercise 2, Chapter 8]). $\square$

Remark 5.4.8.

- The contour γ depends on x, but we do not write this dependence for clarity.
- By construction the saddlepoint defined in (4.2.21) satisfies $\mathtt{i}u^*(x) = \gamma(0)$.
- The map $s \mapsto k_{\mathtt{i}}(s) = \Im(\gamma(s))$ is even, i.e. γ is symmetric around the imaginary axis.

We now prove Theorem 5.1.1 in the two special cases $x \in \left\{-\theta/2, \bar{\theta}/2\right\}$. In these cases, we need a result similar to Proposition 5.4.6, since Lemma 5.4.4 still holds, i.e. we need the asymptotic behaviour of the integral in (5.4.10) for the two special cases. The problem with these special cases is that $(\mathtt{i}k - k^2)^{-1}$ in the integrand in (5.4.10) has a pole at the saddlepoint, so we need to deform the contour using Cauchy's integral theorem and take the real part to remove the singularity, before we can use a saddlepoint expansion.

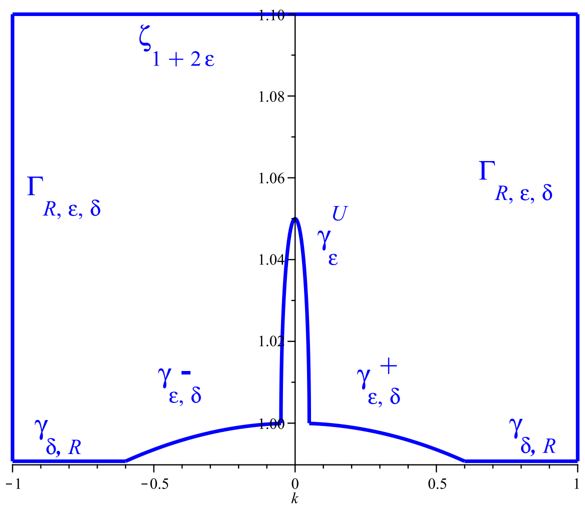


Figure 5.3: We plot the closed contour of integration in (5.4.12) for $\varepsilon = 0.05,\ \delta = 0.6,\ R = 1$.

5.4.4 Proof of the call price expansion for the special cases

We here prove Theorem 5.1.1 in the case $x = \bar{\theta}/2$ for which $u^*\left(\bar{\theta}/2\right) = 1$, $\Lambda^*\left(\bar{\theta}/2\right) = \bar{\theta}/2$, and for simplicity we also assume that $\kappa < \left(\sigma - 2\rho^2\sigma\right)/(2\rho)$ (the other cases follow similarly). From (5.4.11), we see that in this case, γ lies below the horizontal contour $\gamma_H : \mathbb{R} \to \mathbb{C}$ such that $\gamma_H(s) := s + \mathtt{i}$ (in the other case, γ lies above γ_H). We want to construct a new contour leaving the pole outside. Let $\varepsilon > 0$ and $\gamma_\varepsilon : (-\pi, \pi] \to \mathbb{C}$ denote the clockwise oriented circular keyhole contour parameterised as $\gamma_\varepsilon(\theta) := \mathtt{i} + \varepsilon\, \mathrm{e}^{\mathtt{i}\theta}$ around the pole. To leave the pole outside the new contour of integration, we need to follow γ on $\mathbb{R}_-$, switch to the keyhole contour as soon as we touch it, follow it clockwise (above the pole), and get back to γ on $\mathbb{R}_+$. As γ is below γ_H, it intersects γ_ε on its lower half, which can be analytically represented as $\gamma_\varepsilon^- : [-\varepsilon, \varepsilon] \to \mathbb{C}$ such that $\gamma_\varepsilon^-(s) := \mathtt{i} + s - \mathtt{i}\sqrt{\varepsilon^2 - s^2}$. From (5.4.11), the two contours intersect at $s^* = \pm\varepsilon$. Choose now $0 < \varepsilon < \delta < R$ (Lemma 5.4.9 makes the choice of δ precise and ε must be such that $1 + 2\varepsilon < u_+$) and define the following contours (they are all considered anticlockwise, see Figure 5.3)

- $\gamma_{\delta,R} : [-R, -\delta] \cup [\delta, R] \to \mathbb{C}$ given by $\gamma_{\delta,R}(u) = u + \mathtt{i}k_{\mathtt{i}}(\delta)$, with $k_{\mathtt{i}}(\delta)$ defined in (5.4.11);
- $\gamma_{\varepsilon,\delta}$ is the restriction of γ to the union of the intervals $[-\delta, -\varepsilon] \cup [\varepsilon, \delta]$;
- γ_ε^U is the portion of the circular keyhole contour γ_ε which lies above γ, i.e. the upper half keyhole contour as well as the two sections of γ_ε between γ and γ_H;
- $\Gamma^{\pm}_{R,\varepsilon,\delta}$ are the two vertical strips joining $\pm R + \mathtt{i}(1 + 2\varepsilon)$ to $\pm R - \mathtt{i}k_{\mathtt{i}}(\delta)$.

By Cauchy's integral theorem, we now have

$$\left(\int_{\gamma_{\delta,R}} + \int_{\gamma_{\varepsilon,\delta}} + \int_{\gamma_\varepsilon^U} + \int_{\Gamma^{\pm}_{R,\varepsilon,\delta}} - \int_{\zeta_{1+2\varepsilon}}\right) \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) \mathrm{d}k = 0, \tag{5.4.12}$$

which we can rewrite as

$$\left(\int_{\gamma_{1+2\varepsilon}} + \int_{\zeta_{1+2\varepsilon}}\right) \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \mathrm{e}^{\mathtt{i}k\bar{\theta}t/2}\mathrm{d}k = \left(\int_{\gamma_{\delta,R}} + \int_{\Gamma^{\pm}_{R,\varepsilon,\delta}} + \int_{\gamma_{1+2\varepsilon}} + \int_{\gamma^U_\varepsilon}\right) \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \mathrm{e}^{\mathtt{i}k\bar{\theta}t/2}\mathrm{d}k$$
$$+ \int_{\gamma_{\varepsilon,\delta}} \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) \mathrm{d}k, \tag{5.4.13}$$

where the curves $\zeta_{1+2\varepsilon}$ and $\gamma_{1+2\varepsilon}$ are defined in (5.4.3) and (5.4.2). The integral on the left is equal to the normalised call price $2\pi S_0^{-1}\mathrm{e}^{-\bar{\theta}t/2}\mathbb{E}\left(S_t - S_0\mathrm{e}^{\bar{\theta}t/2}\right)_+$ by [71, Theorem 5.1], which is independent of ε (this holds since $1 + 2\varepsilon < u_+$). For k close to $\mathtt{i}$, we have

$$\frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) = \left(\frac{\mathtt{i}}{k - \mathtt{i}} + \mathcal{O}(1)\right) \exp\left(-\bar{\theta}t/2\right),$$

so that

$$\int_{\gamma^U_\varepsilon} \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) \mathrm{d}k = \left(\pi + \mathcal{O}(\varepsilon)\right) \exp\left(-\bar{\theta}t/2\right). \tag{5.4.14}$$

Lemma 5.4.9 gives the behaviour of the last integral on the right-hand side of (5.4.13) as ε tends to 0 for δ small enough.

The other integrals can be bounded as follows. By Lemma 5.4.4, the integral along $\gamma_{1+2\varepsilon}$ is $\mathcal{O}\left(\mathrm{e}^{-\bar{\theta}t/2}\right)$, for $t > t^*(m)$, $R > R(m)$, as ε tends to 0. The curves $\Gamma^{\pm}_{R,0,\delta}$ are both vertical strips of length δ and therefore their images are compact sets. Applying the tail estimate of Lemma 5.4.4 along $\Gamma^{\pm}_{R,0,\delta}$, we know that for any $m > \bar{\theta}/2$, there exist $t(m)$ and $R(m)$ such that

$$\int_{\Gamma^{\pm}_{R,0,\delta}} \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) \mathrm{d}k = \mathcal{O}\left(\mathrm{e}^{-mt}\right), \quad \text{for all } t > t(m), |k| > R(m).$$

Lemma 5.4.3 implies that the function $k_r \mapsto \Re\left(-\mathtt{i}\left(k_r + \mathtt{i}k_\mathtt{i}\right)\bar{\theta}/2 - \Lambda(-\mathtt{i}(k_r + \mathtt{i}k_\mathtt{i}))\right)$ attains its global minimum at zero for any fixed $k_\mathtt{i} \in (u_-, u_+)$ and is strictly decreasing (resp. increasing) for $k_r < 0$ (resp. $k_r > 0$). Therefore the function $u \mapsto \Re\left(-\mathtt{i}\gamma_{\delta,R}(u)\bar{\theta}/2 - \Lambda(-\mathtt{i}\gamma_{\delta,R}(u))\right)$, for $u \in [-R, -\delta] \cup [\delta, R]$, attains its minimum value $g(\delta) := \Re\left(\left(k_\mathtt{i}(\delta) - \mathtt{i}\delta\right)\bar{\theta}/2 - \Lambda(k_\mathtt{i}(\delta) - \mathtt{i}\delta)\right)$ at the points $u = \pm\delta$, where $k_\mathtt{i}(\delta)$ is defined in (5.4.11). It can be checked directly that $g(0) = \bar{\theta}/2$, $g'(0) = 0$ and $g''(0) > 0$ and hence for every $\delta > 0$ there exists $\varepsilon_0 > 0$ such that the inequality

$$\Re\left(-\mathtt{i}\gamma_{\delta,R}(u)\bar{\theta}/2 - \Lambda(-\mathtt{i}\gamma_{\delta,R}(u))\right) > \bar{\theta}/2 + \varepsilon_0, \qquad \text{for all } u \in [-R, -\delta] \cup [\delta, R]$$

holds. Therefore Lemma 5.4.1 yields the following inequality

$$\left|\int_{\gamma_{\delta,R}} \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \mathrm{e}^{\mathtt{i}k\bar{\theta}t/2}\mathrm{d}k\right| \leq \mathrm{e}^{-(\bar{\theta}/2+\varepsilon_0)t} \int_{\gamma_{\delta,R}} \left|\frac{U(-\mathtt{i}k)(1 + \varepsilon(k,t))}{\mathtt{i}k - k^2}\right| \mathrm{d}k = \mathcal{O}\left(\exp\left(-\left(\bar{\theta}/2 + \varepsilon_0\right)t\right)\right).$$

We now prove the following lemma about the integral along $\gamma_{\varepsilon,\delta}$ as ε tends to 0.

Lemma 5.4.9. *For $\delta > 0$ and sufficiently small we have*

$$\lim_{\varepsilon \searrow 0} \int_{\gamma_{\varepsilon,\delta}} \frac{\phi_t(-k)}{\mathtt{i}k - k^2} \exp\left(\mathtt{i}k\bar{\theta}t/2\right) \mathrm{d}k = \sqrt{\frac{2\pi}{\Lambda''(1)t}} \mathrm{e}^{-\bar{\theta}t/2} \left(-1 - \frac{1}{6}\frac{\Lambda'''(1)}{\Lambda''(1)} + U'(1)\right) \left(1 + \mathcal{O}(1/t)\right),$$

where U is given by (5.1.3) *and Λ by* (4.2.19).

Proof. Recall that γ is the contour of steepest descent defined in Lemma 5.4.7 and that the curve $\gamma_{\varepsilon,\delta}$ is its restriction to the intervals $[-\delta,-\varepsilon]\cup[\varepsilon,\delta]$. Since the function $s\mapsto\Re\left(\frac{\phi_t(-\gamma(s))\gamma'(s)}{\mathtt{i}\gamma(s)-\gamma(s)^2}\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2}\right)$ is even and the function $s\mapsto\Im\left(\frac{\phi_t(-\gamma(s))\gamma'(s)}{\mathtt{i}\gamma(s)-\gamma(s)^2}\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2}\right)$ is odd, we obtain

$$\int_{\gamma_{\varepsilon,\delta}}\frac{\phi_t(-k)}{\mathtt{i}k-k^2}\mathrm{e}^{\mathtt{i}k\bar{\theta}/2t}\mathrm{d}k=\int_{[-\delta,-\varepsilon]\cup[\varepsilon,\delta]}\Re\left(\frac{\phi_t\left(-\gamma(s)\right)\gamma'(s)}{\mathtt{i}\gamma(s)-\gamma(s)^2}\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2}\right)\mathrm{d}s. \tag{5.4.15}$$

From (5.4.11) we have $\left(\mathtt{i}\gamma(s)-\gamma(s)^2\right)^{-1}=\mathtt{i}/s-1+\mathcal{O}\left(s\right)$, so $\Re\left(\left(\mathtt{i}\gamma(s)-\gamma(s)^2\right)^{-1}\right)=-1+\mathcal{O}\left(s\right)$, for s around 0, i.e. taking the real part removes the singularity at $k=\mathtt{i}$. Lemma 5.4.1 implies

$$\int_{-\delta}^{\delta}\Re\left(\frac{\phi_t(-\gamma(s))\gamma'(s)}{\mathtt{i}\gamma(s)-\gamma(s)^2}\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2}\right)\mathrm{d}s=\int_{-\delta}^{\delta}\Re\left(q(s)\right)\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2+\Lambda(-\mathtt{i}\gamma(s))t}\mathrm{d}s+\mathcal{O}\left(\mathrm{e}^{-mt}\right),$$

for some $m>0$ large enough, where we define the function $q:\mathbb{R}\backslash\{0\}\to\mathbb{C}$ by

$$q(s):=\frac{U(-\mathtt{i}\gamma(s))\gamma'(s)}{\mathtt{i}\gamma(s)-\gamma(s)^2},\quad\text{for all } s\in\mathbb{R}.$$

Then, from (5.4.11), we have the following expansion

$$q(s)=\left(\frac{\mathtt{i}}{s}-\left(\frac{\Lambda'''(1)}{6\Lambda''(1)}+1\right)\right)\left(1-\mathtt{i}U'(1)s\right)+\mathcal{O}\left(s^3\right). \tag{5.4.16}$$

We can therefore extend the function q to the map $q:B_\delta(0)\backslash\{0\}\to\mathbb{C}$ for some $\delta>0$, where $B_\delta(0):=\{z\in\mathbb{C}:|z|<\delta\}$ is an open disc of radius δ. Note that for $s\in\mathbb{R}$, we have $\Re\left(q(s)\right)=-1-\Lambda'''(1)/\left(6\Lambda''(1)\right)+U'(1)+\mathcal{O}\left(s\right)$ and hence the function $\Re(q):[-\delta,\delta]\to\mathbb{R}$ does not have a singularity at $s=0$. Recall that if a function $G:B_\delta(0)\backslash\{0\}\to\mathbb{C}$ has a Laurent series expansion

$$G(z)=\frac{\mathtt{i}a_{-1}}{z}+\sum_{n=0}^{\infty}a_nz^n,\quad\text{for } z\in B_\delta(0)\setminus\{0\},$$

with $a_{-1}\in\mathbb{R}$, then the function $\Re(G):\mathbb{R}\cap B_\delta(0)\to\mathbb{R}$ has an analytic continuation on the whole disc $B_\delta(0)$. It follows from (5.4.16) that there exists a holomorphic function $Q:B_\delta(0)\to\mathbb{C}$ such that $Q(s)=\Re(q(s))$ for any $s\in(-\delta,\delta)$. Thus by [81, Theorem 7.1, Chapter 4], we have

$$\begin{aligned}\int_{-\delta}^{\delta}\Re\left(q(s)\right)\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2+\Lambda(-\mathtt{i}\gamma(s))t}\mathrm{d}s&=\int_{-\delta}^{\delta}Q(s)\mathrm{e}^{\mathtt{i}\gamma(s)\bar{\theta}t/2+\Lambda(-\mathtt{i}\gamma(s))t}\mathrm{d}s\\&=\sqrt{\frac{2\pi\,\mathrm{e}^{-\bar{\theta}t}}{\Lambda''(1)t}}\left(-1-\frac{1}{6}\frac{\Lambda'''(1)}{\Lambda''(1)}+U'(1)\right)\left(1+\mathcal{O}\left(1/t\right)\right).\end{aligned}$$

□

Letting ε go to 0 in equation (5.4.13), applying Lemma 5.4.9 and the bounds developed above for the other integrals in (5.4.13) concludes the proof of the theorem follows in the case $x=\bar{\theta}/2$. The case $x=-\theta/2$ is analogous.

5.5 Additional proofs

5.5.1 Proof of Lemma 5.1.4

The proof is analogous to that of Theorem 5.1.1. The residue in Theorem 5.1.1 is equal to 1 (arising from the $\mathbb{1}_{\{-\theta/2<x<\bar{\theta}/2\}}$ term), and from (5.4.1), the integral part reads for R large enough

$$\frac{\exp(xt)}{2\pi}\Re\left(\left(\int_{\zeta_{u^*(0)}}+\int_{\gamma_{u^*(0)}}\right)\frac{\phi_t(-k)}{\mathtt{i}k-k^2}\mathrm{e}^{\mathtt{i}kxt}\mathrm{d}k\right).$$

The behaviour of these integrals follows exactly the lines of the proof of Theorem 5.1.1.

5.5.2 Proof of Proposition 5.1.5

Consider a squared volatility of the form $\hat{\sigma}_t^2=\Sigma^2+a_1/t>0$. Then

$$\frac{C_{\mathrm{BS}}\left(xt,t,\hat{\sigma}_t\right)}{S_0}=\mathcal{N}\left(\frac{-x+\left(\Sigma^2+a_1/t\right)/2}{\sqrt{\Sigma^2+a_1/t}}\sqrt{t}\right)-\mathrm{e}^{xt}\mathcal{N}\left(\frac{-x-\left(\Sigma^2+a_1/t\right)/2}{\sqrt{\Sigma^2+a_1/t}}\sqrt{t}\right), \tag{5.5.1}$$

holds. Define $z_\pm:=\left(-x\pm\frac{1}{2}\left(\Sigma^2+a_1/t\right)\right)\sqrt{t}/\sqrt{\Sigma^2+a_1/t}$ and recall that (see [81])

$$\mathcal{N}(-z)=\frac{\exp\left(-z^2/2\right)}{z\sqrt{2\pi}}\left(1+\mathcal{O}\left(\frac{1}{z^2}\right)\right),\quad\text{as } z \text{ tends to infinity.} \tag{5.5.2}$$

The case $x>\Sigma^2/2$. Since $\Sigma^2/2=\lim_{t\to\infty}\hat{\sigma}_t^2/2$, there exists $t^*>0$ such that for all $t>t^*$, $x>\hat{\sigma}_t^2/2=(\Sigma+a_1/t)^2/2$. From (5.5.1), we have, using a Taylor expansion for $z_\pm$,

$$\begin{aligned}\frac{C_{\mathrm{BS}}\left(xt,t,\hat{\sigma}_t\right)}{S_0} &= \frac{\mathrm{e}^{\frac{1}{8}a_1\left(4x^2/\hat{\Sigma}^4-1\right)}}{\sqrt{2\pi t}}\left(\frac{\Sigma}{x-\Sigma^2/2}-\frac{\Sigma}{x+\Sigma^2/2}\right)\exp\left(-\frac{\left(x-\Sigma^2/2\right)^2t}{2\Sigma^2}\right)\left(1+\mathcal{O}\left(\frac{1}{t}\right)\right)\\ &= \frac{A_{\mathrm{BS}}(x,\Sigma,a_1)}{\sqrt{2\pi t}}\exp\left(-\left(\Lambda^*_{\mathrm{BS}}(x,\Sigma)-x\right)t\right)\left(1+\mathcal{O}\left(\frac{1}{t}\right)\right).\end{aligned}$$

The cases $x<-\Sigma^2/2$ and $-\Sigma^2/2<x<\Sigma^2/2$ follow likewise.

The case $x=\Sigma^2/2$. From (5.5.1), we have

$$\begin{aligned}\frac{C_{\mathrm{BS}}\left(\Sigma^2t/2,t,\hat{\sigma}_t\right)}{S_0} &= \mathcal{N}\left(\frac{a_1/2}{\sqrt{\Sigma^2t+a_1}}\right)-\frac{\mathrm{e}^{\Sigma^2t/2}\sqrt{\Sigma^2t+a_1}}{xt+(\Sigma^2t+a_1)/2}\exp\left(-\frac{(xt+(\Sigma^2t+a_1)/2)^2}{2(\Sigma^2t+a_1)}\right)\left(1+\mathcal{O}\left(\frac{1}{t}\right)\right)\\ &= \frac{1}{2}+\frac{a_1/2}{\Sigma\sqrt{2\pi t}}-\frac{1}{\Sigma\sqrt{2\pi t}}(1+\mathcal{O}\left(1/t\right))=\frac{1}{2}+\frac{A_{\mathrm{BS}}\left(\Sigma^2/2,\Sigma,a_1\right)}{\sqrt{2\pi t}}\left(1+\mathcal{O}\left(\frac{1}{t}\right)\right).\end{aligned}$$

The case $x=-\Sigma^2/2$ is analogous.

Part III

The small-maturity behaviour

Introduction

Large deviations theory provides a natural framework for approximating the exponentially small probabilities associated with the behaviour of a diffusion process over a small time interval. In the context of Mathematical Finance, large deviations theory arises in the computation of small-maturity, out-of-the-money call or put option prices, or the probability of reaching a barrier or default level in a small time period (see the recent overview by Pham [83]). In recent years there has been an explosion of literature on small-time asymptotics for stochastic volatility models ([11], [12], [39], [42], [50], [56], [57], [61], [82]). All these articles characterise the behaviour of the implied volatility for European options in the small-maturity limit, and they are essentially applications and/or higher order corrections to the seminal work of Varadhan [102] and the Freidlin-Wentzell theory of large deviations for stochastic differential equations [45].

Berestycki, Busca & Florent [12] proved that for a stochastic volatility model with coefficients satisfying certain growth conditions, the implied volatility in the small-maturity limit can be characterised in terms of the viscosity solution to an non-linear eikonal partial differential equation. The solution to this eikonal equation is the length of the shortest geodesic from the initial value of the diffusion to a given line under a metric characterised by the coefficients of the model. Henry-Labordère [57] formally computed a small-time expansion for the local volatility and the implied volatility for a general stochastic volatility model, using an expansion of the heat kernel on a Riemannian manifold. In the same spirit Paulot [82] developed a Taylor expansion for the implied volatility at all strikes when the maturity is small and gave a precise example for the SABR model. More recently, Gatheral et al. [50] looked at small-time asymptotics for one-dimensional local volatility models and, conditioning with a bridge process, derived a small-time expansion for the transition density of the process and for the corresponding implied volatility.

