Scuola Normale Superiore

Tesi di Perfezionamento in Matematica per la Finanza

Thèse présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité: Mathématiques

par Stefano De Marco

# of Diffusions and Financial Models with non-globally smooth coefficients

dirigée par Vlad Bally et Maurizio Pratelli

Rapporteurs: Emmanuel Gobet et Arturo Kohatsu-Higa.

Soutenue le 23 novembre 2010 devant le jury composé de:

Vlad Bally	Directeur de thèse
Emmanuel Gobet	Rapporteur
Giorgio Letta	Examinateur
Stefano Marmi	Examinateur
Claude Martini	Examinateur
Maurizio Pratelli	Directeur de thèse

Acknowledgments

The achievement of this thesis would not have been possible without the guidance and constant advice of my supervisors, Prof. Vlad Bally and Prof. Maurizio Pratelli. To them goes my deep and sincere gratitude, for the time and knowledge (and patience) they have shared with me.

I am grateful to Prof. Emmanuel Gobet and Prof. Arturo Kohatsu-Higa, who accepted to report on this thesis. Their capability of understanding and summarizing the sense and motivation of the problems I have tackled, together with their precious remarks and suggestions, have been an important help. They allowed me to correct and clarify the presentation of the results, and to look for new directions of research and development.

I am thankful to Prof. Giorgio Letta and Prof. Stefano Marmi, for accepting to take part in my Ph.D defense: I am honoured by their presence. I would like to express the same gratitude to Claude Martini. His presence at the concluding moment of my doctoral work is particularly significant to me.

My work has surely benefited from the friendly and stimulating atmosphere at Scuola Normale Superiore and at the Laboratoire d'Analyse Mathématiques Appliquées at Université Paris-Est Marne-la-Vallé. I have been kindly welcomed by the latter since the year of my Master in Mathématiques et Applications (which was a great experience. I have to thank Prof. Damien Lamberton for his amazing teaching skills in stochastic calculus - to the end of my Ph.D. The same environment I found wiithin the C.E.R.M.I.C.S. research group, who hosted me during the months preceding my defense. I will have the opportunity to benefit from their experience, and to develop a fruitful collaboration with them in the framework of my ongoing post-doc project.

I shall not forget how much I owe to Zeliade System's team: from their knowledge of mathematical finance and information technology, to their kindness and the friendly and energizing atmosphere they always contribute to create. For their ability and patience in making the basis of quantitative financial modeling more clear to me, and for helping me with Python and C# (and I did need help with that), I say a great thank you to all of them. I shared with them some very pleasant moments. Thank you Steve for the never-lacking coffee!

I shall always be truly grateful to my parents, for their ever-present care and encouragement. Among others, my interest for sciences would perhaps have never come to light if it had not been for them. I am thankful to my brother, and to the rest of my family (one of the best I could wish for!), for being always there.

My days in France have been so much bright since when I met Lola. I would like to say a big thank you also to her family, who gave me such a warm welcome.

During the preparation of this thesis in Italy and France, I had the great luck to meet and be surrounded by so many kind and smart people. I say thank you to all the friends in

Padova, to the perfezionandi of Scuola Normale, to the Ph.D candidates of L.A.M.A., to the Italian friends in Paris, to the French friends.

Apart from the Scuola Normale Superiore, my work in France has been supported by the following institutions: the **Fondazione "Ing. Aldo Gini"** with a scholarship for foreign studies in 2007, the **Università Italo-Francese - Université Franco-Italienne** in the framework of the Vinci project in 2008, and the **European Science Foundation** in the framework of the AMaMeF research activity in 2009. Their support is gratefully acknowledged.

Abstract

In this Ph.D. dissertation we deal with some issues of regularity and estimation of probability laws for diffusions with non-globally smooth coefficients, with particular focus on financial models.

The analysis of probability laws for the solutions of Stochastic Differential Equations (SDEs) driven by the Brownian motion is among the main applications of the Malliavin calculus on the Wiener space: typical issues involve the existence and smoothness of a density, and the study of the asymptotic behaviour of the distribution's tails. The classical results in this area are stated assuming global regularity conditions on the coefficients of the SDE: an assumption which fails to be fulfilled by several financial models, whose coefficients involve square-root or other non-Lipschitz continuous functions. Then, in the first part of this thesis (chapters 2, 3 and 4) we study the existence, smoothness and space asymptotics of densities when only local conditions on the coefficients of the SDE are considered. Our analysis is based on Malliavin calculus tools and on tube estimates for Itô processes, namely estimates on the probability that an Itô process remains around a deterministic curve up to a given time. We give applications of our results to general classes of option pricing models, including generalisations of CIR and CEV processes and some Local Stochastic Volatility models. In the latter case, the estimates we derive on the law of the underlying price have an impact on moment explosion and, consequently, on the large-strike asymptotic behaviour of the implied volatility.

Implied volatility modeling, in its turn, makes the object of the second part of this thesis (chapters 5 and 6). We deal with some issues related to the problem of an efficient and economical parametric modeling of the volatility surface. We focus on J. Gatheral's SVI model, first tackling the problem of its calibration to the market smile. We propose an effective quasi-explicit calibration procedure and display its performances on financial data. Then, we analyse the capability of SVI to generate efficient time-dependent approximations of symmetric smiles in general continuous models, building an explicit time-dependent parameterization. We provide and test the numerical application to the uncorrelated Heston model (without and with displacement), generating semi-closed expressions for the smile.

Keywords: SDEs, Smoothness of densities, Local regularity, Tail asymptotics, Malliavin calculus, Tube estimates for Itô processes, Law of the Stock price, Implied Volatility, SVI, Heston, Calibration.

Sommario

In questa tesi di perfezionamento, trattiamo dei problemi di regolarità e di stima di leggi di probabilità per le diffusioni a coefficienti non globalmente regolari, con una particolare attenzione per le applicazioni ai modelli finanziari.
Lo studio delle distribuzioni delle soluzioni di Equazioni Differenziali Stocastiche dirette dal moto Browniano è una dei principali settori di applicazione del calcolo di Malliavin sullo spazio di Wiener: delle problematique tipiche in quest'area sono l'esistenza e la regolarità della densità e il comportamento asintotico delle code della distribuzione. I risultati classici in questo settore sono formulati sotto condizioni di regolarità globale sui coefficienti dell'equazione, un'ipotesi che risulta violata nel caso di numerosi modelli finanziari, i cui coefficienti fanno intervenire delle radici quadrate o altre funzioni non globalmente lipschitziane. Di conseguenza, nella prima parte di questa tesi (capitoli 2, 3 e 4) studiamo l'esistenza, la regolarità e il comportamento asintotico spaziale delle densità nel caso in cui si assumano solo delle condizioni locali sui coefficienti dell'equazione. La nostra analisi è basata sugli strumenti del calcolo di Malliavin e su delle stime per i processi di Itô confinati a restare attorno ad una curva deterministica ("*tube estimates*"). Forniamo applicazione di questi risultati a delle classi generali di modelli per la valutazione delle opzioni, includendo delle estensioni dei processi CIR e CEV e dei modelli a volatilità locale-stocastica (LSV). Per questi ultimi, le stime che otteniamo hanno un impatto sull'esplosione dei momenti dell'attivo sottostante, e quindi sul comportamento asintotico in *strike* della volatilità implicita.
La modellizzazione della volatilità implicita, a sua volta, constituisce l'oggetto della seconda parte della tesi (capitoli 5 e 6), nella quale affrontiamo alcune questioni legate alla costruzione di una parametrizzazione economica ed efficiente della superficie di volatilità. Consideriamo in particolare il modello SVI di J. Gatheral, per il quale proponiamo una nuova strategia di calibrazione semi-esplicita, illustrandone le prestazioni su dei dati di mercato. Quindi, analizziamo la capacità del modello SVI di generare delle approssimazioni parametriche per gli *smiles* simmetrici, estendendolo ad un modello a coefficienti dipendenti dal tempo. In particolare, ne formuliamo e implementiamo l'applicazione numerica ad un modello di Heston (senza e con *displacement*), generando delle approssimazioni semi-esplicite dello smile di volatilità.

Parole chiave: Equazioni differenziali stocastiche, Regolarità della densità, Code della distribuzione, Calcolo di Malliavin, *Tube estimates* per i processi di Itô, Distribuzione del prezzo del sottostante, Volatilità implicita, SVI, Heston, Calibrazione.

Résumé

Dans cette thèse, nous traitons des problèmes de régularité et d'estimation de lois pour des diffusions avec coefficients non globalement réguliers, avec une attention particulière pour les modèles financiers.

L'étude des lois des solutions d'Equations Différentielles Stochastiques (EDS) dirigées par le mouvement Brownien est un des principaux secteurs d'application du calcul de Malliavin sur l'espace de Wiener : des problématiques typiques dans ce domaine concernent l'existence et la régularité d'une densité et l'étude du comportement asymptotique des queues de la distribution. Les résultats classiques sur ce sujet requièrent des conditions de régularité globale sur les coefficients de l'EDS, une condition qui n'est pas satisfaite par plusieurs modèles financiers, dont les coefficients font intervenir des racines carrées ou d'autre fonctions non globalement lipschitziennes. Par conséquent, dans la première partie de cette thèse (chapitres 2, 3 et 4), nous étudions l'existence, la régularité et l'asymptotique en espace de densités lorsqu'on n'impose que des conditions locales sur les coefficients de l'EDS. Notre analyse dans cette partie se base sur les outils du calcul de Malliavin et sur des estimations pour les processus d'Ito confinés dans un tube autour d'une courbe déterministe ("*tube estimates*"). Nous appliquons ces résultats à des classes générales de modèles pour l'évaluation d'options, comprenant des généralisations des processus CIR et CEV et des modèles à volatilité locale-stochastique (LSV). Dans ce deuxième cas, les estimations que nous obtenons pour la loi du sous-jacent entraînent l'explosion des moments et ont ainsi un impact sur le comportement asymptotique en strike de la volatilité implicite.

La modélisation de la volatilité implicite, à son tour, fait l'objet de la deuxième partie de cette thèse (chapitres 5 et 6), où nous abordons des questions liées au problème d'une modélisation paramétrique efficace et économique de la surface de volatilité. Nous considérons en particulier le modèle SVI de J. Gatheral, pour lequel nous proposons une nouvelle stratégie de calibration quasi-explicite, dont nous illustrons les performances sur des données de marché. Ensuite, nous analysons la capacité du SVI à générer de bonnes approximations paramétriques pour les smiles symétriques, en le généralisant à un modèle dépendant du temps. Nous en formulons et testons l'application à un modèle de Heston (sans et avec déplacement), en générant des approximations semi-fermées du smile de volatilité.

Mots clé : Equations différentielles stochastiques, Régularité des densités, Asymptotiques des queues, Calcul de Malliavin, *Tube estimates* pour les processus d'Itô, loi du Stock, Volatilité implicite, SVI, Heston, Calibration.

Contents

Chapter 1

Introduction

1.1 Analysis of probability laws for SDEs

The smoothness and tail behaviour of probability laws for the solutions of Stochastic Differential Equations (SDEs) driven by the Brownian motion is one of the central topics in the study of the applications of the Malliavin calculus on the Wiener space (cf. [54], Chapter 2, and references therein). With the word "tail" we refer here to the behaviour of the complementary cumulative distribution function $\mathbb{P}(|X| > y)$ for large values of y, and eventually, if the law of X admits a continuous density p_X, to the behaviour of p_X at $\pm\infty$.

Let us be more precise: fix a $T > 0$ and consider the SDE

$$X_t^x = x + \int_0^t b(X_s^x)ds + \sum_{j=1}^d \int_0^t \sigma_j(X_s^x)dW_s^j, \quad t \leq T \tag{1.1.1}$$

where $x \in \mathbb{R}^m$; $b, \sigma_j : \mathbb{R}^m \to \mathbb{R}^m$ for every $j = 1, \dots, d$ and $W = ((W_t^1, \dots, W_t^d), t \leq T)$ is a standard d-dimensional Brownian motion. Let $(\sigma)_{ij} = \sigma_j^i$ and consider the following conditions

- $b, \sigma_j \in C_b^\infty(\mathbb{R}^m; \mathbb{R}^m)$ for all $j = 1, \dots, d$;
- there exists a $c > 0$ such that $\sigma\sigma^*(y) \geq c\, I_m \quad \forall y \in \mathbb{R}^m$,

where * stands for matrix transposition and I_m is the $m \times m$ identity matrix. Under these conditions, it is obvious that eq. (1.1.1) admits a unique strong solution for every initial condition x. When looking at the fixed-time marginal laws of this solution, we have the following classical result:

Theorem 1.1.1. *For every $x \in \mathbb{R}^m$ and every $t > 0$, the law of X_t^x admits a density $p_t(x, \cdot)$. p_t is infinitely differentiable with respect to each variable. Moreover, there exist strictly positive constants c_1, c_2 and functions $C_1, C_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

$$C_1(t)\exp\Big(-\frac{|y-x|^2}{c_1 t}\Big) \le p_t(x,y) \le C_2(t)\exp\Big(-\frac{|y-x|^2}{c_2 t}\Big), \tag{1.1.2}$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^m$.

In a word, X_t^x admits a smooth density and this density can be "sandwiched" between two Gaussian bounds as in (1.1.2). This density is smooth with respect to both the state variable y and the initial condition x: in the sequel, we will be interested in particular in the dependence (smoothness and bounds) with respect to the state variable y, since this is what determines European Option prices when X^x models the price (or the log-returns) of a financial security. Theorem 1.1.1 can be proven relying either on analytical methods derived the from classical PDE's theory (cf. for example [3]) or on probabilistic tools, involving Malliavin's stochastic calculus of variations.
Let us remark that a complete discussion of the subject introduced here would require several extensions and comments. First of all, let us mention that an upper bound analogous to the one in (1.1.2) holds for the partial derivatives of p_t with respect to y: namely, for every $k \in \mathbb{N}$ there exist positive functions $C_k : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\big|\frac{\partial^\alpha p_t(x,y)}{\partial y_{\alpha_1}\cdots\partial y_{\alpha_k}}\big| \le C_k(t)\exp\Big(-\frac{|y-x|^2}{c_2 t}\Big)$, where $\alpha \in \{1,\dots,m\}^k$ denotes a multi-index of lenght k. Secondly, although the conditions of regularity and uniform ellipticity on the coefficients b and σ are quite demanding, these hypotheses can be relaxed in many directions. The uniform ellipticity condition can be replaced by a much weaker assumption, the well know Hörmander's condition, allowing for degenerate diffusion matrix $\sigma\sigma^*$ (cf. the groundbreaking article by Malliavin [51] and Kusuoka and Stroock's subsequent work [43], [44]). Since this condition involves the partial derivatives of any order of the coefficients, the C^∞ condition must hold when Hörmander's condition is assumed; moreover, the results in this framework are typically stated for globally Lipschitz-continuous b and σ. On the other hand, PDE methods apply for much less regular coefficients, allowing to prove Theorem 1.1.1 assuming that b and σ are just bounded measurable functions. Nevertheless, in this case the uniform ellipticity condition must hold (cf. again [3]). Last but not least, with PDE methods it is more difficult, in general, to explicitly identify the constants appearing in estimate (1.1.2).

Summing up, even if other partial refinements are possible, few results are available for the case when *all* the global conditions (smoothness, boundedness, Lipschitz-continuity, ellipticity) are dropped at once. To our knowledge, the main result in this direction is provided by Kusuoka and Stroock in [44] (see the discussion of their results at the end of §1.3.1). Before presenting our contribution on this topic, we would like to introduce one of the main motivations for this work, that is the existence of stochastic models for financial securities whose coefficients are not globally "well-behaved". This is a situation where tail asymptotics heavier than Gaussian are possible, or better still, as we are going to see, expressly sought.

1.2 A motivation for extensions: financial models

1.2.1 A framework of non globally well-behaved coefficients

Many popular models for option pricing rely on SDEs whose coefficients do not satisfy global regularity assumptions, in particular they are not globally Lipschitz-continuous or uniformly elliptic.

Let us recall the main examples of such models. The celebrated Cox-Ingersoll-Ross (CIR) process considers a SDE with square root diffusion coefficient:

$$dX_t = (a - bX_t)dt + \sigma\sqrt{X_t}dW_t \tag{1.2.1}$$

where $a \geq 0, b \in \mathbb{R}, \sigma > 0$. The mean-reverting character of the drift coefficient (for $b > 0$) allows to control the large-time value of the solution while the square root coefficient has the advantage of making the characteristic function of the process analytically tractable. The CIR process has been originally employed as a model for short interest rates [16] and subsequently by Heston [38] to model the stochastic variances of assets. Under the risk-neutral measure, the Heston model for the forward price F_t of an asset is given by the two-dimensional SDE

$$\begin{aligned} dF_t &= F_t\sqrt{V_t}dW_t \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dZ_t \\ d\langle W, Z\rangle_t &= \rho dt, \end{aligned} \tag{1.2.2}$$

where $F_0, V_0 > 0$; $\rho \in [-1, 1]$ is the correlation parameter and we have considered the typical reparameterisation of the drift of the instantaneous variance V introducing the factors $k, \theta > 0$. The Heston model shares the analytical properties of the CIR process (it is indeed an Affine Stochastic Volatility Model, cf. the definition in [41]) and is particularly tractable: in particular, it allows for semi-explicit solutions for the prices of Vanillas. This is indeed one of the most widely exploited models for option pricing and hedging in Equity and FX markets. The uncorrelated Heston model (obtained setting $\rho = 0$) will be the main example of application of some results on the implied volatility we present in Chapter 6. Yet in the framework of models for stock price and forward rate dynamics, the main example of a local volatility model is given by the Constant Elasticity of Variance (CEV) model [15], which can be seen as an extension of the Black-Scholes model to a two parameter family:

$$dF_t = \alpha F_t^\beta dW_t, \tag{1.2.3}$$

where F_t is again the forward price of an asset or the forward rate and the parameters $\alpha > 0$ and $\beta \in (0, 1)$ allow to tune respectively the level of the volatility and the asymmetry of option prices with respect to the at-the-money strike price $K = F_0$ (the so-called "skewness" of the implied volatility smile, see below). In this example, the "singularity" of the model is once again given by the non globally Lipschitz-continuous (and not uniformly elliptic) diffusion coefficient f^β, $\beta < 1$. The CEV model has a straightforward generalisation to a stochastic volatility model originally proposed by Hagan, Kumar, Lesniewski and Woodward in [34] and known as the SABR (Stochastic-Alpha Beta Rho) model:

$$\begin{aligned} dF_t &= \sigma_t F_t^\beta dW_t \\ d\sigma_t &= \nu \sigma_t dZ_t, \quad \sigma_0 = \alpha \\ d\langle W, Z\rangle_t &= \rho dt, \end{aligned} \tag{1.2.4}$$

where the log-normal process σ_t with initial value α replaces the constant volatility and the additional parameters ν and ρ allow for a finer tuning of option prices.

It is important to recall, of course, that the key for the popularity of these models, which is their analytical tractability, basically amounts to the fact that they allow for a complete or almost complete knowledge of the law of the underlying process. Let us be more specific: for the CIR and CEV processes, the fixed-time density is known explicitly, eventually under the form of the sum of a series. Precisely, if X is a CIR process (1.2.1) starting at x, the density of X_t on the open positive real line $(0, \infty)$ is given by (cf. for exemple [45]):

$$\begin{aligned} p_t^{CIR}(y) &= \frac{e^{bt}}{2c_t}\Big(\frac{ye^{bt}}{x}\Big)^{\nu/2} \exp\Big(-\frac{x + ye^{bt}}{2c_t}\Big) I_\nu\Big(\frac{1}{c_t}\sqrt{xye^{bt}}\Big) \\ &= \frac{e^{bt(\nu+1)}}{(2c_t)^{\nu+1}}\, y^\nu \exp\Big(-\frac{x + ye^{bt}}{2c_t}\Big) \sum_{n=0}^\infty \Big(\frac{x}{c_t}\Big)^n \frac{\Big(\frac{ye^{bt}}{4c_t}\Big)^n}{n!\Gamma(\nu+n+1)}, \quad y > 0, \end{aligned} \tag{1.2.5}$$

where $c_t = (e^{bt} - 1)\sigma^2/4b$, $\nu = 2a/\sigma^2 - 1$ and I_ν denotes the modified Bessel function of index ν, and in the second line we have expanded I_ν with its classical series representation. Using the asymptotic property $I_\nu(z) \sim \frac{1}{\sqrt{2\pi z}} e^z$ for large z, we can observe that this density has exponential (but not gaussian) decay at infinity. Also, we incidentally remark that p_t^{CIR} is bounded or unbounded in the neighborhood of zero according to the value of ν. The density for the CEV process (1.2.3) starting at F_0, as provided in e.g. Davydov and Linetsky [18], is

$$p_t^{CEV}(y) = F_0^{1/2} \frac{y^{-2\beta+1/2}}{\alpha^2(1-\beta)t} \exp\Big(-\frac{F_0^{2(1-\beta)} + y^{2(1-\beta)}}{2\alpha^2(1-\beta)^2 t}\Big) I_\nu\Big(\frac{F_0^{1-\beta} y^{1-\beta}}{\alpha^2(1-\beta)^2 t}\Big), \quad y > 0 \tag{1.2.6}$$

now with $\nu = \frac{1}{2(1-\beta)}$ (recall that we are considering $\beta \in (0,1)$). Using again the asymptotic properties of I_ν, it is possible to see (as observed in [23], Theorem 1.6) that $\log p_t(y)$ is asymptotic to $-\frac{y^{2(1-\beta)}}{2\alpha^2(1-\beta)^2 t}$, for large values of y.

For the two-dimensional models (1.2.2) and (1.2.4), the density of the underlying price F_t is not explicitly accessible. For the Heston model (1.2.2), the affine property allows to compute the characteristic function $\mathbb{E}[e^{iuX_t}]$ of the log price $X_t = \log(F_t/F_0)$. This is obviously sufficient to characterise the law of X_t (hence of F_t) and to evaluate European options via Fourier inversion methods (cf. e.g. [14]), nevertheless no closed-form expression of the density is provided in the literature. It is still possible to exploit the knowledge of the characteristic function in order to show that the law of X_t actually admits a smooth density (cf. [28], Remark 10) and to derive sharp estimates of this density at $\pm\infty$ (cf. [21] and again [28]). For the SABR model (1.2.4), some approximations of the density of F_t close to F_0 can be obtained relying on perturbation techniques for PDEs as in [35].

If the structure of the model coefficients is slightly changed (e.g. the affine structure of the Heston model is broken), explicit computation are no longer possible. This is the case if, for example, a skew function is added to a stochastic volatility model, yielding a so called local-stochastic volatility model as in [37] (this kind of situation will be our main concern in Chapter 3). The methods based on the knowledge of the characteristic function or in general on affine principles break down, hence accessing the law of the process becomes a more difficult challenge. As addressed above, this is a situation where the classical methods for the evaluation of the density based on Malliavin calculus techniques do not apply, because of the irregularity of the coefficients of the SDE.

1.2.2 Connections between tail asymptotics and Implied Volatility

In this section we present some non-trivial connections between the law of the underlying price and one of the main financial quantities popping up in the context of derivative pricing. It is a generic assertion that the tail behaviour of the law of the underlying price affects the prices of European options: this relation can be made much more precise, in particular when switching from option prices to another remarkable financial variable - the Black-Scholes implied volatility.

Let us briefly recall the definition and properties of the implied volatility: let $C(k,T)$ denote the price of a European call option maturing at T and with strike price F_0e^k, then the implied volatility of the option is the non-negative solution $\sigma(T,k)$ to the equation

$$C(k,T) = C_{BS}(k,T,\sigma(T,k)) \tag{1.2.7}$$

where $C_{BS}(k,T,\sigma)$ is the price of a Black-Scholes call option of strike F_0e^k, maturity T and volatility σ. The Black-Scholes price being strictly monotone with respect to σ, $\sigma(T,k)$ is well defined and unique. The model-implied volatility is obtained replacing the lhs in (1.2.7) by the option price yield by the model, $C(k,T) = \mathbb{E}[(F_0e^{X_T} - F_0e^k)^+]$. The market-implied volatility $\sigma_{market}(T,k)$ is obtained in the same way using market option prices for different values of k and T. The function $k \mapsto \sigma(T,k)$ (and sometimes the whole surface $(k,T) \mapsto \sigma(T,k)$) is referred to as to the smile. Unless differently specified, in the sequel by implied volatility we will mean the model-implied one.

It is clear that the definition of the implied volatility in a given model only requires the price of the option to be an arbitrage price. While it can be checked that in the aforementioned models the risk-neutral price F of the underlying is an integrable martingale, it can happen that some moment of F_T of order $p > 1$ becomes infinite. This phenomenon actually does appear in several stochastic volatility models, such as the Heston model (1.2.2), and has been attentively studied by several authors (cf. e.g. Lions and Musiela [49], Andersen and Piterbarg [2] and Keller-Ressel [41]). If on the one hand the lack of moment stability calls for additional care (if, for example, one wants to manipulate variances), on the other hand the explosion of moments plays a crucial role in determining the asymptotic properties of the smile for large values of $|k|$. In order to make this dependence precise, let the critical exponents $p^*_T(F)$ and $q^*_T(F)$ of F_T be given by

$$p^*_T(F) = \sup\{p \geq 1 : \mathbb{E}[F_T^p] < \infty\}, \qquad q^*_T(F) = \sup\{q \geq 0 : \mathbb{E}[F_T^{-q}] < \infty\}.$$

Lee [47] has proven, in a complete model-free manner, a formula relying the critical exponents to the asymptotic slopes of the implied variance (the square of implied volatility), which reads:

$$\limsup_{k\to\infty} \frac{T\sigma(T,k)^2}{k} = \varphi(p^*_T(F) - 1), \qquad \limsup_{k\to\infty} \frac{T\sigma(T,-k)^2}{k} = \varphi(q^*_T(F)), \tag{1.2.8}$$

where $\varphi(x) := 2 - 4(\sqrt{x^2+x} - x), \varphi(\infty) := 0$. The relations in (1.2.8) are known as the "right" and "left" moment formula. As pointed out by Lee, Eq. (1.2.8) is useful for *model selection* purposes: since the market-implied variance smiles usually display "wings" (i.e. $k \to \sigma_{market}(T,k)^2$ has left and right asymptotes), so has to do the model-implied volatility, hence the exponential moments of the underlying process F *must* explode (otherwise $p^*_T(F) = q^*_T(F) = \infty$ and (1.2.8) states that the implied volatility is flat for large values of $|k|$). Pushing things further, the moment formula can help *model calibration*: the values of the slopes at the left hand sides of (1.2.8) can be observed from market data for values of $|k|$ large enough; if on the other hand the critical exponents are known functions of the model parameters, the use of (1.2.8) can provide reasonable initial guesses of parameters values. Let us recall that Benaim and Friz [9] subsequently refined Lee's result, showing how the asymptotics of the distribution function of log-returs can be directly translated to the asymptotics of the implied volatility smile. Their result reads as follows: let $P_T(y) := \mathbb{P}(X_T \leq y)$ be the cumulative distribution function of $X_T = \log(F_T/F_0)$, then under a mild regularity condition on P_T (namely, the "log-tails" $-\log(1-P_T(y))$ and $-\log(P_T(-y))$ are regularly varying at $+\infty$),

$$\begin{aligned} &\text{if } \exists \epsilon > 0 : \mathbb{E}[F_T^{1+\epsilon}] < \infty, \quad \text{then} \quad \frac{T\sigma(T,k)^2}{k} \sim \varphi\Big(-\frac{\log(1-P_T(k))}{k} - 1\Big) \\ &\text{if } \exists \epsilon > 0 : \mathbb{E}[F_T^{-\epsilon}] < \infty, \quad \text{then} \quad \frac{T\sigma(T,-k)^2}{k} \sim \varphi\Big(-\frac{\log(P_T(-k))}{k}\Big) \end{aligned} \tag{1.2.9}$$

where $f(k) \sim g(k)$ means $\lim_{k\to\infty} f(k)/g(k) = 1$. The relations in (1.2.9) are known as the right and left "tail-wing" formula. As pointed out by the authors, the tail-wing formula contains the full asymptotics of the smile: other than providing conditions under which the lim sup in the moment formula is actually a true limit, the formula allows to identify the behaviour of sublinear smiles. Indeed, in models where $1 - P_T(y) \sim \exp(-cy^{1+\gamma})$ for some $\gamma > 0$ (the Black-Scholes model corresponding to $\gamma = 1, c = \frac{1}{2\sigma^2 T}$), all the positive moments of F_T are finite[1], hence $p_T^*(F) = \infty$ and $\varphi(p_T^*(F) - 1) = 0$ and the moment formula only tells that the implied variance is sublinear. Using the easily proven property $\varphi(x) \sim \frac{1}{2x}$, the tail-wing formula allows to find the explicit asymptotics $\sigma(T,k) \sim \frac{1}{2Tc}k^{1-\gamma}$ (which yields, trivially, $\sigma(T,k) \sim \sigma^2$ for the Black-Scholes model).
It is again a trivial statement that the properties of regularity of the law of the underlying determine the regularity of option prices as functions of the strike price. More precisely, it is easy to see that the density of F_T (when existing) is the second derivative of the option price with respect to the strike: it is sufficient to differentiate twice with respect to K the identity $C(T,K) = \int_0^\infty (y-K)^+ p_T(y)dy$ in order to obtain

$$\frac{\partial^2 C(T,K)}{\partial K^2} = p_T(K).$$

Then, the positivity and smoothness of the density come into play when studying the convexity and smoothness of option prices as functions of the strike price (last but not least, the previous relation allows to design "sanity checks" for option pricing formulae: any approximation formula for the price of a call should display a positive second derivative).