In this part of the thesis we prove some new results in this direction. In Chapter 6 we first digress slightly from our affine stochastic volatility framework to adopt a geometric approach for an uncorrelated local-stochastic volatility model. The reason why we outline these results is that they bridge the gap between the geometric approach developed in [82], [50] and [57], the PDE approach of [12] and sample path large deviations theory. This is clearly a very first step towards the full picture and future research will have to be carried out to extend the scope of these results.

Chapter 7 concentrates on the extended Heston model (with $a \neq 0$) and derives a small-time large deviations principle, from which we are able to precisely deduce the behaviour of the implied volatility for small maturities. Chapter 8 parallels Chapter 5 in Part II and refines the small-time asymptotics of the implied volatility smile in the Heston model using complex saddlepoint methods. Finally in Chapter 8.5.3 we explain the difficulties that arise as soon as the models is allowed to jump and we point out recent preliminary results in this direction. This is an area of active research and we shall simply aim at giving an overview of the different possible directions to be followed in the future.

Chapter 6

The geometric approach

In this chapter we momentarily forget about affine stochastic volatility models to cast a glance at a different approach which does not require the knowledge of the Laplace transform. This is essentially the approach followed in [12] and [57]. We first concentrate on an uncorrelated stochastic volatility model generating symmetric smiles, and we then add a local volatility component in order to take into account the asymmetry observed on the markets. This approach will shed some light on the connection between the large deviations approach of Chapter 7 and the results in [12]. Henry-Labordère [58] has developed a similar approach where he considers a general stochastic volatility model for which he computes a heat kernel expansion for the probability density of the share price process. This knowledge enables him first to compute the corresponding expansion for the local volatility, from which he deduces a small-time asymptotic expression for the implied volatility. Our approach here is more direct since we do not need these different steps and we directly compute the limiting implied volatility as the maturity of the option tends to zero, via the limiting probability distribution of the process. The drawback of our approach is that we need to impose some conditions on the coefficients of the stochastic volatility models under consideration.

We consider a share price process $(S_t)_{t\geq 0}$ and we assume that its logarithm $(X_t)_{t\geq 0}$ satisfies the following system of stochastic differential equations:

$$\begin{aligned} \mathrm{d}X_t &= -\tfrac{1}{2} f(V_t)^2 \mathrm{d}t + f\left(V_t\right) \mathrm{d}W_t, \qquad && X_0 = x_0 \in \mathbb{R}, \\ \mathrm{d}V_t &= \mu\left(V_t\right) \mathrm{d}t + \alpha\left(V_t\right) \mathrm{d}Z_t, && V_0 = v_0 \in \mathbb{R}, \end{aligned} \tag{6.0.1}$$

where $(W_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$ are two independent Brownian motions. In order to streamline the presentation of this chapter we will assume that the share price is normalised so that $x_0 = 0$. We also make the following assumptions on the coefficients of the SDEs:

Assumption 6.0.1. The function μ is bounded and uniformly Lipschitz continuous, and there exists a strictly positive constant m_f such that the inequalities $0 < m_f^{-1} \leq f(y) \wedge \alpha(y) \leq f(y) \vee$

$\alpha(y) \leq m_f < \infty$ hold for all $y \in \mathbb{R}$. Assume further that the functions α and f are strictly increasing $C^\infty(\mathbb{R})$ diffeomorphisms, and converge to some finite values $f_{\pm\infty}$ and $\alpha_{\pm\infty}$ as y tends to (plus or minus) infinity. Finally, assume that there exists $y_c \in \mathbb{R}$ such that α and f are both strictly concave on $[y_c, \infty)$, strictly convex outside and that their derivatives are bounded. We also impose that the function $u \mapsto u^{-1}\alpha(f^{-1}(u))$ is non-increasing.

This model is similar to the one used in [44] where the instantaneous volatility process is bounded away from zero. Let us now define the Riemannian manifold $M = \mathbb{R}^2$ equipped with the metric g whose coefficients are given by $\mathrm{g}_{11} = f(y)^{-2}$, $\mathrm{g}_{12} = \mathrm{g}_{21} = 0$ and $\mathrm{g}_{22} = \alpha(y)^{-2}$. We shall define the function d on this manifold as the distance under g from the point $(0, v_0)$ to the vertical line l_x of constant abscissa x, and we shall use the notation $d(x)$ whenever there is no doubt on the starting point $(0, v_0)$. The conditions in the theorem ensure that $\pm\infty$ are natural boundaries for the process $(V_t)_{t\geq 0}$. The existence and uniqueness of a strong solution for the process V is standard, and the process X is a stochastic integral of V, hence is well defined. The function f is bounded, so that $\exp(X_t)$ is a martingale by the Novikov criterion. Note that the condition on the function $u \mapsto u^{-1}\alpha(f^{-1}(u))$ is trivially satisfied if $\alpha(y) = \xi f(y)$ for some $\xi > 0$, or if α is just a positive constant. Finally note that for a general stochastic volatility model with non-zero correlation where $\mathrm{d}\langle W_t, X_t\rangle = \rho\,\mathrm{d}t$, Lewis [74] has noted that the shortest geodesic γ^* comes perpendicular to the line l_{x_1} under the metric g associated with the correlated model.

We represent a curve γ in the plane parameterised by time by $\gamma(t) := (x(t), y(t))$, where x and y are two functions of time representing the abscissa and the ordinate of the curve. For such a curve we shall use the notation $\dot{\gamma}(t) = (\dot{x}(t), \dot{y}(t))$ in place of $\mathrm{d}\gamma/\mathrm{d}t$. In the following the notation x might stand for either a real number or a curve. The context should make it clear which one it is which should avoid any possible confusion. The following theorem characterises the small-time transition probabilities of the process $(X_t)_{t\geq 0}$ under Assumption 6.0.1 in terms of the distance function d. We shall state and prove some corollaries of this theorem before actually proving it in order to highlight its role in determining the asymptotic behaviour of the implied volatility.

Theorem 6.0.2. *Under Assumption 6.0.1 the distribution of the process X satisfies*

$$-\lim_{t\to 0} t \log \mathbb{P}(X_t > x) = \frac{1}{2} d(x)^2, \quad \text{for all } x \neq 0. \tag{6.0.2}$$

If in addition $v_0 > y_c$ then there is a unique geodesic γ^* from the point $(0, v_0)$ to the line l_x, such that γ^* is perpendicular to the vertical axis at the point $(x, y^*(x))$ where it intersects with the line l_x. In this case we can calculate the distance $d(x)$ explicitly as

$$d(x) = f(y^*(x)) \int_{v_0}^{y^*(x)} \left(f(y^*(x))^2 - f(y)^2\right)^{-1/2} \frac{\mathrm{d}y}{\alpha(y)}, \tag{6.0.3}$$

where the real number $y^*(x)$ is defined implicitly by the relation

$$x = \int_{v_0}^{y^*(x)} \left(f(y^*(x))^2 - f(y)^2\right)^{-1/2} \frac{f(y)^2 \mathrm{d}y}{\alpha(y)}. \tag{6.0.4}$$

A useful corollary of this result are the following small-time limits for European option prices. The proof is analogous to the proof of Corollary 7.1.3 further below and we therefore omit it here.

Corollary 6.0.3. *For the uncorrelated stochastic volatility model* (6.0.1)*, the following limits hold*

$$
\begin{aligned}
-\lim_{t\to 0} t\log\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^{x}\right)_+ &= \frac{1}{2}d\left(x\right)^2, \quad \text{if } x>0,\\
-\lim_{t\to 0} t\log\mathbb{E}\left(\mathrm{e}^{x}-\mathrm{e}^{X_t}\right)_+ &= \frac{1}{2}d\left(x\right)^2, \quad \text{if } x<0,
\end{aligned}
$$

and for all $x\in\mathbb{R}^*$*,*

$$
\lim_{t\to 0} t\log\Big(\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^{x}\right)_+ - \left(1-\mathrm{e}^{x}\right)_+\Big) = \lim_{t\to 0} t\log\Big(\mathbb{E}\left(\mathrm{e}^{x}-\mathrm{e}^{X_t}\right)_+ - \left(\mathrm{e}^{x}-1\right)_+\Big) = \frac{1}{2}d\left(x\right)^2.
$$

Remark 6.0.4. Note that $d(0)=0$, so these estimates are very crude for at-the-money options.

The following corollary computes the asymptotic behaviour of the implied volatility, the proof of which follows from Corollary 6.0.3 and is analogous to that of Theorem 7.1.4 below.

Corollary 6.0.5. *As the maturity* t *tends to zero, we have* $\lim_{t\to 0}\sigma_t(x) = |x|/d\left(x\right)$ *for all* $x\neq 0$*.*

Proof. of Theorem 6.0.2. We divide the proof of the theorem into a series of steps.

Part 1: Freidlin-Wentzell theory for the uncorrelated stochastic volatility model. Let $\varepsilon>0$ and $(X_t^\varepsilon, V_t^\varepsilon)_{t\geq 0}$ be the unique strong solution of the stochastic differential equations

$$
\begin{aligned}
\mathrm{d}X_t^\varepsilon &= -\tfrac{1}{2}\varepsilon f(V_t^\varepsilon)^2\mathrm{d}t + \sqrt{\varepsilon}f\left(V_t^\varepsilon\right)\mathrm{d}W_t, \qquad && X_0^\varepsilon = 0,\\
\mathrm{d}V_t^\varepsilon &= \varepsilon\mu\left(V_t^\varepsilon\right)\mathrm{d}t + \sqrt{\varepsilon}\alpha\left(V_t^\varepsilon\right)\mathrm{d}Z_t, && V_0^\varepsilon = v_0.
\end{aligned}
$$

Then the two processes $(X_t^\varepsilon)_{t\geq 0}$ and $(V_t^\varepsilon)_{t\geq 0}$ have the same drift and diffusion coefficient terms as $(X_{\varepsilon t})_{t\geq 0}$ and $(V_{\varepsilon t})_{t\geq 0}$ respectively so they have the same law. Let $C\left([0,1],\mathbb{R}^2\right)$ denote the space of continuous curves from $[0,1]$ to $\mathbb{R}^2$. It is well known [102, Theorem 6.3] that the joint process $(X^\varepsilon, V^\varepsilon)$ satisfies a large deviations principle as ε tends to zero with lower semicontinuous rate function $I:\mathbb{R}\times\mathbb{R}_+$ characterised by

$$
I(\gamma) = \frac{1}{2}\int_0^1 \Big(f(y(t))^{-2}\dot{x}(t)^2 + \alpha(y(t))^{-2}\dot{y}(t)^2\Big)\,\mathrm{d}t.
$$

By Theorem 1.1.4 we know that X_1^ε (and hence X_ε) satisfy a small-time LDP with the lower semicontinuous rate function $\iota:\mathbb{R}\to\mathbb{R}_+\cup\{+\infty\}$ defined by

$$
\iota(x_1) := \inf\Big\{I\left(x,y\right) : (x,y)\in C\left([0,1],\mathbb{R}^2\right) : x(0)=0,\ y(0)=v_0,\ x(1)=x_1\Big\}.
$$

By the definition of the large deviations principle, it follows that

$$
-\inf_{x>x_1}\iota(x_1) \leq \liminf_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\left(X_\varepsilon\leq x_1\right) \leq \limsup_{\varepsilon\to 0}\varepsilon\log\mathbb{P}\left(X_\varepsilon\geq x_1\right) \leq -\inf_{x\geq x_1}\iota(x_1).
$$

We now prove the continuity of the function ι.

Part 2: Converting from a large deviations problem to a differential geometry problem: the distance from a point to a geodesic. The function ι defined above is obtained as the solution to a variable endpoint calculus of variations problem, which is equivalent to the differential geometry problem of finding the infimum of the energy functional $I(\gamma)$ over all piecewise differentiable curves $\gamma = (x(t), y(t))_{t\in[0,1]}$ joining the point $(0, v_0)$ to the line l_{x_1} on the manifold M, furnished with the Riemannian metric g whose coefficients are given by

$$\mathrm{g}_{11} = f(y)^{-2}, \quad \mathrm{g}_{12} = \mathrm{g}_{21} = 0, \quad \text{and } \mathrm{g}_{22} = \alpha(y)^{-2}. \tag{6.0.5}$$

Without loss of generality we assume that $x_1 > 0$ since g is independent of the x-component of the curve, so that the shortest geodesic to the line l_{-x_1} is just the reflection of the shortest geodesic to l_{x_1} about the vertical axis. By the Hopf-Rinow theorem [29, Theorem 2.8, Chapter 7] a Riemannian manifold is complete as a metric space if and only if it is geodesically complete, which holds if and only if for any two points p and q in M, there exists a geodesic γ joining p and q. The manifold $\mathbb{R}^2$ is complete so that using the bounds on the function f, we can easily compare Cauchy sequences on M to $\mathbb{R}^2$, and verify that M is complete as a metric space. Therefore [29, Lemma 2.3, Chapter 9] implies that $I(\gamma) = l(\gamma)^2/2$, where $l(\gamma)$ denotes the length of γ under the metric g, and γ is the continuous distance-minimising curve from p to q for both the distance and the energy functional.

Part 3: Integrating the geodesic equations. The non-zero Christoffel symbols (defined in Part I, Section 1.3) for the metric g are given by

$$\Gamma^2_{11} = \alpha(y)^2 f'(y)/f(y)^3, \qquad \Gamma^1_{21} = \Gamma^1_{12} = -f'(y)/f(y), \qquad \Gamma^2_{22} = -\alpha'(y)/\alpha(y).$$

Thus the Euler-Lagrange equations (1.3.1) read (we drop the t-dependence for clarity)

$$\ddot{x} - 2f(y)^{-1} f'(y)\,\dot{x}\dot{y} = 0, \tag{6.0.6}$$

$$\ddot{y} + \alpha(y)^2 f(y)^{-3} f'(y)\dot{x}^2 - \alpha(y)^{-1}\alpha'(y)\dot{y}^2 = 0. \tag{6.0.7}$$

Assume that y is constant in time then $\dot{x} = 0$, i.e. the function x is constant as well and hence we cannot have horizontal geodesics. We can divide (6.0.6) by $f(y)^2$ and integrate to obtain

$$f(y)^{-2}\dot{x} = M, \tag{6.0.8}$$

where M is the x-component of momentum under g, which is conserved along geodesics i.e. is constant. Here we only need to consider geodesics for which $\dot{x} > 0$ so that we can assume $M > 0$. Since $0 < \dot{x} < \infty$ the derivative $\dot{y}$ is null if and only if the ratio $\mathrm{d}y/\mathrm{d}x$ is null as well. If $\dot{y}$ is not null we can multiply (6.0.7) by $\dot{y}$ and divide by $\alpha(y)^2$ to obtain

$$\begin{aligned}\alpha(y)^{-2}\ddot{y}\dot{y} + \dot{x}^2 f(y)^{-3} f'(y)\dot{y} - \alpha'(y)\alpha(y)^{-3}\dot{y}^3 &= \alpha(y)^{-2}\ddot{y}\dot{y} + M^2 f'(y) f(y)\dot{y} - \alpha'(y)\alpha(y)^{-3}\dot{y}^3 \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\alpha(y)^{-2}\dot{y}^2 + M^2 f(y)^2\right) = 0.\end{aligned}$$

Therefore there exists a real constant E such that $M^2 f(y)^2 + \alpha(y)^{-2}\dot{y}^2 = E$ or equivalently

$$\dot{y}^2 = \alpha(y)^2 M^2 \left(c - f(y)^2\right), \tag{6.0.9}$$

where we define the constant $c := E/M^2$. We can clearly assume that c is chosen such that $f(v_0)^2 < c < f_\infty^2$, so that $\dot{y} > 0$ at $t = 0$.[1] Then $\dot{y} = 0$ at the point $\overline{y} := f^{-1}\left(\sqrt{c}\right) > v_0$ and the derivative $\dot{y}$ is strictly positive for all $y < \overline{y}$. We can integrate this ODE directly to obtain

$$t(y) = \frac{1}{M}\int_{v_0}^{y} \frac{\mathrm{d}u}{\alpha(y)\sqrt{c - f(u)^2}}. \tag{6.0.10}$$

The function t is henceforth invertible on $(v_0, \overline{y})$ and the length of the geodesic γ is then given by

$$l\left(\gamma\right) := \int_0^{t(y)} \left|\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right| \mathrm{d}t = \sqrt{c}\int_{v_0}^{y} \frac{\mathrm{d}u}{\alpha(u)\sqrt{c - f(u)^2}}.$$

For all $y > \overline{y}$ we take the other root when solving the ODE (6.0.9), and the solution is given by

$$t(y) = \frac{1}{M}\int_{v_0}^{f^{-1}(\sqrt{c})} \frac{\mathrm{d}y}{\alpha(y)\sqrt{c - f(y)^2}} + \frac{1}{M}\int_{y}^{f^{-1}(\sqrt{c})} \frac{\mathrm{d}y}{\alpha(y)\sqrt{c - f(y)^2}}.$$

The spurious solution given by the horizontal line $x(t) = M^2 f\left(\overline{y}\right)^2 t$ and $y(t) = \overline{y}$ also satisfies (6.0.8) and (6.0.9), but fails to satisfy the original geodesic equations (6.0.6) and (6.0.7), and thus is not a true geodesic. This is an artefact of multiplying by $\dot{y}$ when we integrated the geodesic equations. For given M and E, there is a unique solution to (6.0.9) which is a valid geodesic. [29, Lemma 2.3, Chapter 9] tells us that any distance-minimising piecewise differentiable curve is necessarily a geodesic. Since we have established existence and uniqueness for the geodesics, we see that the inverse of the solution $t(y)$ in (6.0.10) is the y-component of a geodesic. Using (6.0.9), we can compute the x-component $x(t)$ by integrating (6.0.8) explicitly to obtain

$$x(t) = M\int_0^t f(y(s))^2 \mathrm{d}s = \int_{v_0}^{y(t)} \frac{f(y)^2 \mathrm{d}y}{\alpha(y)\sqrt{c - f(y)^2}}, \quad \text{for all } t \leq t\left(\overline{y}\right). \tag{6.0.11}$$

From (6.0.11), we see that if the condition

$$0 < x_1 < \int_{v_0}^{y_1} \frac{f(y)^2 \mathrm{d}y}{\alpha(y)\sqrt{f(y_1)^2 - f(y)^2}}$$

is satisfied for some x_1 and $y_1 > v_0$, then there is a unique $c > f(y_1)^2$ such that the equality

$$x_1 = \int_{v_0}^{y_1} \frac{f(y)^2 \mathrm{d}y}{\alpha(y)\sqrt{c - f(y)^2}},$$

holds, since x_1 is a continuous, strictly monotonically decreasing function of c. This means that we can find a unique geodesic joining $(0, v_0)$ to (x_1, y_1).

[1] For this variable endpoint problem we can restrict our attention to geodesics satisfying $\dot{y}(0) > 0$ since any geodesic starting from $(0, v_0)$ with $\dot{y}(0) < 0$ which passes through the line l_{x_1} will be longer than the segment of the horizontal line $\{y = v_0\}$ between $x = 0$ and $x = x_1$ because the functions $\mathrm{g}_{11} : y \mapsto f(y)^{-2}$, and $\mathrm{g}_{22} : y \mapsto \alpha(y)^{-2}$ are both strictly decreasing.

We now return to the variable endpoint problem of finding a geodesic from the point $p = (0, v_0)$ to the vertical line l_{x_1} with $x_1 > 0$. Using the bounds on the functions f and α, we see that the distance from $(0, v_0)$ to (x_1, y_1) tends to infinity as y_1 tends to infinity, so that any geodesic γ^* for the variable endpoint problem must be finite at x_1. Let $\varepsilon > 0$ and consider now a (proper) variation (see Definition 1.3.9) $f_s^\gamma(t) : [0, 1] \to M$ for $s \in (-\varepsilon, \varepsilon)$ of a geodesic γ which starts at $(0, v_0)$ and also passes through the line l_{x_1} which satisfies $f(s, 0) = (0, v_0)$ and $f(s, 1) = (x_1, y_1)$ for all $s \in (-\varepsilon, \varepsilon)$. We know that a geodesic γ satisfies the Euler-Lagrange equations above. It also has to satisfy the equality $\frac{\mathrm{d}}{\mathrm{d}s} I\left(f_s^\gamma\right)\big|_{s=0} = 0$, where I is the energy functional and f_s^γ is a proper variation of γ. The first variation of the energy has a well-known closed-form expression involving an integral term—null by the Euler-Lagrange equations—and boundary terms (the transversality condition). Since the initial point and the abscissa of the terminal one are fixed, we can write this transversality condition $\dot{y}\,(1) = 0$. We refer the interested reader to [29, Proposition 2.4] for more details on this step. From (6.0.9) if we define $c = c^* := f(y_1)^2$, then the equality $\dot{y}(1) = 0$ holds. Equation (6.0.11) shows that this turning point occurs at the value

$$x^*\left(c^*\right) = \int_{v_0}^{f^{-1}\left(\sqrt{c^*}\right)} \frac{f\left(y\right)^2 \mathrm{d}y}{\alpha(y)\sqrt{c^* - f\left(y\right)^2}}. \tag{6.0.12}$$

Let $z := c^* - f(y)^2$, then $\mathrm{d}z/\mathrm{d}y = -2f(y)f'(y) = -2\sqrt{c^* - z}\, f'\left(f^{-1}\left(\sqrt{c^* - z}\right)\right)$, and (6.0.12) reads

$$x^*\left(c^*\right) = \frac{1}{2}\int_0^{c^* - f(v_0)^2} \frac{z^{-1/2}\sqrt{c^* - z}}{\alpha\left(f^{-1}\left(\sqrt{c^* - z}\right)\right)} \frac{\mathrm{d}z}{f'(f^{-1}(\sqrt{c^* - z}))}$$

By the growth condition on the function $\alpha \circ f^{-1}$ and the fact that the function f' is strictly decreasing on the interval (y_c, ∞), we see that the integrand is non-decreasing in c^*, and

$$\mathrm{d}x^*/\mathrm{d}c^* > 0. \tag{6.0.13}$$

Moreover, from (6.0.12) the function x^* maps the interval $\left(f(v_0)^2, f_\infty^2\right)$ into $\mathbb{R}_+^*$. Thus the function x^* is differentiable and strictly increasing, with $x^*\left(f(v_0)^2\right) = 0$ and $x^*(c^*)$ diverges to infinity as c^* tends to f_∞^2, so we can not have two distinct geodesics corresponding to two different values of c^* which both start at $(0, v_0)$ and are both perpendicular to the y-axis when they cross the vertical line l_{x_1}. Note that (6.0.13) implies the inequality

$$\frac{\mathrm{d}y\,(1)}{\mathrm{d}x^*} = \left(\frac{\mathrm{d}x^*}{\mathrm{d}c^*}\frac{\mathrm{d}c^*}{\mathrm{d}y(1)}\right)^{-1} = \left(2\frac{\mathrm{d}x^*}{\mathrm{d}c^*} f\left(y\,(1)\right) f'\left(y\,(1)\right)\right)^{-1} > 0,$$

and hence for any $x_1 > 0$ there is a unique geodesic from $(0, v_0)$ to the line l_{x_1} which is perpendicular to the y-axis at l_{x_1}, and this is the shortest geodesic to this vertical line. Let $y_{x_1}(1)$ denote the y-value where this geodesic intercepts l_{x_1}. As mentioned above we hence obtain the following two equations

$$d\left(x_1\right)=f\left(y_{x_1}(1)\right)\int_{v_0}^{y_{x_1}(1)}\frac{\mathrm{d}y}{\alpha(y)\sqrt{f\left(y_{x_1}(1)\right)^2-f\left(y\right)^2}}, \tag{6.0.14}$$

$$x_1=\int_{v_0}^{y_{x_1}(1)}\frac{f\left(y\right)^2\mathrm{d}y}{\alpha\left(y\right)\sqrt{f\left(y_{x_1}(1)\right)^2-f\left(y\right)^2}} \tag{6.0.15}$$

which help us calculate the length of the shortest geodesic. We first solve for the unique $y_{x_1}(1)$ associated with x_1 in (6.0.15), and then substitute this into (6.0.14) to compute the distance. Clearly the function d is increasing. Since the distance function d is continuous (see [29, Corollary 2.7, Chapter 7]), the theorem follows. □

In [12], Berestycki et al. showed that for a general stochastic volatility model subject to certain growth conditions, the function d is a classical solution to the eikonal equation

$$f(y)^2d_x^2+\alpha(y)^2d_y^2=1, \quad \text{for all } x>0. \tag{6.0.16}$$

In our framework we could in principle give an explicit representation for d_x and d_y using a similar analysis as in the proof of Theorem 6.0.2, i.e. writing and expression for the first variation of the energy in terms of the distance function d and then use the fact that the curve γ is a geodesic to establish some transversality conditions. We do not carry out the whole analysis here and refer the reader to [12] for a complete proof of this property.

We now move on to a generalisation of the previous model, and we consider the following local-stochastic volatility model for a logarithmic price process $(X_t)_{t\geq 0}$:

$$\begin{aligned}\mathrm{d}X_t &= -\tfrac{1}{2}\sigma\left(X_t\right)^2f\left(V_t\right)^2\mathrm{d}t+\sigma(X_t)f\left(V_t\right)\mathrm{d}W_t,\\ \mathrm{d}V_t &= \mu\left(V_t\right)\mathrm{d}t+\alpha\left(V_t\right)\mathrm{d}Z_t,\end{aligned} \tag{6.0.17}$$

with $X_0=0\in\mathbb{R}$, $V_0=v_0>y_c>0$, the functions f, μ and α satisfy Assumption 6.0.1 and W and Z are two independent Brownian motions. We further assume that the function $\sigma:\mathbb{R}\to\mathbb{R}_+$ is uniformly Lipschitz continuous such that for any $x\in\mathbb{R}$, the inequalities $0<\sigma_{\min}<\sigma(x)<\sigma_{\max}<\infty$ hold. Using the triangle inequality and the Lipschitz property of the functions σ, f and α, we can prove that the drift and the diffusion satisfy the global Lipschitz and linear growth conditions as required in [67, Theorem 2.9, Chapter 5], so that a unique strong solution to the SDEs (6.0.17) exists. The model in (6.0.1) was an uncorrelated stochastic volatility and it is well known (see [84] for instance) that such a model only produces symmetric smiles, which is not consistent with equity market data. The uncorrelated local stochastic volatility model in (6.0.17) can internalise non-symmetric implied volatility smiles as proved in [18]. Let us now define the function $d^{\mathrm{loc}}:(0,\infty)\to\mathbb{R}_+$ as

$$d^{\mathrm{loc}}(x):=d\left(\left|\int_0^x\sigma(u)^{-1}\mathrm{d}u\right|\right), \quad \text{for all } x\neq 0, \tag{6.0.18}$$

where the function d is given in (6.0.3). The following theorem characterises the small-time behaviour of the implied volatility for the model (6.0.17).

Theorem 6.0.6. *The implied volatility has the following behaviour as the maturity t tends to zero:*

$$\sigma_0(x) := \lim_{t\to 0} \sigma_t(x) = |x|/d^{\mathrm{loc}}(x), \quad \text{for all } x \neq 0, \tag{6.0.19}$$

where the function d^{loc} is defined in (6.0.18).

Remark 6.0.7. As in (6.0.16) we can show that the function d^{loc} defined in (6.0.18) satisfies the eikonal equation

$$\sigma(x)^2 f(y)^2 \left(d_x^{\mathrm{loc}}\right)^2 + \alpha(y)^2 \left(d_y^{\mathrm{loc}}\right)^2 = 1.$$

Proof. The Riemannian metric induced by the diffusion coefficient of (6.0.17)is given by $\mathrm{g}^{\mathrm{loc}}$, with $\mathrm{g}_{11}^{\mathrm{loc}} = \sigma(x)^{-2} f(y)^{-2}$, $\mathrm{g}_{12}^{\mathrm{loc}} = 0$ and $\mathrm{g}_{22}^{\mathrm{loc}} = \alpha(y)^{-2}$. Let $d^{\mathrm{loc}}(x)$ denote the distance from the point $(0, v_0)$ to the line l_x under this metric. Consider the diffeomorphism $\Psi : \mathbb{R}^2 \to \mathbb{R}^2$, such that $\Psi(x,y) = \left(q^{-1}(x), y\right)$, where $q(x) := \left|\int_0^x \sigma(u)^{-1}\mathrm{d}u\right|$ for any $x > 0$. Then Ψ induces a Riemannian structure g on $\mathbb{R}^2$, with

$$\mathrm{g}(u,v)_p = \mathrm{g}^{\mathrm{loc}}\left(\mathrm{d}\Psi(u), \mathrm{d}\Psi(v)\right)_{\Psi(p)} = f(y)^{-2} u_1 v_1 + \alpha(y)^{-2} u_2 v_2,$$

where $u = (u_1, u_2)$ and $v = (v_1, v_2)$, at a point $p \in \mathbb{R}^2$, so we see that g is the same metric as in (6.0.5), and the mapping Ψ is an isometry (see [29, Example 2.5, page 39]). Thus there is a unique geodesic joining the point $(0, v_0)$ and the line l_x under the metric $\mathrm{g}^{\mathrm{loc}}$, whose length $d^{\mathrm{loc}}(x)$ is equal to $d\left(q(x)\right)$, where d is the distance function in (6.0.2), and the function d^{loc} is continuous. Moreover the Freidlin-Wentzell theorem and the contraction principle imply that the limit

$$-\lim_{t\to 0} t \log \mathbb{P}\left(X_t > x\right) = d^{\mathrm{loc}}\left(x\right)^2/2 = d\left(q(x)\right)^2/2$$

holds for all $x \neq 0$, and the result follows using a similar argument to the one for Corollary 6.0.5. □

Remark 6.0.8. We can write (6.0.18) and (6.0.19) as $q(x) = d^{-1}\left(x/\sigma_0(x)\right)$ and $\sigma(x) = q'(x)^{-1}$ to back out a function σ that makes the model consistent with the observed behaviour of the implied volatility smile in the small-maturity limit as long as the function σ that we extract satisfies all the assumptions of Theorem 6.0.6. This has also been noted by Henry-Labordère [57].

The above formulae are interesting from a theoretical point of view but might not be very easy to use in practice. Tedious but straightforward Taylor series expansions can be performed in order to obtain closed-form approximations for the at-the-money implied volatility in the model defined above. This is the purpose of the following corollary which we state without proof.

Corollary 6.0.9. *For the local-stochastic volatility model defined by* (6.0.17) *and under the same assumptions as in Theorem 6.0.6, the asymptotic implied volatility σ_0 defined in* (6.0.19) *has the* *following expansion around the at-the-money*

$$\begin{aligned}\sigma_0(x) &= \sigma(0)f(v_0) + \tfrac{1}{2}\sigma'(0)f\left(v_0\right)x \\ &+\frac{1}{6}\left(\sigma''(0) + \frac{f'\left(v_0\right)^2\alpha\left(v_0\right)^2}{f\left(v_0\right)^4\sigma(0)} - \frac{1}{2}\frac{\sigma'(0)^2}{\sigma(0)}\right)f\left(v_0\right)x^2 + \mathcal{O}\left(x^3\right).\end{aligned} \tag{6.0.20}$$

We can match the observed level, slope and convexity of the implied volatility in the small-maturity limit by an appropriate choice of $\sigma(0)$, $\sigma'(0)$ and $\sigma''(0)$, for any functions f and α satisfying all the conditions required for the theorem to hold. When the function f is identically equal to one we recover the pure local volatility case and the formula $\sigma_0'(0) = \sigma'(0)/2$, which is proved in [33]. For the pure uncorrelated stochastic volatility case $\sigma(x) = 1$ for all x, Equation (6.0.20) reduces to

$$\sigma_0(x) = f(v_0) + \frac{1}{6}\frac{f'\left(v_0\right)^2\alpha\left(v_0\right)^2}{f\left(v_0\right)^3}x^2 + \mathcal{O}\left(x^4\right),$$

and the function σ_0 is symmetric in the log-moneyness x.

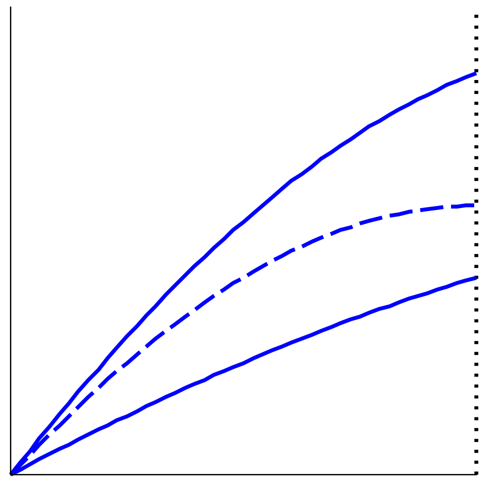


Figure 6.1: This figure is a graphical help to the proof of Theorem 6.0.2. The dotted vertical line represents l_{x_1} and the optimal curve (dashed) in the (x, y) plane comes in perpendicular to l_{x_1}.

Chapter 7

The large deviations approach

In this chapter we first derive a small-time large deviations principle in order to study the limit of the implied volatility as the maturity tends to zero. We first state the a model-free result, which expresses call options and the implied volatility asymptotics under a large deviations assumption. We then prove that the Heston process $(X_t)_{t\geq 0}$ satisfying the system of SDEs (3.2.1) as well as the Black-Scholes model both satisfy such a LDP which enables us to compute small-time asymptotics of call prices, and deduce the asymptotic behaviour of the implied volatility.

7.1 Small-time LDP for the implied volatility

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geq 0}$. We consider a share price process $(S_t)_{t\geq 0}$ assumed to be a true martingale with respect to the filtration $\mathcal{F}$, and we denote $X_t := \log(S_t)$ its logarithmic for all $t \geq 0$. Define its limiting logarithmic Laplace transform Λ as

$$\Lambda(p) := \lim_{t\to 0} t \log \mathbb{E}\left(\mathrm{e}^{p(X_t - x_0)/t}\right),$$

for all $p \in \mathbb{R}$ such that the limit exists as an extended real number. We also define the corresponding Fenchel-Legendre transform by

$$\Lambda^*(x) := \sup_{p\in\mathcal{D}_\Lambda} \{px - \Lambda(p)\}, \quad \text{for all } x \in \mathbb{R}, \tag{7.1.1}$$

where again $\mathcal{D}_\Lambda := \{p \in \mathbb{R} : \Lambda(p) < \infty\}$ denotes the effective domain of the function Λ. In this section we assume that the random process $(X_t - X_0)_{t\geq 0}$ satisfies a large deviations principle with good rate function Λ^*. Furthermore the function Λ^* has a unique minimum attained at zero. We here use the same notations Λ and Λ^* as in the large-time analysis of Part II. Since we will not use these results to study the small-time asymptotics, we hope this will not lead to any confusion. Let us first state and prove the following lemma which will be needed in the course of this section.

Lemma 7.1.1. *Assume that* $X_0 = 0$ *and let* p_1 *and* p_2 *be two real numbers such that* $|p_1| < |p_2|$ *and* $\mathrm{sgn}(p_1) = \mathrm{sgn}(p_2)$. *Then there exists* t^* *such that the inequality* $\mathbb{E}\left(\exp\left(\frac{p_1 X_t}{t}\right)\right) \leq \mathbb{E}\left(\exp\left(\frac{p_2 X_t}{t}\right)\right)$ *holds for all* $0 \leq t < t^*$.

Proof. We first consider $0 < p_1 < p_2 < \infty$. For t small enough the monotonicity of the $\mathcal{L}^p$ norm implies the inequality $\mathbb{E}\left(\mathrm{e}^{p_1 X_t/t}\right) < \mathbb{E}\left(\mathrm{e}^{p_2 X_t/t}\right)^{p_1/p_2}$. For t sufficiently small the inequality $p_2/t > 1$ holds, so that Jensen's inequality implies that $\mathbb{E}\left(\mathrm{e}^{p_2 X_t/t}\right) \geq \mathbb{E}\left(\mathrm{e}^{X_t}\right)^{p_2/t} = 1$ holds, and hence $\mathbb{E}\left(\mathrm{e}^{p_2 X_t/t}\right)^{p_1/p_2} \leq \mathbb{E}\left(\mathrm{e}^{p_2 X_t/t}\right)$. A similar argument holds in the case $p_2 < p_1 < 0$ using the fact the process $\left(\mathrm{e}^{-X_t}\right)_{t\geq 0}$ is a submartingale, and the lemma follows. □

The following corollary is an immediate consequence of the fact that the process $(X_t - X_0)_{t\geq 0}$ satisfies a large deviations principle with good rate function Λ^*.

Corollary 7.1.2. *The following limits hold as t tend to zero.*

$$-\lim_{t\to 0} t\log\mathbb{P}(X_t - x_0 > x) = -\lim_{t\to 0} t\log\mathbb{P}(X_t - x_0 \geq x) = \inf_{y\geq x}\Lambda^*(y) = \Lambda^*(x), \qquad \text{for all } x \geq 0,$$
$$-\lim_{t\to 0} t\log\mathbb{P}(X_t - x_0 < x) = -\lim_{t\to 0} t\log\mathbb{P}(X_t - x_0 \leq x) = \inf_{y\leq x}\Lambda^*(y) = \Lambda^*(x), \qquad \text{for all } x \leq 0.$$

We can now translate these tail probabilities into out-of-the-money call and put options for small maturities. As before we consider European call and put options maturing at time t, and $x := \log(K/S_0)$ denote the log-moneyness.

Corollary 7.1.3. *The following small-time behaviours hold for out-of-the-money options:*

$$-\lim_{t\to 0} t\log\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+ = \Lambda^*(x), \quad \text{for all } x \geq 0,$$
$$-\lim_{t\to 0} t\log\mathbb{E}\left(S_0\mathrm{e}^x - \mathrm{e}^{X_t}\right)_+ = \Lambda^*(x), \quad \text{for all } x \leq 0,$$

and the following limits hold for all real number x:

$$\lim_{t\to 0} t\log\left(\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+ - \left(S_0 - S_0\mathrm{e}^x\right)_+\right) = \lim_{t\to 0} t\log\left(\mathbb{E}\left(S_0\mathrm{e}^x - \mathrm{e}^{X_t}\right)_+ - \left(S_0\mathrm{e}^x - S_0\right)_+\right) = \Lambda^*(x).$$

Proof. We only prove the small-time behaviour of out-of-the-money call options as the small-time behaviour for put options follow the exact same lines. Furthermore it is clear that the last part of the corollary follows immediately from the put-call parity and the first two statements, so that we do not add any detail on this. For clarity we can assume $S_0 = 1$ since this does not entail any loss of generality. We first deal with the lower bound. We see that for any $\delta > 0$, we have $\mathbb{E}\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+ \geq \delta\,\mathbb{P}\left(\mathrm{e}^{X_t} > \mathrm{e}^x + \delta\right)$, and Corollary 7.1.2 implies

$$\liminf_{t\to 0} t\log\mathbb{E}\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+ \geq \liminf_{t\to 0}\left(t\log(\delta) + t\log\mathbb{P}\left(\mathrm{e}^{X_t} > \mathrm{e}^x + \delta\right)\right) \geq -\Lambda^*\left(\log\left(\mathrm{e}^x + \delta\right)\right).$$

Let now δ tend to zero. The continuity of the function Λ^* implies the desired lower bound. For the upper bound, Hölder's inequality implies that for any $p, q > 1$ satisfying $p^{-1} + q^{-1} = 1$,

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+ &= \mathbb{E}\left(\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+ \mathbb{1}_{\{X_t\geq x\}}\right) \leq \left(\mathbb{E}\left(\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+\right)^p\right)^{1/p}\mathbb{E}\left(\mathbb{1}^q_{\{X_t\geq x\}}\right)^{1/q}\\ &= \left(\mathbb{E}\left(\left(\mathrm{e}^{X_t} - \mathrm{e}^x\right)_+\right)^p\right)^{1/p}\mathbb{P}\left(\mathrm{e}^{X_t} \geq \mathrm{e}^x\right)^{1/q} \leq \mathbb{E}\left(\mathrm{e}^{pX_t}\right)^{1/p}\mathbb{P}\left(\mathrm{e}^{X_t} \geq \mathrm{e}^x\right)^{1-1/p}.\end{aligned}$$

Taking logarithms and multiplying by t we obtain

$$t\log\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^x\right)_+ \leq p^{-1}t\log\mathbb{E}\left(\mathrm{e}^{pX_t}\right)+t\left(1-1/p\right)\log\mathbb{P}\left(X_t\geq x\right). \tag{7.1.2}$$

We now claim that the function $t\mapsto t\log\mathbb{E}\left(\mathrm{e}^{pX_t}\right)$ converges to zero as t tends to zero. Indeed, consider $\delta>0$. For t sufficiently small, we have $\delta/t>p$ and hence

$$\limsup_{t\to0} t\log\mathbb{E}\left(\mathrm{e}^{pX_t}\right)\leq\limsup_{t\to0}t\log\mathbb{E}\left(\mathrm{e}^{\delta X_t/t}\right)=\Lambda(\delta),$$

by Lemma 7.1.1, and

$$\liminf_{t\to0} t\log\mathbb{E}\left(\mathrm{e}^{pX_t}\right)\geq\liminf_{t\to0}t\log\left(\mathbb{E}\left(\mathrm{e}^{X_t}\right)\right)^p=0=\Lambda(0),$$

by Jensen's inequality. Since the function Λ is continuous the claim follows. If we then take the limit as p tends to infinity on both sides of (7.1.2), then Corollary 7.1.2 implies the upper bound $\limsup_{t\to0}t\log\mathbb{E}\left(\mathrm{e}^{X_t}-\mathrm{e}^x\right)_+\leq-\Lambda^*(x)$. □

Using these results for the short-time behaviour of call options, we can now translate them into implied volatility asymptotics.