The results presented above largely motivate the interest of studying and enhancing the existing results on tail asymptotics for the solutions of SDEs, in particular for the classes of SDEs arising in financial models. From the mathematical point of view, the challenge is to settle appropriate techniques in order to apply powerful tools (Malliavin calculus, localization arguments, tube estimates for Itô processes) in this non-standard framework.

1.3 Outline of results : Part I

1.3.1 Upper and lower estimates under local assumptions

In the first part of this thesis, we focus on the problem of formulating a statement analogous to Theorem 1.1.1 when every global assumption on the coefficients of the SDE is dropped and replaced by the corresponding local one. In particular, we are interested in developing tools to study the smoothness of the law and the tail behaviour of the cumulative distribution

[1] If this statement does not appear obvious, the following argument can be used: for every $p > 0$, we have $\mathbb{E}[F_T^p]/F_0^p = \mathbb{E}[e^{pX_T}] \le 1 + \sum_{k\ge 0} \mathbb{E}[e^{pX_T} 1_{\{k<X_T\le k+1\}}] \le 1 + e^p \sum_{k\ge 0} e^{pk}\mathbb{P}(X_T > k)$ and the last series is seen to converge if $\mathbb{P}(X_T > k) = 1 - P_T(x) \sim \exp(-cx^{1+\gamma})$.

function (cdf) and the density which be effective when applied to the classes of equations presented in section 1.2.

In **Chapter 2** we state our main results on the smoothness of the density and we provide upper bounds. We start by placing ourselves in an abstract framework, in order to consider general local hypotheses: to do so, let us fix some $y_0 \in \mathbb{R}^m$ and $0 < R \leq 1$ which represent - roughly speaking - the point at which the density is evaluated and the size of the neighborhood where the local assumptions are given. Then, we consider the SDE (1.1.1) under the following assumptions on the coefficients:

(H1) (*local smoothness*) $b, \sigma_j \in C_b^\infty(B_{5R}(y_0); \mathbb{R}^m)$;

(H2) (*local ellipticity*) $\sigma\sigma^*(y) \geq c_{y_0,R}\, I_m$ for every $y \in B_{3R}(y_0)$, for some $0 < c_{y_0,R} < 1$,

where $B_{R'}(y_0)$ denotes the open ball $B_{R'}(y_0) = \{y \in \mathbb{R}^m : |y - y_0| < R'\}$. Moreover, we assume that existence of strong solutions holds for the couple (b, σ) and state our result for a strong solution to the equation. Dropping the dependence w.r.t to the initial condition x for the ease of notation, our main result reads as follows:

Theorem 1.3.1 (Theorem 2.2.4, Chapter 2)**.** *Assume* **(H1)** *and* **(H2)** *and let* $(X_t; t \in [0, T])$ *be a strong solution to (1.1.1). Then, for any initial condition* $x \in \mathbb{R}^m$ *and any* $0 < t \leq T$*, the random vector* X_t *admits an infinitely differentiable density* p_{t,y_0} *on* $B_R(y_0)$*. Furthermore, for any integer* $k \geq 3$ *there exists a function* $\Lambda_k : \mathbb{R}^m \to [0, \infty)$ *depending also on* R, T, m, d *and on the coefficients of equation (2.2.1) such that, setting*

$$P_t(y) = \mathbb{P}\big(\inf\{|X_s - y| : s \in [(t-1) \vee t/2, t]\} \leq 3R\big),$$

then one has

$$p_{t,y_0}(y) \leq P_t(y_0)\left(1 + \frac{1}{t^{m3/2}}\right)\Lambda_3(y_0)$$

for any $y \in B_R(y_0)$*. Analogously, for every* $\alpha \in \{1, \ldots, m\}^k$*,* $k \geq 1$*,*

$$|\partial_\alpha p_{t,y_0}(y)| \leq P_t(y_0)\left(1 + \frac{1}{t^{m(2k+3)/2}}\right)\Lambda_{2k+3}(y_0)$$

for every $y \in B_R(y_0)$*.*

Two facts are of central importance in the statement of Theorem 1.3.1: first of all, the functions Λ_k appearing in the upper bounds are known *explicitly*. These functions contain the dependence with respect to the bounds on the coefficients and their derivatives and to the ellipticity constant $c_{y_0,R}$. The expression of the Λ_k's is provided in a more detailed version of this theorem in Chapter 2 (Theorem 2.2.4), which we do not give here not to burden the notation. As a second important fact, we succeed in keeping the probability term $P_t(y_0)$

as a multiplicative factor in the bounds for the density and its derivatives: this feature is crucial for the estimation of $p_{t,y_0}(y)$ for large values of y. As a consequence, the application of Theorem 1.3.1 in order to evaluate the density of the law in concrete cases requires the simultaneous evaluation of the functions Λ_k - which basically reduces to a computational step - and of the probability term $P_t(y_0)$. We show how to deal with both these terms in situations of interest. The third factor appearing in the upper bounds is close to one for large values of t and explodes as t goes to zero: let us just mention here that the dependence with respect to time of this factor is actually not optimal (the negative powers of t are too high: for $m = 1$, we would expect a factor $1/t^{1/2}$ rather than $1/t^{3/2}$). Since we are mainly interested in the fixed-time tail asymptotics of the law, this factor will not be a problem for the core of our study.

In section 2.2, besides setting up the tools we need in order to prove Theorem 1.3.1, we show that the evaluation of the functions Λ_k - whose original expression is rather involved - can be simplified under some additional assumptions on the growth of the coefficients (precisely, under a condition of polynomial growth: cf. assumptions **(H1')** and **(H4)** in §2.2.1). This is the object of Theorem 2.2.5, which follows from Theorem 1.3.1 and shows how the upper bounds provided therein can be translated into actual tail bounds on the density and its derivatives. We devote the last section of the chapter, section 2.3, to the application of the previous results to a class of equations embedding the CIR and CEV processes, namely to SDEs of the form

$$X_t = x + \int_0^t (a(X_s) - b(X_s)X_s)ds + \int_0^t \gamma(X_s)X_s^\alpha dW_t, \tag{1.3.1}$$

where $\alpha \in [1/2, 1)$ and a, b and γ are C_b^∞ functions. After discussing the existence and uniqueness of positive strong solutions to such equations, we study in details the properties of the density that results from the previous theorems, analysing its behaviour at infinity (Proposition 2.3.3) and at zero (Proposition 2.3.4), i.e. the point where the diffusion coefficient is singular. The explicit expressions of the densities of the classical CIR (1.2.5) and CEV processes (1.2.6) show that the resulting estimates are in the good range on the log-scale (that is, they have the good exponential decay at infinity). We point out that this is a case where the treatment of the functions Λ_k is easily achieved; on the other hand, to deal with the probability term P_t we rely on some specifically designed tools that involve comparison of the SDE to Brownian motion via a Lamperti-type transformation and invoking the existence of quadratic exponential moments for suprema of the Brownian motion (Fernique's theorem). In Chapter 4 we start a deeper study of the techniques that can be used in order to upper bound the probability term P_t in general situations, focusing on the application of tube estimates for Itô processes (see below).

Let us close the outline of Chapter 2 by comparing our main result, Theorem 1.3.1, to the similar results by Kusuoka and Stroock mentioned in section 1.1. In the same spirit of our study, in section "4. Localization" of [44], the authors consider how to obtain local conclusions on the density of a diffusion under local assumptions on the coefficients. Their main statements, Theorem 4.5 and the subsequent Corollary 4.10, already allow to say that under hypotheses **(H1)** and **(H2)**, a strong solution to (1.1.1) admits a smooth density $p_t(x,\cdot)$ on $B_R(y_0)$ (actually, this density is automatically defined on the whole set where the coefficients are regular, namely $B_{3R}(y_0)$ in our framework, but the radius of the set is of course not the main point). In this sense, the first part of Theorem 1.3.1 is not new. In addition, that Theorem 4.5 in [44] proves the regularity of the map $x \mapsto p_t(x,\cdot), x \in B_R(y_0)$ and holds under the more general Hörmander's condition rather than the ellipticity assumption **(H2)**. But, the upper bounds provided therein for $p_t(x,y)$ and its partial derivatives are meaningful only on the diagonal, that is when the initial condition x belongs to $B_R(y_0)$ and y is close to x: up to multiplicative time-depending constants, the main factor appearing in the upper bounds (cf. (4.7) and (4.9) in [44]) is $\exp\big(-(|y-x| \wedge (R \wedge 1))^2/Ct\big)$, which looses its interest when we want to consider large values of $|y-x|$. The upper bounds provided in Theorem 1.3.1 have the advantage of being suitable to obtain tail estimates, as we have shown in the subsequent Theorem 2.2.5 and in the study of eq. (1.3.1). Moreover, we allow for a more precise identification of the bounding constants: see the Introduction of Chapter 2 and §2.2.3.

In **Chapter 3** we tackle the problem of obtaining *lower* bounds on the cumulative distribution function and (when existing) on the density of the law, under local assumptions on the coefficients of the SDE. We work in the same spirit of Chapter 2, that is, keeping in mind the financial models in §1.2.1 as the main target for applications. Typically, to obtain significant lower bounds for the densities of diffusions is a more challenging task: in Chapter 3, we take up this problem working within the class of SDEs appearing in the so-called Local-Stochastic Volatility (LSV) models, namely SDEs of the form:

$$\begin{aligned} dX_t &= -\frac{1}{2}\eta(t,X_t)^2 f(V_t)^2 dt + \eta(t,X_t) f(V_t) dW_t^1, \quad X_0 = 0 \\ dV_t &= \beta(t,V_t)dt + \sigma(t,V_t)dW_t^2, \end{aligned} \tag{1.3.2}$$

where W^1 and W^2 are two correlated Brownian motions. The function f is chosen to be positive, so that $f(V)$ represents a variance, and the real-valued process X models the logarithm of the forward price of an asset. Applying Itô's formula, it is easy to see that $F_t := F_0 e^{X_t}$ satisfies $dF_t = \eta(t,X_t) f(V_t) F_t dW_t^1$, hence LSV models embed stochastic volatility models (when $\eta \equiv 1$) and in particular the Heston model (when in addition $f(v) = \sqrt{v}$ and V is a CIR process). LSV models have gained a great importance in the context of option pricing in recent years, in particular when the appearance of derivatives whose value depends on the dynamics of the implied volatility (volatility derivatives) has demanded the introduction of more elaborate models. They have been studied and supported in particular by Henry-Labordère

[37, 36] or Lipton [50], who considered the problem of an efficient calibration strategy to the market smile, while other authors focused on the asymptotic properties of the implied volatility, as Forde & Jacquier [24] for the small-time asymptotics in the uncorrelated case.

Giving sharp tail estimates of the cumulative distribution and the density of the log-price X_t in this "case-study" yet calls for the solution of non-trivial problems and for the development of non-standard techniques. Let us discuss this point more in detail. Since it satisfies the equation $F_t = F_0 + \int_0^t \eta(s, X_s) f(V_s) F_s dW_s^1$, the forward price $F = (F_t; t \geq 0)$ is seen to be a positive local martingale. Then, a simple application of Fatou's Lemma shows that F is actually an *integrable* supermartingale. From the integrability property of F_T, $T > 0$, a simple application of Markov's inequality shows that, for every $y > 0$, $\mathbb{P}(X_T > y) = \mathbb{P}(e^{X_T} > e^y) \leq e^{-y}\mathbb{E}[e^{X_T}]$ that is, up to multiplicative constants, the right tail of the distribution of X_T admits the exponential *upper* bound e^{-y}. It has been seen in §1.2.2 that in models such as (1.3.2), the moments of F_T of exponent greater than one or smaller than zero can explode: of course, moment explosion occurs when the tails of the distribution of X_T are sufficiently heavy. More precisely, if the law of X_T admits a density and this density behaves as $e^{-c|y|}$ for $|y| \to \infty$ for some constant $c > 1$, then positive and negative exponential moments of X_t of order p will explode for $p \geq c$. Dragulescu and Yakovenko [21] showed that the density of the log-price does behave as $e^{-c|y|}$ in the Heston model (1.2.2), exploiting the analytical computations that can be carried for the characteristic function of X_T. Let us mention that the work of [21] on the stock price distribution in the Heston model has been extended and sharpened with the addition of higher-order terms to the leading $e^{-c|y|}$, first by Gulisashvili and Stein [33] in the case of zero correlation and subsequently by Friz et al. [28].

The main aim of Chapter 3 is to show that the cumulative distribution of the log forward price X and, when existing, its density, behave as $\exp(-c|y|)$ for large $|y|$ in the following class of LSV models:

$$\begin{aligned} dX_t &= -\frac{1}{2}\eta(t, X_t)^2 V_t dt + \eta(t, X_t)\sqrt{V_t} dW_t^X \\ dV_t &= \beta(t, V_t)dt + \sigma(t, V_t)\sqrt{V_t} dW_t^V, \end{aligned} \tag{1.3.3}$$

obtained from (1.3.2) setting $f(v) = \sqrt{v}$. This class contains the Heston model and the "universal volatility model" (without the jump part) considered in [50], but is much wider, allowing for general coefficients β, σ in the SDE of the variance. While on the one hand we consider reasonable Lipschitz, boundedness and ellipticity conditions on the coefficients η and σ (but we allow the drift β to be any measurable function with sub-linear growth), on the other hand we emphasize that the square-root factors in (1.3.3) impose the necessity to work under local regularity assumptions. Let us also remark that, at this level, we are not concerned with the (possibly intricate) discussion on the existence and/or uniqueness of solutions to (1.3.3):

our results indeed hold for *any* couple of processes $(X,V) = (X_t, V_t; t \in [0,T])$ satisfying (1.3.3). The situation where the diffusion coefficient of the second SDE in (1.3.3) is replaced by $\sigma(t,v)v^p$ for a $p > 0$ (thus embedding the class of models considered by Andersen & Piterbarg in [2]) is at the basis of the motivation for the further study we begin to develop in Chapter 4. Let us introduce the main tool of our analysis: we rely on an estimate involving the trajectory of the couple (X,V) up to time T obtained by Bally, Fernandez & Meda in [8], which we refer to as to a "tube" estimate. In [8], the authors provide estimates for the probability that an Itô process remains in a tube of given radius around a given deterministic curve, under some conditions of local Lipschitz-continuity, local boundedness and local ellipticity on the coefficients of the process. As a result, the probability of staying in the tube is lower bounded by an integral functional of the curve itself, of the deterministic radius and the coefficients of the SDE. The work we carry out in Chapter 3 is to cast this functional in a simple form, and then to optimize over the possible choices of curves and radii. This formulation leads to the solution of an Euler-Lagrange optimization problem: the explicit computations that follow allow us to obtain a lower bound which is in the desired asymptotic range. To present our main result in this direction, let us introduce the following objects: for $y \in \mathbb{R}$, define the point $\overline{y}$ and the one-dimensional curves $\tilde{x}_t, \tilde{v}_t, \tilde{R}_t$, $t \in [0,T]$ by

$$\overline{y} = |y| + V_0; \qquad \phi(t) = \frac{\sinh(t/2)}{\sinh(T/2)};$$

$$\tilde{v}_t = V_0\Big(\sqrt{\frac{\overline{y}}{V_0}}\phi(t) - e^{-T/2}\phi(t) + e^{-t/2}\Big)^2; \quad \tilde{x}_t = \mathrm{sign}(y)(\tilde{v}_t - V_0); \quad \tilde{R}_t = \frac{1}{2}\sqrt{(V_0 \wedge 1)\tilde{v}_t}$$

where $\mathrm{sign}(x) = 1$ if $x \geq 0$ and $\mathrm{sign}(x) = -1$ if $x < 0$. Our main result is the following estimate:

$$\mathbb{P}\big(|(X_t, V_t) - (\tilde{x}_t, \tilde{v}_t)| \leq \tilde{R}_t, 0 \leq t \leq T\big) \geq \exp\big(-c_T \psi(\rho_\perp) \times |y|\big) \tag{1.3.4}$$

which holds for $|y|$ large enough, where ψ is an explicit function and c_T is a strictly positive constant depending on the model parameters and explicitly on T, but not on y nor on the correlation parameter ρ (in §3.2.1, Theorem 3.2.1, we precise how large $|y|$ must be and give the expression of ψ and c_T). The curves $\tilde{x}_\cdot, \tilde{v}_\cdot, \tilde{R}_\cdot$ are the product of the optimization procedure we set up, appearing as the solution to the Euler-Lagrange equations in section 3.3.1. We remark that the curve $\tilde{x}_\cdot$ ends up at $\tilde{x}_T = y$ while the terminal radius $\tilde{R}_T$ is proportional to $\sqrt{|y|}$: hence, dropping the multiplicative constants for simplicity and writing $\mathbb{P}(|X_T - y| \leq \sqrt{|y|}) \geq \mathbb{P}(|(X_T, V_T) - (\tilde{x}_T, \tilde{v}_T)| \leq \tilde{R}_T)$, then using (1.3.4), we obtain the desired lower bound for the terminal distribution (this argument is made rigorous in Corollary 2 in §3.2.2). This result already allows us to state our main conclusion on the asymptotic behaviour of the implied volatility, namely: the implied volatility always displays wings (equivalently, is never flat) in the class of models (1.3.3) (cf. estimate (3.2.11) in Corollary 2 for the precise statement on the asymptotic slopes of the implied variance). Notice that the

fact that the tube estimate (1.3.4) is given for the couple (X, V) is crucial in our framework: indeed, in order to estimate the behaviour of X_T we need to have a control on the variance V_t for all $t \in [0, T]$.
We would like to notice that the results we obtain in this chapter on the law of the underlying and on moment explosion are significant by themselves, as generalisations of the existing results on stochastic volatility models (that is to say, apart from the fact that we employ special techniques in order to circumvent the singular coefficients).
As a second part of our study, we extend the previous estimates to the density of the law. More precisely, under some additional regularity hypotheses on the coefficients of the SDE, we discuss the existence of a density for the law of X_T and show that the exponential lower bound holds for the density as well. This last step requires to work out some "small balls" estimates (cf. Proposition 3.2.1 in §3.2.3) and to employ the integration by parts formula of Malliavin calculus. As done in Chapter 2, this tool needs to be coupled with an appropriate localization procedure. We rely here on the decomposition of X_T into a Gaussian term plus a perturbation, following the idea of Bally & Caramellino in [6]: then, the desired lower bound on the density follows from an operation of balance between the two terms of the decomposition. This operation involves a sharp estimation of the Sobolev norms of the perturbation term, for which we take advantage of the estimates of the norms of a diffusion derived in Chapter 2, §2.2.3. Our final estimate on the density p_{X_T} of X_T reads

$$p_{X_T}(y) \geq \frac{1}{M_T} \exp\left(-e_T \psi(\rho_\perp)|y|\right)$$

for $|y| > M_T$, where M_T and e_T are constants depending on model parameters and explicitly on T (cf. Theorem 3.2.2).

In **Chapter 4**, we pursue our study in the area of tube estimates for Itô processes and their applications. First of all, we give a new formulation of the tube estimate itself. Actually, the starting point for the study we develop in this chapter dates back to a discussion that Vlad Bally had with Emmanuel Gobet during a conference in Helsinky in 2008[2]. The tube estimate in [8], which is one of the core tools for the analysis carried in Chapter 3, is obtained by construction of an appropriate time-grid on the interval $[0, T]$, then exploiting the short time behaviour of Itô processes. E. Gobet suggested that martingale time-change techniques could be a powerful (yet simple) instrument to obtain similar results in this area. Here we take up this problem, and the results show that Gobet's intuition was correct: coupling elementary time-change techniques for martingales with the appropriate localization arguments, we manage to obtain similar results to [8], but the current estimates are sharper and the hypothesis are weaker. Last but not least, the machinery we have to settle here is considerably lighter, as a consequence the proofs of the main results are easier. Besides tube

[2]I am grateful to Prof. Vlad Bally for sharing with me these useful insights.

estimates, exploiting the same tools we derive upper bounds for suprema of Itô processes, namely we find functions $C(y)$ such that the estimate $\mathbb{P}\big(\sup_{t\le T}|X_t| > y\big) \le C(y)$ is significant in the asymptotic range. We formulate our basic estimate involving the time-change argument considering general Itô processes and tubes of simple geometry in Theorem 4.2.1, while the subsequent Propositions 4.3.2 and 4.3.3 specialise the result to diffusions and arbitrary deterministic tubes.

The first main application of these results, in the spirit of the work of the previous chapters, is to derive tail estimates for fixed-time distribution functions and - when existing - densities. The lower bounds for the probability to stay inside a tube up to time T translate into lower bounds for the terminal distribution, using for example the techniques already settled in Chapter 3. In its turn, the upper bound on the supremum of the process over $[0,T]$ provides the corresponding upper bound on the law at time T: while the tails of the cdf can be trivially estimated by $\mathbb{P}(|X_T| > y\big) \le \mathbb{P}\big(\sup_{t\le T}|X_t| > y\big)$, to estimate the density we can apply Theorem 1.3.1 in Chapter 2. In particular, the current upper bound on $\sup_{t\le T}|X_t|$ provides a way to estimate, in a general framework, the factor P_t appearing in the upper bounds for the density and its derivatives in Theorem 1.3.1 (cf. Remark 4.3.1). We give applications of these results to the class of one-dimensional SDEs considered in Chapter 2, namely to CIR/CEV-type processes with local coefficients as (1.3.1), showing that the tube estimates provided here are effective and allow to obtain significant tail estimates from above and below (cf. Propositions 4.3.1 and 4.3.4). The application to general diffusions is currently under study, and we pledge to fill this gap very quickly.

1.4 Outline of results : Part II

The results in the second part of this thesis are motivated by numerical applications. We deal with the calibration of a parametric model to the market smile and with approximate pricing formulae for the implied volatility. As a consequence, the mathematical tools that we employ here are somehow less sophisticated than the ones in the first part. Nevertheless, the results we provide are of straight interest for applications and the formulae and algorithms we propose can be directly implemented following our presentation. The computational and calibration performances are illustrated by numerical tests.

1.4.1 A closer look to implied volatility: Heston and SVI

The square root coefficient in (1.2.2) is the central source of difficulties for the Heston model. The non-Lipschitz character of this coefficient makes several manipulations of financial interest, such as the computation of Greeks or the implementation of simulation schemes for the SDE, more intricate, and obstructs the direct application of Malliavin calculus tools. This last limitation is what led us to the study developed in Part I of this thesis. At the same time, the square root coefficient lies at the basis of the analytical tractability of the model in the context of option pricing. Thanks to nothing but the fact that $(\sqrt{x})^2 = x$, as pointed out in §1.2.1, the Heston model is an Affine Stochastic Volatility Model (cf. [41]) and, in particular, the characteristic function of the log-price $X_t = \log(F_t/F_0)$ is known explicitly. As a consequence, the evaluation of European Vanilla options in the Heston model can be put under the form of a Fourier inversion problem: the solution to this problem yields the well known semi-closed formula for Vanillas (which is known since the original article of Heston [38]), and the pricing problem eventually boils downs to the computation of a complex integral. In recent years, several authors have been concerned with the efficient numerical evaluation of this integral, proposing some effective and robust implementation solutions. A survey of such recent advances on the Heston model can be found in [60].

In its turn, the implied volatility is usually recovered from the prices of Vanillas by numerical inversion of the Black-Scholes formula. Due to the high non linearity of the BS formula with respect to the volatility parameter, in general no closed or semi-closed formula is available to *directly* compute the implied volatility (i.e. the numerical inversion step cannot be overcome), and the Heston model makes no exception. The behaviour of the implied volatility in limiting cases (extreme strikes, short and large maturities) can still be attentively studied using different asymptotic techniques, as we have seen in Chapter 3 for the large strike limits, but it is indeed a challenge to analytically describe the whole implied volatility surface at once. In Chapter 6 in Part II, we remark that, exploiting the knowledge of characteristic functions in the Heston model, semi-closed formulae for the implied volatility (i.e., formulae involving the integral of an explicit function) can be actually obtained at some very specific points. It this basis point is sufficiently meaningful (this will be, of course, the At-The-Money point), one can exploit this semi-closed formula to build efficient analytical approximations of the whole smile. To fix ideas, fix a maturity T and recall the implied volatility $\sigma(T, x)$ as a function of log-moneyness $x := \log(K/F_0)$ from (1.2.7). Suppose that a formula is available for the value of σ at the At-The-Money (ATM) point $x = 0$: since the slopes of the function $k \mapsto \sigma^2(T, x)$ for $x \to \pm\infty$ are known in semi-closed form using (1.2.8), to restore the whole time T-slice of the smile we need to make an *interpolation* choice between the ATM value and the behaviour for large values of $|x|$. To this scope, an approach is to choose a suitable parametric family of curves. It turns out that a choice which is *very* well suited for this kind of operation within the Heston model is given by the Stochastic Volatility Inspired (SVI) parametric model proposed by J. Gatheral in [29], which we take into consideration in the second part this thesis.

Chapter 5, then, is devoted to the SVI model. We overview the main properties of the model and deal with the problem of the calibration to the market smile.
Parametric models are of common use in the treatment of the volatility surface. Apart from the extrapolation of smile points, they provide a smoothing of the market smile and the

consequent facilities in the calibration of stochastic models for the underlying (including the reconstruction of a local volatility surface via Dupire's formula). SVI corresponds to the following parametric family of curves:

$$\sigma_{SVI}(x)^2 = a + b\big(r(x-m) + \sqrt{(x-m)^2+\gamma^2}\big). \tag{1.4.1}$$

where σ_{SVI}^2 is the implied variance at fixed time-to-maturity T, x the log-moneyness $x = log(K/F_0)$ and a, b, ρ, m, γ are the model parameters. (5.1.1) is known as the Stochastic Volatility Inspired (SVI) model since the functional form has been inspired by the results on the large-time asymptotics of implied variance in the Heston model. It is widely known that the SVI parameterisation (5.1.1) proves to have outstanding performances in the calibration to single-maturity slices of the implied smile on Equity indexes. Nevertheless, it is also common knowledge that the least square calibration of (5.1.1) is typically affected by the presence of several local minima. To our experience, even when SVI is calibrated to simulated data, i.e. a smile produced by SVI itself, local minima that are difficult to sort out (least square objective $\approx 10^{-8}$ for reasonable volatilities, $\sigma \approx 20\% - 40\%$) are found far away from the global one (objective = 0). This unpleasant feature brings some difficulties, when one wants to design a reliable and robust parameter identifications strategy in SVI model. Because of the local minima, the same smile can be - remarkably well - calibrated with sets of parameters that are totally different one from the other. As a result, the solution yield by a least square optimizer usually displays a strong dependence on the input starting point. The big issue, then, becomes the stability of calibrated parameters through time-to-maturity. This feature comes into play in a significant way when one is trying to parameterize the whole volatility surface.

In Chapter 5, we present a procedure providing a trustworthy and robust calibration of the SVI parametric form (5.1.1), having the pleasant feature to be almost insensitive to the initial parameter guess. We rely on some simple observations on the symmetries of the functional form (5.1.1) (cf. §5.3.1) to downsize the minimization problem from dimension 5 (the number of parameters in (5.1.1)) to dimension 2 (namely, m and γ), while the optimization over the remaining 3 is performed exactly (except for a few minimum searches in dimension 1, which can nevertheless be performed accurately and fast). Last but not least, the method yields an optimal parameter set which is automatically consistent with the arbitrage constraint on the slopes of implied variance. The performances of this calibration strategy are displayed by some numerical tests in section 5.4.

In **Chapter 6** we implement the approach to the approximation of implied volatility we have outlined at the beginning of this section. In particular, we provide semi-closed formulae for the ATM curve $t \to \sigma(t, 0)$. These formulae do not apply only to Heston, but to any continuous model that allows for an explicit computation of the Laplace transform of the quadratic variation of the log price X_t. We formulate our results under the assumption that the smile is symmetric, namely $\sigma(t,x) = \sigma(t,-x)$ for all $t \geq 0$ and $x \in \mathbb{R}$. Of course, this assumption introduces a considerable simplification and is not realistic in financial markets, whose smile are typically skewed, nevertheless the results we provide can be straightforwardly applied to a skewed smile by considering a displaced model, as we show in section 6.4. We consider symmetric smiles since they enjoy some useful properties. In particular, Tehranchi [62] recently showed that for any price process which is a continuous martingale, a symmetric smile characterises the law of the log-price X_t conditional to its quadratic variation $\langle X\rangle_t$, precisely

$$\mathbb{P}(X_t \in dy|\langle X\rangle_t) = \mathcal{N}\Big(-\frac{1}{2}\langle X\rangle_t, \langle X\rangle_t\Big)(dy) \tag{1.4.2}$$

where $\mathcal{N}(\mu,\sigma^2)$ is the Gaussian law of mean μ and variance σ^2. Then, the main idea is to combine this result with the identities relating the log price cumulative distribution and density on the one hand and, on the other, the implied volatility and its second derivative at the ATM point $x = 0$:

$$\mathbb{P}(X_t > 0) = N\Big(-\frac{t\sigma(t,0)}{2}\Big), \qquad p_t(0) = \frac{\exp\Big(-\frac{t\sigma^2(t,0)}{8}\Big)}{\sqrt{2\pi t\,\sigma^2(t,0)}}\big(1 + t\,\partial_x^2\sigma^2(t,x)|_{x=0}\big) \tag{1.4.3}$$

where N is the standard normal cdf and $p_t(\cdot)$ is the density of X_t (we remark that we are in a situation where the law of X_t, $t > 0$, actually admits a density, thanks to (1.4.2)). In the first part of this thesis, we have focused on the estimation of $\mathbb{P}(X_t > y)$ and $p_t(y)$ for large values of y, in general situations and classes of models. Now, exploiting the consequence (1.4.2) of the symmetric smile assumption, together with some properly designed integral representations of the normal density and cumulative distribution (cf. Lemmas 6.2.2 and 6.2.1), we can provide an *exact* computation of these two quantities, for any value of y. The result reads as follows:

Proposition 1.4.1 (Proposition 6.2.4, Chapter 6). *Under the assumption of a symmetric smile, for every $t > 0$ the density p_t and the cdf $\mathbb{P}(X_t > \cdot)$ of the log forward price X_t satisfy*

$$\begin{aligned} p_t(y) &= \frac{e^{-\frac{y}{2}}}{\pi\sqrt{2}}\int_0^\infty \frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}}\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz, \\ \mathbb{P}(X_t > y) &= \frac{1}{4\sqrt{2}\pi}e^{-\frac{y}{2}}\int_0^\infty \frac{\cos(\sqrt{2z}y) - 2\sqrt{2z}\sin(\sqrt{2z}y)}{\sqrt{z}(z+\frac{1}{8})}\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz \end{aligned} \tag{1.4.4}$$

for every $y \in \mathbb{R}$.

It is clear, then, that in models where the Laplace transform of the quadratic variation $\langle X\rangle_t$ is explicitly known, the integral representations (1.4.4) translate into semi-closed formulae for the density and the cdf of X_t. Using this formulae for $y = 0$ and inverting the relations (1.4.3) with respect to $\sigma(t, 0)$ and $\partial_x^2 \sigma^2(t, x)|_{x=0}$, we get the corresponding formulae for the ATM implied volatility level and curvature. This is an important consequence, since it allows us to access detailed information about the time dependence (the Term Structure) of the smile in the neighborhood of the ATM point, which is usually difficult to work out explicitly. Since the asymptotic slopes of the smile are provided by the moment formula (1.2.8), we are only left with the interpolation of the at-the-money and "far-from-the-money" structures. We make use of a symmetric SVI model with time dependent parameters, namely

$$\sigma_{SVI}(x)^2 = a(t) + b(t)\sqrt{x^2 + \gamma(t)^2}. \tag{1.4.5}$$

It is a matter of computation to define $a(t), b(t)$ and $\gamma(t)$ so that (1.4.5) match the ATM level and curvature and the asymptotic slopes of the implied volatility. We eventually obtain an approximation of the smile in any given symmetric model (with known Laplace transform) involving semi closed-form parameters $a(t), b(t)$ and $\gamma(t)$. Let us mention that, since (1.4.4) hold for every $y \in \mathbb{R}$, as an additional consequence of Proposition 1.4.1 we obtain a pricing formula for Vanillas in a general symmetric model.
The Heston model (1.2.2) generates symmetric smiles in the case of zero correlation between the two driving Brownian motions (this is a general feature of stochastic volatility models, cf. for example Renault and Touzi [56]). In section 6.3, we implement the integral representation formulae discussed so far in the framework of the uncorrelated Heston model. In particular, we concentrate our efforts on the manipulation of the resulting semi-closed formulae in order to make the numerical implementation as stable and straightforward as possible. Starting from the formulation in (1.4.5), we remove all the possible singularities of the integrands and squeeze the integration over the bounded interval $[0, 1]$. The final formulae, presented in Propositions 6.3.1 and 6.3.4, only involve the integrals of *bounded* functions over the *fixed* interval $[0, 1]$, a feature which is extremely convenient for numerical purposes. Last but not least, as a result of working only with Laplace transforms and not with Fourier transforms, in the final formulae we avoid complex integrals, which are typically more difficult to deal with. Some numerical experiments in section 6.5 show the validity of our results for the Heston model: the formula for the ATM implied volatility proves to be extremely accurate, and the time-dependent SVI approximation displays considerable performances in a wide range of maturities and strikes (cf. the discussion in section 6.5, Tables 6.1-6.2 and Figures 6.2-6.3).

The body of this document is organized as follows. Each chapter contains an Introduction, the presentation of the main results with detailed proofs of all the involved technical lemmas and possibly a conclusion, and can therefore be read as a stand-alone part of the document. The Introduction of each chapter, then, borrows some parts of the current introductory section, adding more technical information and a more specific presentation of the material.

Part I

Tail Asymptotics for SDEs with locally smooth coefficients

Chapter 2

Smoothness and upper bounds on densities for SDEs

Abstract

We study smoothness of densities for the solutions of SDEs whose coefficients are smooth and nondegenerate only on an open domain D. We prove that a smooth density exists on D and give upper bounds for this density. Under some additional conditions (mainly dealing with the growth of the coefficients and their derivatives), we formulate upper bounds that are suitable to obtain asymptotic estimates of the density for large values of the state variable ("tail" estimates). These results specify and extend some results by Kusuoka and Stroock in [44], but our approach is substantially different and based on a technique to estimate the Fourier transform inspired from Fournier [27] and Bally [5]. This study is motivated by existing models for financial securities which rely on SDEs with non-Lipschitz coefficients. Indeed, we apply our results to a square root-type diffusion (CIR or CEV) with coefficients depending on the state variable, i.e. a situation where standard techniques for density estimation based on Malliavin calculus do not apply. We establish the existence of a smooth density, for which we give exponential estimates and study the behaviour at the origin (the singular point).

Keywords: Smoothness of Densities, Stochastic Differential Equations, Tail estimates, Irregular Coefficients, Malliavin calculus

Note

The results in this chapter are due to appear in the article [19] in the *Annals of Applied Probability.* I would like to thank my advisor Prof. Vlad Bally for many stimulating discussions on the subject of this chapter, Prof. Alexander Yu. Veretennikov of University of Leeds for insights on equations with square root coefficients and an anonymous referee who read the first version of [19] and helped me improving the presentation.

2.1 Introduction

It is well known that Malliavin calculus is a tool which allows, among other things, to prove that the law of a diffusion process admits a smooth density. More precisely, if one assumes that the coefficients of an SDE are bounded C^∞ functions with bounded derivatives of any order and that, on the other hand, the Hormandër condition holds, then the solution is a smooth functional in Malliavin's sense and it is nondegenerate at any fixed positive time. Then, the general criterion given by Malliavin [51] allows to say that the law of such a random variable is absolutely continuous with respect to the Lebesgue measure and its density is a smooth function (see [54] for a general presentation of this topic).

The aim of this chapter is to relax the aforementioned conditions on the coefficients: roughly speaking, we assume that the coefficients are smooth only on an open domain D and have bounded partial derivatives therein. Moreover, we assume that the nondegeneracy condition on the diffusion coefficient holds true on D only. Under these assumptions, we prove that the law of a strong solution to the equation admits a smooth density on D (Theorem 2.2.1). Furthermore, when D is the complementary of a compact ball and the coefficients satisfy some additional assumptions on D (mainly dealing with their growth and the one of derivatives), we give upper bounds for the density for large values of the state variable (Theorem 2.2.2). We will occasionally refer to these aymptotic estimates of the density as to "tail estimates" or estimates on the density's "tails". As we have pointed out in the Introduction of this thesis, local results have already been obtained by S. Kusuoka and D. Stroock in [44], section 4. Let us recall that here the authors work under local regularity and nondegeneracy hypotheses, too, but the upper bounds they provide on the density are mostly significant on the diagonal (i.e. close to the starting point), but they are not appropriate for tail estimates. Moreover, the constants appearing in their estimates are not explicit (cf. (4.7)-(4.9) in Theorem 4.5 in [44] and the corresponding estimates in Corollary 4.10). In the present chapter, we provide upper bounds that are suitable for tail estimates and we find out the explicit dependence of the bounding constants with respect to the coefficients of the SDE and their derivatives. Our bounds turn out to be applicable to the case of diffusions with tails stronger than Gaussian. This is what happens in the framework of square-root diffusions, which is our major example of interest (see further below and section 2.3). Also, our approach is substantially different from the one in [44]. In particular, we rely on a Fourier transform argument, employing a technique to estimate the Fourier transform of the process inspired from the work of N. Fournier in [27] and of V. Bally in [5] and relying on specifically-designed Malliavin calculus techniques. We estimate the density $p_t(y)$ of the diffusion at a point $y \in D$ performing an integration by parts that involves the contribution of the Brownian noise only on an arbitrarily small time interval $[t-\delta, t]$: this allows us to gain a free parameter δ that we can eventually optimize, and the appropriate choice of δ proves to be a key point in our argument. We do not study here the regularity with respect to initial condition (which may be the subject of future work).

As in the Introduction of this thesis, our study is motivated by applications in Finance, in particular by the study of models for financial securities which rely on SDEs with non-Lipschitz coefficients. The main examples of such models (CIR, CEV, SABR, Heston) are recalled in §1.2.1. In this chapter, we apply our results to one dimensional SDEs of the form

$$X_t = x + \int_0^t (a(X_s) - b(X_s)X_s)ds + \int_0^t \gamma(X_s)X_s^{\alpha}dW_t, \tag{2.1.1}$$

where $\alpha \in [1/2, 1)$ and a, b and γ are C_b^∞ functions. When the coefficients a, b, σ are constant, the solutions to this class of equations include the classical CIR process ($\alpha = 1/2$) and a subclass of the CEV local volatility diffusions (when $a \equiv b \equiv 0$). As pointed out by M. Bossy and A. Diop in [12], SDEs with square-root terms and coefficients depending on the level of the state variable arise as well in the context of the modelization of turbulent flows in fluid mechanics. Recall from the Introduction that for the CIR and CEV processes the density of X_t is known explicitly: then, the main contribution of our results lies in the fact that they apply to the more general framework of SDEs whose coefficients are functions of the state variable, thus when explicit computations are no longer possible. Theorem 2.2.2 directly applies and tells that X_t admits a smooth density on $(0, +\infty)$: under some additional conditions on the coefficients (mainly dealing with their asymptotic behaviour at ∞ and zero) we give exponential-type upper bounds for the density at infinity (Proposition 2.3.3) and study the explosive behaviour of the density at zero (Proposition 2.3.4). The explicit expressions of the densities of the classical CIR (1.2.5) and CEV processes (1.2.6) show that our estimates are in the good range.

The Chapter is organized follows: in section 2.2 we present our main results on SDEs with locally smooth coefficients (§2.2.1) and we collect all the technical elements we need to give their proofs. In particular, in §2.2.2 we recall the basic tools of Malliavin calculus on the Wiener space, which will be used in §2.2.3 to obtain some explicit estimates of the L^2-norms of the weights involved in the integration by parts formula. This is done following some standard techniques of estimation of Sobolev norms and inverse moments of the determinant of the Malliavin matrix (as in [54] and [17], section 4), but in our computations we explicitly pop out the dependence with respect to the coefficients of the SDE and their derivatives. This further allows to obtain the explicit estimates on the density tails. section 2.2.4 is devoted to the proof of the theorems stated in §2.2.1. We employ the Fourier transform argument and the optimized integration by parts we have discussed above. Finally, in section 2.3 we apply our results to the solutions of (2.1.1).

2.2 Smoothness and upper bounds on densities for SDEs with locally smooth coefficients

2.2.1 Main results

Some notation. In what follows, b and σ_j are measurable functions from $\mathbb{R}^m$ into $\mathbb{R}^m$, $j = 1, \dots, d$. For $y_0 \in \mathbb{R}^m$ and $R > 0$, we denote by $B_R(y_0)$ (resp. $\overline{B}_R(y_0)$) the open (resp. closed) ball $B_R(y_0) = \{y \in \mathbb{R}^m : |y - y_0| < R\}$ (resp. $\overline{B}_R(y_0) = \{y \in \mathbb{R}^m : |y - y_0| \leq R\}$), where $|\cdot|$ stands for the Euclidean norm. We denote $C_b^\infty(A)$ the class of infinitely differentiable functions on the open set $A \subseteq \mathbb{R}^m$ which are bounded together with their partial derivatives of any order. For a multi-index $\alpha \in \{1, \dots, m\}^k$, $k \geq 1$, ∂_α denotes the partial derivative $\frac{\partial^k}{\partial_{x_{\alpha_1}} \dots \partial_{x_{\alpha_k}}}$.

Let $0 < R \leq 1$ and $y_0 \in \mathbb{R}^m$ be given. We consider the SDE

$$X_t^i = x^i + \int_0^t b^i(X_s)ds + \sum_{j=1}^d \int_0^t \sigma_j^i(X_s)dW_s^j, \quad t \in [0, T], \quad i = 1, \dots, m \tag{2.2.1}$$

for a finite $T > 0$ and $x \in \mathbb{R}^m$, and assume that the following hold:

(H1) (*local smoothness*) $b, \sigma_j \in C_b^\infty(B_{5R}(y_0); \mathbb{R}^m)$;

(H2) (*local ellipticity*) $\sigma\sigma^*(y) \geq c_{y_0,R}\, I_m$ for every $y \in B_{3R}(y_0)$, for some $0 < c_{y_0,R} < 1$;

(H3) existence of strong solutions holds for the couple (b, σ).

Let then $(X_t; t \in [0, T])$ denote a strong solution of (2.2.1). Our first main result follows.

Theorem 2.2.1. *Assume* **(H1)**, **(H2)** *and* **(H3)**. *Then for any initial condition $x \in \mathbb{R}^m$ and any $0 < t \leq T$, the random vector X_t admits an infinitely differentiable density p_{t,y_0} on $B_R(y_0)$. Furthermore, for any integer $k \geq 3$ there exists a positive constant Λ_k depending also on y_0, R, T, m, d and on the coefficients of equation (2.2.1) such that, setting*

$$P_t(y) = \mathbb{P}\big(\inf\{|X_s - y| : s \in [(t-1) \vee t/2, t]\} \leq 3R\big),$$

then one has

$$p_{t,y_0}(y) \leq P_t(y_0)\left(1 + \frac{1}{t^{m3/2}}\right)\Lambda_3 \tag{2.2.2}$$

for any $y \in B_R(y_0)$. Analogously, for every $\alpha \in \{1, \dots, m\}^k$, $k \geq 1$,

$$|\partial_\alpha p_{t,y_0}(y)| \leq P_t(y_0)\left(1 + \frac{1}{t^{m(2k+3)/2}}\right)\Lambda_{2k+3} \tag{2.2.3}$$

for every $y \in B_R(y_0)$.

The functional dependence of Λ_k with respect to y_0, R, T and the bounds on the coefficients b and σ is known *explicitly*. We provide the expression of Λ_k in section 2.2.4, in a more detailed version of Theorem 2.2.1 (Th. 2.2.4) which we do not give here for the simplicity of notation.

When the coefficients of (2.2.1) are smooth outside a compact ball and have polynomial growth together with their derivatives therein, by Theorem 2.2.1 a smooth density exists outside the same ball, and one can deduce some more easily-read upper bounds. More precisely, we consider the following assumptions:

(H1') there exist $\eta \geq 0$ such that b, σ_j are of class C^∞ on $\mathbb{R}^m \setminus \overline{B}_\eta(0)$, and **(H2)** holds for any $R > 0$ and y_0 such that $\overline{B}_{3R}(y_0) \subset \mathbb{R}^m \setminus \overline{B}_\eta(0)$;

(H4) there exist $q, \overline{q} > 0$ and positive constants $0 < C_0 < 1$ and C_k, $k \geq 0$, such that for any $\alpha \in \{1, \dots, m\}^k$

$$|\partial_\alpha b^i(y)| + |\partial_\alpha \sigma^i_j(y)| \leq C_k(1 + |y|^q) \tag{2.2.4}$$

and

$$\sigma\sigma^*(y) \geq C_0 |y|^{-\overline{q}}\, I_m \tag{2.2.5}$$

hold for $|y| > \eta$.

Theorem 2.2.2. *Assume* **(H1')** *and* **(H3)**.

(a) *For any initial condition $x \in \mathbb{R}^m$ and for any $0 < t \leq T$, X_t admits a smooth density on $\mathbb{R}^m \setminus \overline{B}_\eta(0)$.*

(b) *Assume* **(H4)** *as well. Then estimates (2.2.2) and (2.2.3) hold with $R = 1$ and*

$$\Lambda_k = \Lambda_k(y_0) := C_{k,T}\big(1 + |y_0|^{q'_k(q)}\big), \tag{2.2.6}$$

for every $|y_0| > \eta + 5$. The value of the exponent $q'_k(q)$ is explicitly known (and provided in Theorem 2.2.5).

(c) If moreover $\sup_{0 \leq s \leq t} |X_s|$ has finite moments of all orders, then for every $p > 0$ and $k \geq 1$ there exist positive constants $C_{k,p,T}$ such that

$$\begin{aligned} |p_t(y)| &\leq C_{3,p,T}\left(1 + \frac{1}{t^{m3/2}}\right)|y|^{-p} \\ |\partial_\alpha p_t(y)| &\leq C_{k,p,T}\left(1 + \frac{1}{t^{m(2k+3)/2}}\right)|y|^{-p}, \quad \alpha \in \{1, \dots, m\}^k \end{aligned} \tag{2.2.7}$$

for every $0 < t \leq T$ and every $|y| > \eta + 5$.

In the previous, the $C_{k,p,T}$ are positive constants depending on k, p, T and also on m, d and on the bounds (2.2.4) and (2.2.5) on the coefficients.

The proofs of these results will be given in section 2.2.4.

Additional notation. Through the rest of the chapter, $< \cdot, \cdot >$ will denote the Euclidean scalar product in $\mathbb{R}^m$, while the notation $|\cdot|$ will be used both for the absolute value of real numbers and for the Euclidean norm in $\mathbb{R}^m$. Furthermore, when $\Theta = \theta_1, \ldots, \theta_\nu$ is a family of parameters, by C_Θ we denote a constant depending on the θ_i's but not on any of the other existing variables, unless different specified. All constants of such type may vary from line by line, but always depend only on the θ_i's. For functions of one variable, the k-th derivative will be denoted by $\frac{d^k f}{dx^k}$ or $f^{(k)}$. Finally, we will follow the convention of summation over repeated indexes, wherever present.

2.2.2 Elements of Malliavin Calculus

We recall hereafter some elements of Malliavin calculus on the Wiener space, following [54]. Let $W = (W_t^1, \ldots, W_t^d; t \geq 0)$ be a d-dimensional Brownian motion defined on the canonical space $(\Omega, \mathcal{F}, \mathbb{P})$. For fixed $T > 0$, let $\mathcal{H}$ be the Hilbert space $\mathcal{H} = L^2([0,T]; \mathbb{R}^d)$. For any $h \in \mathcal{H}$ we set $W(h) = \sum_{j=1}^d \int_0^T h^j(s) dW_s^j$, and consider the family $\mathcal{S} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ of smooth random variables defined by

$$\mathcal{S} = \left\{ F : F = f(W(h_1), \ldots, W(h_n)); h_1, \ldots, h_n \in \mathcal{H}; f \in C^\infty_{\text{pol}}(\mathbb{R}^n); n \geq 1 \right\},$$

where C^∞_{pol} denotes the class of C^∞ functions which have polynomial growth together with their derivatives of any order.
The Malliavin derivative of $F \in \mathcal{S}$ is the d-dimensional stochastic process $DF = (D_r^1 F, \ldots, D_r^d F; r \in [0,T])$ defined by

$$D_r^j F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_n)) h_i^j(r), \quad j = 1, \ldots, d.$$

For any positive integer k, the k-th order derivative of F is obtained by iterating the derivative operator: for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_k) \in \{1, \ldots, d\}^k$ and $(r_1, \ldots, r_k) \in [0,T]^k$, we set $D^{\alpha_1, \ldots, \alpha_k}_{r_1, \ldots r_k} F := D^{\alpha_1}_{r_1} \cdots D^{\alpha_k}_{r_k} F$. Given $p \geq 1$ and positive integer k, for every $F \in \mathcal{S}$ we define the seminorm

$$||F||_{k,p} = \left(\mathbb{E}[|F|^p] + \sum_{h=1}^k \mathbb{E}\left[||D^{(h)} F||^p_{\mathcal{H}^{\otimes h}} \right] \right)^{1/p},$$

where

$$||D^{(k)} F||_{\mathcal{H}^{\otimes k}} = \left(\sum_{|\alpha|=k} \int_{[0,T]^k} \left| D^{\alpha_1, \ldots, \alpha_k}_{r_1, \ldots r_k} F \right|^2 dr_1 \cdots dr_k \right)^{1/2},$$

and the sum is taken over all the multi-indexes $\alpha = (\alpha_1, \ldots, \alpha_k) \in \{1, \ldots, d\}^k$. We denote with $\mathbb{D}^{k,p}$ the completion of $\mathcal{S}$ with respect to the seminorms $|| \cdot ||_{k,p}$, and we set $\mathbb{D}^\infty = \cap_{p \geq 1} \cap_{k \geq 1} \mathbb{D}^{k,p}$. We may occasionally refer to $||F||_{k,p}$ as the stochastic Sobolev norm of F.

In a similar way, for any separable Hilbert space V we can define the analogous spaces $\mathbb{D}^{k,p}(V)$ and $\mathbb{D}^\infty(V)$ of V-valued random variables with the corresponding $||\cdot||_{k,p,V}$ semi-norms (the smooth functionals being now of the form $F=\sum_{j=1}^n F_j v_j$, where $F_j \in \mathcal{S}$ and $v_j \in V$). In particular, for any $\mathbb{R}^d$-valued process $(u_s; s \leq t)$ such that $u_s \in \mathbb{D}^{k,p}$ for all $s \in [0,t]$ and

$$||u||_{\mathcal{H}} + \sum_{h=1}^k ||D^{(h)}u||_{\mathcal{H}^{\otimes h+1}} < \infty \quad \mathbb{P}\text{-a.s.},$$

we have

$$||u||_{k,p,\mathcal{H}} = \left(\mathbb{E}[||u||_{\mathcal{H}}^p] + \sum_{h=1}^k \mathbb{E}\left[||D^{(h)}u||^p_{\mathcal{H}^{\otimes h+1}}\right]\right)^{1/p}.$$

Finally, we denote by δ the adjoint operator of D.

One of the main applications of Malliavin calculus consists in showing that the law of a nondegenerate random vector $F=(F^1,\dots,F^m) \in (\mathbb{D}^\infty)^m$ admits an infinitely differentiable density. The property of nondegeneracy, understood in the sense of the Malliavin covariance matrix , is introduced in the following

Definition 2.2.1. *A random vector $F=(F^1,\dots,F^m) \in (\mathbb{D}^\infty)^m$, $m \geq 1$, is said to be nondegenerate if its Malliavin covariance matrix σ_F, defined by*

$$(\sigma_F)_{i,j} = < DF^i, DF^j >_{\mathcal{H}}, \quad i,j=1,\dots,m,$$

is invertible a.s. and moreover

$$\mathbb{E}[\det(\sigma_F)^{-p}] < \infty$$

for all $p \geq 1$.

The key tool to prove smoothness of the density for a nondegenerate random vector is the following integration by parts formula (cf. [54]).

Proposition 2.2.1. *Let $F=(F^1,\dots,F^m) \in (\mathbb{D}^\infty)^m$, $m \geq 1$, be a nondegenerate random vector. Let $G \in \mathbb{D}^\infty$ and $\phi \in C^\infty_{pol}(\mathbb{R}^m)$. Then for any $k \geq 1$ and any multi-index $\alpha=(\alpha_1,\dots,\alpha_k) \in \{1,\dots,m\}^k$ there exists a random variable $H_\alpha(F,G) \in \mathbb{D}^\infty$ such that*

$$\mathbb{E}\left[\partial_\alpha \phi(F) G\right] = \mathbb{E}\left[\phi(F) H_\alpha(F,G)\right], \tag{2.2.8}$$

where the $H_\alpha(F,G)$ are recursively defined by

$$H_\alpha(F,G) = H_{(\alpha_k)}(F, H_{(\alpha_1,\dots,\alpha_{k-1})}(F,G)),$$
$$H_{(i)}(F,G) = \sum_{j=1}^m \delta\left(G(\sigma_F^{-1})_{i,j} DF^j\right).$$

2.2.3 Explicit bounds on integration by parts formula for diffusion processes

The notation of this section is somehow cumbersome, as we try to keep our bounds as general and as accurate as possible. The framework will nevertheless considerably simplify in section 2.2.4, when we will give the proofs of the results stated in section 2.2.1.

Throughout this section, $X = (X_t; t \geq 0)$ will denote the unique strong solution of the SDE

$$X_t^i = x^i + \int_0^t B^i(X_s)ds + \sum_{j=1}^d \int_0^t A_j^i(X_s)dW_s^j, \quad t \geq 0, \quad i = 1, \dots, m, \tag{2.2.9}$$

where $x \in \mathbb{R}^m$ and $B^i, A_j^i \in \mathcal{C}_b^\infty(\mathbb{R}^m)$ for all $i = 1, \dots, m$ and $j = 1, \dots, d$. We assume that the diffusion matrix A satisfies the following ellipticity condition at starting point x:

(E) $$A(x)A(x)^* \geq c_* I_m,$$

for some $c_* > 0$, where $\cdot^*$ stays for matrix transposition. Without loss of generality, we will suppose $c_* < 1$.