Theorem 7.1.4. *As the maturity t tends to zero, we have*

$$\sigma_0(x):=\lim_{t\to0}\sigma_t(x)=\frac{|x|}{\sqrt{2\Lambda^*(x)}},\quad \textit{for all } x\neq0.$$

Proof. We first assume that $x>0$. We first establish the lower bound for $\sigma_t(x)$. Using the classical notation for the Black-Scholes formula, we set $d_\pm(x):=\left(-x\pm\frac{1}{2}\sigma_t^2(x)t\right)/\left(\sigma_t(x)\sqrt{t}\right)$. By Corollary 7.1.3 and the definition of the implied volatility, we know that for all $\delta,C>0$, there exists $t_l^*(\delta)>0$ such that for all $t<t_l^*(\delta)$ we have

$$\begin{aligned}\exp\Big(-\left(\Lambda^*(x)+\delta\right)/t\Big)&\leq\mathbb{E}\left(\mathrm{e}^{X_t}-S_0\mathrm{e}^x\right)_+=S_0\mathcal{N}(d_+(x))-S_0\mathrm{e}^x\mathcal{N}(d_-(x))\\&\leq S_0\left(1-\mathcal{N}(-d_+(x))\right)\leq|d_+|^{-1}S_0\,n\left(d_+(x)\right)\leq CS_0\,n\left(d_+(x)\right).\end{aligned}$$

The last two lines follow from Lemma 2.0.10 and the fact that the total implied variance $\sigma_t^2(x)t$ converges to zero as t tends to zero. We now take the logarithm on both sides, multiply by t, and hence there exists some $0<t_l^{**}(\delta)\leq t_l^*(\delta)$ such that

$$-\left(\Lambda^*(x)+\delta\right)\leq t\log(S_0C)-\frac{t}{2}\log(2\pi)-\frac{1}{2\sigma_t^2(x)}\left(x-\frac{1}{2}\sigma_t^2(x)t\right)^2\leq-\frac{x^2}{2\sigma_t^2(x)}+\delta$$

for all $t<t_l^{**}(\delta)$, and the lower bound follows. For the upper bound consider $\delta>0$, then Corollary 7.1.3 implies that there exists $t_u^*(\delta)>0$ such that for all $t<t_u^*(\delta)$ we have

$$\begin{aligned}\exp\Big(-t^{-1}\left(\Lambda^*(x)-\delta\right)\Big)&\geq\mathbb{E}\left(\mathrm{e}^{X_t}-S_0\mathrm{e}^x\right)_+=S_0\mathcal{N}\left(d_+(x)\right)-S_0\mathrm{e}^x\mathcal{N}\left(d_-(x)\right)\\&=S_0\left(1-\mathcal{N}\left(-d_+(x)\right)\right)-S_0\mathrm{e}^x\left(1-\mathcal{N}(-d_-(x))\right).\end{aligned}$$

Set $d_\delta(x) := -\left(x+\delta+\frac{1}{2}\sigma_t^2(x)t\right)/(\sigma_t(x)\sqrt{t})$. Similar to Corollary 7.1.3, we can now use the fact that one call option with strike $S_0\mathrm{e}^x$ is worth more than $S_0\mathrm{e}^x\left(\mathrm{e}^\delta-1\right)$ digital call options of strike $S_0\mathrm{e}^{x+\delta}$, and invoke the lower bound on the Normal distribution in Lemma 2.0.10 and the fact that $\sigma_t^2(x)t$ converges to zero as t tends to zero to see that

$$\begin{aligned}\exp\Big(-t^{-1}\left(\Lambda^*(x)-\delta\right)\Big) &\geq S_0\mathrm{e}^x(\mathrm{e}^\delta-1)\Big(1-\mathcal{N}(-d_\delta(x))\Big)\\ &\geq S_0\mathrm{e}^x\left(\mathrm{e}^\delta-1\right)n\left(d_\delta(x)\right)\left(|d_\delta(x)|^{-1}+|d_\delta(x)|\right)^{-1}\\ &\geq S_0\mathrm{e}^x\left(\mathrm{e}^\delta-1\right)n\left(d_\delta(x)\right)\frac{\sigma_t(x)\sqrt{t}}{x+\delta}(1-\delta)\\ &\geq S_0\mathrm{e}^x\left(\mathrm{e}^\delta-1\right)n\left(d_\delta(x)\right)(1-\delta)\exp\left(-\delta\frac{x+\delta}{\sigma_t(x)\sqrt{t}}\right).\end{aligned}$$

We can further find $t_u^{**}(\delta)\leq t_u^*(\delta)$ such that for all $t<t_u^{**}(\delta)$ we have

$$\begin{aligned}-(\Lambda^*(x)-\delta) &\geq t\log\left(S_0\mathrm{e}^x\right)+t\log\left(\mathrm{e}^\delta-1\right)-\frac{t\log(2\pi)}{2}-\frac{(x+\delta+\frac{1}{2}\sigma_t^2(x)t)^2}{2\sigma_t^2(x)}-\sqrt{t}\delta\frac{(x+\delta)}{\sigma_t(x)}+t\log(1-\delta)\\ &\geq -\frac{(x+\delta)^2}{2\sigma_t^2(x)}-\delta.\end{aligned}$$

Combining these two bounds, for any $\delta>0$, we can define $t^{**}(\delta):=t_l^{**}(\delta)\wedge t_u^{**}(\delta)$ such that for all $t<t^{**}(\delta)$, the theorem follows for $x>0$. We proceed similarly for $x<0$, and hence for all non zero x and all $\delta>0$, there exists $\tilde{t}\,(\delta)$ such that for all $0\leq t\leq\tilde{t}\,(\delta)$ the inequality $\left|\sigma_t(x)-|x|/\sqrt{2\Lambda^*(x)}\right|<\delta$ holds and the theorem follows. □

7.2 Small-time LDP for the extended Heston model

We now return to the generalised Heston model where the process $(X_t)_{t\geq 0}$ follows the SDE (3.2.1). Let us define the two real numbers p_- and p_+ by

$$\begin{aligned}p_- &:= \frac{2}{\sigma\bar{\rho}}\arctan\left(\frac{\bar{\rho}}{\rho}\right)\mathbb{1}_{\{\rho<0\}}-\frac{\pi}{\sigma}\mathbb{1}_{\{\rho=0\}}+\frac{2}{\sigma\bar{\rho}}\left(\arctan\left(\frac{\bar{\rho}}{\rho}\right)-\pi\right)\mathbb{1}_{\{\rho>0\}},\\ p_+ &:= \frac{2}{\sigma\bar{\rho}}\left(\arctan\left(\frac{\bar{\rho}}{\rho}\right)+\pi\right)\mathbb{1}_{\{\rho<0\}}+\frac{\pi}{\sigma}\mathbb{1}_{\{\rho=0\}}+\frac{2}{\sigma\bar{\rho}}\arctan\left(\frac{\bar{\rho}}{\rho}\right)\mathbb{1}_{\{\rho>0\}}.\end{aligned}\tag{7.2.1}$$

The following proposition provides an explicit representation for the limiting logarithmic Laplace transform Λ defined in (7.1.1) as well as the properties of its Fenchel-Legendre transform Λ^*.

Proposition 7.2.1. *The function Λ is well defined on (p_-,p_+) and*

$$\Lambda(p)=\frac{vp}{\sigma\left(\bar{\rho}\cot\left(\sigma p\bar{\rho}/2\right)-\rho\right)}+\frac{a}{2}p^2,\quad \text{for all } p\in(p_-,p_+),\tag{7.2.2}$$

and is infinite outside. Furthermore, the Fenchel-Legendre transform Λ^* is a good rate function, i.e. is strictly convex on $\mathbb{R}$ and has a unique minimum attained at zero.

Remark 7.2.2. Note that the function Λ does not depend on the drift terms κ or θ.

Theorem 7.2.3. *The process $(X_t - x_0)_{t\geq 0}$ satisfies a large deviations principle as t tends to zero with the good rate function Λ^*.*

We prove Proposition 7.2.1 and Theorem 7.2.3 altogether.

Proof. From (3.2.2) we know the logarithmic Laplace transform of the process $(X_t - x_0)_{t\geq 0}$:

$$\log \mathbb{E}\left(\mathrm{e}^{p(X_t - X_0)}\right) = C\left(t, -\mathtt{i}p\right) + D\left(t, -\mathtt{i}p\right) v + \frac{a}{2} p\left(p-1\right) t.$$

Let us now set $p \mapsto p/t$. For t sufficiently small it is easy to see that the inequality $(\kappa - \rho\sigma p/t)^2 + \sigma^2 p\left(1 - p/t\right)/t < 0$ holds. Therefore the explosion time defined in (3.2.7) reads

$$t^*(p) = 2\left(\sigma^2 p\left(p-1\right) - \left(\kappa - \rho\sigma p\right)^2\right)^{-1/2}\left(\pi \mathbb{1}_{\{\rho\sigma p - \kappa < 0\}} + \arctan\left(\frac{\sqrt{\sigma^2 p\left(p-1\right) - \left(\kappa - \rho\sigma p\right)^2}}{\rho\sigma p - \kappa}\right)\right)$$

and we have, as t tends to zero

$$t^*\left(p/t\right) \sim \begin{cases} \dfrac{2t}{\sigma\bar{\rho}|p|}\left(\pi \mathbb{1}_{\{\rho p \leq 0\}} + \operatorname{sgn}(p) \arctan\left(\dfrac{\bar{\rho}}{\rho}\right)\right), & \text{for } \rho \neq 0 \text{ and } p \neq 0, \\ \pi t/\left(\sigma|p|\right), & \text{for } \rho = 0 \text{ and } p \neq 0, \end{cases}$$

and $t^*\left(0\right)$ is infinite. (7.2.1) implies that for any $p \in (p_-, p_+)$ the inequality $t^*\left(p/t\right) > t$ holds for t sufficiently small so that $\Lambda_t(p)$ as characterised in (3.2.6) is well defined. Let now t tend to zero and define $\zeta := \sigma\bar{\rho}p$, we then have

$$\begin{aligned}
\lim_{t\to 0} d\left(-\mathtt{i}p/t\right) t &= \mathtt{i}\sigma\bar{\rho}|p|, \qquad \lim_{t\to 0} g\left(-\mathtt{i}p/t\right) = \frac{\rho + \mathtt{i}\bar{\rho}\operatorname{sgn}(p)}{\rho - \mathtt{i}\bar{\rho}\operatorname{sgn}(p)}, \\
\lim_{t\to 0} D\left(t, -\frac{\mathtt{i}p}{t}\right) t &= -\frac{vp}{\sigma^2}\left(\rho\sigma + \mathtt{i}\sigma\bar{\rho}\right)\left(1 - \mathrm{e}^{-\mathtt{i}\zeta}\right)\left(1 - \frac{\rho + \mathtt{i}\bar{\rho}}{\rho - \mathtt{i}\bar{\rho}}\mathrm{e}^{-\mathtt{i}\zeta}\right)^{-1} = \Lambda(p), \\
\lim_{t\to 0} C\left(t, -\frac{\mathtt{i}p}{t}\right) &= \frac{\kappa\theta}{\sigma^2}\left[-\left(\rho + \mathtt{i}\bar{\rho}\right)\sigma p - 2\log\left(\left(1 - \frac{\rho + \mathtt{i}\bar{\rho}}{\rho - \mathtt{i}\bar{\rho}}\mathrm{e}^{-\mathtt{i}\zeta}\right)\left(1 - \frac{\rho + \mathtt{i}\bar{\rho}}{\rho - \mathtt{i}\bar{\rho}}\right)^{-1}\right)\right].
\end{aligned}$$

For the two limits $D\left(t, -\mathtt{i}p/t\right) t$ and $C\left(t, -\mathtt{i}p/t\right) t$ we have assumed $p > 0$. The final limits hold when $p \leq 0$ but some signs are flipped in the intermediate calculations. Therefore the limit $\lim_{t\to 0} t \log \mathbb{E}\left(\exp\left(t^{-1} p\left(X_t - x_0\right)\right)\right) = \Lambda(p)$ holds for $p \in (p_-, p_+)$ defined in (7.2.1), and Λ in (7.2.2). The function Λ clearly tends to infinity as p converges up to p_+ and down to p_-. The function Λ is smooth in the interval (p_-, p_+) and

$$\begin{aligned}
\Lambda'(p) &= \frac{v}{\sigma\left(\bar{\rho}\cot\left(\zeta/2\right) - \rho\right)} + \frac{\sigma v p \bar{\rho}^2 \csc^2\left(\zeta/2\right)}{2\sigma\left(\bar{\rho}\cot\left(\zeta/2\right) - \rho\right)^2} + ap, \\
\Lambda''(p) &= \frac{v\bar{\rho}^2\csc^2\left(\zeta/2\right)\left(1 - \zeta\cot\left(\zeta/2\right)/2\right)}{\left(\bar{\rho}\cot\left(\zeta/2\right) - \rho\right)^2} + \frac{v\zeta\bar{\rho}^3\csc^4\left(\zeta/2\right)}{2\left(\bar{\rho}\cot\left(\zeta/2\right) - \rho\right)^3} + a.
\end{aligned}$$

for all $p \in (p_-, p_+)$. From Proposition 7.2.1 and (7.2.1) $\Lambda''(p)$ is strictly positive for all $p \in (p_-, p_+)$, and that $|\Lambda'(p)|$ diverges up to infinity as p tends to p_- from above or to p_+ from below. Henceforth the function Λ is convex, essentially smooth and lower semi-continuous, and hence the Gärtner-Ellis theorem implies that the random variable $(X_t - x_0)_{t>0}$ satisfies the large deviations principle with rate function equal to the Fenchel-Legendre transform of $\Lambda(p)$. By the essential smoothness property of Λ and the positivity of Λ'', the equation $\partial_p(px - \Lambda(p)) = 0$ has a unique solution $p^*(x)$ in (p_-, p_+) for all $x \in \mathbb{R}$, which equivalently solves $x = \Lambda'(p^*(x))$. The unique minimum of Λ^* occurs at $x^* = (\Lambda^{*'})^{-1}(0) = \Lambda'(0) = 0$ and $\Lambda^*(0) = 0$, and the theorem follows. $\square$

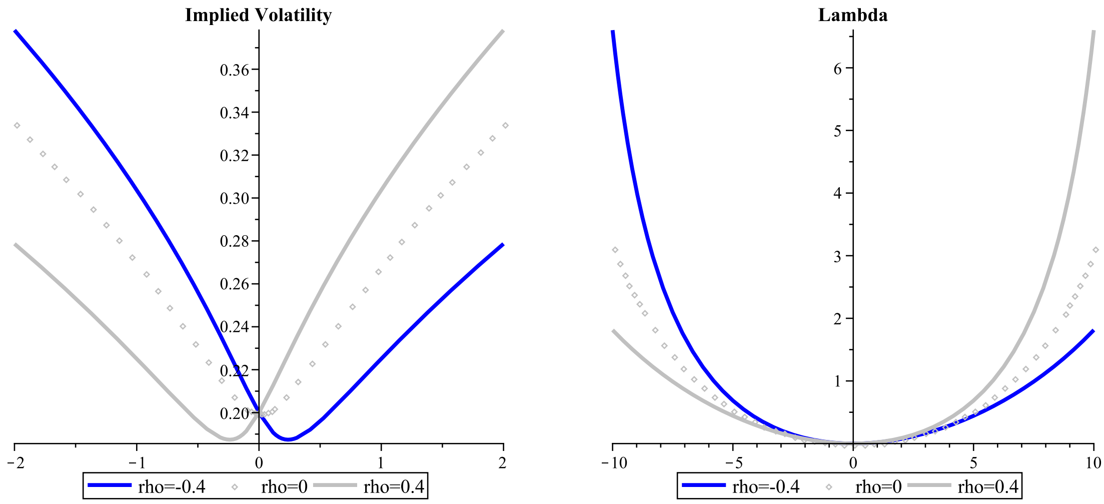


Figure 7.1: We plot here the functions Λ and σ_0 for the Heston model with the parameters $v = \theta = 0.04$, $\sigma = 0.2$ and $\rho = -0.4, 0, 0.4$.

Remark 7.2.4. In Section 4.3.1 we discussed the extension of the standard Heston model ($a = 0$) to the Schöbel-Zhu model. From the representation (4.3.1) it is easy to prove that only the function B_v plays a role in the limit as t tends to zero, and the function λ is well defined on the set $\mathcal{D}_\lambda$ with

$$\lambda(u) = \frac{u}{2\xi}\left(\bar{\rho}\cot(\xi\bar{\rho}u) - \rho\right)^{-1}, \quad \text{for all } u \in \mathcal{D}_\lambda,$$

where $\bar{\rho} := \sqrt{1-\rho^2}$. Note that this function is the same as in the Heston model, which implies that the two smiles are exactly the same in the short-maturity limit.

Note that Theorem 7.1.4 does not deal with the at-the-money case $x = 0$, because the large deviations bounds do not provide us with any information in this case. Here, we can either first set x equal to zero, and then let t tend to zero, or vice versa. We here prove the other way round, i.e. consider $\sigma_0(x)$ as determined above for x different from zero and let x tend to zero.

Corollary 7.2.5. *Around the at-the-money $x = 0$ the asymptotic implied volatility satisfies*

$$\sigma_0(x) = \sqrt{a+v} + \frac{1}{4}\frac{\sigma\rho v x}{(a+v)^{3/2}} + \frac{\sigma^2 v}{48}\frac{\left(2-5\rho^2\right)v + 2a\left(1+2\rho^2\right)}{(a+v)^{7/2}}x^2 + \mathcal{O}\left(x^3\right).$$

In the standard Heston case when $a = 0$, this is equivalent to (see also [73] and [33])

$$\sigma_0^2(x) = v + \frac{1}{2}\rho\sigma x + \left(4 - 7\rho^2\right)\frac{\sigma^2 x^2}{48 v^2} + \mathcal{O}\left(x^3\right).$$

Proof. A Taylor expansion of the function Λ near $p = 0$ gives $\Lambda(p) = \frac{1}{2}\left(a+v\right)p^2 + \frac{\sigma}{4}v\rho p^3 + \frac{\sigma^2}{24}v\left(1+2\rho^2\right)p^4 + \mathcal{O}\left(p^5\right)$. Therefore since $p^*\left(x\right)$ solves the equation $\Lambda'\left(p^*\left(x\right)\right) = p^* x$ for any real number x, we have $p^{*'}(0) = 1/\Lambda''(\Lambda'^{-1}(0)) = \left(a+v\right)^{-1}$ and $p^{*''}(0) = -\Lambda'''(0)/\Lambda''(0)^3 = -3\rho\sigma v/\left(2\left(a+v\right)^2\right)$, and the corollary follows. □

Remark 7.2.6. Consider for simplicity the case $a = 0$. It might look strange at first sight that the at-the-money curvature $\partial^2\sigma_0(0)$ is strictly negative for $\rho^2 > 2/5$. This does not however mean that the formula is inconsistent with no-arbitrage theory. A simple way to see this is by considering the following operator (see [49] for instance):

$$\mathcal{L}\sigma_0^2 := 1 - \frac{x}{\sigma_0^2}\partial_x\sigma_0^2 + \frac{1}{4}\left(\frac{x^2}{\sigma_0^4} - \frac{1}{4} - \frac{1}{\sigma_0^2}\right)\left(\partial_x\sigma_0^2\right)^2 + \frac{1}{2}\partial_{xx}\sigma_0^2.$$

A condition for no-arbitrage is that $\mathcal{L}\sigma_0^2\left(x\right) \geq 0$ for all $x \in \mathbb{R}$, and some calculations show that this condition holds, and hence the at-the-money approximation is not inconsistent with the absence of arbitrage.

Chapter 8

Back to the steepest descent curve

8.1 Small-time behaviour of European call options

8.1.1 The Heston model

The main result of this section is Theorem 8.1.1, which describes the asymptotic behaviour of call option prices in the small-time limit, and we postpone its proof to Section 8.4 for ease of clarity.

Theorem 8.1.1. *For the Heston model in* (3.2.1)*, the following asymptotic behaviour*

$$\frac{\mathbb{E}\left(\mathrm{e}^{X_t}-S_0\mathrm{e}^x\right)_+}{S_0}=\begin{cases}(1-\mathrm{e}^x)_++\exp\left(-\dfrac{\Lambda^*(x)}{t}\right)\left(\dfrac{A(x)}{\sqrt{2\pi}}t^{3/2}+\mathcal{O}\left(t^{5/2}\right)\right), & \text{if } x\neq 0,\\[2ex] \sqrt{\dfrac{vt}{2\pi}}-A_0t^{3/2}+\mathcal{O}\left(t^{5/2}\right), & \text{if } x=0,\end{cases}$$

holds as the maturity t tends to zero, where

$$A(x):=\frac{\mathrm{e}^x U\left(p^*(x)\right)}{p^*(x)^2\sqrt{\Lambda''\left(p^*(x)\right)}}, \tag{8.1.1}$$

$$U(p):=\exp\Bigg(\frac{p^2v}{2}\frac{2\alpha\left(\rho\sigma-\kappa\right)\cos\alpha+2\sigma\left(p\kappa\rho+p\left(\frac{1}{2}-\rho^2\right)\sigma\right)\sin\alpha+2\kappa\alpha\left(1-\rho\sigma p\right)+\left(p\sigma-2\rho\right)\sigma\alpha}{\alpha\left(\left(\rho^2\sigma^2p^2-\alpha^2\right)\cos\alpha+2\rho\sigma p\alpha\sin\alpha-\rho^2\sigma^2p^2-\alpha^2\right)}$$

$$+\frac{\kappa\theta}{\sigma^2}\Big\{2\log\left(2|\alpha|\right)-\rho\sigma p-\log\left(\left(2\alpha^2-2\rho^2\sigma^2p^2\right)\cos\alpha+2\rho^2\sigma^2p^2+2\alpha^2-4\rho\sigma p\alpha\sin\alpha\right)\Big\}\Bigg),$$

$$A_0:=\frac{1}{48}\sqrt{\frac{2}{v\pi}}\left(\sigma^2\left(1-\frac{\rho^2}{4}\right)+v\left(v-3\rho\sigma\right)-6\kappa(\theta-v)\right),$$

where the function Λ was characterised in Proposition 7.2.1, $\bar\rho:=\sqrt{1-\rho^2}$, $\alpha:=-\sigma\bar\rho p$ and $p^*(x)\in\mathbb{R}$ is the unique solution to the equation $\Lambda'\left(p^*(x)\right)=x$.

The function Λ is real-valued on (p_-,p_+), where p_+ and p_- are defined in (7.2.1). and the function Λ^* is the Fenchel-Legendre transform of Λ. Proposition 7.2.1 implies that $\Lambda^*(x)\geq 0$ and $\Lambda''(p^*(x))>0$ hold for all real number x.

8.1.2 The Black-Scholes model

We derive here the corresponding asymptotic small-time expansion for call prices under the Black-Scholes model. This result is stated below and was also derived in [89, Proposition 5.1]. The following more general proposition proved in Section 8.5.1 will be important in the next section.

Proposition 8.1.2. *Let $\Sigma > 0$, define $\sigma_t := \sqrt{\Sigma^2 + ct}$ for $t > 0$ and assume that $t \in \left(0, \Sigma^2/|c|\right)$ if $c < 0$. Then the following behaviour holds as the maturity t tends to zero,*

$$\frac{C_{\mathrm{BS}}\left(x,t,\sigma_t\right)}{S_0} = \begin{cases} \left(1-\mathrm{e}^x\right)_+ + \exp\left(-\dfrac{x^2}{2\Sigma^2 t} + \dfrac{x}{2} + \dfrac{cx^2}{2\Sigma^4}\right)\left(\dfrac{\Sigma^3}{x^2\sqrt{2\pi}}\mathcal{O}\left(t^{3/2}\right) + \mathcal{O}\left(t^{5/2}\right)\right), & \text{if } x \neq 0, \\ \dfrac{\Sigma t^{1/2}}{\sqrt{2\pi}} + \mathcal{O}\left(t^{3/2}\right), & \text{if } x = 0. \end{cases}$$

The following corollary follows immediately and we omit its proof.