We recall that the first-variation process of X is the matrix-valued process

$$(Y_t)_{i,j} = \frac{\partial X_t^i}{\partial x_j}, \quad i, j = 1, \dots, m$$

which satisfies the following equation, written in matrix form:

$$dY_t = I_m + \int_0^t \partial B(X_s)Y_s ds + \sum_{l=1}^d \int_0^t \partial A_l(X_s)Y_s dW_s^l,$$

where ∂B and ∂A_l are respectively the $m \times m$ matrices of components $(\partial B)_{i,j} = \partial_j B^i$ and $(\partial A_l)_{i,j} = \partial_j A_l^i$. By means of Itô's formula, one shows that Y_t is invertible and that the inverse $Z_t := Y_t^{-1}$ satisfies the equation

$$Z_t = I_m - \int_0^t Z_s \Big\{ \partial B(X_s) - \sum_{l=1}^d (\partial A_l(X_s))^2 \Big\} ds - \sum_{l=1}^d \int_0^t Z_s \partial A_l(X_s) dW_s^l. \tag{2.2.10}$$

Additional notation. For $k \geq 0$, we define

$$\begin{aligned} |B|_k &= 1 + \sum_{i=1}^m \sum_{0 \leq |\alpha| \leq k} \sup_{x \in \mathbb{R}^m} |\partial_\alpha B^i(x)|, \\ |A|_k &= 1 + \sum_{i,j} \sum_{0 \leq |\alpha| \leq k} \sup_{x \in \mathbb{R}^m} |\partial_\alpha A_j^i(x)|, \end{aligned} \tag{2.2.11}$$

where $|\alpha|$ is the length of the multi-index α. Then, for $p \geq 1$ and $t \geq 0$ we set

$$e_p(t) := e^{t^{p/2}(t^{1/2}|B|_1+|A|_1)^p} \tag{2.2.12}$$

and

$$e_p^Z(t) := e^{t^{p/2}\left(t^{1/2}(|B|_1+|A|_1^2)+|A|_1\right)^p}. \tag{2.2.13}$$

The constants in (2.2.12) and (2.2.13) naturally arise when estimating the moments of the random variables X_t, Y_t and Z_t. Indeed, the results given in the following proposition can be easily obtained from (2.2.9) and (2.2.10) applying Burkholder's inequality and Gronwall's lemma.

Proposition 2.2.2. *For every $p > 1$ there exists a positive constant $C_{p,m}$ depending on p and m but not on the bounds on B and A and their derivatives such that, for every $0 \leq s \leq t \leq T$,*

$$i)\quad \mathbb{E}\Big[\sup_{s\leq r\leq t}|X_r^i - X_s^i|^p\Big] \leq C_{p,m}(t-s)^{p/2}\Big((t-s)^{1/2}|B|_0 + |A|_0\Big)^p, \tag{2.2.14}$$

$$ii)\quad \sup_{s\leq t}\mathbb{E}\left[|(Z_s)_{i,j}|^p\right] \leq C_{p,m}e_p^Z(t)^{C_{p,m}} \tag{2.2.15}$$

for all $i, j = 1, \dots, m$.

For any $t > 0$, the iterated Malliavin derivative of X_t is the solution of a linear SDE. The coefficients of this equation are bounded, hence it is once again a straightforward application of Gronwall's lemma to show that the random variables $D^{\alpha_1,\dots,\alpha_k}_{r_1,\dots,r_k}X_t$ have moments of any order which are finite and uniformly bounded in $r_1, \dots, r_k$. This is indeed the content of [54], Th. 2.2.2. The following lemma highlights the explicit constants appearing in the estimates of the L^p-norms of the iterated derivative, expressing them in terms of the bounds (2.2.11) on A and B.

Lemma 2.2.1. *For every $k \geq 1$ and every $p > 1$ there exist a positive integer $\gamma_{k,p}$ and a positive constant $C_{k,p}$ depending on k, p but not on the bounds on B and A and their derivatives such that, for any $t > 0$,*

$$\sup_{r_1,\dots,r_k\leq t}\mathbb{E}\left[\left|D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X_t^i\right|^p\right] \leq C_{k,p}|A|_{k-1}^{kp}\left(t^{1/2}|B|_k + |A|_k\right)^{(k+1)^2 p}e_p(t)^{\gamma_{k,p}}, \tag{2.2.16}$$

for all $i = 1, \dots, m$ and $(j_1, \dots, j_k) \in \{1, \dots, d\}^k$.

The proof of this result is based on some standard but rather cumbersome computations, hence we skip it to Appendix 2.4.1. We rather give hereafter the proof of some estimates which follow easily from Lemma 2.2.1 and will be useful in the following sections.

Corollary 1. *For any $k \geq 1$ and $p > 1$, there exists a positive constant $C_{k,p}$ depending only on k and p such that, for any $t > 0$,*

$$i) \quad \mathbb{E}\left[||D^{(k)}X_t^i||^p_{\mathcal{H}^{\otimes k}}\right]^{1/p} \leq C_{k,p}\, t^{k/2}|A|^k_{k-1}\left(t^{1/2}|B|_k + |A|_k\right)^{(k+1)^2} e_p(t)^{\gamma_{k,p}}; \tag{2.2.17}$$

$$ii) \quad ||\phi(X_t)||_{k,p} \leq C_{k,p}\,|\phi|_k\left(1 + \left(t \vee t^k\right)^{1/2}\right)|A|^k_{k-1}\left(t^{1/2}|B|_k + |A|_k\right)^{(k+2)^2} e_p(t)^{k\gamma_{k,p}}, \tag{2.2.18}$$

where i) holds for $i = 1,\dots,m$ and ii) for any $\phi \in C^\infty(\mathbb{R}^m)$.

Proof. i) Employing the definition of $||\cdot||_{\mathcal{H}^{\otimes k}}$ and Lemma 2.2.1, a simple computation holds:

$$\mathbb{E}\left[||D^{(k)}X_t^i||^p_{\mathcal{H}^{\otimes k}}\right]^{1/p} \leq C_{k,p}\left\{t^{k\left(\frac{p}{2}-1\right)}\int_{[0,t]^k}\mathbb{E}\Big[\sup_{|\alpha|=k}\left|D^{\alpha_1,\dots,\alpha_k}_{r_1,\dots r_k}X_t^i\right|^p\Big]dr_1\cdots dr_k\right\}^{1/p}$$

$$\leq C_{k,p}\, t^{k/2}|A|^k_{k-1}\left(t^{1/2}|B|_k + |A|_k\right)^{(k+1)^2} e_p(t)^{\gamma_{k,p}/p},$$

hence we get bound (2.2.17).
ii) We start from the definition of $||\cdot||_{k,p}$ and write:

$$\begin{aligned} ||\phi(X_t)||_{k,p} &= \left(\mathbb{E}[|\phi(X_t)|^p] + \sum_{h=1}^k \mathbb{E}\left[||D^{(h)}\phi(X_t)||^p_{\mathcal{H}^{\otimes h}}\right]\right)^{1/p} \\ &\leq ||\phi||_0 + \sum_{h=1}^k \mathbb{E}\left[||D^{(h)}\phi(X_t)||^p_{\mathcal{H}^{\otimes h}}\right]^{1/p}. \end{aligned} \tag{2.2.19}$$

Using the notation introduced in the proof Lemma 2.2.1, we have :

$$D^{(h)}\phi(X_t) = \sum_{I_1,\dots,I_\nu=\{1,\dots,h\}} \partial_{k_1}\cdots\partial_{k_\nu}\phi(X_t)\prod_{l=1}^{\nu} D^{(card(I_l))}X_t^{k_l},$$

where, with a slight abuse of notation, we have now written $D^{(h)}$ for the generic derivative of order h. Repeatedly applying Holder's inequality for Sobolev norms and using bound (2.2.17), we get:

$$\mathbb{E}\left[||D^{(h)}\phi(X_t)||^p_{\mathcal{H}^{\otimes h}}\right] \leq c_{h,p}\sum_{\substack{h_1,\dots,h_\nu=1,\dots,h\\ h_1+\cdots+h_\nu=h}}\mathbb{E}\left[\left\|\partial_{k_1}\cdots\partial_{k_\nu}\phi(X_t)\prod_{l=1}^{\nu}D^{(h_l)}X_t^{k_l}\right\|^p_{\mathcal{H}^{\otimes h}}\right]$$

$$\leq c_{h,p}||\phi||_h^p\sum_{\substack{h_1,\dots,h_\nu=1,\dots,h\\ h_1+\cdots+h_\nu=h}}\sup_{i=1,\dots,m}\prod_{l=1}^{\nu}\mathbb{E}\left[||D^{(h_l)}X_t^i||^{2^l p}_{\mathcal{H}^{\otimes h_l}}\right]^{1/2^l}$$

$$\leq c_{h,p}||\phi||_h^p\left\{t^{h/2}|A|^h_{h-1}\left(t^{1/2}|B|_h + |A|_h\right)^{(h+2)^2}\right\}^p e_p(t)^{h\gamma_{h,p}}.$$

By means of this bound, from (2.2.19) we get the desired estimate when setting $C_{k,p} \geq \max\{c_{h,p} : h \leq k\}$. □

We need a last preliminary result on the inverse moments of the determinant of the Malliavin covariance matrix of X_t. Once again, this result is again achieved with some standard arguments but, as happens for Lemma 2.2.1, the next lemma finds out the explicit constants appearing in the estimate of the L^p-norms of $\det(\sigma_{X_t})^{-1}$.

Lemma 2.2.2. *For every $p > 1$ and $t > 0$,*

$$\mathbb{E}\left[|\det \sigma_{X_t}|^{-p}\right]^{1/p} \leq C_{p,m,d}\, e^Z_{4(mp+1)}(t)^{C_{p,m,d}} K_m(t, c_*), \tag{2.2.20}$$

where

$$K_m(t, c_*) = 1 + \left(\frac{4}{tc_*} + 1\right)^m + \frac{1}{c_*^{2(m+1)}} |A|_1^{2(m+1)} \left(t^{1/2}|B|_0 + |A|_0\right)^{2(m+1)}$$

for some positive constant $C_{p,m,d}$ depending on p, m and d but not on the bounds on B and A and their derivatives.

The proof is once again postponed to Appendix 2.4.1.

We now come to the main result of this section. We give an estimate of the L^2-norm of the random variables H_α involved in the integration by parts formula (2.2.8), when $F = X_t$. The proof follows the arguments of [17], proof of Lemma 4.11, but is given in the general setting of an integration by parts of order $k \in \mathbb{N}$, and moreover it takes advantage of the explicit bounds which have been obtained in Corollary 1 and Lemma 2.2.2.
We give this result employing some slightly more compact notation, defining:

$$P_k(t) = t^{1/2}|B|_k + |A|_k;$$
$$P_k^A(t) = |A|_k P_{k+1}(t).$$

Theorem 2.2.3. *For every $k \geq 1$ there exists a positive constant $C_k = C_{k,m,d}$ such that, for any multi-index $\alpha \in \{1, \ldots, m\}^k$, any $G \in \mathbb{D}^\infty$ and $t > 0$,*

$$\begin{aligned} ||H_\alpha(X_t, G)||_{0,2} \leq C_k\, ||G||_{k,2^{k+1}} \left(t^{-\frac{k}{2}} \vee t^{\frac{k(k-1)}{2}}\right) \left(t^m K_m(t, c_*)\right)^{\frac{k(k+3)}{2}} \\ \times \left(P_k^A\right)^{\phi_k} \left(e_8(t) \vee e_{2^{k+2}}(t)\right)^{C_k} \left(e^Z_{32m+4}(t) \vee e^Z_{2^{k+4}m+4}(t)\right)^{C_k} \end{aligned} \tag{2.2.21}$$

where $K_m(t, c_*)$ has been defined in Lemma 2.2.2, and

$$\phi_k = 3m(k+4)^2.$$

Comment 2.2.1. *Estimate (2.2.21) is rather involved. For our purposes, the most important* *elements are: the dependence with respect to time of the factor* $t^{-\frac{k}{2}} \vee t^{\frac{k(k-1)}{2}}$ *and the coefficient* P_k^A *containing the bounds on the derivatives of the coefficients. In particular, let us remark that the factor* $t^m K_m(t, c_*)$ *is bounded for* t *close to zero. Moreover, when* $t < 1$*, the factor* $t^{-\frac{k}{2}} \vee t^{\frac{k(k-1)}{2}}$ *reduces to* $t^{-\frac{k}{2}}$*.*

Proof. We write $\sigma_t = \sigma_{X_t}$ for simplicity of notation. We first use the continuity of δ (see [17], Proposition 4.5) and Holder's inequalities for Sobolev norms to obtain :

$$\begin{aligned} ||H_{(\alpha_1,\dots,\alpha_k)}(X_t, G)||_{0,2} &= ||H_{(\alpha_k)}(X_t, H_{(\alpha_1,\dots,\alpha_{k-1})}(X_t, G))||_{0,2} \\ &= \left\| \sum_{j=1}^m \delta\left(H_{(\alpha_1,\dots,\alpha_{k-1})}(X_t, G)(\sigma_t^{-1})_{\alpha_k,j} DX_t^j\right) \right\|_{0,2} \\ &\le C_m ||H_{(\alpha_1,\dots,\alpha_{k-1})}(X_t, G)||_{1,4} \sum_{j=1}^m ||(\sigma_t^{-1})_{\alpha_k,j}||_{1,8} ||DX_t^j||_{1,8,\mathcal{H}}. \end{aligned} \tag{2.2.22}$$

To estimate the last factor we can directly use the definition of $||\cdot||_{k,p,\mathcal{H}}$ and apply Corollary 1. The major part of the efforts in the rest of the proof will be targeted on the estimation of $||(\sigma_t^{-1})_{i,j}||_{k,p}$.

We claim that for any $k \ge 1$, $p > 1$ and for all $i, j = 1, \dots, m$,

$$\begin{aligned} ||(\sigma_t^{-1})_{i,j}||_{k,p} \le c_{k,p} \left(t^{-1} \vee t^{\frac{k}{2}-1}\right) (t^m K_m(t, c_*))^{1+k} \\ \times P_k^A(t)^{\phi_k' + 2(k+4)^2} e_p(t)^{c_{k,p}} e^Z_{4(mp+1)}(t)^{c_{k,p}}, \end{aligned} \tag{2.2.23}$$

where

$$\phi_k' = 2(k+1)(m-1)$$

and $c_{k,p}$ is a positive constant depending also on m, d but not on t and on the bounds on B and A and their derivatives. Iterating process (2.2.22) and repeatedly using estimates

(2.2.23) and (2.2.17), one easily obtains the desired estimate:

$$
\begin{aligned}
||H_{(\alpha_1,\dots,\alpha_k)}(X_t,G)||_{0,2} \le\ & C_{k,m,d}\, ||G||_{k,2^{k+1}} \left(t^m K_m(t,c_*)\right)^k \\
& \times \prod_{h=1}^{k} \left(t^{-1} \vee t^{\frac{h}{2}-1}\right) \left(t \vee t^h\right)^{\frac{1}{2}} \left(t^m K_m(t,c_*)\right)^h \\
& \times P_k^A(t)^{k(\phi'_k+2(k+4)^2+(k+1)^2)} \\
& \times \prod_{h=1}^{k} e_{2^{h+2}}(t)^{c_{h,m,d}}\, e^Z_{4(2^{h+2}m+1)}(t)^{c_{h,m,d}} \\
\le\ & C_{k,m,d}\, ||G||_{k,2^{k+1}} \left(t^m K_m(t,c_*)\right)^{\frac{k(k+3)}{2}} \\
& \times \left(t^{-\frac{k}{2}} \vee t^{\frac{k(k-1)}{2}}\right) P_k^A(t)^{\phi_k} \\
& \times \left(e_8(t) \vee e_{2^{k+2}}(t)\right)^{C_{k,m,d}} \left(e^Z_{32m+4}(t) \vee e^Z_{2^{k+4}m+4}(t)\right)^{C_{k,m,d}}.
\end{aligned}
$$

Proof of (2.2.23). We follow [17], proof of Lemma 4.11. We start from the definition of $||\cdot||_{k,p}$ and write:

$$
||(\sigma_t^{-1})_{i,j}||_{k,p} = \left(\mathbb{E}\left[|(\sigma_t^{-1})_{i,j}|^p\right] + \sum_{h=1}^{k} \mathbb{E}\left[||D^{(h)}(\sigma_t^{-1})_{i,j}||^p_{\mathcal{H}^{\otimes h}}\right]\right)^{1/p}. \tag{2.2.24}
$$

For the first term, we simply use Cramer's formula for matrix inversion,

$$
|(\sigma_t^{-1})_{i,j}| = (\det \sigma_t)^{-1} \sigma_t^{(i,j)},
$$

where $\sigma_t^{(i,j)}$ denotes the (i,j) minor of σ_t. We then apply Holder's inequality and bounds (2.2.17) and (2.2.20) and get

$$
\begin{aligned}
\mathbb{E}[|(\sigma_t^{-1})_{i,j}|^p] &\le c^{(1)}_{p,m} \left\{\mathbb{E}\left[\det(\sigma_t)^{-2p}\right] \mathbb{E}\left[|\sigma_t^{(i,j)}|^{-2p}\right]\right\}^{1/2} \\
&\le c^{(1)}_{p,m} \left\{\mathbb{E}\left[\det(\sigma_t)^{-2p}\right]\ \mathbb{E}\left[\sup_i ||DX^i_t||^{4(m-1)p}_{\mathcal{H}}\right]\right\}^{1/2} \\
&\le c^{(1)}_{p,m,d}\, t^{-p} \left(t^m K_m(t,c_*)\right)^p \\
&\quad \times \left\{|A|_0 \left(t^{1/2}|B|_1 + |A|_1\right)^4\right\}^{2(m-1)p} \\
&\quad \times e_p(t)^{c^{(1)}_{p,m,d}}\, e^Z_{4(mp+1)}(t)^{c^{(1)}_{p,m,d}},
\end{aligned} \tag{2.2.25}
$$

where $K_m(t,c_*)$ is the constant defined in Lemma 2.2.2. To estimate the second term, as done in[17], proof of Lemma 4.11, we iterate the chain rule for D:

$$D(\sigma_t^{-1})_{i,j} = -\sum_{a,b=1}^{m} (\sigma_t^{-1})_{i,a} D(\sigma_t)_{a,b} (\sigma_t^{-1})_{b,j}.$$

We take advantage of the notation introduced in the proof of Lemma 2.2.1 and for $(\beta_1, \ldots, \beta_k) \in \{1, \ldots, m\}^k$, $k \geq 1$, we write:

$$\begin{aligned}
&|D^{\beta_1,\ldots,\beta_k}_{r_1,\ldots,r_k}(\sigma_t^{-1})_{i,j}| \leq \\
&\qquad \sum_{I_1\cup\cdots\cup I_\nu=\{1,\ldots,k\}} \sum_{\substack{a_1,\ldots,a_\nu=1\\ b_1,\ldots,b_\nu=1}}^{m} |(\sigma_t^{-1})_{i,a_1}(\sigma_t^{-1})_{b1,a_2}\cdots(\sigma_t^{-1})_{b_{\nu-1},a_\nu}(\sigma_t^{-1})_{b_\nu,j}| \\
&\qquad\qquad \times |D^{\beta(I_1)}_{r(I_1)}(\sigma_t)_{a_1,b_1}\cdots D^{\beta(I_\nu)}_{r(I_\nu)}(\sigma_t)_{a_\nu,b_\nu}|. \qquad (2.2.26)
\end{aligned}$$

We repeatedly apply Holder's inequality for Sobolev norms to (2.2.26) and get:

$$\begin{aligned}
\mathbb{E}\left[||D^{(k)}(\sigma_t^{-1})_{i,j}||^p_{\mathcal{H}^{\otimes k}}\right] &\leq c^{(2)}_{k,p,m} \sum_{\substack{k_1,\ldots,k_\nu=1,\ldots,k\\ k_1+\cdots+k_\nu=k}} \left\{ \sup_{\substack{a,a_1,\ldots,a_\nu=1,\ldots,m\\ b,b_1,\ldots,b_\nu=1,\ldots,m}} \mathbb{E}\left[\left|\left|(\sigma_t^{-1})^{\nu+1}_{a,b}\right.\right.\right.\right. \\
&\qquad \left.\left.\left.\left. \times\ D^{(k_1)}(\sigma_t)_{a_1,b_1}\cdots D^{(k_\nu)}(\sigma_t)_{a_\nu,b_\nu}\right|\right|^p_{\mathcal{H}^{\otimes k}}\right]\right\} \\
&\leq c^{(2)}_{k,p,m} \sup_{\substack{k_1,\ldots,k_\nu=1,\ldots,k\\ k_1+\cdots+k_\nu=k}} \left\{ \sup_{a,b=1,\ldots,m} \mathbb{E}\left[|(\sigma_t^{-1})_{a,b}|^{(\nu+1)p}\right]\right. \\
&\quad \left. \times \prod_{l=1}^{\nu} \sup_{a_l,b_l=1,\ldots,m} \mathbb{E}\left[\left|\left|D^{(k_l)}(\sigma_t)_{a_l,b_l}\right|\right|^{2^l p}_{\mathcal{H}^{\otimes k_l}}\right]^{1/2^l}\right\}, \qquad (2.2.27)
\end{aligned}$$

where, as in the proof of Corollary 1, we have written $D^{(k_l)}$ for the generic derivative of order k_l. To estimate $D^{(k_l)}(\sigma_t)_{a_l,b_l}$ we use bound (2.2.17) and get:

$$\begin{aligned}
\mathbb{E}\left[||D^{(k)}(\sigma_t)_{i,j}||^p_{\mathcal{H}^{\otimes k}}\right] &\leq \mathbb{E}\left[\left|\left|\sum_{h=0}^{k}\binom{k}{h}\int_0^t D^{(h)}D_sX^i_t \cdot D^{(k-h)}D_sX^j_t\right|\right|^p_{\mathcal{H}^{\otimes k}}\right] \\
&\leq c^{(3)}_{k,p}\sum_{h=0}^{k}\mathbb{E}\left[||D^{(h)}DX^i_t||^{2p}_{\mathcal{H}^{\otimes h+1}}\right]^{1/2} \\
&\quad \times \mathbb{E}\left[||D^{(k-h)}DX^j_t||^{2p}_{\mathcal{H}^{\otimes k-h+1}}\right]^{1/2} \qquad (2.2.28)\\
&\leq c^{(3)}_{k,p}\, t^{\left(\frac{k}{2}+1\right)p} \\
&\quad \times |A|_k^{(k+2)p}\left(t^{1/2}|B|_{k+1}+|A|_{k+1}\right)^{2(k+2)^2p} e_p(t)^{2\gamma_{k+1,p}},
\end{aligned}$$

where we have once again applied Holder's inequality for Sobolev norms in the second step.

Using (2.2.28) together with (2.2.27), bound (2.2.25) and (2.2.24) and observing that $t^m K_m(t)$ is greater than one for all the values of t, we finally obtain

$$||(\sigma_t^{-1})_{i,j}||_{k,p} \leq c_{k,p} \left(t^{-1} \vee t^{\frac{k}{2}-1}\right) (t^m K_m(t, c_*))^{1+k}$$

$$\times |A|_k^{\phi'_k+k(k+2)} (t^{1/2}|B|_{k+1} + |A|_{k+1})^{\phi'_k+2(k+4)^2}$$

$$\times e_p(t)^{c_{k,p}} e^Z_{4(mp+1)}(t)^{c_{k,p}},$$

for a positive constant $c_{k,p}$ depending also on m, d. Estimate (2.2.23) follows. □

2.2.4 Proof of Theorems 2.2.1 and 2.2.2

We now come to the proof of the results stated in section 2.2.1. We recall that an $\mathbb{R}^m$-valued random vector X is said to admit a density on an open set $A \in \mathbb{R}^m$ if $\mathcal{L}_X|_A$ possesses a density, $\mathcal{L}_X$ being the law of X. This is equivalent to say that

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)p(x)dx \tag{2.2.29}$$

holds for all $f \in C_b(\mathbb{R})$ such that $supp(f) \subset A$, for some positive $p \in L^1(A)$.

We refer to the setting of section 2.2.1. We recall that $X = (X_t; t \in [0,T])$ denotes a strong solution of

$$X_t^i = x^i + \int_0^t b^i(X_s)ds + \sum_{j=1}^d \int_0^t \sigma_j^i(X_s)dW_s^j, \quad t \in [0,T], \quad i = 1, \ldots, m, \tag{2.2.30}$$

where b and σ satisfy the assumptions **(H1)-(H3)**. For $k \geq 1$ and $f \in C^k(\mathbb{R}^m)$, we denote

$$|f|_{k,B_R(y_0)} = 1 + \sum_{|\alpha| \leq k} \sup_{x \in B_R(y_0)} |\partial_\alpha f(x)|, \tag{2.2.31}$$

where the sum is taken over all the multi-index $\alpha \in \{1, \ldots, m\}^k$. Let us define the following "local" version of the constants appearing in the estimates of the previous section:

$$
\begin{aligned}
&P_k(t,y_0)=t^{1/2}|b|_{k,B_{5R}(y_0)}+|\sigma|_{k,B_{5R}(y_0)}, \quad P^\sigma_k(t,y_0)=|\sigma|_{k,B_{5R}(y_0)}P_{k+1}(t,y_0)\\
&P^Z_1(t,y_0)=t^{1/2}\big(|b|_{1,B_{5R}(y_0)}+|\sigma|^2_{1,B_{5R}(y_0)}\big)+|\sigma|_{1,B_{5R}(y_0)}\\
&P^C_m(t,y_0)=|\sigma|^{2(m+1)}_{1,B_{5R}(y_0)}\Big(t^{1/2}|b|_{0,B_{5R}(y_0)}+|\sigma|_{0,B_{5R}(y_0)}\Big)^{2(m+1)}\\
&C_m(t,y_0)=t^m+\frac{4^m}{c_{y_0}^{2(m+1)}}\Big(1+P^C_m(t,y_0)\Big)\\
&e_p(t,y_0)=\exp\Big(t^{p/2}P_1(t,y_0)^p\Big), \quad e^Z_p(t,y_0)=\exp\Big(t^{p/2}P^Z_1(t,y_0)^p\Big).
\end{aligned}
$$

In order to prove Theorem 2.2.1, we simplify this rather heavy notation introducing a constant that contains the factors appearing in estimate (2.2.21) in Theorem 2.2.3 (recall the constant ϕ_k defined there):

$$
\begin{aligned}
\Theta_k(t,y_0,\gamma)=&\,C_m(t,y_0)^{\frac{mk(mk+3)}{2}}P^\sigma_{mk}(t,y_0)^{\phi_{mk}+(mk+2)^2}\\
&\times\Big(e_8(t,y_0)\vee e_{2^{mk+2}}(t,y_0)\Big)^\gamma\Big(e^Z_{32m+4}(t,y_0)\vee e^Z_{2^{mk+4}m+4}(t,y_0)\Big)^\gamma.
\end{aligned}
$$

As addressed in section 2.2.1, the following theorem is a more detailed version of Theorem 2.2.1. In particular, it provides the explicit expression of the constant Λ_k appearing in estimates (2.2.2) and (2.2.3).

Theorem 2.2.4. *Assume* **(H1)**, **(H2)** *and* **(H3)**. *Then, for any initial condition $x\in R^m$ and any $0<t\le T$, the random vector X_t admits an infinitely differentiable density p_{t,y_0} on $B_R(y_0)$. Furthermore, for every $k\ge 1$ there exists a positive constant $C_k=C_{k,m,d}$ such that, setting*

$$\Lambda_k(t,y_0)=C_k\,R^{-mk}\Big(P_0(t,y_0)^{mk}+\Theta_k(t,y_0,C_k)\Big) \tag{2.2.32}$$

and

$$P_t(y)=\mathbb{P}\Big(\inf\{|X_s-y|:s\in[(t-1)\vee t/2,t]\}\le 3R\Big),$$

then one has

$$p_{t,y_0}(y)\le P_t(y_0)\bigg(1+\frac{1}{t^{m3/2}}\bigg)\Lambda_3(t\wedge 1,y_0) \tag{2.2.33}$$

for every $y\in B_R(y_0)$. Analogously, for any $\alpha\in\{1,\dots,m\}^k$, $k\ge 1$,

$$|\partial_\alpha p_{t,y_0}(y)|\le P_t(y_0)\bigg(1+\frac{1}{t^{m3/2}}\bigg)\Lambda_{2k+3}(t\wedge 1,y_0) \tag{2.2.34}$$

for every $y\in B_R(y_0)$.

To prove this result we rely on the following classical criterion for smoothness of laws based on a Fourier transform argument (cf. [54], Lemma 2.1.5).

Proposition 2.2.3. *Let μ be a probability law on $\mathbb{R}^m$, and $\widehat{\mu}(\xi) = \int_{\mathbb{R}} e^{i<\xi,y>}\mu(dy)$ its characteristic function. If $\widehat{\mu}$ is integrable, then μ is absolutely continuous w.r.t. the Lebesgue measure and*

$$p(y) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{-i<\xi,y>}\widehat{\mu}(\xi)d\xi \tag{2.2.35}$$

is a continuous version of its density. If moreover

$$\int_{\mathbb{R}^m} |\xi|^k|\widehat{\mu}(\xi)|d\xi < \infty \tag{2.2.36}$$

holds for any $k \in \mathbb{N}$, then p is of class C^∞ and for any multi-index $\alpha = (\alpha_1, \dots, \alpha_k) \in \{1, \dots, m\}^k$,

$$\partial_\alpha p(y) = (-i)^k \int_{\mathbb{R}} \Big(\prod_{j=1}^k \xi^{\alpha_j}\Big) e^{-i<\xi,y>}\widehat{\mu}(\xi)d\xi.$$

Proof (of Theorem 2.2.4). Step 1 ("localized" characteristic function). Fix a t in $(0,T]$. Let $\phi_R \in C_b^\infty(\mathbb{R}^m)$ be such that $1_{B_R(0)} \le \phi_R \le 1_{B_{2R}(0)}$ and $|\phi_R|_k \le 2^k R^{-k}$. We first observe that if $m_0 = \mathbb{E}[\phi_R(X_t - y_0)]$ is zero, then it just follows that $p \equiv 0$ is a density for X_t on $B_R(y_0)$. Otherwise, we consider $\mathcal{L}_{t,y_0}$ the law on $\mathbb{R}^m$ such that

$$\int_{\mathbb{R}^m} f(y)\mathcal{L}_{t,y_0}(dy) = \frac{1}{m_0}\mathbb{E}\left[f(X_t)\phi_R(X_t - y_0)\right], \tag{2.2.37}$$

for all $f \in C_b(\mathbb{R}^m)$. If $\mathcal{L}_{t,y_0}$ possesses a density, say p'_{t,y_0}, it follows that $p_{t,y_0}(y) := m_0 p'_{t,y_0}$ is a density for X_t on $B_R(y_0)$. Indeed, for any $f \in C_b$ such that $supp(f) \subset B_R(y_0)$, (2.2.37) implies

$$\begin{aligned}\int_{\mathbb{R}^m} f(y)p_{t,y_0}(y)dy &= \int_{\mathbb{R}^m} f(y)m_0 p'_{t,y_0}(y)dy \\ &= m_0\int_{\mathbb{R}^m} f(y)\mathcal{L}_{t,y_0}(dy) \\ &= \mathbb{E}\left[f(X_t)\right].\end{aligned}$$

If the characteristic function of $\mathcal{L}_{t,y_0}$

$$\widehat{p}_{t,y_0}(\xi) = \int_{\mathbb{R}^m} e^{i<\xi,y>}\mathcal{L}_{t,y_0}(dy) = \frac{1}{m_0}\mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t - y_0)\right]$$

is integrable, then by Proposition 2.2.3 $\mathcal{L}_{t,y_0}$ admits a density. Hence, we focus on the integrability of $\widehat{p}_{t,y_0}$; in particular, we show that condition (2.2.36) of Proposition 2.2.3 holds true for all $k \in \mathbb{N}$.

Moreover, the inversion formula (2.2.35) yields the representation for p_{t,y_0}

$$\begin{aligned} p_{t,y_0}(y) := m_0\, p'_{t,y_0}(y) &= \frac{m_0}{(2\pi)^m}\int_{\mathbb{R}^m} e^{-i<\xi,y>}\widehat{p}_{t,y_0}(\xi)d\xi \\ &= \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{-i<\xi,y>}\mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t-y_0)\right]d\xi. \end{aligned} \tag{2.2.38}$$

Step 2 (localization). We define the coefficients

$$\begin{aligned} \overline{b}^i(y) &= b^i(\psi(y-y_0)), \\ \overline{\sigma}^i_j(y) &= \sigma^i_j(\psi(y-y_0)), \end{aligned} \tag{2.2.39}$$

where $\psi \in C^\infty(\mathbb{R}^m;\mathbb{R}^m)$ (a *truncation* function) is defined by

$$\psi(y) = \begin{cases} y & \text{if } |y| \le 4R \\ 5\frac{y}{|y|} & \text{if } |y| \ge 5R \end{cases}$$

and $\psi(y) \in \overline{B}_{5R}(0)$ for all $y \in \mathbb{R}^m$. ψ can be defined in such a way that, for all $i = 1,\dots,m$, $||\psi^i||_1 \le 1$ and $||\psi^i||_k \le 2^{k-2}R^{-(k-1)}$ for all $k \ge 2$. As a consequence of **(H1)**, the $\overline{b}$ and $\overline{\sigma}$ defined in this way are C^∞_b-extensions of $b|_{B_{4R}(y_0)}$ and $\sigma|_{B_{4R}(y_0)}$. Furthermore, there exist constants $c = c_{k,m}$ such that

$$\begin{aligned} |\overline{b}^i|_k &\le c_{k,m}R^{-(k-1)}\,|b^i|_{k,B_{5R}(y_0)}, \\ |\overline{\sigma}^i_j|_k &\le c_{k,m}R^{-(k-1)}\,|\sigma^i_j|_{k,B_{5R}(y_0)} \end{aligned} \tag{2.2.40}$$

and by $(H2)$, for any $y \in B_{3R}(y_0)$ the matrix $\overline{\sigma}(y)$ is elliptic:

$$\overline{\sigma}\,\overline{\sigma}^*(y) \ge c_{y_0,R}I_m, \qquad y \in B_{3R}(y_0). \tag{2.2.41}$$

For $y \in \mathbb{R}^m$ we denote by $\overline{X}(y) = (\overline{X}_s(y); 0 \le s \le t)$ the unique strong solution of the equation

$$\overline{X}^i_s(y) = y^i + \int_0^s \overline{b}^i(\overline{X}_u(y))du + \sum_{j=1}^d \int_0^s \overline{\sigma}^i_j(\overline{X}_u(y))dW^j_u, \quad 0 \le s \le t, \quad i = 1,\dots,m. \tag{2.2.42}$$

Let now $0 < \delta < t/2 \wedge 1$. We employ a downcrossing argument in order to estimate the increments of X in the neighborhood of y_0, by replacing them with the increments of $\overline{X}$.

More precisely, let $\nu = \nu_{t,\delta}$ and $\tau = \tau_{t,\delta}$ be the stopping times defined by

$$\begin{aligned} \nu_{t,\delta} &= \inf\left\{s \geq t-\delta : X_s \in \overline{B}_{3R}(y_0)\right\}; \\ \tau_{t,\delta} &= \inf\left\{s \geq \nu_{t,\delta} : X_s \notin B_{4R}(y_0)\right\} \end{aligned} \tag{2.2.43}$$

and $\inf\{\emptyset\} = \infty$. Suppose that $\phi_R(X_t - y_0) > 0$, so that $X_t \in B_{2R}(y_0)$ and $\nu < t$. On this set, if $\nu > t-\delta$, then $|X_{t\wedge\tau} - X_\nu| \geq R$. This implies $|\overline{X}_{t\wedge\tau-\nu}(X_\nu) - X_\nu| = |X_{t\wedge\tau} - X_\nu| \geq R$. Here we are employing the fact that on the interval $[\nu, \tau]$, X stays in $B_{4R}(y_0)$, hence in the region where the truncated coefficients $\overline{b}, \overline{\sigma}$ coincide with the original ones b, σ. On this interval, both X and $\overline{X}$ satisfy equation (2.2.42) for which pathwise uniqueness holds, hence we can replace X by $\overline{X}$ and employ the flow property for $\overline{X}$. Notice that the flow property may not hold true for X (due to possible lack of uniqueness for the couple (b, σ)), but it always does for $\overline{X}$.
Analogously, if $\nu = t-\delta$ and $\tau \leq t$, then $|X_\tau - X_\nu| = |\overline{X}_{\tau-\nu}(X_\nu) - X_\nu| \geq R$. In both cases, $\sup_{0\leq s\leq\delta} |\overline{X}_s(X_\nu) - X_\nu| \geq R$. We conclude that

$$\begin{aligned} \{\phi_R(X_t - y_0) > 0\} = \{&\phi_R(X_t - y_0) > 0, t-\delta = \nu < t < \tau, X_{t-\delta} \in \overline{B}_{3R}(y_0)\} \\ &\cup \left\{\phi_R(X_t - y_0) > 0, \sup_{0\leq s\leq\delta} |\overline{X}_s(X_\nu) - X_\nu| \geq R\right\}, \end{aligned}$$

hence $\widehat{p}_{t,y_0}$ rewrites as:

$$\begin{aligned} m_0\widehat{p}_{t,y_0}(\xi) &= \mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t - y_0)1_{\{\phi_R(X_t-y_0)>0,\, t-\delta=\nu<t<\tau, X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}\right] \\ &+ \mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t - y_0)1_{\{\phi_R(X_t-y_0)>0,\, \sup_{0\leq s\leq\delta}|\overline{X}_s(X_\nu)-X_\nu|\geq R\}}\right]. \end{aligned}$$

Writing $1_{\{t-\delta=\nu<t<\tau, X_{t-\delta}\in\overline{B}_{3R}(y_0)\}} = \left(1 - 1_{\overline{\{t-\delta=\nu<t<\tau\}}}\right)1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}} = 1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}$ $-1_{\{t-\delta=\nu,\tau\leq t\}}1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}$, the first term can be estimated as follows:

$$\begin{aligned} &\mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t - y_0)1_{\{\phi_R(X_t-y_0)>0,\, t-\delta=\nu<t<\tau, X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}\right] \\ &\qquad\leq \left|\mathbb{E}\left[e^{i<\xi,\overline{X}_\delta(X_{t-\delta})>}\phi_R(\overline{X}_\delta(X_{t-\delta}) - y_0)1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}\right]\right| \\ &\qquad\qquad + \mathbb{P}\Big(X_{t-\delta} \in \overline{B}_{3R}(y_0), \sup_{0\leq s\leq\delta} |\overline{X}_s(X_{t-\delta}) - X_{t-\delta}| \geq R\Big) \end{aligned}$$

Now, for the first term of the previous estimate we have

$$\begin{aligned} &\left|\mathbb{E}\left[e^{i<\xi,\overline{X}_\delta(X_{t-\delta})>}\phi_R(\overline{X}_\delta(X_{t-\delta}) - y_0)1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}\right]\right| \\ &\qquad= \left|\mathbb{E}\left[\mathbb{E}\left[e^{i<\xi,\overline{X}_\delta(X_{t-\delta})>}\phi_R(\overline{X}_\delta(X_{t-\delta}) - y_0)\Big|X_{t-\delta}\right]1_{\{X_{t-\delta}\in\overline{B}_{3R}(y_0)\}}\right]\right| \\ &\qquad\leq \mathbb{P}(|X_{t-\delta} - y_0| < 3R) \times \sup_{y\in B_{3R}(y_0)}\left|\mathbb{E}\left[e^{i<\xi,\overline{X}_\delta(y)>}\phi_R(\overline{X}_\delta(y) - y_0)\right]\right|. \end{aligned} \tag{2.2.44}$$

On the other hand, we claim that for all $q>0$, the following estimate holds:

$$\mathbb{P}\Big(X_{t-\delta}\in \overline{B}_{3R}(y_0),\ \sup_{0\leq s\leq \delta}|\overline{X}_s(X_{t-\delta})-X_{t-\delta}|\geq R\Big)\vee \mathbb{P}\Big(\phi_R(X_t-y_0)>0,\ \sup_{0\leq s\leq \delta}|\overline{X}_s(X_\nu)-X_\nu|\geq R\Big)$$
$$\leq c_{q,m}R^{-q}\,\delta^{q/2}P_0(\delta,y_0)^q\,\mathbb{P}\Big(\inf_{t-\delta\leq s\leq t}|X_s-y_0|\leq 3R\Big), \quad (2.2.45)$$

for some positive constant $c_{q,m}$. Estimate (2.2.45) will be proved later on.

Step 3 (integration by parts). We apply integration by parts formula (2.2.8) to estimate the last term in (2.2.44). By (2.2.40), (2.2.41) and Lemma 2.2.2, $\overline{X}_\delta(y)$ is a smooth and nondegenerate random vector for any $\delta>0$ and $y\in B_{3R}(y_0)$. Then, for a given $k\geq 1$ we define the multi-index

$$\alpha=(\underbrace{1,\ldots,1}_{k\text{ times}},\ldots,\underbrace{m,\ldots,m}_{k\text{ times}}),$$

such that $|\alpha|=km$. Hence, recalling that $\partial_{x_k}e^{i<\xi,x>}=i\xi^k e^{i<\xi,x>}$,

$$\begin{aligned}&\Big|\mathbb{E}\Big[e^{i<\xi,\overline{X}_\delta(y)>}\phi_R(\overline{X}_\delta(y)-y_0)\Big]\Big|\\ &\leq \frac{1}{\prod_{i=1}^m|\xi^i|^k}\Big|\mathbb{E}\Big[\partial_\alpha e^{i<\xi,\overline{X}_\delta(y)>}\phi_R(\overline{X}_\delta(y)-y_0)\Big]\Big|\\ &\leq \frac{1}{\prod_{i=1}^m|\xi^i|^k}\mathbb{E}\Big[\big|H_\alpha(\overline{X}_\delta(y),\phi_R(\overline{X}_\delta(y)-y_0))\big|\Big],\end{aligned} \quad (2.2.46)$$

for any $y\in B_{3R}(y_0)$.

We need to separately estimate $||\phi(\overline{X}_\delta(y)-y_0)||_{|\alpha|,2^{|\alpha|+1}}$. By Corollary 1, this is given by

$$\begin{aligned}||\phi(\overline{X}_\delta(y)-y_0)||_{mk,2^{mk+1}}&\leq c_{k,m}R^{-mk}\big(1+\delta^{1/2}\big)|\sigma|^{mk}_{mk-1,B_{5R}(y_0)}\\ &\quad\times P_{mk}(y_0,\delta)^{(mk+2)^2}\,e_{2^{mk+1}}(\delta)^{c_{k,m}}\\ &\leq c_{k,m}R^{-mk}P^\sigma_{mk}(y_0,\delta)^{(mk+2)^2}\,e_{2^{mk+1}}(\delta)^{c_{k,m}}\end{aligned}$$

for some positive constant $c_{k,m}$. Then, from (2.2.45), (2.2.44), (2.2.46) and Theorem 2.2.3 it follows that

$$m_0|\widehat{p}_{t,y_0}(\xi)|\leq C_{k,q}P_R(\delta,t,y_0)I_{k,q}(\xi,\delta,y_0) \quad (2.2.47)$$

for some constant $C_{k,q}$ depending also on m and d, with

$$P_R(\delta,t,y_0)=\mathbb{P}\Big(\inf_{t-\delta\leq s\leq t}|X_s-y_0|\leq 3R\Big)$$

and

$$I_{k,q}(\xi,\delta,y_0)=R^{-q}\delta^{q/2}P_0(\delta,y_0)^q+\frac{R^{-mk}}{\prod_{i=1}^m|\xi^i|^k}\,\delta^{-mk/2}\,\Theta_k(\delta,y_0,C_{k,q}).$$

Estimate (2.2.47) holds simultaneously for any $\xi\in\mathbb{R}^m$, $0<\delta<t/2\wedge 1$, $q>0$ and $k\geq 1$. The constant $\Theta_k(\delta,y_0,C_{k,q})$ appears when applying estimate (2.2.21).

Step 4 (optimization). We show that for any ξ and any $l \geq 1$, δ can always be chosen so that there exist q and k such that $I_{k,q}(\xi, \delta, y_0)$ goes to zero at ∞ faster than $\left(\prod_{i=1}^m |\xi^i|\right)^{-(l+2)}$. Denoting $||\xi|| = \prod_{i=1}^m |\xi^i|$, we set

$$\delta := \delta(\xi) = t/2 \wedge 1 \wedge ||\xi||^{-a}$$

for some $a > 0$ that is to be identified hereafter. For this choice of δ,

$$P_R(\delta(\xi), t, y_0) \leq \mathbb{P}\Big(\inf_{t/2 \vee (t-1) \leq s \leq t} |X_s - y_0| \leq 3R\Big) = P_t(y_0)$$

and

$$\begin{aligned} I_{k,q}(\xi, \delta(\xi), y_0) \leq R^{-q}\Big(||\xi||^{-\frac{qa}{2}} \wedge (t \wedge 1)^{\frac{q}{2}}\Big) P_0(t \wedge 1, y_0)^q \\ + R^{-mk}\Big(||\xi||^{-k(1-\frac{ma}{2})} \vee ||\xi||^{-k}(t \wedge 1)^{-\frac{mk}{2}}\Big)\Theta_k(t \wedge 1, y_0, C_{k,q}), \quad (2.2.48) \end{aligned}$$

since $\delta \to P_0(\delta, y_0)$ and $\delta \to \Theta_k(\delta, y_0, C_{k,q})$ are increasing, hence $P_0(\delta(\xi), y_0) \leq P_0(t \wedge 1, y_0)$ and the same holds for Θ_k.

We consider the leading terms determining the decay of $I_{k,q}(\xi, \delta(\xi), y_0)$ with respect to ξ and impose

$$\frac{qa}{2} = k(1 - \frac{ma}{2}). \quad (2.2.49)$$

Setting $a = 1/m$, (4.2.7) yields $q = mk$, hence $\frac{qa}{2} = k(1 - \frac{ma}{2}) = \frac{k}{2}$. Therefore, we get the bound

$$\begin{aligned} I_{k,q_k^*}(\xi, \delta(\xi), y_0) \leq R^{-mk}\Big(||\xi||^{-k/2} \wedge (t \wedge 1)^{mk/2}\Big) P_0(t \wedge 1, y_0)^{mk} \\ + R^{-mk}\Big(||\xi||^{-k/2} \vee (t \wedge 1)^{-mk/2}||\xi||^{-k}\Big)\Theta_k(t \wedge 1, y_0, C_{k,q_k^*}) \quad (2.2.50) \end{aligned}$$

with $q_k^* = mk$. Estimate (2.2.50) holds for any $k \geq 1$ and $\xi \neq 0$, hence it proves that the function $p_{t,y_0}(y)$ defined in (2.2.38) is in fact well defined and infinitely differentiable with respect to y.

Let us come to estimate (2.2.33). We take (2.2.38) and cut off the integration over a region I of finite Lebesgue measure on which $||\xi|| = \prod_{i=1}^m |\xi^i|$ remains smaller than a given constant. That is, we write:

$$\begin{aligned}
p_{t,y_0}(y) &= \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{-i<\xi,y>}\mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t-y_0)\right]d\xi \\
&\leq \frac{1}{(2\pi)^m}\Big[\int_I \mathbb{E}\left[\phi_R(X_t-y_0)\right]d\xi \\
&\qquad + \int_{I^c} e^{-i<\xi,y>}\mathbb{E}\left[e^{i<\xi,X_t>}\phi_R(X_t-y_0)\right]d\xi\Big] \\
&\leq \frac{1}{(2\pi)^m}\Big[\mathbb{P}(|X_t-y_0|<2R)\lambda_m(I) \\
&\qquad + C_{k,q_k^*}P_t(y_0)\int_{I^c} I_{k,q_k^*}(\xi,\delta(\xi),y_0)d\xi\Big],
\end{aligned}$$

where λ_m denotes the Lebesgue measure on $\mathbb{R}^m$. As we have seen, the last term is such that

$$\begin{aligned}
\int_{I^c} I_{k,q_k^*}(\xi,\delta(\xi),y_0)d\xi &\leq R^{-mk}P_0(t\wedge 1,y_0)^{mk}\int_{I^c}|\xi|^{-k/2}d\xi \\
&\quad + R^{-mk}\Theta_k(t\wedge 1,y_0,C_{k,q_k^*})\Big((t\wedge 1)^{-mk/2}\int_{I^c\cap\{\xi:|\xi|<(t\wedge 1)^{-m}\}}|\xi|^{-k}d\xi \\
&\quad + \int_{I^c\cap\{\xi:|\xi|\geq(t\wedge 1)^{-m}\}}|\xi|^{-k/2}d\xi\Big).
\end{aligned}$$

Now, since

$$\int_{I^c\cap\{\xi:|\xi|\geq(t\wedge 1)^{-m}\}}|\xi|^{-k/2}d\xi \leq \int_{I^c}|\xi|^{-k/2}d\xi = c_k^{(1)} < \infty$$

and

$$(t\wedge 1)^{-mk/2}\int_{I^c\cap\{\xi:|\xi|<(t\wedge 1)^{-m}\}}|\xi|^{-k}d\xi \leq (t\wedge 1)^{-mk/2}c_k^{(2)} < \infty$$

hold for any $k\geq 3$, we then take $k=3$ and get the estimate

$$p_{t,y_0}(y)\leq C^*\,P_t(y_0)\Big[1+R^{-3m}\Big(P_0(t\wedge 1,y_0)^{3m}+(t\wedge 1)^{-3m/2}\Theta_3(t\wedge 1,y_0,C_{m,d})\Big)\Big]$$

for every $y\in B_R(y_0)$, for a positive constant C^*: estimate (2.2.33) then follows. For estimate (2.2.34) on the derivatives we proceed in the same way, observing that for $\alpha\in\{1,\ldots,m\}^l$, $|\xi|^{-k/2}\times\prod_{j=1}^l|\xi^{\alpha_j}|$ is integrable at ∞ as soon as $k\geq 2l+3$.

Proof of (2.2.45). Let μ be a stopping time such that $t-\delta\leq\mu\leq t$. For any $q>0$, we have

$$\mathbb{P}\Big(X_\mu\in\overline{B}_{3R}(y_0),\ \sup_{0\leq s\leq\delta}|\overline{X}_s(X_\mu)-X_\mu|\geq R\Big)$$

$$
\begin{aligned}
&= \mathbb{E}\Big[\mathbb{E}\Big[1_{\{\sup_{0\le s\le\delta}|\overline{X}_s(X_\mu)-X_\mu|\ge R\}}\Big|X_\mu\Big]1_{\{X_\mu\in\overline{B}_{3R}(y_0)\}}\Big] \\
&\le \mathbb{P}(X_\mu\in\overline{B}_{3R}(y_0))\times\sup_{y\in\overline{B}_{3R}(y_0)}\mathbb{P}\Big(\sup_{0\le s\le\delta}|\overline{X}_s(y)-y|^q\ge R^q\Big) \\
&\le R^{-q}\,\mathbb{P}(X_\mu\in\overline{B}_{3R}(y_0))\times\sup_{y\in\overline{B}_{3R}(y_0)}\mathbb{E}\Big[\sup_{0\le s\le\delta}|\overline{X}_s(y)-y|^q\Big].
\end{aligned}
\tag{2.2.51}
$$

Using boundedness of coefficients of (2.2.42), it is easy to show that

$$
\mathbb{E}\Big[\sup_{0\le s\le\delta}|\overline{X}_s(y)-y|^q\Big]\le c_{q,m}\,\delta^{q/2}P_0(\delta,y_0)^q
$$

for some positive constant $c_{q,m}$. Estimate (2.2.45) follows applying (2.2.51) at first for $\mu=t-\delta$ and then, using $\{\phi_R(X_t-y_0)>0\}\subseteq\{\nu\le t\}$, for $\mu=\nu\wedge t$. □

As addressed in section 2.2.1, the constants appearing in the definition of Λ_k (2.2.32) can be considerably simplified under assumptions **(H1)'** and **(H4)**, resulting in some polynomial-type bounds. The following result corresponds to Theorem 2.2.2: in the current statement, we explicitly give the expression of the exponent $q'_k(q)$ appearing in bound (2.2.6).

Theorem 2.2.5. *Assume* **(H1)'** *and* **(H3)**.

(a) For any initial condition $x\in\mathbb{R}^m$ and any $0<t\le T$, X_t admits a smooth density on $\mathbb{R}^m\setminus\overline{B}_\eta(0)$.

*(b) Assume **(H4)** as well. Then the constant Λ_k defined in Theorem 2.2.4 is such that*

$$
\Lambda_k(t,y_0)\le C_{k,T}\big(1+|y_0|^{q'_k(q)}\big),
\tag{2.2.52}
$$

for every $0<t\le T$ and every $|y|>\eta+5$. The exponent $q'_k(q)$ is worth

$$
q'_k(q)=mk(\overline{q}+2q)(m+1)(mk+3)+2q(\phi_{mk}+(mk+2)^2).
$$

(c) If moreover $\sup_{0\le s\le t}|X_s|$ has finite moments of all orders, then for every $p>0$ and every $k\ge1$ there exist positive constants $C_{k,p,T}$ such that

$$
\begin{aligned}
|p_t(y)| &\le C_{3,p,T}\left(1+\frac{1}{t^{m3/2}}\right)|y|^{-p} \\
|\partial_\alpha p_t(y)| &\le C_{k,p,T}\left(1+\frac{1}{t^{m(2k+3)/2}}\right)|y|^{-p},\qquad \alpha\in\{1,\dots,m\}^k
\end{aligned}
\tag{2.2.53}
$$

for every $0<t\le T$ and every $|y|>\eta+5$.

The $C_{k,p,T}$ are positive constants depending also on m,d and on bounds (2.2.4) and (2.2.5) on the coefficients.

Proof. *(a)* We do not need any more to distinguish between y_0 and the (close) point y where the density is evaluated, hence we just set $y = y_0$ and consider suitable radii. For $|y| > \eta$, we set $R_y = \frac{1}{10} dist(y, \overline{B}_\eta(0)) \wedge 1$. By **(H1)**', b and σ are of class C_b^∞ on $B_{5R_y}(y)$ and satisfy (H2) on $B_{3R_y}(y)$. From Theorem 2.2.4 it follows that X_t admits a smooth density on $B_{R_y}(y)$. This holds true for every ball $B_{R_y}(y)$ with center y in $\mathbb{R}^m \setminus \overline{B}_\eta(0)$, hence statement (a) follows.

(b) Without loss of generality, we take $R = 1$. As a consequence of (2.2.4), the constants introduced before Th. 2.2.4 can be bounded as follows, for $0 \le t \le T$ and $|y| > \eta + 5$:

$$\begin{aligned}
&P_k(t,y) \le c_k^{(1)}(1 + (|y| + 5)^q) \le c_k^{(1)}|y|^q \\
&P_k^\sigma(t,y) \vee P_1^Z(t,y) \le c_k^{(1)}|y|^{2q} \\
&P_m^C(t,y) \le c^{(1)}|y|^{2q(m+1)}(|y|^q + |y|^q)^{2(m+1)} \le c^{(1)}|y|^{4q(m+1)} \\
&C_m(t,y) \le \frac{c^{(1)}}{C_0^{2(m+1)}}|y|^{2\overline{q}(m+1)}|y|^{4q(m+1)} \le c^{(1)}|y|^{2(\overline{q}+2q)(m+1)},
\end{aligned} \tag{2.2.54}$$

for some constants $c^{(1)}$ and $c_k^{(1)}$ depending also on m, q, T and on the bounds on b, σ and their derivatives in (2.2.4) and (2.2.5).

The exponential factors e and e^Z must be treated on a specific basis. Indeed, $e_\cdot(t,y)$ and $e_\cdot^Z(t,y)$ may explode when $|y| \to +\infty$. Nevertheless, explosion can be avoided stepping further into the optimization procedure set up in the proof of Th. 2.2.4. More precisely, we restart from step 4 and force the state variable y to appear in the choice of δ, setting

$$\delta(\xi, y) = t/2 \wedge 1 \wedge |\xi|^{-a} \wedge |y|^{-4q}.$$

Now, whatever the value of p is, e_p and e_p^Z are reduced to

$$\begin{aligned}
e_p(\delta(\xi,y)) \le e_p(1 \wedge |y|^{-4q}, y) &\le \exp\left(c_1^{(1)}\left(1 \wedge |y|^{-2qp}\right)\left(1 + 1 \wedge |y|^{-2q}\right)^p |y|^{qp}\right) \\
&\le \exp\left(c_p\left(1 \wedge |y|^{-2qp}\right)|y|^{qp}\right) \\
&\le \exp\left(c_p\right)
\end{aligned}$$

and

$$\begin{aligned}
e_p^Z(\delta(\xi,y)) \le e_p^Z(1 \wedge |y|^{-4}, y) &= \exp\left(c_1^{(1)}\left(1 \wedge |y|^{-2qp}\right)\left(1 + 1 \wedge |y|^{-2q}\right)^p |y|^{2qp}\right) \\
&\le \exp\left(c_p\right).
\end{aligned}$$

We then perform the integration over ξ as done in the last step of the proof of Th. 2.2.4, and employing (2.2.54) we obtain estimate (2.2.52) for Λ_k, for $|y| > \eta + 5$. The value of $q_k'(q)$ is obtained from the definition of Θ_k and (2.2.54).