Corollary 8.1.3. *In the case $c = 0$, we have $\sigma_t = \Sigma$ and*

$$\frac{C_{\mathrm{BS}}\left(x,t,\Sigma\right)}{S_0} = \begin{cases} \left(1-\mathrm{e}^x\right)_+ + \dfrac{A_{\mathrm{BS}}(x,\Sigma)}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2\Sigma^2 t}\right)t^{3/2} + \mathcal{O}\left(t^{5/2}\right), & \text{if } x \neq 0, \\ \dfrac{\Sigma}{\sqrt{2\pi}}t^{1/2} + \mathcal{O}\left(t^{3/2}\right), & \text{if } x = 0, \end{cases}$$

where

$$A_{\mathrm{BS}}\left(x,\Sigma\right) := \frac{\Sigma^3}{x^2}\exp\left(\frac{x}{2}\right). \tag{8.1.2}$$

8.2 Small-time behaviour of implied volatility

In Theorem 7.1.4 we have obtained the limiting behaviour σ_0 of the implied volatility. In this section we derive the higher-order small-time asymptotic behaviour for implied volatility using Theorem 8.1.1. We first prove the asymptotic behaviour of the implied volatility for out-of-the-money options and then for the at-the-money ones. Let us define the function $a : \mathbb{R}^* \to \mathbb{R}$ by

$$a\left(x\right) := \frac{2\sigma_0^4\left(x\right)}{x^2}\log\left(\frac{A\left(x\right)}{A_{\mathrm{BS}}\left(x,\sigma_0\left(x\right)\right)}\right), \qquad \text{for all } x \neq 0, \tag{8.2.1}$$

where the functions σ_0, A and A_{BS} are defined in Theorem 7.1.4, in (8.1.1) and in (8.1.2).

Theorem 8.2.1. *The asymptotic expansion $\sigma_t^2\left(x\right) = \sigma_0^2\left(x\right) + a\left(x\right)t + o\left(t\right)$ holds for all $x \neq 0$,*

For sake of clarity we postpone the proof of this theorem to Section 8.5.2. The following corollary—proved in Section 8.5.3—is an immediate consequence of the previous theorem. In Corollary 7.2.5, we have derived a Taylor expansion for the limiting behaviour $\sigma_0(x)$ around the at-the-money $x = 0$. We do the same here for the correction term.

Corollary 8.2.2. *The following expansion for the function a holds when x is close to zero,*

$$a(x) = \left(-\frac{1}{12}\sigma^2\left(1-\frac{1}{4}\rho^2\right)+\frac{1}{4}v\rho\sigma+\frac{1}{2}\kappa\left(\theta-v\right)\right)+\frac{\rho\sigma}{24v}\left(\sigma^2\bar{\rho}^2-2\kappa\left(\theta+v\right)+v\rho\sigma\right)x$$
$$+\frac{\sigma^2}{7680}\left(176\sigma^2-480\kappa\theta-712\rho^2\sigma^2+521\rho^4\sigma^2+40v\rho^3\sigma+1040\kappa\theta\rho^2-80v\kappa\rho^2\right)\frac{x^2}{v^2}+\mathcal{O}\left(x^3\right).$$

This corollary enables one to fit v, ρ and σ to an observed level, slope and convexity for $\sigma_t(x)$ at the at-the-money point $x=0$. In principle we then choose exogenously choose κ (resp. θ) and then solve for θ (resp. κ) so as to be consistent with a small-time term structure slope $\partial_t\sigma_t|_{t=0}$ (see Durrleman [34] for more on this). Note also that the two functions σ_0 and a are symmetric when $\rho=0$. This reflects the fact that the smile is known to be symmetric at all maturities when $\rho=0$ (see [18]). The following corollary explains why the correction term is important.

Corollary 8.2.3. *The following small-time approximation for call options holds as t tends to zero:*

$$\mathbb{E}\left(S_t-S_0\mathrm{e}^x\right)_+ = C_{\mathrm{BS}}\left(x,t,\sqrt{\sigma_0^2(x)+a\left(x\right)t}\right)\left(1+\mathcal{O}\left(t\right)\right). \tag{8.2.2}$$

Proof. This just follows from the way we equate call prices under the Heston model and the Black-Scholes formula with a time-dependent implied volatility. □

Note that (8.2.2) is false if we remove the correction term $a(x)$, hence the leading order term alone is not sufficient to estimate call option prices for small maturities. The following expansion gives us precise information on the behaviour of the at-the-money implied volatility for small maturities. The proof (which we omit) is analogous to that of the other expansions for the smile comparing the small-time behaviour of options under the Heston model and the Black-Scholes one.

Corollary 8.2.4. *The small-time at-the-money implied volatility has the asymptotic behaviour*

$$\sigma_t^2\left(0\right) = v+\left(-\frac{1}{12}\sigma^2\left(1-\frac{\rho^2}{4}\right)+\frac{1}{4}v\rho\sigma+\frac{1}{2}\kappa\left(\theta-v\right)\right)t+o\left(t\right),\quad \textit{as } t \textit{ tends to } 0. \tag{8.2.3}$$

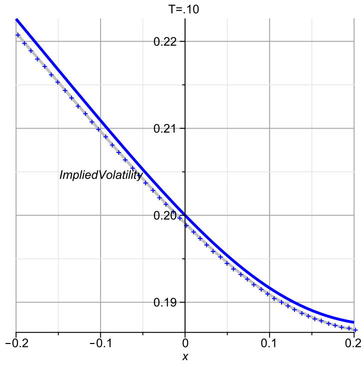
This is consistent with the small-time term structure slope of the implied volatility derived in [33, Section 3.1.2]. We can actually prove—in a very analogous way as for its large-time behaviour—that the limiting implied volatility σ_0 is continuous on the whole real line.

8.3 Numerical application

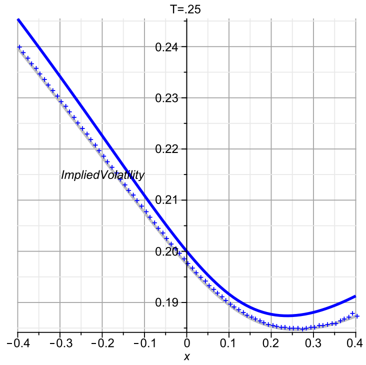
We test below the validity of the small-time expansions for implied volatility against the values obtained numerically (using an inverse Fourier transform quadrature[1]) for the parameters $\kappa=1.15$, $\sigma=0.2$, $v=\theta=0.04$, $\rho=-0.4$. We observe that our approximation and the generated data are very close for $t=0.1$ and $t=0.25$ years, but begin to differ for $t=0.5$ as we would expect since the $\mathcal{O}\left(t^2\right)$ term in the implied volatility expansion is no longer negligible. The correction

[1]Computed using the Zeliade Quant Framework by Zeliade Systems.

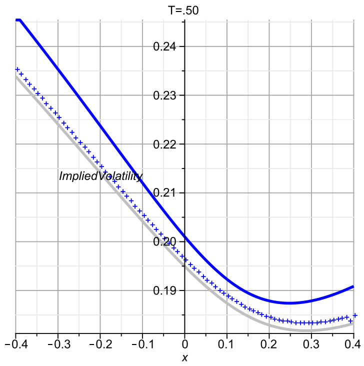
term is essentially the smile-flattening effect which is a stylised feature of implied volatility surfaces observed in the market. Our results also provide a good way of checking for the presence of jumps: if we add an exponential Lévy process driving the share price process, then the small-time implied volatility smile will explode as t tends to zero (see Chapter 8.5.3 for details). We also plot the exact correction term $a(x)$ given by the formula in (8.2.1).



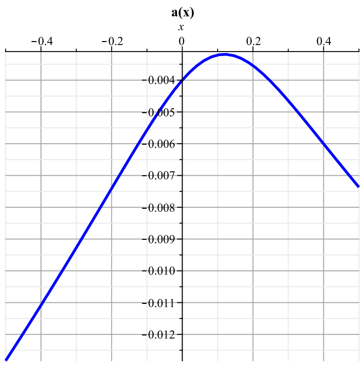
(a) t=0.1



(b) t=0.25



(c) t=0.5



(d) a(x)

Figure 8.1: Figures (a)-(c) represent the leading order smile $\sigma(x)$ (solid blue), the corrected implied volatility $\sqrt{\sigma\left(x\right)^{2}+a\left(x\right)t}$ (solid grey), and the implied volatility obtained numerically using Fourier transforms (blue crosses) for $\kappa = 1.15, \theta = 0.04, \sigma = 0.2, \rho = -0.4$ and $t = 0.1, 0.25$ and 0.5. Figure (d) represents $a(x)$ analytically using (8.2.1).

For sake of clarity in this section we introduce the notation $\alpha := \kappa\theta$. Given implied variances for five contracts, our objective is to find explicit formulae to calibrate the five Heston parameters $(v, \sigma, \rho, \alpha, \kappa)$ to the five implied variances. For some configurations of contracts, one cannot expect to solve this problem. For example, a set of contracts (all at the same expiry) would be uninformative in regard to the term structure of implied variance, hence uninformative in regard to the mean-reversion parameter κ. We therefore choose configurations which involve multiple expiries. In such a setting our maturity-dependent formula plays a crucial role in capturing the maturity effects needed to calibrate the full set of Heston parameters. Recall that Theorem 8.2.1

and Corollary 8.2.2 imply $\sigma_t^2(x) = H(x,t) + \mathcal{O}(x^3) + o(t)$, as x and t tend to zero, where

$$\begin{aligned}
H(x,t) &:= H(x,t;v,\rho,\sigma,\alpha,\kappa) := v\left(1 + \frac{1}{2}\rho\frac{\sigma x}{v} + \left(1 - \frac{7}{4}\rho^2\right)\frac{\sigma^2 x^2}{12 v^2}\right) \\
&+ \left(\left(\frac{\rho\sigma}{2}v - \frac{\sigma^2}{6}\left(1 - \frac{1}{4}\rho^2\right) + \alpha - \kappa v\right) + \frac{\rho\sigma}{12 v}\left(\sigma^2\left(1-\rho^2\right) + v\rho\sigma - 2\alpha - 2v\kappa\right)x\right)\frac{t}{2} \\
&+ \frac{\sigma^2}{7680 v^2}\Big(\left(176 - 712\rho^2 + 521\rho^4\right)\sigma^2 + 40 v\rho^3\sigma + 80\left(13\rho^2 - 6\right)\alpha - 80 v\rho^2\kappa\Big)x^2 t.
\end{aligned} \tag{8.3.1}$$

Let us consider a (squared) volatility skew $\mathsf{V} : \mathbb{K} \to \mathbb{R}$, where each point in the configuration $\mathbb{K} \subset \mathbb{R} \times [0,\infty]$ represents a (log moneyness, expiration). We shall say that $(v,\rho,\sigma,\alpha,\kappa)$ calibrates H to the skew $\mathsf{V} : \mathbb{K} \to \mathbb{R}$ if the equality $H(x,t;v,\rho,\sigma,\alpha,\kappa) = \mathsf{V}(x,t)$ holds for all $(x,t) \in \mathbb{K}$. This definition of calibration demands exact fitting of H to the given volatility skew at all points in $\mathbb{K}$. Consider now the configuration $\mathbb{K} = \{(0,0),(x_0,t_1),(-x_0,t_1),(x_0,t_2),(-x_0,t_2)\}$ where $0 < t_1 < t_2$ and $x_0 > 0$. Given $\mathsf{V} : \mathbb{K} \to \mathbb{R}$, define

$$\begin{aligned}
\mathsf{V}_0 &:= \mathsf{V}(0,0), \qquad \mathsf{S} := \frac{\mathsf{V}_+ - \mathsf{V}_-}{2x_0}, \qquad \mathsf{C} := \frac{\mathsf{V}_+ - 2\mathsf{V}_0 + \mathsf{V}_-}{2x_0^2}, \\
\mathsf{V}_\pm &:= \frac{t_2}{t_2 - t_1}\mathsf{V}(\pm x_0, t_1) - \frac{t_1}{t_2 - t_1}\mathsf{V}(\pm x_0, t_2).
\end{aligned}$$

Theorem 8.3.1. *The parameter choices $(\tilde{y}_0,\tilde{\rho},\tilde{\sigma},\tilde{\alpha},\tilde{\kappa})$ calibrate H to the skew $\mathsf{V} : \mathbb{K} \to \mathbb{R}$, where*

$$\begin{pmatrix} \tilde{y}_0 \\ \tilde{\sigma} \\ \tilde{\rho} \end{pmatrix} := \begin{pmatrix} \mathsf{V}_0 \\ \sqrt{7\mathsf{S}^2 + 12\mathsf{V}_0\mathsf{C}} \\ 2\mathsf{S}/\sqrt{7\mathsf{S}^2 + 12\mathsf{V}_0\mathsf{C}} \end{pmatrix} \tag{8.3.2}$$

$$\begin{pmatrix} \tilde{\alpha} \\ \tilde{\kappa} \end{pmatrix} := \mathbf{M}^{-1}(\mathbf{q} - \mathbf{r}) \tag{8.3.3}$$

where

$$\mathbf{q} := \frac{1}{2t_1}\begin{pmatrix} \mathsf{V}(x_0,t_1) - \mathsf{V}(-x_0,t_1) - \mathsf{V}_+ + \mathsf{V}_- \\ \mathsf{V}(x_0,t_1) + \mathsf{V}(-x_0,t_1) - \mathsf{V}_+ - \mathsf{V}_- \end{pmatrix}$$

and

$$\mathbf{r} := \begin{pmatrix} \dfrac{\tilde{\rho}\tilde{\sigma}^3\left(1-\tilde{\rho}^2\right) + \tilde{y}_0\tilde{\rho}^2\tilde{\sigma}^2}{24\tilde{y}_0}x_0 \\[2ex] \dfrac{12\tilde{y}_0\tilde{\rho}\tilde{\sigma} + \left(\tilde{\rho}^2 - 4\right)\tilde{\sigma}^2}{48} + \dfrac{(176 - 712\tilde{\rho}^2 + 521\tilde{\rho}^4)\tilde{\sigma}^4 + 40\tilde{y}_0\tilde{\rho}^3\tilde{\sigma}^3}{7680\tilde{y}_0^2}x_0^2 \end{pmatrix}$$

and

$$\mathbf{M} := \begin{pmatrix} -\dfrac{\tilde{\rho}\tilde{\sigma}}{12\tilde{y}_0}x_0 & -\frac{\tilde{\rho}\tilde{\sigma}}{12}x_0 \\[2ex] \dfrac{1}{2} + \dfrac{(13\tilde{\rho}^2 - 6)\tilde{\sigma}^2}{96\tilde{y}_0^2}x_0^2 & -\frac{\tilde{y}_0}{2} - \frac{\tilde{\rho}^2\tilde{\sigma}^2}{96\tilde{y}_0}x_0^2 \end{pmatrix}$$

provided that $\tilde{y}_0 \neq 0$, $\tilde{\rho} \neq 0$ and $\tilde{\rho}^2 \neq \frac{3}{7}\left(1 - 16\tilde{y}_0^2/\left(x_0^2\tilde{\sigma}^2\right)\right)$.

Proof. Directly substitute (8.3.2), (8.3.3), and each $(x,t) \in \mathbb{K}$, into (8.3.1). $\square$

8.4 Proof of Theorem 8.1.1

In the proof of Theorem 8.1.1, we split the cases $x \neq 0$ and the at-the-money case $x = 0$. Sections 8.4.1 to 8.4.5 deal with the general case $x \neq 0$ and the proof in the special case $x = 0$ is postponed to Section 8.4.6. We refer the reader to this very section for the reasons of such a separation of the cases. From [71], we know that the call price can be written as an inverse Fourier transform (see (8.4.1)). We rescale this integral (subsection 8.4.2) and move the horizontal contour of integration so it passes through the saddlepoint of the small-time approximation of the integrand (Lemmas 8.4.1 and 8.4.3). The idea is then to split up the inverse Fourier transform into three parts; we prove that two of them are negligible (Corollary 8.4.6) and finally that the remaining integral provides the asymptotic expansion for the call price (Proposition 8.4.8). There are obvious similarities between the proof of this theorem and the proof of Theorem 5.1.1 in Section 5.4. However there are also substantial differences, in particular due to the fact that the saddlepoint is not available in closed-form anymore.

8.4.1 The Fourier inversion formula for call options

As in Section 5.4 let us define the sets $A_{t,X}$ and $\Lambda_{t,X}$ for each $t \geq 0$ by

$$A_{t,X} := \{\nu \in \mathbb{R} : \mathbb{E}\left(\exp\left(\nu(X_t - x_0)\right)\right) < \infty\}, \quad \text{and} \quad \Lambda_{t,X} := \{z \in \mathbb{C} : -\Im(z) \in A_{t,X}\}.$$

Let us define the characteristic function $\phi : \mathbb{C} \to \mathbb{C}$ of the log return $X_t - x_0$ by

$$\phi_t(z) := \mathbb{E}\left(\mathrm{e}^{\mathtt{i}z(X_t - x_0)}\right), \quad \text{for all } z \in \Lambda_{t,X}.$$

By [71, Theorem 5.1], we know that for any $\alpha \in \mathbb{R}$ such that $\alpha + 1 \in A_{t,X}$ and $\alpha \neq 0$, we have the following Fourier inversion formula for the price of a call option

$$\begin{aligned}\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+}{S_0} &= \phi_t\left(-\mathtt{i}\right)1\!\!1_{\{-1<\alpha<0\}} + \left(\phi_t\left(-\mathtt{i}\right) - \mathrm{e}^x\phi_t\left(0\right)\right)1\!\!1_{\{\alpha<-1\}} + \left(\phi_t\left(-\mathtt{i}\right) - \frac{\mathrm{e}^x}{2}\phi_t\left(0\right)\right)1\!\!1_{\{\alpha=-1\}} \\ &+ \frac{1}{\pi}\int_{0-\mathtt{i}\alpha}^{+\infty-\mathtt{i}\alpha} \Re\left(\mathrm{e}^{-\mathtt{i}zx}\frac{\phi_t\left(z-\mathtt{i}\right)}{\mathtt{i}z - z^2}\right)\mathrm{d}z.\end{aligned}$$

The first three terms on the right hand side are complex residue terms that arise when we cross the pole of $\left(\mathtt{i}z - z^2\right)^{-1}$ at $z = 0$ and at $z = \mathtt{i}$. Setting $u = \mathtt{i} - z$ and using the fact that S_t is a true martingale (see [3, Proposition 2.5]) we obtain

$$\begin{aligned}\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+}{S_0} &= 1\!\!1_{\{-1<\alpha<0\}} + \left(1 - \mathrm{e}^x\right)1\!\!1_{\{\alpha<-1\}} + \left(1 - \frac{\mathrm{e}^x}{2}\right)1\!\!1_{\{\alpha=-1\}} \\ &+ \frac{\mathrm{e}^x}{\pi}\int_{0+\mathtt{i}(\alpha+1)}^{+\infty+\mathtt{i}(\alpha+1)} \Re\left(\mathrm{e}^{\mathtt{i}ux}\frac{\phi_t\left(-u\right)}{\mathtt{i}u - u^2}\right)\mathrm{d}u.\end{aligned} \tag{8.4.1}$$

8.4.2 Rescaling the variable of integration

Let $u := k/t$ and set $\alpha+1 := p^*(x)/t$ in (8.4.1) (see Theorem 8.1.1 for a definition of $p^*(x)$). Then

$$\frac{1}{S_0}\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+ = \left(1-\mathrm{e}^x\right)\mathbb{1}_{\{x<0\}} + \frac{\mathrm{e}^x}{\pi}\Re\left(\int_{0+\mathtt{i}p^*(x)}^{+\infty+\mathtt{i}p^*(x)} \exp\left(\frac{\mathtt{i}xk}{t}\right)\frac{\phi_t(-k/t)}{\mathtt{i}k/t - k^2/t^2}\frac{\mathrm{d}k}{t}\right), \quad (8.4.2)$$

holds for t sufficiently small. The first residue term corresponds to the intrinsic value of the call option. For $k \neq 0$, we have $t^{-1}\left(\mathtt{i}\frac{k}{t} - \frac{k^2}{t^2}\right)^{-1} = -\frac{t}{k^2}\left(1+\mathcal{O}(t)\right)$ therefore we can rewrite (8.4.2) as

$$\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+}{S_0} = \left(1-\mathrm{e}^x\right)\mathbb{1}_{\{x<0\}} - t\frac{\mathrm{e}^x}{\pi}\Re\left(\int_{\mathtt{i}p^*(x)}^{+\infty+\mathtt{i}p^*(x)} \exp\left(\frac{\mathtt{i}xk}{t}\right)\frac{\phi_t(-k/t)}{k^2}\left(1+\mathcal{O}(t)\right)\mathrm{d}k\right).$$

In Section 7.2 we have shown that for any $p \in (p_-, p_+)$ the explosion time $t^*(p/t)$ is strictly larger than t for t sufficiently small, so that $p^*(x)/t \in \Lambda_{t,X}$ for t sufficiently small. We also note that

$$\Re\left(\mathrm{e}^{\mathtt{i}kx/t}\frac{\phi_t(-k/t)}{k^2}\right) = \Re\left(\frac{\exp(\mathtt{i}kx/t)}{k^2}\mathbb{E}\left(\mathrm{e}^{-\mathtt{i}k(X_t-x_0)/t}\right)\right) = \mathbb{E}\left(\Re\left(\frac{\exp(\mathtt{i}kx/t)}{k^2}\mathrm{e}^{-\mathtt{i}k(X_t-x_0)/t}\right)\right),$$

and we can easily show that this expression is an even function of $\Re(k)$ and an odd function of $\Im(k)$. Thus, we can rewrite the normalised call price as

$$\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+}{S_0} = \left(1-\mathrm{e}^x\right)\mathbb{1}_{\{x<0\}} - t\frac{\mathrm{e}^x}{2\pi}\Re\left(\int_{-\infty+\mathtt{i}p^*(x)}^{+\infty+\mathtt{i}p^*(x)} \mathrm{e}^{\mathtt{i}xk/t}\phi_t(-k/t)\left(\frac{1}{k^2}+\mathcal{O}(t)\right)\mathrm{d}k\right).$$

Let R be a strictly positive real number. We can then further break up this expression as

$$\frac{\mathbb{E}\left(S_t - S_0\mathrm{e}^x\right)_+}{S_0} = \left(1-\mathrm{e}^x\right)\mathbb{1}_{\{x<0\}} - \frac{\mathrm{e}^x t}{2\pi}\Re\left(\left(\int_{\gamma_x} + \int_{\zeta_x}\right)\mathrm{e}^{\mathtt{i}xk/t}\phi_t(-k/t)\left(\frac{1}{k^2}+\mathcal{O}(t)\right)\mathrm{d}k\right),$$

where we define the contours γ_x and ζ_x by

$$\gamma_\alpha : (-\infty, -R] \cup [R, +\infty) \to \mathbb{C} \text{ such that } \gamma_\alpha(u) := u + \mathtt{i}p^*(x), \quad (8.4.3)$$

and

$$\zeta_\alpha : (-R, R) \to \mathbb{C} \text{ such that } \zeta_\alpha(u) := u + \mathtt{i}p^*(x), \quad (8.4.4)$$

and we drop the dependence on the constant R for clarity. Note that these contours are the same as those defined in (5.4.2) and (5.4.3) where $\alpha = p^*(x)$. We shall see later how to choose R precisely.