(c) From boundedness of moments of $\sup_{s \le t}|X_s|$, for any interval $I_t \subseteq [0,t]$ one can easily deduce the estimate

$$\mathbb{P}\big(\inf\{|X_s - y| : s \in I_t\} \le 3\big)$$

$$
\begin{aligned}
&\le \mathbb{P}\big(\sup\{|X_s| : s \in I_t\} \ge |y| - 3\big) \\
&\le \frac{1}{(|y|-3)^r}\mathbb{E}\Big[\sup_{s\le t}|X_s|^r\Big] \\
&\le c^{(2)}_{r,T}\frac{1}{|y|^r},
\end{aligned} \tag{2.2.55}
$$

for any $r > 0$, $0 \le t \le T$ and $|y|$ greater than, say, 5. It is then easy to obtain the desired estimate on p_t with Theorem 2.2.4: for a given $p > 0$, we employ (2.2.33) with $y_0 = y$ and (2.2.55) with $r > p + q'_3(q)$. Similarly, to obtain the estimate on derivatives one employs (2.2.34) and (2.2.55) with $r > p + q'_{2k+3}(q)$.

2.3 Application to square root-type diffusions: a CIR/CEV process with local coefficients

We apply our results to the solution of equation (2.1.1) We will be able to refine the polynomial estimate on the density at $+\infty$ giving exponential-type upper bounds. Under some additional assumptions on the coefficients, we also study the asymptotic behaviour of the density at zero, i.e. the point where the diffusion coefficient is singular.

We first collect some basic facts concerning the solution of (2.1.1). Let us recall the SDE

$$
\begin{cases}
dX_t = (a(X_t) - b(X_t)X_t)dt + \gamma(X_t)X_t^\alpha dW_t, \quad t \ge 0, \quad \alpha \in [1/2, 1) \\
X_0 = x \ge 0.
\end{cases} \tag{2.3.1}
$$

When $\alpha = 1/2$ and a, b and γ are constant, the solution to (2.3.1) is the celebrated Cox-Ingersoll-Ross process (see [16]), appearing in finance as a model for short interest rates. It is well known that, in spite of the lack of globally Lipschitz-continuous coefficients, existence and uniqueness of strong solutions hold for the equation of a CIR process. If $a \ge 0$, the solution stays a.s. in $\mathbb{R}_+ = [0, +\infty)$; furthermore, a solution starting at $x > 0$ stays a.s. in $\mathbb{R}_> = (0, +\infty)$ if the Feller condition $2a \ge \gamma^2$ is achieved (cf. [45] for details). The following proposition gives the (straightforward) generalization of the previous statements to the case of coefficients a, b, γ that are functions of the underlying process. The proof is left to Appendix 2.4.2.

Proposition 2.3.1. *Assume*

(s0) $\alpha \in [1/2, 1)$*;* a, b *and* $\gamma \in C^1_b$ *with* $a(0) \ge 0$ *and* $\gamma(x)^2 > 0$ *for every* $x > 0$*.*

Then, for any initial condition $x \ge 0$ *there exist a unique strong solution to (2.3.1) which is such that* $\mathbb{P}(X_t \ge 0; t \ge 0) = 1$*. Let then* $x > 0$ *and* $\tau_0 = \inf\{t \ge 0 : X_t = 0\}$*, with* $\inf\{\emptyset\} = \infty$*.*

- *If* $\alpha > 1/2$ *and*

(s1)' *$a(0) > 0$ and $z \mapsto \frac{1}{\gamma^2(z) z^{2\alpha-1}}$ is integrable at 0^+,*

then

$$\mathbb{P}(\tau_0 = \infty) = 1. \tag{2.3.2}$$

- *If $\alpha = 1/2$ and*

  (s1) *$\frac{1}{\gamma^2}$ is integrable at zero,*

  (s2) *there exists $\overline{x} > 0$ such that $\frac{2a(x)}{\gamma(x)^2} \geq 1$ for $0 < x < \overline{x}$,*

  then the same conclusion on τ_0 holds.

When X is a CIR process, the moment-generating function of X_t can be computed explicitly, leading to the knowledge of the density. Setting $L_t = (1 - e^{-bt})\gamma^2/4b$, then X_t/L_t follows a non central chi-square law with $\delta = 4a/\gamma^2$ degrees of freedom and parameter $\zeta_t = 4xb/(\gamma^2(e^{bt} - 1))$ (recall that here x is the initial condition). The density of X_t is then given by (cf. [45]):

$$p_t(y) = \frac{e^{-\zeta_t/2}}{2^{\delta/2} L_t} e^{-y/(2L_t)} \left(\frac{y}{L_t}\right)^{\delta/2-1} \sum_{n=0}^{\infty} \frac{\left(\frac{y}{4L_t}\right)^n}{n!\Gamma(\delta/2 + n)} \zeta_t^n, \quad y > 0.$$

We incidentally remark that p_t is in general unbounded, since $y^{\delta/2-1}$ diverges at zero when $\delta/2 - 1 = 2a/\gamma^2 - 1$ is negative (in fact, fixed a value of $\delta/2 - 1$, there exists a $n \geq 0$ such that $\frac{d^n}{dy^n} p_t$ is unbounded).

The standard techniques of Malliavin calculus cannot be directly applied to study the existence of a smooth density for the solution of (2.3.1), as the diffusion coefficient in general is not (depending on γ) globally Lipschitz continuous. Actually, E. Alos and C. O. Ewald [1] have shown that if X is CIR process, then X_t, $t > 0$, belongs to $\mathbb{D}^{1,2}$ when the Feller condition $2a \geq \gamma^2$ is achieved. Higher order of differentiability (in the Malliavin sense) can be proven, requiring a stronger condition on a and γ, and the authors apply these results to option pricing within the Heston model. If we are interested in density estimation, the results of the previous sections allow to overcome the problems related to the singular behaviour of the diffusion coefficient and to directly establish the existence of a smooth density, independently from any Feller-type condition (provided that (s0) is satisfied). More precisely, we can give the following preliminary result:

Proposition 2.3.2 (Preliminary result)**.** *Assume (s0) and let a, b, γ be of class C_b^∞. Let $X = (X_t; t \geq 0)$ be the strong solution of (2.3.1) starting at $x \geq 0$. For any $t > 0$, X_t admits a smooth density p_t on $(0, +\infty)$. p_t is such that $\lim_{y\to\infty} p_t(y) y^p = 0$ for any $p > 0$.*

Proof. It is easy so see that, under the current assumptions, the drift and diffusions coefficients of (2.3.1) satisfy **(H1)'** with $\eta = 0$ and **(H4)** with $q = 1$. **(H3)** holds as well, by Proposition

2.3.1. As the coefficients have sub-linear growth, for any $t > 0$, $\sup_{s\le t} X_s$ has finite moments of any order. The conclusion follows from Theorem 2.2.5 (c).

2.3.1 Exponential decay at ∞

In order to further develop our study of the density, we could take advantage of some of the generalized-chaining tools settled by Viens and Vizcarra in [63]. In particular notice that, in order to estimate the density by means of Theorem 2.2.5, we need to deal with the probability term $P_t(y)$ appearing therein. For our present purposes, we can rely on alternative strategies involving time-change arguments and the existence of quadratic exponential moments for suprema of Brownian motions (Fernique's theorem) in the current section, and a detailed analysis of negative moments of the process X in section 2.3.2.

From now on, condition (s0) is assumed, the coefficients a, b, γ are of class C_b^∞ and $(X_t; t \le T)$ denotes the unique strong of (2.3.1) on $[0, T]$, $T > 0$. We make explicit the dependence with respect to the initial condition denoting $p_t(x, \cdot)$ the density at time t of X starting at $x \ge 0$. The following result improves Proposition 2.3.2 in the estimate of the density for $y \to \infty$.

Proposition 2.3.3. *Assume that*

$$\underline{\lim}_{x\to\infty} b(x)x^{1-\alpha} > -\infty. \tag{2.3.3}$$

Then there exist positive constants γ_0 and $C_k(T)$, $k \ge 3$, such that

$$p_t(x, y) \le C_3(T)\Big(1 + \frac{1}{t^{3/2}}\Big) \exp\Big(-\gamma_0 \frac{(y - x)^{2(1-\alpha)}}{2Ct}\Big) \tag{2.3.4}$$

and

$$p_t^{(k)}(x, y) \le C_k(T)\Big(1 + \frac{1}{t^{(2k+3)/2}}\Big) \exp\Big(-\gamma_0 \frac{(y - x)^{2(1-\alpha)}}{2Ct}\Big) \tag{2.3.5}$$

for every $y > x + 1$, with $C = 2^{3-2\alpha} + 2|\gamma|_0^2(1 - \alpha)^2$. The $C_k(T)$ also depend on α and on the coefficients a, b and γ.

Remark 2.3.1. *In the case of constant coefficients and $a = b = 0$, the bound (2.3.4) can be compared to the density of the CEV process as provided for example in [23], Th. 1.6 (see also the references therein). The comparison shows that our estimate is in the good range on the log-scale.*

Proof. In the spirit of Lamperti's change-of-scale argument (cf. [40], pag 294), let $\varphi \in C^2((0, \infty))$ be defined by

$$\varphi(x) = \int_0^x \frac{1}{|\gamma|_0 y^\alpha} dy = \frac{1}{|\gamma|_0(1 - \alpha)} x^{1-\alpha},$$

so that $\varphi'(x) = \frac{1}{|\gamma|_0 x^\alpha}$. Let moreover $\theta \in C_b^\infty(\mathbb{R})$ be such that $1_{[2,\infty)} \leq \theta \leq 1_{[1,\infty)}$ and $\theta' \leq 1$. We set

$$\rho(x) = \begin{cases} \theta(x)\varphi(x) & x > 0 \\ 0 & x \leq 0, \end{cases}$$

so that ρ is of class $C^2(\mathbb{R})$. We define the auxiliary process $Y_t = X_t - x$, $t \geq 0$, which is such that $\mathbb{P}(Y_t \geq -x) = 1$, $t \geq 0$. An application of Itô's formula yields:

$$\rho(Y_t) = \int_0^t f(Y_s)ds + M_t,$$

where

$$f(y) = \rho'(y)(a(y) - b(y)y) + \frac{1}{2}\rho''(y)\gamma(y)^2 y^{2\alpha}$$

and

$$M_t = \int_0^t \rho'(Y_s)\gamma(Y_s)Y_s^\alpha dW_s.$$

The key point is the fact that f is bounded from above on $(-x, \infty)$ and, on the other hand, M is a martingale with bounded quadratic variation. Indeed, f is continuous, it is zero for $y \leq 1$ and for $y > 2$ one has:

$$f(y) = \frac{a(y)}{|\gamma|_0 y^\alpha} - \frac{b(y)}{|\gamma|_0} y^{1-\alpha} - \frac{\alpha}{2}\frac{\gamma(y)^2}{|\gamma|_0} y^{\alpha-1},$$

hence, recalling that a is bounded, $\overline{\lim}_{y\to\infty} f(y) < \infty$ is ensured by condition (2.3.3). Then we set $C_1 = \sup_{y \geq 0} f(y)$. For M, one has:

$$\begin{aligned} \langle M \rangle_t &= \int_0^t \rho'(Y_s)^2 \gamma(Y_s)^2 Y_s^{2\alpha} ds \\ &\leq \int_0^t \big(\varphi(2) + \varphi'(Y_s)\big)^2 \gamma(Y_s)^2 Y_s^{2\alpha} ds \\ &\leq 2(\varphi(2)^2 + 1)t, \end{aligned} \tag{2.3.6}$$

hence we set $C_2 = 2(\varphi(2)^2 + 1) = 2\big(\frac{2^{2(1-\alpha)}}{|\gamma|_0^2(1-\alpha)^2} + 1\big)$. Now, since ρ is strictly increasing, $\{Y_t > y\} = \{\rho(Y_t) > \rho(y)\}$ for any $y > 0$. Moreover, $\{\rho(Y_t) > \rho(y)\} \subseteq \{M_t + C_1 t > \rho(y)\} \subseteq \{2M_t^2 + 2C_1^2 t^2 > \rho(y)^2\}$. We set $I_t = [(t-1) \vee t/2, t]$ and $\tau = \inf\{s \geq 0 : Y_s \geq 3/2\}$. The quadratic variation of M is strictly increasing after τ, since $(\rho(Y_{\tau\vee t})\gamma(Y_{\tau\vee t})Y_{\tau\vee t}^\alpha)^2 > 0$: hence, by Dubins & Schwartz theorem (cf. Th. 3.4.6 in [40]) there exists a 1-dimensional Brownian motion $(b_t; t \geq 0)$ such that $M_{\tau\vee t} = b_{\langle M \rangle_{\tau\vee t}}$. Clearly one has $\{Y_t > 2\} \subseteq \{\tau < t\}$, so that for

$y > 2$

$$
\begin{aligned}
\overline{P}_t(y) = \mathbb{P}(\exists s \in I_t : Y_s > y) &\le \mathbb{P}(\exists s \in I_t : Y_s > y, \tau < s) \\
&\le \mathbb{P}(\exists s \in I_t : 2M_s^2 + 2C_1^2 s^2 > \rho(y)^2, \tau < s) \\
&\le \mathbb{P}\Big(\sup_{\tau < s \le t} \Big(2M_s^2 + 2C_1^2 s^2\Big) > \rho(y)^2 \Big) \\
&\le \mathbb{P}\Big(\sup_{0 < s \le t} b^2_{\langle M \rangle_s} + C_1^2 t^2 > \frac{1}{2}\rho(y)^2 \Big) \\
&\le \mathbb{P}\Big(\sup_{s \le C_2 t} b_s^2 + C_1^2 t^2 > \frac{1}{2}\rho(y)^2 \Big).
\end{aligned}
$$

We now employ the scaling property for the Brownian motion $(b_s; s \ge 0) \sim (\sqrt{a}\, b_{s/a}; s \ge 0)$, $a > 0$, and Fernique's Theorem (cf. [39], pg 402). The latter tells that there exists a positive constant γ_0 such that $\exp(\gamma_0 \sup_{s \le 1} b_s^2)$ is integrable, hence

$$
\begin{aligned}
\overline{P}_t(y) &\le \mathbb{P}\Big(\gamma_0 \sup_{s \le 1} b_s^2 + \gamma_0 \frac{C_1^2}{C_2} t > \frac{\gamma_0}{2C_2 t}\rho(y)^2 \Big) \\
&\le \exp\Big(-\gamma_0 \frac{\rho(y)^2}{2C_2 t} \Big) \mathbb{E}\Big[e^{\gamma_0 \frac{C_1^2}{C_2} t + \gamma_0 \sup_{s \le 1} b_s^2} \Big] \\
&\le C_0 \exp\Big(-\gamma_0 \frac{\rho(y)^2}{2C_2 t} + \gamma_0 \frac{C_1^2}{C_2} t \Big),
\end{aligned}
$$

where $C_0 = \mathbb{E}[\exp(\gamma_0 \sup_{s \le 1} b_s^2)]$ is a universal constant. The estimates on the density of X_t and its derivatives now follow from Theorem 2.2.4 (estimates (2.2.33) and (2.2.34)) and Theorem 2.2.5 (b), using $X_t - x = Y_t$, the value of the constant C_2 and taking e.g. $R = 1/6$.

□

2.3.2 Asymptotics at 0

We have established conditions under which the solution of (2.3.1) admits a smooth density p_t on $(0, +\infty)$. According to Proposition 2.3.1, the process remains almost surely in $\mathbb{R}_+$: this trivially means that for any $t > 0$, X_t has an identically zero density on $(-\infty, 0)$, which can be extended to 0 when $\tau_0 = \infty$ a.s. We are now wondering what are sufficient conditions for p_t to converge to zero at the origin, hence providing the existence of a continuous (eventually differentiable, eventually C^∞) density on the whole real line.

What we have in mind is the application of Theorem 2.2.5 to the inversed process $Y_t = \frac{1}{X_t}$ (considered on the event $\{\tau_0 = \infty\}$). An application of Itô's formula yields

$$
dY_t = J_\alpha(Y_t)dt - \widehat{\gamma}(Y_t) Y_t^{2-\alpha} dW_t, \tag{2.3.7}
$$

where

$$
J_\alpha(Y_t) = -\widehat{a}(Y_t) Y_t^2 + \widehat{b}(Y_t) Y_t + \widehat{\gamma}(Y_t)^2 Y_t^{3-2\alpha}
$$

with the notation $\widehat{f}(y) = f(1/y)$, $y > 0$, for $f = a, b, \sigma$. Equation (2.3.7) has super-linear coefficients, in particular condition (2.2.4) of Theorem (2.2.2) holds with $q = 2$. Willing to apply Theorem 2.2.5 (c), we first need some preliminary result on the moments of Y. Let us start with the following general Lemma.

Lemma 2.3.1. *Let $A = (A_t; t \geq 0)$ be a continuous adapted real-valued process. Let $t > 0$ and $p > 0$. If there exists a positive constant C such that*

$$\mathbb{E}[|A|^p_{\tau\wedge t}] < C \quad \text{for all stopping times } \tau, \tag{2.3.8}$$

then

$$\mathbb{E}\Big[\sup_{s\leq t} |A_s|^q\Big] < \infty, \quad \forall\, 0 < q < p. \tag{2.3.9}$$

Precisely, $\mathbb{E}\big[\sup_{s\leq t} |A_s|^q\big] \leq 1 + C\frac{q}{p-q}$.

Proof. Step 1. We show that

$$\mathbb{P}\Big(\sup_{s\leq t} |A_s| > z\Big) \leq C z^{-p}, \quad \forall z > 0. \tag{2.3.10}$$

Let us define the stopping time $\tau_z = \inf\{s \geq 0 : |A_s| \geq z\}$, for $z > 0$. We have

$$\begin{aligned}\mathbb{P}\Big(\sup_{s\leq t} |A_s| \geq z\Big) &= \mathbb{P}(\tau_z \leq t) = \mathbb{P}(\tau_z \leq t, |A_{\tau_z\wedge t}| \geq z)\\ &\leq \mathbb{P}(|A_{\tau_z\wedge t}| \geq z)\\ &\leq C z^{-p},\end{aligned}$$

where we have applied Markov's inequality in the last step.

Step 2. It is now easy to show that (2.3.10) implies (2.3.9). For example, using the fact that $\mathbb{E}[B] = \int_0^\infty \mathbb{P}(|B| > z)dz$ for any positive random variable B, we have

$$\mathbb{E}\Big[\sup_{s\leq t} |A_s|^q\Big] = \int_0^\infty \mathbb{P}\Big(\sup_{s\leq t} |A_s|^q > z\Big)dz = \int_0^\infty \mathbb{P}\Big(\sup_{s\leq t} |A_s| > z^{1/q}\Big)dz \leq 1 + C\int_1^\infty z^{-p/q}dz$$

and the last integral is seen to be finite and equal to $C\frac{q}{p-q}$ if $p > q$. $\square$

The proof of the next statement is based on the previous lemma and on the techniques used in [11], proof of Lemma 2.1, that we adapt to our framework.

Lemma 2.3.2. 1) *If $\alpha > 1/2$, assume* (s1)'*. Then for any initial condition $x > 0$, for any $t > 0$ and $p > 0$*

$$\mathbb{E}\Big[\sup_{s\leq t} \frac{1}{X_s^p}\Big] \leq C, \tag{2.3.11}$$

for some positive constant C depending on x, p, t, α and on the bounds on the coefficients of equation (2.3.1).

2) *If* $\alpha = 1/2$, *then assume* (s1) *and* (s2) *and let*

$$l^* = \varliminf_{x\to 0} \frac{2a(x)}{\gamma(x)^2} > 1. \tag{2.3.12}$$

Then (2.3.11) holds for $p > 0$ *such that*

$$p + 1 < l^*. \tag{2.3.13}$$

Proof. Let τ be a stopping time, and define τ_n by $\tau_n = \inf\{t \geq 0 : X_t \leq 1/n\}$. Let $\tilde\tau_n = \tau \wedge \tau_n$. The application of Itô's formula to $X^p_{t\wedge\tilde\tau_n}$, $p > 0$, yields

$$\mathbb{E}\Big[\frac{1}{X^p_{t\wedge\tilde\tau_n}}\Big] = \frac{1}{x^p} + \mathbb{E}\int_0^{t\wedge\tilde\tau_n} \varphi(X_s)ds, \tag{2.3.14}$$

where

$$\varphi(x) = p\frac{b(x)}{x^p} + \frac{p}{x^{p+1}} \underbrace{\Big(\frac{p+1}{2}\gamma(x)^2 x^{2\alpha-1} - a(x)\Big)}_{g(x)}, \quad x > 0.$$

It is easy to see that, if

$$\varlimsup_{x\to 0} g(x), < 0 \tag{2.3.15}$$

there exists a positive constant C such that $\varphi(x) < p\frac{|b|_0}{x^p} + C$, for every $x > 0$. If (2.3.15) holds, from (2.3.14) we get

$$\mathbb{E}\Big[\frac{1}{X^p_{t\wedge\tilde\tau_n}}\Big] \leq \frac{1}{x^p} + Ct + p|b|_0 \int_0^t \mathbb{E}\Big[\frac{1}{X^p_{s\wedge\tilde\tau_n}}\Big]ds,$$

hence, by Gronwall's lemma,

$$\mathbb{E}\Big[\frac{1}{X^p_{t\wedge\tilde\tau_n}}\Big] \leq \Big(\frac{1}{x^p} + Ct\Big)e^{p|b|_0 t}. \tag{2.3.16}$$

for all $n \in \mathbb{N}$. We verify (2.3.15), distinguishing the two cases.

Case $\alpha > 1/2$. We simply observe that $\lim_{x\to 0} g(x) = -a(0) < 0$. By taking the limit $n \to \infty$ in (2.3.16) and observing that $\tau_n \to \infty$ by Proposition 2.3.1 under assumption (s1)', we obtain

$$\mathbb{E}\Big[\frac{1}{X^p_{t\wedge\tau}}\Big] \leq \Big(\frac{1}{x^p} + Ct\Big)e^{p|b|_0 t},$$

hence estimate (2.3.11) follows from Lemma 2.3.1.

Case $\alpha = 1/2$. We have $g(x) = \frac{p+1}{2}\gamma^2(x) - a(x)$. If p satisfies (2.3.13), then (2.3.12) ensures that $\varlimsup_{x\to 0} g(x) < 0$. We conclude again by taking the limit $n \to \infty$ and using Proposition 2.3.1 under assumptions (s1) and (s2) and Lemma 2.3.1. □

We are now provided with the tools to prove the following

Proposition 2.3.4. 1) *If $\alpha > 1/2$, assume* (s1)'. *Then for every $t > 0$, every $p \geq 0$ and every $k > 0$ the density p_t of X_t on $(0, +\infty)$ is such that*

$$\begin{aligned} &\lim_{y\to 0^+} y^{-p}|p_t(y)| = 0 \\ &\lim_{y\to 0^+} y^{-p}|p_t^{(k)}(y)| = 0. \end{aligned} \tag{2.3.17}$$

2) *If $\alpha = 1/2$, then assume* (s1) *and* (s2) *and define l^* as in Lemma 2.3.2. If*

$$l^* > 3 + q_3'(2) \tag{2.3.18}$$

(where $q'_{\cdot}(.)$ has been defined in Theorem 2.2.5), then

$$\lim_{y\to 0^+} y^{-p} p_t(y) = 0 \tag{2.3.19}$$

for every $0 \leq p < l^* - (3 + q_3'(2))$. Moreover, if

$$l^* > 2k + 3 + q_{2k+3}'(2), \tag{2.3.20}$$

then

$$\lim_{y\to 0^+} y^{-p}|p_t^{(k)}(y)| = 0 \tag{2.3.21}$$

for every $0 \leq p < l^* - (2k + 3 + q_{2k+3}'(2))$.

Proof. We apply Theorem 2.2.2 to $Y = 1/X$. For simplicity of notation, we write p for p_t and p_Y for the density of Y_t. As Y satisfies equation (2.3.7), from Theorem 2.2.2 (b) it follows that the bound (2.2.52) on Λ_k holds with $q_k'(q) = q_k'(2)$. Hence, from Theorem 2.2.2 (c) it follows that

$$\lim_{y'\to +\infty} p_Y(y')|y'|^{p'} = 0 \tag{2.3.22}$$

and, for a given $k \geq 0$,

$$\lim_{y'\to +\infty} p_Y^{(k)}(y')|y'|^{p'} = 0 \tag{2.3.23}$$

if $\sup_{s\leq t} Y_s$ has a finite moment of order $r > p' + q_{2k+3}'(2)$.
Now, it is easy to see that

$$p(y) = \frac{1}{y^2} p_Y\Big(\frac{1}{y}\Big) \tag{2.3.24}$$

hence, after some rather straightforward computations,

$$|p^{(k)}(y)| \leq C_k \Big(\frac{1}{y}\Big)^{2(k+1)} \sum_{j=0}^{k} \sum_{\nu=1}^{j} \frac{d^\nu}{dy^\nu} p_Y\Big(\frac{1}{y}\Big), \quad 0 < y < 1. \tag{2.3.25}$$

Once again, we distinguish the two cases.

Case $\alpha > 1/2$. If $1/2 < \alpha < 1$, by Lemma 2.3.2, (2.3.22) and (2.3.23) hold for any $p' > 0$. Then (2.3.17) easily follows from (2.3.24) and (2.3.25).

Case $\alpha = 1/2$. By Lemma 2.3.2, (2.3.18) is the condition for $\sup_{s\le t} Y_s$ to have finite moment of order strictly greater than $2 + q_3'(2) + p$, with $p < l^* - (3 + q_3'(2))$. By (2.3.24), in this case (2.3.22) holds true with $k = 0$ and $p' = 2 + p$, hence (2.3.19) holds. Similarly, by (2.3.25) estimate (2.3.21) holds if (2.3.22) holds with $p' = p + 2(k+1)$. The latter condition is achieved if $\sup_{s\le t} Y_s$ has finite moment of order strictly greater than $2(k+1) + q'_{2(k+3)}(2) + p$, which is in turn ensured by (2.3.20). □

Remark 2.3.2. *Proposition 2.3.3 states that p_t decays exponentially at infinity for any value of α, as far as condition (2.3.3) holds true. When $\alpha > 1/2$, Proposition 2.3.4 states that p_t and all its derivatives tend to zero at the origin, while the price to pay for the same conclusion to hold is higher when $\alpha = 1/2$ (cf. conditions (2.3.18) and (2.3.20), which become rapidly strong for growing values of k). With regard to this behaviour at zero, we recall that Proposition 2.3.4 only provides sufficient conditions for estimates (2.3.19) and (2.3.21) to hold. We do not give any conclusion on the behaviour of the density at zero when condition (2.3.18) (or (2.3.20) for the derivatives) fail to hold.*

2.4 Technical proofs

We collect here the proofs of the more technical results of this chapter.