8.4.3 The saddlepoint and the Legendre transform

Let us define the set $\mathcal{Z}$ by

$$\mathcal{Z} := \left\{k \in \mathbb{C} : \Im(k) \in (p_-, p_+)\right\},$$

where p_- and p_+ are defined in (7.2.1). Let us further define the function $F : \mathcal{Z} \to \mathbb{C}$ by

$$F(k) := -\mathtt{i}kx - \Lambda(-\mathtt{i}k), \quad \text{for all } k \in \mathcal{Z}, \quad (8.4.5)$$

where the function Λ is defined in Proposition 7.2.1.

Lemma 8.4.1. *For all real number x the function F has a saddlepoint at $k^*(x) = \mathtt{i}p^*(x)$, where $p^*(x) \in (p_-, p_+)$ is the unique solution to the equation*

$$\Lambda'(p^*(x)) = x. \tag{8.4.6}$$

Remark 8.4.2. It is straightforward to check that the identities $p^*(x) > 0$ for $x > 0$, $p^*(x) < 0$ for $x < 0$ and $p^*(0) = 0$ hold.

Proof. From Theorem 7.2.3 we already know that for any real number x, the equality in (8.4.6) has a unique solution $p^*(x) \in (p_-, p_+)$, which implies that $F'(k^*(x)) = -\mathtt{i}x + \mathtt{i}\Lambda'(p^*(x)) = 0$. □

8.4.4 Small-time asymptotics for the rescaled characteristic function

In the following lemma we characterise the asymptotic behaviour of the rescaled characteristic function when the maturity t is small.

Lemma 8.4.3. *For all $k \in \mathbb{C}$ such that $-k_{\mathtt{i}} \in (p_-, p_+) \setminus \{0\}$ the function ϕ_t satisfies*

$$\phi_t\left(-\frac{k}{t}\right) = U(-\mathtt{i}k)\exp\left(\frac{\Lambda(-\mathtt{i}k)}{t}\right)(1 + \mathcal{O}(t)), \quad \text{as } t \text{ tends to zero.}$$

Proof. Recall the closed-form expression for ϕ_t given in (3.2.2). We further know that ϕ_t is an even function of d so that we can take $\sqrt{k^2} = k$ in all calculations, where the square root is the principal branch. Recall that k is non null. We hence have the following asymptotic behaviour for $d(k/t)$ and $g(k/t)$ as the ratio t/k tends to zero:

$$d\left(-\frac{k}{t}\right) = t^{-1}\left(\sigma^2\bar{\rho}^2k^2 + \left(2\kappa\rho\sigma - \sigma^2\right)\mathtt{i}kt + \kappa^2t^2\right)^{1/2} = \frac{k}{t}d_0 + d_1 + \mathcal{O}(t/k),$$

$$g\left(-\frac{k}{t}\right) = \frac{\kappa t + \rho\sigma\mathtt{i}k - \left(\sigma^2\bar{\rho}^2k^2 + \left(2\kappa\rho\sigma - \sigma^2\right)\mathtt{i}kt + \kappa^2t^2\right)^{1/2}}{\kappa t - \rho\sigma\mathtt{i}k + \left(\sigma^2\bar{\rho}^2k^2 + \left(2\kappa\rho\sigma - \sigma^2\right)\mathtt{i}kt + \kappa^2t^2\right)^{1/2}} = g_0 + \frac{t}{k}g_1 + \mathcal{O}\left(\frac{t^2}{k^2}\right),$$

where we define the following quantities:

$$d_0 := \sigma\bar{\rho}, \qquad d_1 := \frac{2\kappa\rho - \sigma}{2\bar{\rho}}\mathtt{i}, \qquad g_0 := \frac{\sigma\bar{\rho} - \mathtt{i}\rho\sigma}{\sigma\bar{\rho} + \mathtt{i}\rho\sigma}, \qquad g_1 := \frac{2\kappa - \rho\sigma}{\left(\sigma\bar{\rho} + \mathtt{i}\rho\sigma\right)^2\bar{\rho}}.$$

We know that $|k| \geq |k_{\mathtt{i}}| > 0$ so that

$$d\left(-\frac{k}{t}\right) = d_0\frac{k}{t} + d_1 + \mathcal{O}(t) \qquad \text{and} \qquad g\left(-\frac{k}{t}\right) = g_0 + g_1\frac{t}{k} + \mathcal{O}\left(t^2\right),$$

as t tends to zero. From this we obtain

$$\begin{aligned}
D\left(-\frac{k}{t}, t\right) &= \frac{1}{\sigma^2}\left(\kappa + \rho\sigma\mathtt{i}\frac{k}{t} - d_0\frac{k}{t} - d_1 + \mathcal{O}(t)\right)\frac{1 - \exp\left(-d_0k - d_1t + \mathcal{O}\left(t^2\right)\right)}{1 - g(k)\exp\left(-d_0k - d_1t + \mathcal{O}\left(t^2\right)\right)} \\
&= \frac{v\left(\rho\sigma\mathtt{i} - d_0\right)k}{\sigma^2t}\frac{1 - \mathrm{e}^{-d_0k}}{1 - g_0\mathrm{e}^{-d_0k}} \\
&+ \frac{\exp(-d_0k)}{\left(1 - g_0\mathrm{e}^{-d_0k}\right)\sigma^2}\left(\left(\mathtt{i}\rho\sigma - d_0\right)kd_1 + \left(\kappa - d_1\right)\left(1 - \mathrm{e}^{d_0k}\right) - \frac{\left(\mathtt{i}\rho\sigma - d_0\right)\left(1 - \mathrm{e}^{-d_0k}\right)\left(g_1 - d_1g_0k\right)}{1 - g_0\mathrm{e}^{-d_0k}}\right) \\
&+ \mathcal{O}(t).
\end{aligned}$$

Similarly we have

$$\begin{aligned}
C\left(-\frac{k}{t}\right) &= \frac{\kappa\theta}{\sigma^2}\left(\kappa t + \rho\sigma\mathtt{i}k - d\left(-\frac{k}{t}\right) - 2\log\left(\frac{1-g\left(-k/t\right)\exp\left(-d\left(-k/t\right)t\right)}{1-g\left(-k/t\right)}\right)\right)\\
&= \frac{\kappa\theta}{\sigma^2}\left(\mathtt{i}\rho\sigma k - d_0 k - 2\log\left(\frac{1-g_0\exp\left(-d_0 k\right)}{1-g_0}\right)\right) + \mathcal{O}\left(t\right).
\end{aligned}$$

□

Corollary 8.4.4. *For all $p \in (p_-, p_+)$ we have*

$$\mathbb{E}\left(\mathrm{e}^{p(X_t - x_0)/t}\right) = U\left(p\right)\exp\left(t^{-1}\Lambda\left(p\right)\right)\left(1+\mathcal{O}\left(t\right)\right), \quad \textit{as } t \textit{ tends to } 0.$$

8.4.5 The minimum of $\Re\left(F(k)\right)$ along the horizontal contour

Lemma 8.4.5. *Let $k \in \mathcal{Z}$. Then for any $k_{\mathtt{i}} \in (p_-, p_+)$, the function $k_r \mapsto \Re\left(-\mathtt{i}kx - \Lambda\left(-\mathtt{i}k\right)\right)$ has a unique minimum at zero.*

Proof. Using the double angle formulae for trigonometric functions we have

$$\Re\left(\Lambda\left(p+\mathtt{i}q\right)\right) = \Re\left(\frac{v(p+\mathtt{i}q)}{\sigma\left(\bar{\rho}\cot\left(\frac{1}{2}\sigma\left(p+\mathtt{i}q\right)\bar{\rho}\right)-\rho\right)}\right) = \frac{vM\left(q\right)}{\sigma N\left(q\right)},$$

where the functions M and N are defined by

$$\begin{aligned}
M(q) &:= p\Big(\rho\cos\left(p\bar{\rho}\sigma\right) + \bar{\rho}\sin\left(p\bar{\rho}\sigma\right)\Big) - p\rho\cosh\left(q\bar{\rho}\sigma\right) - q\bar{\rho}\sinh\left(q\bar{\rho}\sigma\right),\\
N(q) &:= \cosh\left(q\bar{\rho}\sigma\right) + \left(1-2\rho^2\right)\cos\left(p\bar{\rho}\sigma\right) - 2\rho\bar{\rho}\sin\left(p\bar{\rho}\sigma\right).
\end{aligned}$$

Note that $N(0) > 0$, and $N'(q) > 0$ for $q > 0$. We need to show that $M(q)/N(q) < M(0)/N(0)$ for $q \neq 0$. By the symmetry $q \mapsto -q$, we can take $q > 0$ and by the symmetry $(p, \rho) \mapsto (-p, -\rho)$, we can take $p \geq 0$. It hence suffices to show that for $q > 0$

$$\frac{M'(q)}{N'(q)} < \frac{M(0)}{N(0)}, \tag{8.4.7}$$

because then, for all $q > 0$ the inequality

$$\frac{M(q)}{N(q)} < \frac{M(0) + M(0)\left(N(q) - N(0)\right)/N(0)}{N(q)} = \frac{M(0)}{N(0)}$$

holds. We have $M'(q)/N'(q) = -1/\sigma - p\rho - q\bar{\rho}\coth\left(q\bar{\rho}\sigma\right) < -p\rho - 2/\sigma$, so (8.4.7) will be satisfied by as soon as

$$-\frac{2}{\sigma} < p\rho + \frac{M(0)}{N(0)}. \tag{8.4.8}$$

If $\rho \geq 0$ then (8.4.8) holds since the ratio $M(0)/N(0)$ is a positive factor multiplied by $\Lambda(p) \geq 0$. Otherwise, for $\rho < 0$, note that the derivative of the right-hand side of (8.4.8) with respect to ρ is equal to a positive factor multiplied by

$$\left(2 - p\rho\sigma\right)\bar{\rho} - 2\rho^2\bar{\rho}\cos\left(p\bar{\rho}\sigma\right) - \rho\bar{\rho}^2\sin\left(p\bar{\rho}\sigma\right) \geq \bar{\rho}\left(2 - p\rho\sigma - 2\rho^2 + p\rho\sigma\right) > 0,$$

since $\left|-\bar{\rho}^2 \sin\alpha\right| \leq \alpha$ for all $\alpha > 0$. So it suffices to verify (8.4.8) in the limit as ρ tends to -1. But we have

$$\lim_{\rho\searrow -1} \left(p\rho + \frac{M(0)}{N(0)}\right) = -\frac{2p\sigma}{\sigma(2+p\sigma)} > -\frac{2}{\sigma},$$

which concludes the proof of the lemma. □

Corollary 8.4.6. *For any $R > 0$ we have*

$$\int_{\gamma_x} \phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{k^2} = \mathcal{O}\left(\mathrm{e}^{t^{-1}\Re(F(R+\mathtt{i}p^*(x)))}\right), \quad \text{as } t \text{ tends to zero,}$$

where the contour γ_x is defined in (8.4.3).

Remark 8.4.7. By Lemma 8.4.5 we know that $\Re\left(F\left(R+\mathtt{i}p^*\left(x\right)\right)\right) > F\left(\mathtt{i}p^*(x)\right)$. Evaluating the exponent $F(k) = -\mathtt{i}kx - \Lambda(-\mathtt{i}k)$ at the saddlepoint $k^* = \mathtt{i}p^*(x)$, we note that $F\left(k^*\right) = F\left(\mathtt{i}p^*(x)\right) = p^*(x)x - \Lambda\left(p^*(x)\right) = \Lambda^*(x)$, so that the equality $F\left(\mathtt{i}p^*(x)\right) = \Lambda\left(p^*(x)\right) > 0$ holds, and we also have $F'(k) = -x + \mathtt{i}\Lambda'(-\mathtt{i}k)$ and $F''(k) = \Lambda''\left(-\mathtt{i}k\right)$.

Proposition 8.4.8. *For any $R > 0$ and $x \neq 0$ the following equality holds as t tends to zero,*

$$I_1 = -\frac{\mathrm{e}^x t}{2\pi}\Re\left(\int_{\zeta_x} \mathrm{e}^{\mathtt{i}kx/t}\phi_t\left(-\frac{k}{t}\right)\left(\frac{1}{k^2}+\mathcal{O}\left(t\right)\right)\mathrm{d}k\right) = -\exp\left(-\frac{\Lambda^*(x)}{t}\right)\frac{A(x)t^{3/2}}{\sqrt{2\pi}}\left(1+\mathcal{O}\left(t\right)\right),$$

where the function A is given in Theorem 8.1.1 and where the contour ζ_x is defined in (8.4.4).

Proof. Since $F(k) = -\mathtt{i}kx - \Lambda\left(-\mathtt{i}k\right)$, Lemma 8.4.3 applied on the interval $[-R, R]$ implies

$$\Re\left(\int_{\zeta_x} \mathrm{e}^{\mathtt{i}kx/t}\phi_t\left(-\frac{k}{t}\right)\left(\frac{1}{k^2}+\mathcal{O}\left(t\right)\right)\mathrm{d}k\right) = \Re\left(\int_{\zeta_x} U\left(-\mathtt{i}k\right)\mathrm{e}^{-F(k)/t}(1+\mathcal{O}\left(t\right))\left(\frac{1}{k^2}+\mathcal{O}\left(t\right)\right)\mathrm{d}k\right). \tag{8.4.9}$$

The functions F and u are both analytic, thus [81, Theorem 7.1, Chapter 4] shows that

$$\begin{aligned}
-\frac{\mathrm{e}^{xt}}{2\pi}\Re\left(\int_{\zeta_x} U\left(-\mathtt{i}k\right)\exp\left(-\frac{F(k)}{t}\right)\frac{\mathrm{d}k}{k^2}\right) &= \frac{\mathrm{e}^x}{\sqrt{\pi}}\exp\left(-\frac{F\left(k^*\right)}{t}\right)\frac{U\left(-\mathtt{i}k^*\right)t^{3/2}}{k^{*2}\sqrt{2F''\left(k^*\right)}}\left(1+\mathcal{O}\left(t\right)\right)\\
&= -\frac{\mathrm{e}^x}{\sqrt{\pi}}\exp\left(-\frac{\Lambda^*(x)}{t}\right)\frac{U\left(p^*\right)t^{3/2}}{p^{*2}\sqrt{2\Lambda''(p^*)}}\left(1+\mathcal{O}\left(t\right)\right)\\
&= -\exp\left(-\frac{\Lambda^*(x)}{t}\right)\frac{A\left(x\right)t^{3/2}}{\sqrt{2\pi}}\left(1+\mathcal{O}\left(t\right)\right),
\end{aligned}$$

as t tends to zero, using the fact that $F(k^*) = \Lambda^*(x)$ and $F''(k^*) = \Lambda''(p^*)$. The $\mathcal{O}\left(t\right)$ terms in (8.4.9) constitute higher order terms which we can neglect at the order we are interested in. □

Combining Corollary 8.4.6 and Proposition 8.4.8, Theorem 8.1.1 follows for $x \neq 0$.

8.4.6 The at-the-money case $x = 0$

We now proove Theorem 8.1.1 for the at-the-money case $x = 0$. We cannot apply the same methodology as before since the horizontal contour now passes through the saddlepoint $k^*(0) = \mathtt{i}p^*(0) = 0$, and the ratio $u(k)/(\mathtt{i}k)$ is not analytic at the origin, so [81, Theorem 7.1, chapter 4] does not apply anymore. To circumvent this problem we use Cauchy's Integral theorem to deform the path of steepest descent for the function F, along which it is real, and we deal with the pole at zero using a limiting keyhole contour argument to compute the residue. We first start with the following two lemmata which we shall need later. The first one, Lemma 8.4.9 computes the asymptotic behaviour of the path of steepest descent around the origin. The second one, Lemma 8.4.11 provides the logarithmic tail behaviour of the characteristic function ϕ_t as the maturity tends to zero.

Lemma 8.4.9. *For $|s|$ sufficiently small the path of steepest descent for the function F in (8.4.5) starting at the saddlepoint $k^* = 0$ is the map $\gamma : s \in \mathbb{R} \to \mathbb{C}$ satisfying $\gamma(s) = s + \mathtt{i}k_{\mathtt{i}}(s)$, where $k_{\mathtt{i}}$ maps $\mathbb{R}$ to $\mathbb{R}$ and has the following expansion around the origin*

$$k_{\mathtt{i}}(s) = \frac{1}{4}\rho\sigma s^2 + \frac{1}{64}\sigma^3\rho^3 s^4 + \mathcal{O}\left(s^6\right). \tag{8.4.10}$$

Remark 8.4.10. Note that for $|s|$ small enough, the contour γ will lie above the horizontal contour $s + \mathtt{i}p^*(0)$ if $\rho > 0$ and below otherwise.

Proof. Around the saddlepoint $k^*(0) = 0$, we have the series expansion

$$F(k) = \frac{1}{2}\Lambda''(0)\,k^2 - \frac{1}{6}\mathtt{i}\Lambda'''(0)\,k^3 - \frac{1}{24}\Lambda^{(4)}(0)\,k^4 + \mathcal{O}\left(k^5\right).$$

We seek a contour of the form $\gamma(s) = s + \mathtt{i}\left(a_1 s + a_2 s^2 + \ldots\right)$ with all coefficients real-valued such that F is real along this contour. Matching the coefficients of the series expansion for $F \circ \gamma$ around the origin to those of the expansion above leads to the expansion given in the lemma. $\square$

Lemma 8.4.11. *For any $k_{\mathtt{i}} \in (p_-, p_+)$, the function ϕ_t has the following behaviour along the horizontal contour $k = k_r + \mathtt{i}k_{\mathtt{i}}$:*

$$\Re\left(\frac{\mathtt{i}kx}{t} + \log\left|\phi_t\left(-\frac{k}{t}\right)\right|\right) \sim -\frac{1}{\sigma}\left(\kappa\theta + \frac{v}{t}\right)\bar{\rho}\,|k_r|, \quad \textit{as } |k_r| \textit{ tends to infinity,}$$

uniformly in t for all $t < 1$.

Proof. For $|k|$ large we have

$$d\left(\frac{k}{t}\right) = \sqrt{\left(\kappa - \rho\sigma\mathtt{i}\frac{k}{t}\right)^2 + \sigma^2\left(\mathtt{i}\frac{k}{t} + \frac{k^2}{t^2}\right)} \sim \frac{1}{t}\sigma\bar{\rho}\sqrt{k^2}, \quad \text{as } |k| \text{ tends to infinity,}$$

uniformly in t for all $t < 1$ and the lemma follows. $\square$

We now prove Theorem 8.1.1 for the special case $x = 0$. We shall assume that $\rho < 0$ for clarity, and the arguments in the case $\rho \geq 0$ are analogous. Let us note the following facts

- From Remark 8.4.10 we know that the contour γ lies below the horizontal contour $\gamma_H : \mathbb{R} \to \mathbb{C}$ defined by $\gamma(s) = s + \mathtt{i}p^*(0) = s$, for all $s \in \mathbb{R}$. We wish to deform the contour so that the pole remains outside the new contour, so that we can invoke Cauchy's theorem.

- Let ε be a strictly positive real number and define the contour $\gamma_\varepsilon : (-\pi, \pi] \mapsto \mathbb{C}$ as the clockwise oriented circular keyhole contour parameterised by $\gamma_\varepsilon(\theta) = \varepsilon e^{\mathtt{i}\theta}$ around the pole. To leave the pole outside the contour we follow the steepest descent contour γ for $s < 0$ until it touches γ_ε at the point $s = -\varepsilon + \mathcal{O}(\varepsilon^3)$. We then follow the contour γ_ε clockwise around the pole, and finally switch back to the right part of $\gamma(s)$.

- The contour γ is below γ_H so that γ intersects γ_ε on its lower half ($\theta \in (-\pi, 0]$), which can be analytically represented as $\gamma_\varepsilon^- := s \in [-\varepsilon, \varepsilon] \mapsto s - \mathtt{i}\sqrt{\varepsilon^2 - s^2}$. Equating the Taylor series approximations of the path of steepest descent γ and γ_ε^-, we have

$$\gamma(s) = s + \frac{1}{4}\mathtt{i}\rho\sigma s^2 + \mathcal{O}\left(s^3\right) = s - \mathtt{i}\sqrt{\varepsilon^2 - s^2},$$

  and we see that the two curves intersect at the point $s^* = \pm\varepsilon + \mathcal{O}\left(\varepsilon^3\right)$ (see Figure 8.4.6 for a graphical representation of these contours).

Choose now $0 < \varepsilon < \delta < 1 < R$, with δ inside the radius of convergence of (8.4.10), and small enough to ensure that the function $k_\mathtt{i}$ is strictly decreasing on $[0, \delta]$ and strictly increasing on $[-\delta, 0]$. We now define the following contours (see also Figure 8.4.6)

- $\gamma_{\varepsilon,\delta}$ is γ restricted to $s \in [-\delta, -\varepsilon] \cup [\varepsilon, \delta]$;
- γ_ε^U is the portion of the circular keyhole contour γ_ε which lies above γ;
- $\Gamma^-_{R,\varepsilon,\delta}$ is the vertical strip which joins $-R + 2\mathtt{i}\varepsilon$ to $-R - \mathtt{i}k_\mathtt{i}(-\delta)$;
- $\Gamma^+_{R,\varepsilon,\delta}$ is the vertical strip which joins $R - \mathtt{i}k_\mathtt{i}(-\delta)$ to $R + 2\mathtt{i}\varepsilon$.