2.4.1 Proofs of Lemmas 2.2.1 and 2.2.2

Proof (of Lemma 2.2.1). We refer to the notation introduced in the proof of [54], Th. 2.2.2, allowing to write the equation satisfied by the k-th Malliavin derivative in a compact form. This is stated as follows: for any subset $K = \{h_1, \ldots, h_\eta\}$ of $\{1, \ldots, k\}$, one sets $j(K) = j_{h_1}, \ldots, j_{h_\eta}$ and $r(K) = r_{h_1}, \ldots, r_{h_\eta}$. Then, one defines

$$\begin{aligned}\alpha^i_{l,j_1,\ldots,j_k}(s, r_1, \ldots, r_k) :&= D^{j_1,\ldots,j_k}_{r_1,\ldots,r_k} A^i_l(X_s)\\ &= \sum \partial_{k_1} \cdots \partial_{k_\nu} A^i_l(X_s)\\ &\quad \times D^{j(I_1)}_{r(I_1)} X^{k_1}_s \cdots D^{j(I_\nu)}_{r(I_\nu)} X^{k_\nu}_s\end{aligned}$$

and

$$\begin{aligned}\beta^i_{j_1,\ldots,j_k}(s, r_1, \ldots, r_k) :&= D^{j_1,\ldots,j_k}_{r_1,\ldots,r_k} B^i(X_s)\\ &= \sum \partial_{k_1} \cdots \partial_{k_\nu} B^i(X_s)\\ &\quad \times D^{j(I_1)}_{r(I_1)} X^{k_1}_s \cdots D^{j(I_\nu)}_{r(I_\nu)} X^{k_\nu}_s,\end{aligned}$$

where in both cases the sum is extended to the set of all partitions of $\{1,\dots,k\}=I_1\cup\cdots\cup I_\nu$. Finally, one sets $\alpha^i_j(s)=A^i_j(X_s)$. Making use of this notation, it is shown that the equation satisfied by the k-th derivative reads as:

$$D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^i_t=\sum_{\epsilon=1}^k\alpha^i_{j_\epsilon,j_1,\dots,j_{\epsilon-1},j_{\epsilon+1},\dots,j_k}(r_\epsilon,r_1,\dots,r_{\epsilon-1},r_{\epsilon+1},\dots,r_k)$$
$$+\int_{r_1\vee\cdots\vee r_k}^t\left(\beta^i_{j_1,\dots,j_k}(s,r_1,\dots,r_k)ds+\alpha^i_{l,j_1,\dots,j_k}(s,r_1,\dots,r_k)dW^l_s\right)\quad(2.4.1)$$

if $t\geq r_1\vee\cdots\vee r_k$, and $D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^i_t=0$ otherwise. We prove (2.2.16) by induction. The estimate is true for $k=1$, with $\gamma_{1,p}=2C_{1,p}$: this simply follows with an application of Burkholder's inequality and Gronwall's lemma to (2.4.1) taken for $k=1$. Let us suppose that (2.2.16) is true up to $k-1$. As done for $k=1$, we apply Burkholder's inequality to (2.4.1) and, setting $r=r_1\vee\cdots\vee r_k$, we get:

$$\begin{aligned}
&\mathbb{E}\left[|D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^i_t|^p\right]\\
&\leq C_{k,p}\Bigg\{\sum_{\epsilon=1}^k\mathbb{E}\Big[|\alpha^i_{j_\epsilon,j_1,\dots,j_{\epsilon-1},j_{\epsilon+1},\dots,j_k}(r_\epsilon,r_1,\dots,r_{\epsilon-1},r_{\epsilon+1},\dots,r_k)|^p\Big]\\
&+(t-r)^{\frac{p}{2}-1}\sum_{\substack{I_1\cup\cdots\cup I_\nu\\ \mathrm{card(I)}\leq k-1}}\int_r^t\mathbb{E}\Big[\Big((t-r)^{\frac12}|\partial_{k_1}\cdots\partial_{k_\nu}B^i(X_s)|+\sum_{l=1}^d|\partial_{k_1}\cdots\partial_{k_\nu}A^i_l(X_s)|\Big)^p\\
&\qquad\times|D^{j(I_1)}_{r(I_1)}X^{k_1}_s\cdots D^{j(I_\nu)}_{r(I_\nu)}X^{k_\nu}_s|^p\Big]ds\\
&+(t-r)^{\frac{p}{2}-1}\int_r^t\mathbb{E}\Big[\Big((t-r)^{\frac12}|\partial_kB^i(X_s)|+\sum_{l=1}^d|\partial_kA^i_l(X_s)|\Big)^p|D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^k_s|^p\Big]ds\Bigg\},\quad(2.4.2)
\end{aligned}$$

where, in the last line, we have isolated the term depending on $D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X$.
To estimate the second term in (2.4.2) we notice that, for any partition $I_1\cup\cdots\cup I_\nu$ of $\{1,\dots,k\}$ such that $\mathrm{card}(I)\leq k-1$, using (2.2.16) up to order $k-1$ we have:

$$\begin{aligned}
&\mathbb{E}\Big[\Big((t-r)^{\frac12}|\partial_{k_1}\cdots\partial_{k_\nu}B^i(X_s)|+\\
&\qquad\sum_{l=1}^d|\partial_{k_1}\cdots\partial_{k_\nu}A^i_l(X_s)|\Big)^p|D^{j(I_1)}_{r(I_1)}X^{k_1}_s\cdots D^{j(I_\nu)}_{r(I_\nu)}X^{k_\nu}_s|^p\Big]\\
&\qquad\qquad\leq C\left\{|A|^k_{k-2}(t^{1/2}|B|_k+|A|_k)^{\chi_k}\right\}^pe_p(t)^{\lambda^{(1)}_{k,p}},\quad(2.4.3)
\end{aligned}$$

where we have defined

$$\lambda^{(1)}_{k,p}:=\sup_{\substack{I_1\cup\cdots\cup I_\nu=\{1,\dots,k\}\\ card(I)\leq k-1}}\left\{\gamma_{card(I_1),p}+\cdots+\gamma_{card(I_\nu),p}\right\}$$

and

$$\chi_k = 1 + \sum_{l=1}^{\nu} (card(I_l) + 1)^2.$$

It is easy to see that

$$\chi_k \leq (k+1)^2,$$

since

$$\sum_{l=1}^{\nu} card(I_l)^2 = (k-1)^2 \sum_{l=1}^{\nu} \frac{card(I_l)^2}{(k-1)^2} \leq (k-1)^2 \sum_{l=1}^{\nu} \frac{card(I_l)}{(k-1)} = (k-1)k,$$

so that

$$\begin{aligned}\chi_k &= 1 + \sum_{l=1}^{\nu} card(I_l)^2 + 2\sum_{l=1}^{\nu} card(I_l) + \nu \\ &\leq 1 + (k-1)k + 2k + \nu \\ &\leq 1 + k^2 - k + 3k = (k+1)^2.\end{aligned}$$

To estimate the first term in (2.4.2), notice that we have as well:

$$\begin{aligned}&\mathbb{E}\left[\left|\alpha^i_{j_\epsilon, j_1, \ldots, j_{\epsilon-1}, j_{\epsilon+1}, \ldots, j_k}(r_\epsilon, r_1, \ldots, r_{\epsilon-1}, r_{\epsilon+1}, \ldots, r_k)\right|^p\right] \\ &\qquad \leq C\left\{|A|^k_{k-1}(r_\epsilon^{\frac{1}{2}}|B|_{k-1} + |A|_{k-1})^{k^2}\right\}^p e_p(t)^{\lambda^{(2)}_{k,p}}, \quad (2.4.4)\end{aligned}$$

with

$$\lambda^{(2)}_{k,p} := \sup_{I_1 \cup \cdots \cup I_\nu = \{1, \ldots, k-1\}} \left\{\gamma_{card(I_1),p} + \cdots + \gamma_{card(I_\nu),p}\right\}.$$

We remark that $\lambda_k^{(1)}$ and $\lambda_k^{(2)}$ are defined by means of the γ's up to order $k-1$.
Collecting (2.4.2), (2.4.3) and (2.4.4), we get

$$\begin{aligned}\mathbb{E}\left[|D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^i_t|^p\right] \le C_{k,p}\Bigg\{ & |A|^{kp}_{k-1}(t^{1/2}|B|_{k-1}+|A|_{k-1})^{k^2p}e_p(t)^{\lambda^{(2)}_{k,p}} \\ & + (t-r)^{\frac{p}{2}}|A|^k_{k-2}(t^{1/2}|B|_k+|A|_k)^{(k+1)^2p}e_p(t)^{\lambda^{(1)}_{k,p}} \\ & + (t-r)^{\frac{p}{2}-1}\left(t^{1/2}|B|_1+|A|_1\right)^p\sum_{k=1}^m\int_r^t\mathbb{E}\left|D^{j_1,\dots,j_k}_{r_1,\dots,r_k}X^i_s\right|^p ds\Bigg\} \\ \le\; & C_{k,p}|A|^{kp}_{k-1}(t^{1/2}|B|_k+|A|_k)^{(k+1)^2p}e_p(t)^{\lambda^{(1)}_{k,p}\vee\lambda^{(2)}_{k,p}} \\ & \times\left(1+C_{k,p}\,t^{p/2}(t^{1/2}|B|_1+|A|_1)^pe_p(t)^{C_{k,p}}\right) \\ \le\; & C_{k,p}|A|^{kp}_{k-1}\left(t^{1/2}|B|_k+|A|_k\right)^{(k+1)^2p}e_p(t)^{\lambda^{(1)}_{k,p}\vee\lambda^{(2)}_{k,p}+2C_{k,p}},\end{aligned}$$

where we have applied Gronwall's lemma to get the second inequality. The constant $C_{k,p}$ may vary from line by line, but never depends on t and on the bounds on B and A. We recursively define $\gamma_{k,p}$ by setting $\gamma_{k,p} := \lambda^{(1)}_{k,p}\vee\lambda^{(2)}_{k,p}+2C_{k,p}$, and we finally obtain (2.2.16). □

Proof (of Lemma 2.2.2). Step 1. We first use the decomposition $D_sX_t = Y_tZ_sA(X_s)$ (see for example [54]) and write

$$\begin{aligned}\sigma_{X_t} &= Y_t\int_0^t Z_sA(X_s)A(X_s)^*Z_s^*ds\,Y_t^* \\ &= Y_tU_tY_t^*,\end{aligned} \tag{2.4.5}$$

where we have set $U_t=\int_0^t Z_sA(X_s)A(X_s)^*Z_s^*ds$. Notice that U_t is a positive operator, and that for any $\xi\in\mathbb{R}^m$ we have

$$<\xi,U_t\xi>=\int_0^t <A(X_s)^*Z_s^*\xi,A(X_s)^*Z_s^*\xi>ds=\sum_{j=1}^d\int_0^t <Z_sA_j(X_s),\xi>^2.$$

From identity (2.4.5) it follows that $\det\sigma_{X_t}=(\det Y_t)^2\det U_t=(\det Z_t)^{-2}\det U_t$. Hence, applying Holder's inequality:

$$\mathbb{E}\left[|\det\sigma_{X_t}|^{-p}\right] \le \left(\mathbb{E}\left[|\det Z_t|^{4p}\right]\mathbb{E}\left[(\det U_t)^{-2p}\right]\right)^{1/2}$$
$$\le C_{p,m}\left(e^Z_{4p}(t)^m\mathbb{E}\left[(\det U_t)^{-2p}\right]\right)^{1/2}, \tag{2.4.6}$$

where in the last step we have used bound (2.2.15) on the entries of Z_t.

Step 2. Let $\lambda_t = \inf_{|\xi|=1} < \xi, U_t\xi >$ be the smallest eigenvalue of U_t, so that $\mathbb{E}\left[(\det U_t)^{-2p}\right] \le \mathbb{E}\left[\lambda_t^{-2mp}\right]$. We evaluate $\mathbb{P}(\lambda_t \le \epsilon)$.

For any ξ such that $|\xi| = 1$, using the elementary inequality $(a+b)^2 \ge a^2/2 - b^2$ we get

$$\sum_{j=1}^d < Z_sA_j(X_s), \xi >^2 \ge \frac{1}{2}\sum_{j=1}^d < A_j(x), \xi >^2 - \sum_{j=1}^d < Z_sA_j(X_s) - A_j(x), \xi >^2$$
$$\ge \frac{1}{2}c_* - \sum_{j=1}^d |Z_sA_j(X_s) - A_j(x)|^2,$$

where in the last step we have used the ellipticity assumption **(E)**. For any $\epsilon > 0$ and $a > 0$ such that $a\epsilon < t$, the previous inequality gives

$$\mathbb{P}(\lambda_t \le \epsilon) \le \mathbb{P}\left(\frac{1}{2}ac_*\epsilon - \sup_{s\le a\epsilon}\left\{a\epsilon\sum_{j=1}^d |Z_sA_j(X_s) - A_j(x)|^2\right\} \le \epsilon\right),$$

thus, if we take $a = 4/c_*$ in order to have $ac_*/2 = 2$ and apply Markov's inequality, we obtain:

$$\mathbb{P}(\lambda_t \le \epsilon) \le \mathbb{P}\left(\sup_{s\le a\epsilon}\left\{\sum_{j=1}^d |Z_sA_j(X_s) - A_j(x)|^2\right\} \ge \frac{c_*}{4}\right)$$
$$\le d^{q-1}\frac{4^q}{c_*^q}\sum_{j=1}^d \mathbb{E}\Big[\sup_{s\le a\epsilon}|Z_sA_j(X_s) - A_j(x)|^{2q}\Big], \tag{2.4.7}$$

where the last holds for all $q > 1$. Now, to estimate the last term we claim that, for all $j = 1, \ldots, d$,

$$\mathbb{E}\Big[\sup_{s\le t}|Z_sA_j(X_s) - A_j(x)|^{2q}\Big] \le C\, t^q|A|_1^{2q}\left(t^{1/2}|B|_0 + |A|_0\right)^{2q} e^Z_{4q}(t)^C, \tag{2.4.8}$$

for a constant C depending on q and m but not on the bounds on B and A. From (2.4.7) and this last estimate, it follows that

$$\mathbb{P}(\lambda_t \le \epsilon) \le C_{q,m,d}\,\frac{\epsilon^q}{c_*^{2q}}|A|_1^{2q}\left(t^{1/2}|B|_0 + |A|_0\right)^{2q} e^Z_{4q}(t)^{C_{q,m,d}},$$

for any ϵ such that $4\epsilon/c_* < 1 \wedge t$.

Step 2. We finally estimate $\mathbb{E}[\lambda_t^{-2mp}]$. We write

$$\begin{aligned}\mathbb{E}[\lambda_t^{-2mp}] &= \mathbb{E}[\lambda_t^{-2mp}1_{\{\lambda_t>1\}}] + \sum_{k=1}^{\infty}\mathbb{E}[\lambda_t^{-2mp}1_{\{1/(k+1)<\lambda_t\leq 1/k\}}] \\ &\leq 1 + \sum_{k=1}^{\infty}(k+1)^{2mp}\,\mathbb{P}(1/(k+1)<\lambda_t\leq 1/k),\end{aligned}$$

and separate the contribution of the sum over $k > \frac{4}{tc_*}$ to obtain:

$$\begin{aligned}\mathbb{E}[\lambda_t^{-2mp}] &\leq 1 + \sum_{1\leq k\leq \frac{4}{tc_*}}(k+1)^{2mp}\mathbb{P}(1/(k+1)<\lambda_t\leq 1/k) \\ &\quad + \sum_{k>\frac{4}{tc_*}}(k+1)^{2mp}\,\mathbb{P}(\lambda_t\leq 1/k) \\ &\leq 1 + \left(\frac{4}{tc_*}+1\right)^{2mp}\mathbb{P}(\lambda_t\leq 1) \\ &\quad + C_{q,m,d}\frac{|A|_1^{2q}}{c_*^{2q}}\Big(t^{1/2}|B|_0+|A|_0\Big)^{2q}e_{4q}^Z(t)^{C_{q,m,d}} \\ &\quad \times \sum_{k>\frac{4}{tc_*}}(k+1)^{2mp}\frac{1}{k^q}.\end{aligned}$$

We finally take $q = 2mp+2$ in order to get convergent series. This last estimate, together with (2.4.6), gives the desired result.

Proof of (2.4.8). Let us remark that

$$\begin{aligned}|Z_sA_j(X_s)-A_j(x)| &\leq m\sup_{i,j}|(Z_s)_{i,j}|\times|A|_1|X_s-x| + m\sup_{i,j}|(Z_s-I_m)_{i,j}|\times|A|_0 \\ &\leq m|A|_1\Big(\sup_{i,j}|(Z_s)_{i,j}||X_s-x| + \sup_{i,j}|(Z_s-I_m)_{i,j}|\Big).\end{aligned}$$

Using the same arguments allowing to prove estimate (2.2.14), it is easy to see that for every $p>1$ there exists a constant $C_{p,m}$ such that $\mathbb{E}[\sup_{i,j}\sup_{s\leq t}|(Z_s-I_m)_{i,j}|^p]\leq C_{p,m}t^{p/2}e_p^Z(t)^{C_{p,m}}$, for all $t\geq 0$. Using these inequalities together with (2.2.14), (2.2.15) and Holder's inequality, we obtain

$$\begin{aligned}\mathbb{E}\Big[\sup_{s\leq t}|Z_sA_j(X_s)-A_j(x)|^{2q}\Big] &\leq m^2 2^{2q-1}|A|_1^{2q}\Big(\mathbb{E}\Big[\sup_{i,j}\sup_{s\leq t}|(Z_s)_{i,j}|^{4q}\Big]^{1/2}\mathbb{E}\Big[\sup_{s\leq t}|X_s-x|^{4q}\Big]^{1/2} \\ &\quad + \mathbb{E}\Big[\sup_{i,j}\sup_{s\leq t}|(Z_s-I_m)_{i,j}|^{2q}\Big]\Big) \\ &\leq C\,t^q|A|_1^{2q}(t^{1/2}|B|_0+|A|_0)^{2q}e_{4q}^Z(t)^C,\end{aligned}$$

where C is a constant depending on m and q. □

2.4.2 Proof of Proposition 2.3.1

We first collect the basic facts we need to give the proof of Proposition 2.3.1. We will start by proving existence and uniqueness of strong solutions for the following equation:

$$X_t = x + \int_0^t (a(X_s) - b(X_s)X_s)ds + \int_0^t \gamma(X_s)|X_s|^\alpha dW_s, \quad t \geq 0, \quad \alpha \in [1/2, 1) \tag{2.4.9}$$

whose coefficients are defined on the whole real line (a, b and γ are the functions appearing in (2.3.1)). Once we have established that the unique strong solution of (2.4.9) is a.s. positive, then (2.4.9) will coincide with the original equation (2.3.1).

The proof of Proposition 2.3.1 is splitted in the following two short Lemmas.

Lemma 2.4.1. *Assume condition (s0) of Proposition 2.3.1. Then, existence and uniqueness of strong solutions hold for (2.4.9). Moreover, for any initial condition $x \geq 0$ the solution is a.s. positive, $\mathbb{P}(X_t \geq 0; t \geq 0) = 1$.*

Proof. Existence of non-explosive weak solutions for (2.4.9) follows from continuity and sub-linear growth of drift and diffusion coefficients. The existence of weak solutions together with pathwise uniqueness imply the existence of strong solutions (cf. [40], Prop. 5.3.20 and Cor. 5.3.23). Pathwise uniqueness follows in its turn from a well-known theorem of uniqueness of Yamada and Watanabe (cf. [40], Prop. 5.2.13). Indeed, as $a, b, \gamma \in C_b^1$ the diffusion coefficient of (2.4.9) is locally Holder-continuous of exponent $\alpha \geq 1/2$ and the drift coefficient is locally Lipschitz-continuous. We apply the standard localization argument for locally Lipschitz coefficients and Yamada-Watanabe's theorem to establish that solutions are pathwise unique up to their exit time from a compact ball, hence pathwise uniqueness holds for (2.4.9). □

Lemma 2.4.2 deals with the second part of Proposition 2.3.1, i.e. the behaviour at zero. The proof is based on Feller's test for explosions of solutions of one-dimensional SDEs (cf. [40], Th. 5.5.29). Letting τ denote the exit time from $(0, \infty)$, that is $\tau = \inf\{t \geq 0 : X_t \notin (0, \infty)\}$ with $\inf \emptyset = \infty$, we have to verify that

$$\lim_{x \to 0} p_c(x) = -\infty \tag{2.4.10}$$

with p_c defined by

$$p_c(x) := \int_c^x \exp\left(-2\int_c^y \frac{a(z) - b(z)z}{\gamma(z)^2 z^{2\alpha}} dz\right) dy, \quad x > 0, \tag{2.4.11}$$

for a fixed $c > 0$. Property (2.4.10) implies that $\mathbb{P}(\tau = \infty) = 1$, then we simply have $\tau \equiv \tau_0$ with τ_0 as defined in Proposition 2.3.1, because the solution of (2.4.9) does not explode at

∞ (cf Lemma 2.4.1). The inner integral in (2.4.11) is well defined and finite for any $y > 0$ because $\gamma(z)^2 > 0$ for any $z > 0$ and γ is continuous.

Remark 2.4.1. The conclusion does not depend on the choice of $c \in (0, \infty)$.

Lemma 2.4.2. *Assume (s0) and let $X = (X_t; t \geq 0)$ denote the unique strong solution of (2.4.9) with initial condition $x > 0$. Then, the assertions of Proposition 2.3.1 on the stopping time τ_0 hold true.*

Proof. We prove (2.4.10), for $c = 1$. We assume without restriction that $x < 1$ and distinguish the two cases.

Case $\alpha > 1/2$. We have $a(z) \geq a(0) - |a|_1 z$, $z > 0$. Then

$$\frac{a(z) - b(z)z}{\gamma(z)^2 z^{2\alpha}} \geq \frac{a(0) - (|a|_1 + b(z))z}{\gamma(z)^2 z^{2\alpha}} \geq \frac{a(0)}{|\gamma|_0^2 z^{2\alpha}} - \frac{|a|_1 + |b|_0}{\gamma(z)^2 z^{2\alpha-1}}.$$

$\frac{1}{\gamma(z)^2 z^{2\alpha-1}}$ is integrable at zero by (s1)', then there exist a positive constant K such that

$$-2\int_1^y \frac{a(z) - b(z)z}{\gamma(z)^2 z^{2\alpha}} \geq \frac{2a(0)}{|\gamma|_0^2}\int_y^1 \frac{dz}{z^{2\alpha}} + K = \frac{2a(0)}{(2\alpha-1)|\gamma|_0^2}\Big(\frac{1}{y^{2\alpha-1}} - 1\Big) + K$$

hence

$$\begin{aligned} p_1(x) &\leq -C\int_x^1 \exp\Big(\frac{2a(0)}{(2\alpha-1)|\gamma|_0^2}\frac{1}{y^{2\alpha-1}}\Big)dy \\ &= -C\int_1^{\frac{1}{x}} \frac{1}{t^2}\exp\Big(\frac{2a(0)}{(2\alpha-1)|\gamma|_0^2}t^{2\alpha-1}\Big)dt \to_{x\to 0^+} -\infty. \end{aligned}$$

Case $\alpha = 1/2$. By (s2),

$$\frac{2a(z)}{\gamma(z)^2 z} \geq \frac{1}{z},$$

for $z < \overline{x}$. Hence

$$2\frac{a(z) - b(z)z}{\gamma(z)^2 z} \geq \frac{1}{z} - 2|b|_0\frac{1}{\gamma(z)^2}$$

and thus, $\frac{1}{\gamma^2}$ being integrable at zero, for $x < \overline{x}$ we have

$$\begin{aligned} p_1(x) &\leq -C\int_x^1 \exp\Big(\int_y^1 \frac{1}{z}dz\Big)dy \\ &= -C\int_x^1 \frac{1}{y}dy \to_{x\to 0^+} -\infty. \end{aligned}$$

$\square$

Chapter 3

Lower bounds on distribution functions and densities: the case of Local-Stochastic Volatility models

Abstract

We show that in a large class of stochastic volatility models with additional skew-functions (local-stochastic volatility models) the tails of the cumulative distribution of the log-returns behave as $\exp(-c|y|)$, where c is a positive constant depending on time and on model parameters. We obtain this estimate proving a stronger result: using some estimates for the probability that Itô processes remain around a deterministic curve from [8], we lower bound the probability that the couple (X, V) remains around a two-dimensional curve up to a given maturity, X being the log-return process and V its instantaneous variance. Then we find the optimal curve leading to the bounds on the terminal cdf. The method we rely on does not require inversion of characteristic functions but works for general coefficients of the underlying SDE (in particular, no affine structure is needed). Even though the involved constants are less sharp than the ones derived for stochastic volatility models with a particular structure ([2, 41, 28]), our lower bounds entail moment explosion, thus implying that Black-Scholes implied volatility always displays wings in the considered class of models. As a second step, using Malliavin calculus techniques, we show that an analogous estimate holds for the density of the log-returns as well.

Keywords: Law of the stock price, Local-Stochastic Volatility, Moment Explosion, Implied Volatility, tube estimates for Itô processes, Malliavin calculus

Note

The results in this chapter have been submitted for publications in *Finance & Stochastics*, in a joint paper with V. Bally. The corresponding preprint can be found at [7].

3.1 Introduction

As outlined in the Introduction of this thesis, in this chapter we will focus on the following class of diffusions:

$$dX_t = -\frac{1}{2}\eta(t,X_t)^2 f(V_t)^2 dt + \eta(t,X_t) f(V_t) dW_t^1$$
$$dV_t = \beta(t,V_t)dt + \sigma(t,V_t)dW_t^2, \tag{3.1.1}$$

where W^1 and W^2 are two correlated Brownian motions on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. The function f usually being positive, the couple $(X, f(V))$ lives in $\mathbb{R}\times\mathbb{R}^+$: when X models the logarithm of the forward price of an asset and V its instantaneous variance, Eq. (3.1.1) defines a so-called local-stochastic volatility model (LSV). The function η is the local volatility (or skew) function; the autonomous process V is the stochastic volatility. Local-stochastic volatility models embed stochastic volatility models (when $\eta \equiv 1$) and have been intensively studied by the financial community in these last years, in particular when the appearance of derivatives whose value depends on the dynamics of the implied volatility demanded the introduction of more elaborate models. Some authors have focused on the problem of how to design an efficient calibration strategy of such a model to the market smile, as Lipton [50] or Henry-Labordère [37, 36], others have given a particular attention to the asymptotics of implied volatility, as Forde & Jacquier in [24] for the small-time asymptotics in the uncorrelated case. This chapter is devoted to the issue of giving asymptotic estimates of the cumulative distribution and (when existing) of the density of the log forward price X.

Let us recall the context of our study and the related work of other authors. Recall from the Introduction 1 that, setting $F_t = F_0 e^{X_t}$, $F_0 > 0$, then $F = (F_t; t \geq 0)$ satisfies

$$F_t = \int_0^t F_s \eta(s, X_s) f(V_s) dW_s^1,$$

hence F is a positive Itô local martingale. Then, a simple application of Fatou's Lemma allows to show that F is actually an *integrable* supermartingale. From the integrability property of F_T, $T > 0$, a simple application of Markov's inequality shows that, for every $y > 0$, $\mathbb{P}(X_T > y) = \mathbb{P}(e^{X_T} > e^y) \leq e^{-y}\mathbb{E}[e^{X_T}]$ that is, up to multiplicative constants, the right tail of the distribution of X_T admits the exponential *upper* bound e^{-y}. It has been seen in §1.2.2 that in models such as (1.3.2), the moments of F_T of exponent greater than one or smaller than zero can explode. (cf. [49], Andersen & Piterbarg [2] and Keller-Ressel [41]). The interest of several authors for the phenomenon of moment explosion is explained - among others - by its connexion with the asymptotic behaviour of implied volatility, which is made rigorous by Lee's Moment Formula [47]. Let us recall the basic notation we will employ: the definition of the implied volatility is found at (1.2.7), and we define the critical exponents

$p_T^*(X)$ and $q_T^*(X)$ of e^{X_T} by

$$p_T^*(X) = \sup\{p \geq 1 : \mathbb{E}[e^{pX_T}] < \infty\}, \qquad q_T^*(X) = \sup\{q \geq 0 : \mathbb{E}[e^{-qX_T}] < \infty\}.$$

We slightly change the notation with respect to the definition in the Introduction 1, since from now on we want to refer to the log-price X rather than the stock price F. Lee's moment formula (1.2.8) reads:

$$\limsup_{k\to\infty} \frac{T\sigma(T,k)^2}{k} = \varphi(p_T^*(X) - 1), \qquad \limsup_{k\to\infty} \frac{T\sigma(T,-k)^2}{k} = \varphi(q_T^*(X)), \tag{3.1.2}$$

where $\varphi(x) = 2 - 4(\sqrt{x^2+x} - x), \varphi(\infty) = 0$ (cf. §1.2.2 for the refinement of this result by Benaim and Friz [9]). Concerning the impact of (3.1.2) on model calibration, we refer the discussion in §1.2.2. Some authors focus on the explicit computation of the critical exponents, as [2] or [41] for some classes of stochastic volatility models: as a result, the critical exponents are not available in closed form but can be straightforwardly obtained solving (numerically) a simple equation.
Of course, moment explosion occurs when the tails of the distribution of X_T are sufficiently heavy. More precisely, if the law of X_T admits a density and this density behaves as $e^{-c|y|}$ for $|y| \to \infty$ for some constant $c > 1$, then positive and negative exponential moments of X_t of order p will explode for $p \geq c$. Dragulescu and Yakovenko [21] showed that the density of the log-price does behave as $e^{-c|y|}$ in the Heston model (1.2.2), exploiting the analytical computations that can be carried for the characteristic function of X_T. Let us mention that the work of [21] on the stock price distribution in the Heston model has been extended and sharpened with the addition of higher-order terms to the leading $e^{-c|y|}$, first by Gulisashvili and Stein [33] in the case of zero correlation and subsequently by Friz et al. [28].