The symmetry of the contour γ around the imaginary axis implies the symmetry of the function $k_\mathtt{i}$. Motivated by [71, Theorem 5.1], we have the following identity:

$$\left(\int_{\Gamma^\pm_{R,\varepsilon,\delta}} + \int_{-R+\mathtt{i}k_\mathtt{i}(-\delta)}^{-\delta+\mathtt{i}k_\mathtt{i}(-\delta)} + \int_{\gamma_{\varepsilon,\delta}} + \int_{\gamma_\varepsilon^U} + \int_{\delta+\mathtt{i}k_\mathtt{i}(\delta)}^{R+\mathtt{i}k_\mathtt{i}(\delta)} - \int_{-R+2\mathtt{i}\varepsilon}^{R+2\mathtt{i}\varepsilon}\right)\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k} = 0 \qquad (8.4.11)$$

by Cauchy's Integral theorem. For $|k|$ small enough and $t > 0$ fixed, we have $(\mathtt{i}k)^{-1}\phi_t(-k/t) = (\mathtt{i}k)^{-1} + \mathcal{O}(1)$, and hence

$$\int_{\gamma_\varepsilon^U}\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k} = \int_{\pi+\mathcal{O}(\varepsilon)}^{\mathcal{O}(\varepsilon)}\left(\frac{1}{\mathtt{i}\varepsilon e^{\mathtt{i}\theta}} + \mathcal{O}(\varepsilon)\right)\mathtt{i}\varepsilon e^{\mathtt{i}\theta}\mathrm{d}\theta = -\pi + o(1), \quad \text{as } \varepsilon \text{ tends to zero.} \qquad (8.4.12)$$

We further have

$$\int_{\gamma_{\varepsilon,\delta}}\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k} = \int_{[-\delta,-\varepsilon]\cup[\varepsilon,\delta]}\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\mathrm{d}s.$$

As before we can easily verify that the function $s \mapsto \Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)$ is an even function and that $s \mapsto \Im\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)$ is an odd function. Thus

$$\int_{[-\delta,-\varepsilon]\cup[\varepsilon,\delta]}\phi_t\left(-\frac{k}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\mathrm{d}s = \int_{[-\delta,-\varepsilon]\cup[\varepsilon,\delta]}\Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)\mathrm{d}s. \qquad (8.4.13)$$

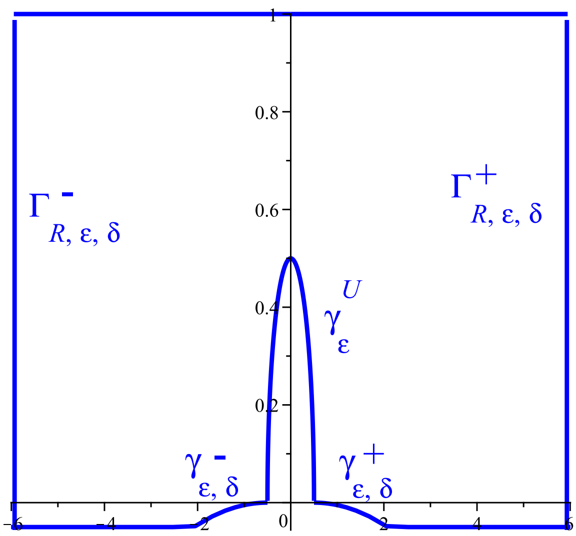


Figure 8.2: We plot here the closed contour of integration in (8.4.11) for $\varepsilon = .5$, $\delta = 2$ and $R = 6$.

But on the contour γ we have $\Re\left(1/\mathtt{i}\gamma(s)\right) = -\rho\sigma/4 + \mathcal{O}(s)$, for $|s|$ small enough, i.e. taking the real part removes the singularity at the origin. Equation (8.4.12) henceforth implies

$$\left(\int_{\Gamma^-_{R,\varepsilon,\delta}} + \int_{-R+\mathtt{i}k_{\mathtt{i}}(-\delta)}^{-\delta+\mathtt{i}k_{\mathtt{i}}(-\delta)} + \int_{\delta+\mathtt{i}k_{\mathtt{i}}(\delta)}^{R+\mathtt{i}k_{\mathtt{i}}(\delta)} + \int_{\Gamma^+_{R,\varepsilon,\delta}} + \int_{-\infty+2\mathtt{i}\varepsilon}^{-R+2\mathtt{i}\varepsilon} + \int_{R+2\mathtt{i}\varepsilon}^{\infty+2\mathtt{i}\varepsilon}\right)\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k}$$
$$+\int_{[-\delta,-\varepsilon]\cup[\varepsilon,\delta]}\Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)\mathrm{d}s - \pi + o(1) = \int_{-\infty+2\mathtt{i}\varepsilon}^{\infty+2\mathtt{i}\varepsilon}\frac{\phi_t(-k)}{\mathtt{i}k}\mathrm{d}k = -\int_{-\infty+2\mathtt{i}\varepsilon}^{\infty+2\mathtt{i}\varepsilon}\frac{\phi_t(k)}{\mathtt{i}k}\mathrm{d}k,$$

as ε tends to zero. But by [71, Theorem 5.1], the integral on the right hand side is equal to minus the normalised digital call price $2\pi\mathbb{P}(S_t > S_0)$, which is clearly independent of ε. Taking the limit as ε goes to zero and dividing by -2π, we obtain

$$-\frac{1}{2\pi}\int_{\Gamma^{\pm}_{R,0,\delta}}\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k} - \frac{1}{2\pi}\left(\int_{-R+\mathtt{i}k_{\mathtt{i}}(-\delta)}^{-\delta+\mathtt{i}k_{\mathtt{i}}(-\delta)} + \int_{\delta+\mathtt{i}k_{\mathtt{i}}(\delta)}^{R+\mathtt{i}k_{\mathtt{i}}(\delta)} + \int_{-\infty}^{-R} + \int_{R}^{\infty}\right)\Re\left(\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k}\right)$$
$$+\frac{1}{2} - \int_{-\delta}^{\delta}\Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)\mathrm{d}s = \mathbb{P}(S_t > S_0). \tag{8.4.14}$$

From Lemma 8.4.9, we know that the contour γ has the following expansion for $|s|$ small enough,

$$\gamma(s) = s + \frac{\mathtt{i}}{4}\rho\sigma s^2 + \frac{\mathtt{i}}{64}\sigma^3\rho^3 s^4 + \mathcal{O}(s^6).$$

Setting $k = \gamma(s)$ and performing a Taylor series expansion of $\Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)\mathrm{e}^{F(\gamma(s))/t}$ around $t = 0$ and $s = 0$, we find that

$$\Re\left[\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right] = \left(q(s) + \left(\frac{1}{4}\kappa(\theta - v)\right)t + \varepsilon(s,t)\right)\exp\left(-\frac{F(\gamma(s))}{t}\right), \tag{8.4.15}$$

where the function q is defined by

$$q(s) := q(k;\kappa,\theta,\rho,\sigma) = \frac{1}{2}v + \frac{1}{4}\rho\sigma + \left[\frac{1}{4}\rho\sigma + \left(\frac{7}{64}\sigma^3\rho^3 - \frac{1}{8}v\kappa\theta + \frac{1}{32}v^2\sigma\rho + \frac{13}{96}\sigma^2 v\rho^2 + \frac{1}{8}v^2\kappa \right.\right.$$
$$\left.\left. - \frac{1}{48}v^3 + \frac{1}{8}\sigma^3\rho + \frac{1}{48}v\kappa\sigma\rho - \frac{5}{48}\kappa\theta\sigma\rho - \frac{1}{12}\sigma^2 v - \left(\frac{1}{64}(\rho\sigma+2v)\right)\sigma^2\rho^2\right)\right]s^2, \tag{8.4.16}$$

and $\varepsilon(s,t) = \mathcal{O}(s^3) + \mathcal{O}(st) + \mathcal{O}(t^2)$. The function $q:\mathbb{R}\mapsto\mathbb{R}$ is just a quadratic so it has an analytic continuation to the complex plane. Thus using [81, Theorem 7.1, Chapter 4] and assuming δ is inside the radius of convergence of (8.4.15), we obtain

$$\begin{aligned}
&\frac{1}{2} - \frac{1}{2\pi}\int_{-\delta}^{\delta}\Re\left(\phi_t\left(-\frac{\gamma(s)}{t}\right)\frac{\gamma'(s)}{\mathtt{i}\gamma(s)}\right)\mathrm{d}s\\
&= \frac{1}{2} - \frac{1}{2\pi}\int_{-\delta}^{\delta}\left(q(s) + \left(\frac{1}{4}\kappa(\theta - v)\right)t + \varepsilon(s,t)\right)\exp\left(-\frac{F(\gamma(s))}{t}\right)\mathrm{d}s\\
&= \frac{1}{2} - \left(\frac{1}{\pi}\sum_{n=0}^{\infty}\Gamma\left(n+\frac{1}{2}\right)a_{2n}t^{n+1/2} + \frac{1}{4\pi}\frac{\kappa(\theta-v)}{\sqrt{2F''(0)}}\Gamma\left(\frac{1}{2}\right)t^{3/2}\right) + \int_{-\delta}^{\delta}\varepsilon(s,t)\exp\left(-\frac{F(\gamma(s))}{t}\right)\mathrm{d}s\\
&= \frac{1}{2} - \frac{q\Gamma(1/2)\,t^{1/2}}{\pi\sqrt{2(F\circ\gamma)''(0)}} - \frac{\Gamma(3/2)}{\pi}\left(2q'' + \left(\frac{5(F\circ\gamma)'''(0)^2}{6(F\circ\gamma)''(0)^2} - \frac{(F\circ\gamma)^{(4)}(0)}{2(F\circ\gamma)''(0)}\right)q\right)t^{3/2} + \mathcal{O}\left(t^{5/2}\right)\\
&= \frac{1}{2} - \frac{v+\rho\sigma/2}{2\sqrt{2\pi v}}t^{1/2} + \sqrt{\frac{2}{\pi}}\frac{v^{-3/2}}{768}\Big(-8v\kappa\sigma\rho + 9\sigma^3\rho^3 + 8v^3 + 2\sigma^2 V_0\rho^2\\
&+ 8\sigma^2 v + 48v^2\kappa - 48v\kappa\theta + 40\kappa\theta\sigma\rho - 12v^2\sigma\rho - 12\sigma^3\rho\Big)t^{3/2} + \mathcal{O}\left(t^{5/2}\right).
\end{aligned}$$

We now bound the remaining terms in (8.4.14).

- By the uniform convergence in Lemma 8.4.3, the absolute value of the integrals along $(-\infty,-R]$ and $[R,\infty)$ in (8.4.14) is of order $\mathcal{O}\left(\exp\left(-t^{-1}\left(\Re(F(-R))\wedge\Re(F(R))\right)\right)\right)$ and note that $\Re(F(-R))\wedge\Re(F(R)) > F(0) = 0$.
- Using Lemmas 8.4.3 and 8.4.5 we bound the integrals $\int_{-R+\mathtt{i}k_\mathtt{i}(-\delta)}^{\gamma(-\delta)} + \int_{\gamma(\delta)}^{R+\mathtt{i}k_\mathtt{i}(\delta)}$ in (8.4.14) as

$$\begin{aligned}
&\left(\int_{-R+\mathtt{i}k_\mathtt{i}(-\delta)}^{\gamma(-\delta)} + \int_{\gamma(\delta)}^{R+\mathtt{i}k_\mathtt{i}(\delta)}\right)\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k}\\
&= \left(\int_{-R+\mathtt{i}k_\mathtt{i}(-\delta)}^{\gamma(-\delta)} + \int_{\gamma(\delta)}^{R+\mathtt{i}k_\mathtt{i}(\delta)}\right)U(-\mathtt{i}k)\exp\left(-\frac{F(k)}{t}\right)(1+\varepsilon(k,t))\frac{\mathrm{d}k}{\mathtt{i}k}\\
&\leq \exp\left(-\frac{1}{t}\min\left\{\Re(F(\gamma(-\delta))),\Re(F(\gamma(\delta)))\right\}\right)\left(\int_{-R+\mathtt{i}k_\mathtt{i}(-\delta)}^{\gamma(-\delta)} + \int_{\gamma(\delta)}^{R+\mathtt{i}k_\mathtt{i}(\delta)}\right)\left|(1+\varepsilon(k,t))\frac{U(-\mathtt{i}k)}{\mathtt{i}k}\right|\mathrm{d}k\\
&= \mathcal{O}\left(\exp\left(-\frac{1}{t}\min\left\{\Re(F(\gamma(-\delta))),\Re(F(\gamma(\delta)))\right\}\right)\right),
\end{aligned}$$

- The contours $\Gamma_{R,0,\delta}^{-}$ and $\Gamma_{R,0,\delta}^{+}$ are both vertical strips of length $k_\mathtt{i}(\delta) > 0$. Using Lemma 8.4.11, for any $m > 0$ we obtain

$$\int_{\Gamma_{R,0,\delta}^{\pm}}\phi_t\left(-\frac{k}{t}\right)\frac{\mathrm{d}k}{\mathtt{i}k} = \mathcal{O}\left(\exp\left(-\frac{m}{t}\right)\right),$$

for t sufficiently small and R sufficiently large.

Combining all the estimates of the integrals in (8.4.14), we obtain

$$\mathbb{P}\left(X_t - x_0 > 0\right) = \frac{1}{2} - \frac{v+\rho\sigma/2}{2\sqrt{2\pi v}}t^{1/2} + \sqrt{\frac{2}{\pi}}\frac{v^{-3/2}}{768}\Big(- 8v\kappa\sigma\rho + 9\sigma^3\rho^3 + 8v^3 + 2\sigma^2 v\rho^2$$
$$+ 8\sigma^2 v + 48V_0^2\kappa - 48v\kappa\theta + 40\kappa\theta\sigma\rho - 12v^2\sigma\rho - 12\sigma^3\rho\Big)t^{3/2} + \mathcal{O}\left(t^{5/2}\right).$$

Similarly, under the Share measure $\mathbb{P}^*$ we have $q^*(k) = q(k,\bar{\kappa};\bar{\theta},\sigma,-\rho)$, where the function q is defined in (8.4.16), and

$$\mathbb{P}^*\left(X_t - x_0 > 0\right) = \frac{1}{2} + \frac{v+\rho\sigma/2}{2\sqrt{2\pi v}}t^{1/2} - \sqrt{\frac{2}{\pi}}\frac{v^{-3/2}}{768}\Big(8v\bar{\kappa}\sigma\rho - 9\sigma^3\rho^3 + 8v^3 + 2\sigma^2 v\rho^2$$
$$+ 8\sigma^2 v + 48v^2\bar{\kappa} - 48v\bar{\kappa}\bar{\theta} - 40\bar{\kappa}\bar{\theta}\sigma\rho + 12v^2\sigma\rho + 12\sigma^3\rho\Big)t^{3/2} + \mathcal{O}\left(t^{5/2}\right),$$

which implies the equality

$$\begin{aligned}\frac{\mathbb{E}\left(S_t - S_0\right)_+}{S_0} &= \mathbb{P}^*\left(X_t - x_0 > 0\right) - \mathbb{P}\left(X_t - x_0 > 0\right)\\ &= \sqrt{\frac{vt}{2\pi}} - \sqrt{\frac{2}{v\pi}}\frac{t^{3/2}}{48}\left(\sigma^2\left(1-\frac{1}{4}\rho^2\right) + \sigma^2 + v^2 - 6\kappa\left(\theta - v\right) - 3v\rho\sigma\right) + \mathcal{O}\left(t^{5/2}\right).\end{aligned}$$

8.5 Additional proofs

8.5.1 Proof of Theorem 8.1.2

As before define $d_\pm := \left(-x \pm \sigma_t^2 t/2\right)/\left(\sigma_t\sqrt{t}\right)$. We first consider the case $x > 0$. Note that $d_\pm$ tends to $-\infty$ as t tends to 0. Substituting the asymptotic series

$$1 - \mathcal{N}(z) = (2\pi)^{-1/2}\mathrm{e}^{-z^2/2}\left(z^{-1} - z^{-3} + \mathcal{O}\left(z^{-5}\right)\right), \quad \text{as } z \text{ tends to infinity,}$$

into the Black-Scholes call option formula with implied volatility σ_t, we obtain

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+ &= S_0\mathcal{N}\left(d_+\right) - S_0\mathrm{e}^x\mathcal{N}\left(d_-\right) = S_0\left(1 - \mathcal{N}\left(-d_+\right) - \mathrm{e}^x + \mathrm{e}^x\mathcal{N}\left(-d_-\right)\right)\\ &= \frac{S_0}{\sqrt{2\pi}}\exp\left(-\frac{d_+^2}{2}\right)\left(-\frac{1}{d_+} + \frac{1}{d_-} + \frac{1}{d_+^3} - \frac{1}{d_-^3} + \mathcal{O}\left(d_+^{-5}\right)\right),\end{aligned} \tag{8.5.1}$$

as t tends to zero, where we used the fact that $-d_-^2/2 + x = -d_+^2/2$. Now, expanding

$$\begin{aligned}\exp\left(-\frac{d_+^2}{2}\right) &= \exp\left(-\frac{x^2}{2\Sigma^2 t} + \frac{x\left(\Sigma^4 + cx\right)}{2\Sigma^4} - \frac{\Sigma^8 + 4x^2c^2}{\Sigma^6}\frac{t}{8} + \mathcal{O}\left(t^2\right)\right)\\ &= \exp\left(-\frac{x^2}{2\Sigma^2 t} + \frac{x(\Sigma^4 + cx)}{2\Sigma^4}\right)\left(1 - \frac{\Sigma^8 + 4x^2c^2}{\Sigma^6}\frac{t}{8} + \mathcal{O}\left(t^2\right)\right),\end{aligned}$$

and $d_-^{-1} - d_+^{-1} + d_+^{-3} - d_-^{-3} = x^{-2}\Sigma^3 t^{3/2} + \mathcal{O}\left(t^{5/2}\right)$, and plugging this expression into (8.5.1), we obtain the desired result. The proof of the case $x < 0$ is analogous.

When $x = 0$ note that z converges to zero as t tends to zero, so that we use the asymptotic series

$\mathcal{N}(z) = 1/2 + (2\pi)^{-1/2}\left(z - \frac{1}{6}z^3 + \mathcal{O}\left(z^5\right)\right)$, and we obtain

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+ &= S_0\mathcal{N}(d_+) - S_0\mathcal{N}(d_-) = \frac{S_0}{\sqrt{2\pi}}\left((d_+ - d_-) - \frac{1}{6}\left(d_+^3 - d_-^3\right) + \mathcal{O}\left(d_+^5\right)\right)\\ &= \frac{S_0}{\sqrt{2\pi}}\left(\Sigma t^{1/2} + \frac{12c - \Sigma^4}{24\Sigma}t^{3/2} + \mathcal{O}\left(t^{5/2}\right)\right),\end{aligned}$$

and the result follows.

8.5.2 Proof of Theorem 8.2.1

Let us first assume $x > 0$. If we equate the leading order and correction terms for the Heston and the Black-Scholes models as

$$\Lambda^*(x) = \Lambda^*_{\mathrm{BS}}(x) = \frac{x^2}{2\sigma_0^2(x)}, \quad \text{and} \quad A(x) = A_{\mathrm{BS}}(x, \sigma_0(x))\exp\left(\frac{1}{2}\frac{a(x)\,x^2}{\sigma_0^4(x)}\right), \tag{8.5.2}$$

we obtain (8.2.1). We now have to make this argument rigorous, since we do not know a priori that σ_t admits an expansion of the form stated in the theorem. However by Theorem 8.1.1 and (8.5.2), we know that for all $\varepsilon > 0$, there exists $t^*(\varepsilon)$ such that for all $t < t^*(\varepsilon)$ we have

$$\frac{\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+}{S_0} \le \frac{A(x)}{\sqrt{2\pi}}t^{3/2}\exp\left(-\frac{\Lambda^*(x)}{t}\right)\mathrm{e}^{\varepsilon} = \frac{A_{\mathrm{BS}}(x, \sigma_0(x))}{\sqrt{2\pi}}\exp\left(\frac{a(x)x^2}{2\sigma_0^4(x)}\right)t^{3/2}\exp\left(-\frac{x^2}{2\sigma_0^2(x)t}\right)\mathrm{e}^{\varepsilon}.$$

The function a is continuous on $(0, \infty)$. Therefore for any $\delta > 0$ sufficiently small we can choose $\varepsilon > 0$ such that the equality

$$\exp\left(\frac{1}{2}\frac{a(x)\,x^2}{\sigma_0^4(x)}\right)\mathrm{e}^{\varepsilon} = \exp\left(\frac{1}{2}\frac{(a(x) + \delta)\,x^2}{\sigma_0^4(x)}\right)\mathrm{e}^{-\varepsilon},$$

holds and hence Theorem 8.1.1 implies the inequalities

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{X_t} - S_0\mathrm{e}^x\right)_+ &\le \frac{A_{\mathrm{BS}}(x, \sigma_0(x))}{\sqrt{2\pi}}\exp\left(\frac{(a(x) + \delta)\,x^2}{2\sigma^4(x)}\right)t^{3/2}\exp\left(-\frac{x^2}{2\sigma_0^2\,(x)\,t}\right)\mathrm{e}^{-\varepsilon}\\ &\le C_{\mathrm{BS}}\left(x, t, \sqrt{\sigma_0^2(x) + (a(x) + \delta)\,t}\right)\end{aligned}$$

for t sufficiently small. Since the Black-Scholes formula is a strictly increasing function of the volatility then the upper bound $\sigma_t^2(x) \le \sigma_0^2(x) + a\,(x)\,t + \delta t$ holds for all $x > 0$. We proceed similarly for the lower bound and for the case $x < 0$, and the theorem follows.

8.5.3 Proof of Corollary 8.2.2

We wish to compute a series expansion for $p^*(x)$ defined as the unique solution to $x = \Lambda'(p^*(x))$. We substitute an ansatz power series expansion for $p^*(x)$ and then recursively we determine the coefficients such that the power series of the composition $\Lambda'(p^*(x))$ equals x. We only indicate below the important auxiliary quantities:

$$\begin{aligned}
\Lambda(p) &= \frac{1}{2}vp^2 + \frac{1}{4}v\rho\sigma p^3 + \frac{1}{24}v\sigma^2\left(1+2\rho^2\right)p^4 + \frac{1}{48}v\rho\sigma^3\left(2+\rho^2\right)p^5 + \mathcal{O}\left(p^6\right),\\
U(p) &= 1 - \frac{v}{2}p + \left(\kappa(\theta - v) - v\sigma\rho + \frac{v^2}{2}\right)\frac{p^2}{4}\\
&\quad + \left(\frac{\kappa\rho\sigma}{3}(\theta - 2v) - \frac{v\sigma^2}{3}\left(1+\rho^2\right) - \frac{v}{2}(\kappa(\theta - v) - \rho\sigma v) - \frac{v^3}{12}\right)\frac{p^3}{4} + \mathcal{O}\left(p^4\right),\\
\Lambda^*(x) &= \frac{x^2}{2v} - \frac{\rho\sigma}{4v^2}x^3 + \frac{\sigma^2}{96v^3}\left(19\rho^2 - 4\right)x^4 + \mathcal{O}\left(x^5\right),\\
\sigma_0(x) &= \sqrt{v} + \frac{\rho\sigma x}{4\sqrt{v}} + \left(1 - \frac{5}{2}\rho^2\right)\frac{\sigma^2 x^2}{24v^{3/2}} + \mathcal{O}\left(x^3\right),\\
A(x) &= \frac{v^{3/2}}{x^2} + \frac{\sqrt{v}}{4x}(3\rho\sigma + 2v) - \frac{1}{96\sqrt{v}}\left(11\rho^2\sigma^2 - 8\sigma^2 - 48v\rho\sigma - 24\kappa\theta + 24v\kappa - 12v^2\right) + \mathcal{O}(x).
\end{aligned}$$

Finally,

$$\begin{aligned}
\frac{A(x)}{A_{\mathrm{BS}}(x,\sigma_0(x))} &= 1 + \frac{\sigma^2\left(\rho^2 - 4\right) + 12v\rho\sigma + 24\kappa(\theta - v)}{96v^2}x^2\\
&\quad - \frac{\rho\sigma}{96v^3}\left(10v\rho\sigma + 3\rho^2\sigma^2 - 6\sigma^2 - 20v\kappa + 28\kappa\theta\right)x^3 + \mathcal{O}\left(x^4\right),
\end{aligned}$$

which implies the expression for $a(x) = x^{-2}2\sigma_0^4(x)\log\left(A(x)/A_{\mathrm{BS}}(x,\sigma_0(x))\right)$ in Corollary 8.2.2.

Discussion: What happens when models jump?

In Theorem 8.1.1—for the Heston model—we have proved the convergence rate of the call option to its intrinsic value $\mathbb{E}\left(S_0 - S_0 \mathrm{e}^x\right)_+$ for all real number x. We have seen that it is of order $\sqrt{t}$ for at-the-money options and of exponential order for in and out-of-the-money options. This is of fundamental importance since this is exactly the same rate as for the Black-Scholes model as is clear by Proposition 8.1.2. From this it follows that the behaviour of the implied volatility as the maturity tends to zero will be given by the rate of convergence of the option price to its intrinsic value at inception. It turns out that only the continuous part of the process is able to provide a non trivial smile. Indeed Tankov [100] proved, using an earlier result by Rüschendorf and Woerner [91], the following proposition.

Proposition 8.5.1. *(Proposition 4 in [100])*
Let X be a Lévy process with Lévy measure supported on the whole real line. Then the implied volatility satisfies the following limit:

$$\lim_{t\to 0} -\frac{2\sigma_t^2(x)\, t\log(t)}{x^2} = 1, \qquad \text{for all } x \neq 0.$$

This proposition clearly indicates that the presence of jumps implies a blow-up of the implied volatility as the maturity tends to zero. For the at-the-money case, Tankov also proves that the implied volatility converges to the diffusion component as the maturity tends to zero. More recently, Muhle-Karbe and Nutz [78] precisely characterised the rate of convergence of the at-the-money call option price to its intrinsic value for general martingales. If the process has a continuous component, then this rate of convergence is or order $\sqrt{t}$. Otherwise it is linear in t if the process has finite variation and between $\sqrt{t}$ and t if the process is of infinite variation. This seems to end the discussion on the small-time asymptotics of the implied volatility. However, two directions of research are currently under investigation:

(i) how can we extend these results to affine stochastic volatility models in which jumps can occur both in the process X and in the variance process?

(ii) How can we rescale the process in such a way that the implied volatility will not be trivial?

Case (ii) is of primary importance since jump processes are heavily used in practice and hence it is fundamental to understand how jumps can be calibrated in the short term. The idea underlying this is similar to the large-time results we have proved: it is well known that the implied volatility becomes flat as the maturity tends to infinity; however if one wants to match a large-time formula to observed data one has to scale the process—consider the process $(X_t/t)_{t\geq 0}$ instead of $(X_t)_{t\geq 0}$— to observe a full smile. In the small-time case one has to rescale the process as well. Recently, Rosenbaum & Tankov [90] have proved that certain Lévy processes (i.e. where the Lévy density is assumed to have a particular form) converge weakly to strictly α-stable Lévy processes when suitably rescaled. Work in this direction is under progress and we henceforth leave it for future research.

Conclusion

It is not obvious that "forecasting" events within an infinite time horizon agrees with Niels Bohr's warning in the introduction. However as we said the purpose of deriving formulae for infinite maturities is merely a tool to obtain information and initial calibration points to feed a minimisation algorithm. This should hopefully save us from Bohr's casting a spell on us from his grave.

The methodology we have developed in this thesis enabled us to derive closed-formed formulae for the implied volatility smile in affine stochastic volatility models. Since our analysis is solely based on the behaviour of the Laplace transform of the model such processes are natural to study in this framework. We have seen that jumps have an influence on the smile for large maturities and that they make the smile "explode" in the short-term. More research has to be carried out to fully understand the influence of the jumps for both short and large maturities. We have here and there provided some hints that hopefully grasp the flavours of the difficulties and the paths to overcome them. From the theoretical point of view properties of cumulant generating functions of infinitely divisible distributions have to be studied in greater detail and large deviations results have to be analysed more precisely when the origin is not in the interior of the domain of the limiting cumulant generating function. For short maturities we need to understand more precisely the rate of convergence of the process in order to obtain a non trivial (null or infinite) implied volatility smile. Finally the geometric approach has to be refined in order to handle more general classes of models.

On the practical side we have presented some numerical evidence about the usefulness of our formulae: even for "not too large or not too small maturities" the asymptotic smile is close to the real one, and hence these formulae can be used in practice as first approximations. We now need to carry out a more statistical and numerical research program in order to precisely assess the validity of such results.

Bibliography

[1] Y. Ait-Sahalia and J. Yu. Saddlepoint Approximations for Continuous-Time Markov Processes. *Journal of Econometrics*, 134, 507-551, 2006.

[2] H. Albrecher, P. Mayer, W. Schoutens and J. Tistaert. The Little Heston Trap. *Wilmott Magazine*, January issue, 83-92, 2007.

[3] L.B.G. Andersen and V.V. Piterbarg. Moment Explosions in Stochastic Volatility Models. *Finance and Stochastics*, 11 (1): 29-50, 2007.

[4] O. E. Barndorff-Nielsen, N. Nicolata and N. Shephard. Some recent developments in stochastic volatility modelling. *Quantitative Finance*, 2: 11-23, 2002.

[5] D. S. Bates. Jump and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9: 69-107, 1996.

[6] S. Benaim and P. Friz. Regular Variation and Smile Asymptotics. *Mathematical Finance*, 19 (1): 1-12, 2006.

[7] S. Benaim and P. Friz. Smile Asymptotics 2: Models with Known Moment Generating Functions. *Journal of Applied Probability*, 45 (1): 16-32, 2008.

[8] E. Benhamou, E. Gobet and M. Miri. Time dependent Heston model. *SIAM Journal on Financial Mathematics*, 1: 289-325, 2010.

[9] B. Bercu, F. Gamboa and M. Lavielle. Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM PS*, 4: 1-24, 2000.

[10] B. Bercu and A. Rouault. Sharp large deviations for the Ornstein-Uhlenbeck process. *SIAM Theory of Probability and its Applications*, 46: 1-19, 2002.

[11] H. Berestycki, J. Busca and I. Florent. Asymptotics and calibration of local volatility models. *Quantitative Finance*, 2: 61-69, 2002.

[12] H. Berestycki, J. Busca and I. Florent. Computing the Implied Volatility in Stochastic Volatility models. *Comm. Pure App. Math.*, 57: 1352-1373, 2004.

[13] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81 (3): 637-654, 1973.

[14] N. Bleistein and R.A. Handelsman. Asymptotic expansions of integrals. New York, London, Holt, Rinehart and Winston, 1975.

[15] H. Buehler. Volatility Markets: Consistent Modelling, Hedging and Practical Implementation. *PhD Dissertation*, Technischen Universität Berlin, 2006.

[16] R.W. Butler. Saddlepoint Approximations with Applications. Cambridge University Press, 2007.

[17] P. Carr and J. Crosby. A Class of Levy Process Models with almost exact calibration of both barrier and vanilla FX options. Forthcoming in *Quantitative Finance*, 2010.

[18] P. Carr and R.W. Lee. Put-call symmetry : Extensions and applications. *Mathematical Finance*, 19 (4): 523-560, 2009.

[19] P. Carr and D. Madan. Option valuation using the Fast Fourier transform. *Journal of Comp. Fin.*, 2 (4): 61-73, 1999.

[20] P. Carr and D. Madan. Saddlepoint Methods for Option Pricing. *Journal of Computational Finance*, 13 (1): 49-61, 2009.

[21] P. Carr and L. Wu. What Type of Process Underlies Options? A Simple Robust Test. *Journal of Finance*, 58 (6): 2581-2610, 2003.

[22] P. Carr and L. Wu. Stochastic Skew for Currency Options. *Journal of Fin. Economics*, 86: 213-247, 2007.

[23] R. Cont and P. Tankov. Financial modelling with jump processes. *Chapman & Hall* / CRC, Boca Raton, FL, 2004.

[24] J.C. Cox, J.E. Ingersoll and S.A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53: 385-407, 1995.

[25] J. Crosby. Practicalities of pricing exotic derivatives. *Lecture given at the Mathematical Institute, Oxford University*, 2009.

[26] S. Del Baño Rollin. Spot inversion in the Heston model. *Preprint available at* `www.crm.es/Publications/08/Pr837.pdf`, 2009.

[27] S. Del Baño Rollin, A. Ferreiro-Castilla and F. Utzet. A new look at the Heston characteristic function. *Preprint available at* `arxiv.org/abs/0902.2154`, 2009.

[28] A. Dembo and O. Zeitouni. Large deviations techniques and applications. Jones and Bartlet publishers, Boston, 1993.

[29] M. Do Carmo. Riemmanian Geometry. Birkhäuser, 1992.

[30] D. Duffie, D. Filipović and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13: 984-1053, 2003.

[31] D. Dufresne. The integrated square-root process. *Research Paper no. 90*, Centre for Actuarial Studies, University of Melbourne, 2001.

[32] B. Dupire. Pricing with a smile. *Risk*, 7: 18-20, 1994.

[33] V. Durrleman. From implied to spot volatilities. *PhD dissertation*, Princeton University, 2004.

[34] V. Durrleman. From implied to spot volatilities. *Finance and Stochastics*, 14 (2): 157-177, 2010.

[35] A. Elices. Models with time-dependent parameters using transform methods: application to Heston's model. *Preprint available at* arxiv.org/abs/0708.2020, 2008.

[36] J. Feng, M. Forde and J.P. Fouque. Short maturity asymptotics for a fast mean-reverting Heston stochastic volatility model. *SIAM Journal on Financial Mathematics*, 1: 126-141, 2010.

[37] J. Figueroa-López. Small-time moment asymptotics for Lévy processes. *Statistics and Probability Letters*, 78: 3355-3365, 2008.

[38] D. Florens-Landais and H. Pham. Large deviations in estimation of an Ornstein-Uhlenbeck model. *Journal of Applied Probability* 36: 60-77, 1999.

[39] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the Heston model. *International Journal of Theoretical and Applied Finance*, 12 (6): 861-876, 2009.

[40] M. Forde and A. Jacquier. The large-maturity smile for the Heston model. Forthcoming in *Finance & Stochastics*, 2010.

[41] M. Forde and A. Jacquier. Robust approximations for pricing Asian options and volatility swaps under stochastic volatility. *Applied Mathematical Finance*, 17 (3): 241-259, 2010.

[42] M. Forde, A. Jacquier and R. Lee. Small-time asymptotics for implied volatility under the Heston model: Part 2. *Preprint available at* www2.imperial.ac.uk/~ajacquie, 2010.

[43] M. Forde, A. Jacquier and A. Mijatović. Asymptotic formulae for implied volatility in the Heston model. Forthcoming in *Proceedings of the Royal Society A*, 2010.

[44] J.P. Fouque, G. Papanicolaou and R.K. Sircar. Derivatives in financial markets with stochastic volatility. Cambridge University Press, Cambridge, 2000.

[45] M.I. Freidlin and A. Wentzell. Random perturbations of dynamical systems, Second Edition. Springer-Verlag, New York, 1998.

[46] P. Friz, S. Gerhold, A. Gulisashvili, S. Sturm. On refined volatility smile expansion in the Heston model. *Preprint available at* `arxiv.org/abs/1001.3003`, 2010.

[47] W. Gander and W. Gautschi. Adaptive quadrature revisited. *Bit Numerical Mathematics*, 40 (1): 84-101, 2000.

[48] J. Gatheral. A parsimonious arbitrage-free implied volatility parameterisation with application to the valuation of volatility derivatives. *Global Derivatives & Risk Management*, 2004.

[49] J. Gatheral. The Volatility Surface: A Practitioner's Guide. John Wiley & Sons, Hoboken, NJ, 2006.

[50] J. Gatheral, E. Hsu, P. Laurence, C. Ouyang and T-H. Wang. Asymptotics of implied volatility in local volatility models. *Preprint available at* `papers.ssrn.com/sol3/papers.cfm?abstract_id=1542077`, 2009.

[51] J. Gatheral and A. Jacquier. Convergence of Heston to SVI. *Preprint available at* `arxiv.org/abs/1002.3633`, 2010.

[52] P. Glasserman and K. Kim. Gamma Expansion of the Heston Stochastic Volatility Model. Forthcoming in *Finance & Stochastics*, 2009.

[53] A. Going-Jaeschke and M. Yor. A survey and some generalisations of Bessel processes. *Bernoulli*, 9 (2): 313-349, 2003.

[54] A. Gulisashvili. Asymptotic Formulas with Error Estimates for Call Pricing Functions and the Implied Volatility at Extreme Strikes. *Preprint available at* `arxiv1.library.cornell.edu/abs/0906.0394`, 2009.

[55] A. Gulisashvili and E. Stein, Asymptotic Behavior of the Stock Price Distribution Density and Implied Volatility in Stochastic Volatility Models. *Preprint available at* `arxiv.org/abs/0906.0392`, 2009.

[56] P. Hagan, D. Kumar, A.S. Lesniewski and D.E. Woodward. Managing Smile Risk. *Wilmott Magazine*, September issue, 84-108, 2002.

[57] P. Henry-Labordère. A general asymptotic implied volatility for stochastic volatility models. *Preprint available at* `arxiv.org/abs/cond-mat/0504317`, 2005.

[58] P. Henry-Labordère. Analysis, Geometry and Modeling in Finance: Advanced Methods in Option Pricing. Chapman & Hall/CRC, Financial Mathematics Series, 2008.

[59] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6: 237-343, 1993.

[60] S.L. Heston and S. Nandi. A closed-form GARCH option valuation model. *Review of Financial Studies*, 13: 585-625, 2000.

[61] O. Islah. Solving SABR in exact form and unifying it with LIBOR market model. *Preprint available at* papers.ssrn.com/sol3/papers.cfm?abstract_id=1489428, 2009.

[62] P. Jaeckel and C. Kahl. Not-so-complex logarithms in the Heston model. *Wilmott Magazine*, 19: 94-103, 2005.

[63] J. Jost. Riemannian Geometry and Geometric Analysis. Springer, 2008.

[64] B. Jourdain. Loss of martingality in asset price models with lognormal stochastic volatility. *Preprint CERMICS 2004-267* : results previously obtained by C.A. Sin, Complications with stochastic volatility models. *Adv. in Appl. Probab.*, 30 (1998) (1): 56–268, 2005.

[65] C. Kahl and R. Lord. Complex Logarithms in Heston-Like Models. Forthcoming in *Mathematical Finance*, 2008.

[66] D. Kainth. Pricing FX options under multi-factor Heston. *Presentation given at the Global Derivatives conference, Paris*, 2007.

[67] I. Karatzas and S. Shreve. Brownian motion and Stochastic Calculus. Springer-Verlag, 1991.

[68] M. Keller-Ressel. Moment Explosions and Long-Term Behavior of Affine Stochastic Volatility Models. Forthcoming in *Mathematical Finance*, 2009.

[69] F. Kilin. Accelerating the calibration of stochastic volatility models. *Preprint available at* papers.ssrn.com/sol3/papers.cfm?abstract_id=965248, 2007.

[70] T. Kluge. Pricing Derivatives in Stochastic Volatility Models using the Finite Difference Method. Diploma thesis, Technical University, Chemnitz, *available at* archiv.tu-chemnitz.de/pub/2003/0008/index.html, 2002.

[71] R. W. Lee. Option Pricing by Transform Methods: Extensions, Unification, and Error Control. *Journal of Computational Finance*, 7 (3): 51-86, 2004.

[72] R. W. Lee. The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance*, 14: 469-480, 2004.

[73] A. Lewis. Option valuation under stochastic volatility. Finance Press, 2000.

[74] A. Lewis. Geometries and smile asymptotics for a class of Stochastic Volatility models, *Slides available at* `www.optioncity.net`, 2007.

[75] V. Gorovoi and V. Linetsky. Black's Model of Interest Rates as Options, Eigenfunction Expansions and Japanese Interest Rates. *Mathematical Finance*, 14: 49-78, 2004.

[76] R. Lugannani and S.O. Rice. Saddlepoint approximations for the distribution of the sum of independent random variables. *Advances in Applied Probability* 12: 475-490, 1980.

[77] E. Lukacs. Characteristic functions, second edition. Griffin, London, 1970.

[78] J. Muhle-Karbe and M. Nutz. Small-Time Asymptotics of Option Prices and First Absolute Moments. *Preprint available at* `arxiv.org/abs/1006.2294`, 2010.

[79] A. Neuberger. Volatility Trading. *London Business School Working Paper*, 1992.

[80] E. Nicolato and E. Venardos. Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type. *Mathematical Finance*, 13 (4): 445-466, 2003.

[81] F.W. Olver. Asymptotics and Special Functions. Academic Press, 1974.

[82] L. Paulot. Asymptotic Implied Volatility at the Second Order With Application to the SABR Model. *Preprint available at* `papers.ssrn.com/sol3/papers.cfm?abstract_id=1413649`, 2009.

[83] H. Pham. Grandes déviations en finance. *Lecture notes.* `proba.jussieu.fr/~pham`, 2010.

[84] E. Renault and N. Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, 6 (3): 279-302, 1996.

[85] R.T. Rockafellar. Convex Analysis. Princeton University Press, 1970.

[86] L.C.G. Rogers and M. Tehranchi. Can the implied volatility surface move by parallel shifts?. *Finance and Stochastics*, 14 (2): 235-248, 2010.

[87] L.C.G. Rogers and O. Zane. Saddlepoint approximations to option prices. *Annals of Applied Probability*, 9: 493-503, 1999.

[88] M. Roper. Implied volatility explosions: European calls and implied volatilities close to expiry in exponential Levy models. *Preprint available at* `arxiv.org/abs/0809.3305`, 2008.

[89] M. Roper and M. Rutkowski. A note on the behaviour of the Black-Scholes implied volatility close to expiry. Forthcoming in *International Journal of Theoretical and Applied Finance*, 2009.

[90] M. Rosenbaum and P. Tankov. Asymptotic results and statistical procedures for time-changed Lévy processes sampled at hitting times. *Preprint available at* `arxiv4.library.cornell.edu/abs/1007.1414`, 2010.

[91] L. Rüschendorf and J.H. Woerner. Expansion of transition distributions of Lévy processes in small time. *Bernoulli*, 8: 81-96, 2002.

[92] K.I. Sato. Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, 1999.

[93] R. Schöbel and J. Zhu. Stochastic volatility with an Ornstein-Uhlenbeck process: an extension. *European Finance Review*, 3: 23-46, 1999.

[94] W. Schoutens. Lévy Processes in Finance: Pricing Financial Derivatives. *Wiley*, 2003.

[95] A. Sepp. Pricing Options on Realized Variance in the Heston Model with Jumps in Returns and Variance. *Journal of Computational Finance*, 11 (4): 33-70, 2008.

[96] H. Shirakawa. Squared bessel processes and their applications for the square root interest rate model. *Financial Engineering and the Japanese Markets*, 9 (3-4): 169-190, 2002.

[97] S. Shreve. Stochastic Calculus for Finance, Volume II. Springer-Verlag, New-York, 2004.

[98] E.M. Stein and R. Sharkarchi. Complex Analysis. Princeton University Press, 2003.

[99] E. Stein and J. Stein. Stock-price distributions with stochastic volatility - an analytic approach. *Review of Financial studies*, 4: 727-752, 1991.

[100] P. Tankov. Pricing and hedging in exponential Lévy models: review of recent results. *Paris-Princeton Lecture Notes in Mathematical Finance*, Springer, 2010.

[101] M. Tehranchi. Asymptotics of implied volatility far from maturity. *Journal of Applied Probability*, 46 (3): 629-650, 2009.

[102] S.R.S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Communications in Pure and Applied Mathematics*, 20: 431-455, 1967.

[103] S.R.S. Varadhan. Diffusion processes in a small time interval. *Communications in Pure and Applied Mathematics*, 20: 659-685, 1967.

[104] M. Zani. Grandes déviations pour des fonctionnelles issues de la statistique des processus. *PhD Dissertation* Orsay, `perso-math.univ-mlv.fr/users/zani.marguerite/THE.ps`, 2000.