The main aim of Chapter 3 is to show that the cumulative distribution of the log forward price X and, when existing, its density, behave as $\exp(-c|y|)$ for large $|y|$ in the following class of LSV models:

$$dX_t = -\frac{1}{2}\eta(t,X_t)^2 V_t dt + \eta(t,X_t)\sqrt{V_t}dW_t^X \tag{3.1.3}$$

$$dV_t = \beta(t,V_t)dt + \sigma(t,V_t)\sqrt{V_t}dW_t^V, \tag{3.1.4}$$

obtained from (3.1.1) setting $f(v) = \sqrt{v}$. Eq. (3.1.4) for the variance process is given on the domain $\mathbb{R}_+ = [0,\infty)$, i.e. a process V satisfying (3.1.4) is such that $\mathbb{P}(V_t \in [0,\infty), t \in [0,T]) = 1$. This class contains the Heston model and the "universal volatility model" (without the jump part) considered in [50], but is much wider, allowing for general coefficients β, σ in the SDE of the variance. While on the one hand we consider reasonable Lipschitz, boundedness and ellipticity conditions on the coefficients η and σ (but we allow the drift β to be any measurable function with sub-linear growth), on the other hand we emphasize that the square-root factors in (3.1.3)-(3.1.4) impose the necessity to work under local regularity assumptions. Let us also remark that, at this level, we are not concerned with the (possibly intricate) discussion on the existence and/or uniqueness of solutions to (3.1.3)-(3.1.4): our results indeed hold for *any* couple of processes $(X, V) = (X_t, V_t; t \in [0, T])$ satisfying (3.1.3)-(3.1.4). The situation where the diffusion coefficient of (3.1.4) is replaced by $\sigma(t, v)v^p$ for a $p > 0$ (thus embedding the class of models considered by Andersen & Piterbarg in [2]) is at the basis of the motivation for the further study we begin to develop in Chapter 4. Let us introduce the main tool of our analysis: we rely on an estimate involving the trajectory of the couple (X, V) up to time T obtained by Bally, Fernandez & Meda in [8], which we refer to as to a "tube" estimate. In [8], the authors provide estimates for the probability that an Itô process remains in a tube of given radius around a given deterministic curve, under some conditions of local Lipschiz-continuity, local boundedness and local ellipticity on the coefficients of the process. As a result, the probability of staying in the tube is lower bounded by an integral functional of the curve itself, of the deterministic radius and the coefficients of the SDE. The work we carry out in this chapter is to cast this functional in a simple form, and then to optimize over the possible choices of curves and radii. This formulation leads to the solution of an Euler-Lagrange optimization problem: the explicit computations that follow allow us to obtain a lower bound which is in the desired asymptotic range. To present our main result in this direction, let us introduce the following objects: for $y \in \mathbb{R}$, define the point $\overline{y}$ and the one-dimensional curves $\tilde{x}_t, \tilde{v}_t, \tilde{R}_t$, $t \in [0, T]$ by

$$\overline{y} = |y| + V_0; \qquad \phi(t) = \frac{\sinh(t/2)}{\sinh(T/2)}; \tag{3.1.5}$$

$$\tilde{v}_t = V_0\Big(\sqrt{\frac{\overline{y}}{V_0}}\phi(t) - e^{-T/2}\phi(t) + e^{-t/2}\Big)^2; \quad \tilde{x}_t = \mathrm{sign}(y)(\tilde{v}_t - V_0); \quad \tilde{R}_t = \frac{1}{2}\sqrt{(V_0 \wedge 1)\tilde{v}_t}$$

where $\mathrm{sign}(x) = 1$ if $x \geq 0$ and $\mathrm{sign}(x) = -1$ if $x < 0$. Our main result is the following estimate:

$$\mathbb{P}\big(|(X_t, V_t) - (\tilde{x}_t, \tilde{v}_t)| \leq \tilde{R}_t, 0 \leq t \leq T\big) \geq \exp\big(-c_T\psi(\rho_\perp) \times |y|\big) \tag{3.1.6}$$

which holds for $|y|$ large enough, where ψ is an explicit function and c_T is a strictly positive constant depending on the model parameters and explicitly on T, but not on y nor on the correlation parameter ρ (in §3.2.1, Theorem 3.2.1, we precise how large $|y|$ must be and give the expression of ψ and c_T). The curves $\tilde{x}_\cdot, \tilde{v}_\cdot, \tilde{R}_\cdot$ in (3.1.5) are the product of the optimization procedure we set up, appearing as the solution to the Euler-Lagrange equations in section 3.3.1. We remark that the curve $\tilde{x}_\cdot$ ends up at $\tilde{x}_T = y$ while the terminal radius $\tilde{R}_T$ is proportional to $\sqrt{|y|}$: hence, dropping the multiplicative constants for simplicity and writing $\mathbb{P}(|X_T - y| \leq \sqrt{|y|}) \geq \mathbb{P}(|(X_T, V_T) - (\tilde{x}_T, \tilde{v}_T)| \leq \tilde{R}_T)$, then using (3.1.6), we obtain the desired lower bound for the terminal distribution (this argument is made rigorous in Corollary 2 in §3.2.2). This result already allows us to state our main conclusion on the asymptotic behaviour of the implied volatility, namely: the implied volatility always displays wings (equivalently, is never flat) in the class of models (3.1.3)-(3.1.4) (cf. (3.2.11) in Corollary 2 for the precise statement on the asymptotic slopes of the implied variance). Notice that the fact that the tube estimate (3.1.6) is given for the couple (X, V) is crucial in our framework: indeed, in order to estimate the behaviour of X_T we need to have a control on the variance V_t for all $t \in [0, T]$.

We would like to notice that the results we obtain in this chapter on the law of the underlying and on moment explosion are significant by themselves, as generalisations of the existing results on stochastic volatility models (that is to say, apart from the fact that we employ special techniques in order to circumvent the singular coefficients).

As a second part of our study, we extend the previous estimates to the density of the law. More precisely, under some additional regularity hypotheses on the coefficients of the SDE, we discuss the existence of a density for the law of X_T and show that the exponential lower bound holds for the density as well. This last step requires to work out some "small balls" estimates (cf. Proposition 3.2.1 in §3.2.3) and to employ the integration by parts formula of Malliavin calculus. As done in Chapter 2, this tool needs to be coupled with an appropriate localization procedure. We rely on the decomposition of X_T as a Gaussian term plus a perturbation, following the idea of Bally & Caramellino in [6]: the desired lower bound on the density follows from an operation of balance between the two terms of the decomposition. This operation involves a sharp estimation of the Sobolev norms of the perturbation term, for which we take advantage of the estimates of the Sobolev norms of a diffusion derived in Chapter 2, §2.2.3. Our final estimates on the density p_{X_T} of X_T reads

$$p_{X_T}(y) \geq \frac{1}{M_T} \exp\bigl(-e_T \psi(\rho_\perp)|y|\bigr)$$

for $|y| > M_T$, where M_T and e_T are constants depending on model parameters and explicitly on T. If the density happens not to be continuous, the inequality is understood almost surely (see Theorem 3.2.2 for the precise statement).

The chapter is organized as follows. In section 3.2 we give our working hypotheses and a detailed presentation of the main results. In particular, in §3.2.1 we prove estimate (3.1.6) and in §3.2.2 we state the corollary for the terminal cdf, the moment explosion and the implied volatility slopes. In §3.2.3 we give our results on the density of X_T. Sections 3.3 and 3.4 are devoted to the proofs of the results stated in section 3.2: Malliavin calculus appears from section 3.4, while the tools employed in the previous are borrowed from stochastic calculus for Itô processes. Finally, section 3.6 contains the proofs of the most technical results.

3.2 Main results

In this section we give our working hypotheses and a detailed presentation of the main results.

Let us consider the class of models (3.1.3)-(3.1.4). For the ease of computations, we decorrelate the driving Brownian motions in the usual way and rewrite (3.1.3)-(3.1.4) as

$$dX_t = -\frac{1}{2}\eta(t,X_t)^2 V_t dt + \eta(t,X_t)\sqrt{V_t}(\rho dW_t^1 + \rho_\perp dW_t^2), \quad t \le T \tag{3.2.1}$$

$$dV_t = \beta(t,V_t)dt + \sigma(t,V_t)\sqrt{V_t}dW_t^1, \quad t \le T \tag{3.2.2}$$

were $(W_t^1, W_t^2; t \le T)$ is a two-dimensional standard Brownian motion. We consider deterministic initial conditions $X_0 = 0$ and $V_0 > 0$, finite time horizon $T > 0$, $\rho \in (-1,1)$ and we denote $\rho_\perp := \sqrt{1-\rho^2}$. Eq. (3.2.2) for the variance process is given on the domain $\mathbb{R}_+ = [0,\infty)$, i.e. a process V satisfying (3.2.2) is such that $\mathbb{P}(V_t \in [0,\infty), t \in [0,T]) = 1$. We assume that the coefficients η, β and σ in (3.2.1)-(3.2.2) satisfy the following conditions, for some $K > 1$:

(R) *(regularity)* $\eta : [0,T]\times\mathbb{R} \to \mathbb{R}$ and $\sigma : [0,T]\times[0,\infty) \to \mathbb{R}$ are Lipschitz-continuous functions, more precisely

$$|\eta(s,x) - \eta(t,y)| \le K(|x-y| + |s-t|)$$

$$|\sigma(s,v) - \sigma(t,u)| \le K(|v-u| + |s-t|)$$

hold for every $(s,t,x,y) \in [0,T]\times[0,T]\times\mathbb{R}\times\mathbb{R}$, respectively every $(s,t,v,u) \in [0,T]\times[0,T]\times[0,\infty)\times[0,\infty)$.

(G) *(growth)* The mesurable function $\beta : [0,T]\times[0,\infty) \to \mathbb{R}$ has sub-linear growth in v, more precisely

$$|\beta(t,v)| \le K(1+v)$$

for every $(t,v) \in [0,T]\times[0,\infty)$. Moreover, there exist constants $0 < \underline{\eta} < 1 < \overline{\eta}$ and $0 < \underline{\sigma} < 1 < \overline{\sigma}$ such that

$$\underline{\eta} \le \eta(t,x) \le \overline{\eta}, \qquad \underline{\sigma} \le \sigma(t,v) \le \overline{\sigma}$$

hold for every $(t,x) \in [0,T]\times\mathbb{R}$, respectively $(t,v) \in [0,T]\times[0,\infty)$.

Remark 3.2.1. Despite of the boundedness condition on η, σ given in (G), obviously none of the drift and diffusion coefficients in the system (3.2.1)-(3.2.2) is bounded, because of the factors V_t and $\sqrt{V_t}$.

Remark 3.2.2. In hypothesis (R), we could replace Lipschitz-continuity with respect to the couple (t, x) (resp. (t, v)) with Lipschitz-continuity with respect to the state variable x (resp. v) and Holder-continuity of exponent $1/2$ with respect to time, and all the results of sections 3.2.1 and 3.2.2 would still hold.

Remark 3.2.3. We are not interested here in discussing the existence and/or uniqueness of solutions to (3.2.1)-(3.2.2) under conditions (R) and (G). All the results we give in this subsection and in the following (but not in subsection 3.2.3, where a new set of hypotheses is considered) indeed hold for *any* couple of processes $(X, V) = (X_t, V_t; t \in [0, T])$ which satisfy (3.2.1)-(3.2.2).

Notation *Sets and filtrations.* As done previously, $|\cdot|$ will still denote the absolute value for real numbers as well as the Euclidean norm for vectors, i.e. $|x| = \sqrt{\sum_i^n x_i^2}$ if $x \in \mathbb{R}^n$. We recall that $B_R(x)$ is the open ball in $\mathbb{R}^n$ of center x and radius R, $B_R(x) = \{y \in R^n : |y - x| < R\}$. Moreover, we denote $(\mathcal{F}_t^i, t \geq 0)$ the completion of the filtration generated by W^i, $i = 1, 2$, and $\mathcal{F}_t = \mathcal{F}_t^1 \vee \mathcal{F}_t^2$. λ_n is the Lebesgue measure on $\mathbb{R}^n$.
Classes of functions, derivatives, norms. $C^1([0, T])$ denotes the class of real functions of $[0, T]$ which have uniformly continuous derivative on $(0, T)$. We will make use of the class $L(\mu, h)$ defined in [8], section 2 : given a fixed time horizon T, $\mu \geq 1$ and $h > 0$, $L(\mu, h)$ is the class of functions $f : [0, T] \to R_+ = [0, \infty)$ such that for every $t, s \in [0, T]$ with $|t - s| < h$ one has

$$f(t) \leq \mu f(s).$$

We denote $C^{0,k}([0, T] \times \mathbb{R}^n)$ (resp. $C_b^{0,k}([0, T] \times \mathbb{R})$) the class of continuous functions of $[0, T] \times \mathbb{R}^n$ which have continuous (resp. bounded continuous) partial derivatives with respect to the second variable up to order k. Let Θ_k be the set of multi-indexes of length k with components in $\{1, \ldots, n\}$, $\Theta_k = \{1, \ldots, n\}^k$. For $\alpha \in \Theta_k$ and $g \in C^{0,k}([0, T] \times \mathbb{R}^n)$, we denote $\partial_\alpha g = \frac{\partial^k g}{\partial_{x^{\alpha_1}} \cdots \partial_{x^{\alpha_k}}}$. We define the norms

$$|g|_k = 1 \vee \sum_{j=0}^{k} \sum_{\alpha \in \Theta_k} \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^n} |\partial_\alpha g(t, x)|.$$

Constants. For a vector of parameters $\Lambda = (\lambda_1, \ldots, \lambda_\nu)$, we shall denote C_Λ (resp. $C_\Lambda(t)$) a positive constant (resp. function of time) depending on the λ_i's but not on any of the other existing variables. As in the previous chapter, constants of such a type may vary from line by line, but always depend only on Λ.

3.2.1 Estimates around a Deterministic Curve

We fix $T > 0$. We consider three one-dimensional curves x, v, R of class $C^1([0,T])$ such that: $R_t > 0$ for any $t \in [0,T]$ and x and v have the same initial values as X and V in (3.2.1)-(3.2.2) (in particular, $x_0 = 0$). We look for a lower bound on the probability that a process $(X_t, V_t) = (X_t, V_t; t \le T)$ satisfying (3.2.1)-(3.2.2) stays in the tube of radius R_t around the deterministic curve (x_t, v_t) up to time T, that is a lower bound on the quantity

$$\mathbb{P}(|(X_t, V_t) - (x_t, v_t)| \le R_t, 0 \le t \le T). \tag{3.2.3}$$

To lower bound (3.2.3) we employ the estimate provided in [8], Theorem 1. The main result of this section, Theorem 3.2.1, makes use of the triplet of curves $\tilde{x}_t, \tilde{v}_t, \tilde{R}_t$ defined in (3.1.5). As addressed in the Introduction, the choice of these particular curves relies on an optimization problem: they indeed appear as the solutions of some Euler-Lagrange equations (see section 3.3.1). Recall the ψ from the Introduction:

$$\psi(r) = \frac{1}{r^6}\Big(\ln\Big(\frac{1}{r}\Big) + 1\Big), \quad r > 0. \tag{3.2.4}$$

Theorem 3.2.1. *Assume conditions* (R) *and* (G) *and let* $(X_t, V_t; 0 \le t \le T)$ *be two processes satisfying (3.2.1)-(3.2.2). Then for every* $y \in \mathbb{R}$ *with* $|y|$ *large enough, precisely*

$$|y| > V_0\big(1 + 2\sinh(T/2)\big)^2, \tag{3.2.5}$$

and with the curves $\tilde{x}_\cdot, \tilde{v}_\cdot, \tilde{R}_\cdot$ *defined in (3.1.5) one has*

$$\mathbb{P}\big(|(X_t, V_t) - (\tilde{x}_t, \tilde{v}_t)| \le \tilde{R}_t, t \in [0,T]\big) \ge \exp\big(-c_T\psi(\rho_\perp) \times |y|\big). \tag{3.2.6}$$

The constant c_T *is given by*

$$c_T = c^*\Big(\frac{1}{T} + 1\Big)e^{c^* T^2}, \tag{3.2.7}$$

where c^* *is strictly positive constant depending on the model parameters* $V_0, K, \underline{\eta}, \underline{\sigma}, \overline{\eta}, \overline{\sigma}$ *given in* (R) *and* (G) *but not on* y *nor on the correlation parameter* ρ.

Remark 3.2.4. Let us discuss the impact of the factor $\psi(\rho_\perp)$ and of the maturity T in the lower bound (3.2.6) a bit further. It is known that the correlation effects moment explosion in stochastic volatility models, a negative correlation bringing - as intuitively clear - a dampening effect (cf. [2], sections 3 and 4). In a Heston model, obtained when the variance process in (3.2.2) has constant parameters and mean-reverting drift, the upper critical moment of e^{X_T} tends to infinity when $\rho \to -1$ and X_T even becomes a bounded random variable when $\rho = -1$, and the behaviour is the opposite when $\rho > 1$. The factor $\psi(\rho_\perp) = \psi(\sqrt{1-\rho^2})$ has the expected explosive behaviour when $\rho \to -1$, but it symmetrically decrements the rhs of

(3.2.6) for $\rho > 0$, making the lower bound significant in particular for $\rho \in (-1, 0)$. The small time asymptotics $\frac{1}{T}$ of the constant c_T is what expected for a diffusion; on the other hand, the large time dependence $e^{c^* T^2}$ makes the bound (3.2.6) not directly applicable to study the large-time asymptotics.

3.2.2 Lower bounds for Cumulative Distribution Function and Moments

Theorem 3.2.1 leads, in particular, to lower bounds on the tails of the complementary cumulative distribution function (complementary cdf in short) of X_T, i.e. $\mathbb{P}(|X_T| > \cdot)$. Indeed, on the one hand we can simply lower bound the probability to be in the tube at the final "time-slice", $\mathbb{P}(|(X_T, V_T) - (\tilde{x}_T, \tilde{v}_T)| \le \tilde{R}_T)$ with the probability to stay in the tube up to time T. On the other hand, the final time radius $\tilde{R}_T$ in (3.1.5) is - roughly speaking - proportional to $\sqrt{|y|}$. Hence, when $y \to \infty$ (resp. $y \to -\infty$) the infimum (resp. the supremum) of the interval $[y - \tilde{R}_T, y + \tilde{R}_T]$ becomes large (resp. small) and this allows to obtain tail estimates for $\mathbb{P}(|X_T| > \cdot)$ that are in the same asymptotic range as (3.2.6). This observation is made rigorous in the proof of the following Corollary, which is indeed a direct consequence of Theorem 3.2.1.

Corollary 2. *Under the assumptions of Theorem 3.2.1, for any $y > 0$ satisfying (3.2.5) and*

$$y > 2(V_0 \vee 1)^2(1 + V_0), \tag{3.2.8}$$

one has

$$\mathbb{P}(X_T > y) \wedge \mathbb{P}(X_T < -y) \ge \exp\big(-c_T \psi(\rho_\perp) \times y\big) \tag{3.2.9}$$

where c_T is the constant given in (3.2.7). In particular, the critical exponents are finite:

$$p_T^*(X) \vee q_T^*(X) \le c_T \psi(\rho_\perp), \tag{3.2.10}$$

hence the implied volatility displays left and right wings, i.e.

$$\begin{aligned} \limsup_{k \to \infty} \frac{T\sigma(T,k)^2}{k} &\ge \varphi(c_T \psi(\rho_\perp) - 1) > 0, \\ \limsup_{k \to -\infty} \frac{T\sigma(T,k)^2}{k} &\ge \varphi(c_T \psi(\rho_\perp)) > 0. \end{aligned} \tag{3.2.11}$$

Remark 3.2.5. As addressed in the Introduction, e^{X_T} is integrable for every $T > 0$. A simple application of Markov's inequality shows that, for every $y > 0$, $\mathbb{P}(X_T > y) = \mathbb{P}(e^{X_T} > e^y) \le e^{-y}\mathbb{E}[e^{X_T}]$. This is not in contradiction with (3.2.9), because on the one hand $\psi(\rho_\perp)$ is greater than or equal to one for any value of $\rho_\perp \in (0, 1]$ (cf. (3.2.4)), and on the other the constant c^* in (3.2.7) is greater than 1, hence $c_T > 1$ for every $T > 0$, too.

Remark 3.2.6. In this chapter we are mainly interested in the law of X_T. Estimate (3.2.6) can of course be applied to derive the analogous lower bound for the *joint* law of X_T and V_T. See Proposition 3.2.1 in the next section for a refined statement in this direction.

3.2.3 Lower bounds for the density

We consider now some stronger regularity conditions on the coefficients of (3.2.1)-(3.2.2):

(R') (*regularity'*) (R) and (G) hold and $\eta \in C_b^{0,2}([0,T] \times \mathbb{R})$, $\sigma \in C_b^{0,2}([0,T] \times [0,\infty))$, $\beta \in C_b^{0,2}([0,T] \times [0,\infty)) \cap Lip([0,T] \times [0,\infty))$ with $|\eta|_2 \vee |\sigma|_2 \leq K$.

Remark 3.2.7. Under condition (R'), the system (3.2.1)-(3.2.2) admits a unique strong solution. Indeed, the existence of a weak solution (X,V) that satisfies (3.6.1) follows from the continuity and sub-linearity of the coefficients. Then, pathwise uniqueness holds for (3.2.2) after a theorem of uniqueness of Yamada and Watanabe (cf. [40], Prop. 5.2.13) and weak existence and pathwise uniqueness together imply strong existence ([40], Cor. 5.3.23). Given the unique solution to (3.2.2), standard arguments allow to prove pathwise uniqueness for (3.2.1).

We will give a lower bound for the density of the law of X_T under hypothesis (R'). Notice first that the law of X_T is absolutely continuous with respect to the Lebesgue measure λ_1 on $\mathbb{R}$. This fact may be proven in (at least) two ways. First, we may look to the law of X_T conditional to $(W_t^1, t \leq T)$. Then X_T appears as a functional of the independent Brownian motion $(W_t^2, t \leq T)$ and, using the Bouleau-Hirsch criterium (cf. [54]), we obtain a density $p_{X_T}(W^1, x)$ for the conditional law. Then, the law of X_T has the density $\mathbb{E}[p_{X_T}(W^1, x)] = p_{X_T}(x)$. A second way would be to use the results in Chapter 2, Theorem 2.2.5, telling that the *couple* (X_T, V_T) admits a density $p_T(x,v)$ on $\mathbb{R} \times (0,\infty)$ (meaning that the law of (X_T, V_T) restricted to $\mathbb{R} \times (0,\infty)$ has the density $p_T(x,v)$). This immediately yields the existence of a density $p_{X_T}(x)$ for the marginal law of X_T. Nevertheless, we remark that none of the above approaches guarantee that the density of X_T is continuous.

Before giving an estimate of the density of X_T itself, we need to work out some estimates for the probability that X_T stays in a ball of "small" radius.

Proposition 3.2.1 (Lower bounds for balls of small radius). *Let $R^{(j)}(y)$ be given by*

$$R^{(j)}(y) = \left(\sqrt{|y|}\right)^{1-j}, \quad j \in \mathbb{N}$$

(so that $R^{(0)}(y) = \sqrt{|y|}$, $R^{(1)}(y) = 1$, $R^{(2)}(y) = \frac{1}{\sqrt{|y|}}$, ...). Assume (R') and let $(X_t, V_t; 0 \leq t \leq T)$ be the unique strong solution to (3.2.1)-(3.2.2). Then, for any y satisfying (3.2.5) and $|y| > 16 \vee 2(V_0 \vee 1)^2(1+V_0)$,

$$\mathbb{P}\left(|(X_T, V_T) - (y, |y| + V_0)| \leq R^{(j)}(y)\right) \geq \exp\left(-(j+1) d_T \psi(\rho_\perp) \times |y|\right). \tag{3.2.12}$$

The constant d_T *is given by* $d_T = 2c^*\Big(\frac{1}{T^2}+1\Big)e^{(c^*+1)T^2}$, c^* *being the constant in (3.2.7).*

Remark 3.2.8. By taking j large enough, the radius $R^{(j)}(y)$ can be made arbitrarily small. Then we would like to make use of (3.2.12) to obtain a lower bound for the density of X_T computed at y, but we cannot pass to the limit with j in (3.2.12) because the rhs tends to zero as $j \to \infty$. Nevertheless, we can obtain a lower bound for the density using (3.2.12) for finite j and the integration by parts formula of Malliavin Calculus. This is what we actually do in order to prove the next theorem.

Here is the main result for this section.

Theorem 3.2.2. *Assume (R') and let* $(X_t, V_t; 0 \le t \le T)$ *be the unique strong solution to (3.2.1)-(3.2.2). Then, there exists a strictly positive constant* M_T *depending on* T *and on the model parameters such that for* λ_1*-a.e.* y *with* $|y| > M_T$,

$$p_{X_T}(y) \ge \frac{1}{M_T}\exp\big(-e_T\psi(\rho_\perp)|y|\big) \tag{3.2.13}$$

where $e_T = 136c^*\Big(\frac{1}{T^2}+1\Big)e^{(c^*+1)T}$. *The inequality (3.2.13) is understood in the sense*

$$\int_{|y|>M_T} f(y)p_{X_T}(y)dy \ge \frac{1}{M_T}\int_{|y|>M_T} f(y)\exp\big(-e_T\psi(\rho_\perp)|y|\big)dy$$

for every $f \in C_b(\mathbb{R})$.

Remark 3.2.9. If the density $p_{X_T}(y)$ is continuous, then (3.2.13) holds for every y with $|y| > M_T$.

We recall that for a Heston model with constant coefficients, the density $p_{X_T}(y)$ is asymptotic to $\exp(-c|y|)$, cf. [21, 28].

3.3 Proof of results in 3.2.1 and 3.2.2

We start by giving a preliminary result that will be used in the proof of Theorem 3.2.1. We consider $x., v.$ in $C^1([0,T])$ and $R, c, \lambda, \gamma, L : [0,T] \to \mathbb{R}_+$ satisfying

$$\begin{aligned} &x_0 = 0; \quad v_0 = V_0; \quad v'_t > 0;\\ &x', v', R, c, \lambda, \gamma, L \in L(\mu, h) \end{aligned} \tag{3.3.1}$$

and we define the stopping time

$$\tau_R = \tau_R(X,V) = \inf\{t \le T : |(X_t, V_t) - (x_t, v_t)| > R_t\}.$$

Moreover, we denote

$$b(t,x,v) = \begin{pmatrix} -\frac{1}{2}\eta(t,x)^2 v \\ \beta(t,v) \end{pmatrix};$$

$$\sigma_1(t,x,v) = \begin{pmatrix} \rho\eta(t,x)\sqrt{v} \\ \sigma(t,v)\sqrt{v} \end{pmatrix}; \quad \sigma_2(t,x,v) = \begin{pmatrix} \rho_\perp\eta(t,x)\sqrt{v} \\ 0 \end{pmatrix}$$

and consider the conditions:

$$|b(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R})| + \sum_{j=1,2} |\sigma_j(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R})| \leq c_t; \tag{3.3.2}$$

$$\lambda_t I_2 \leq \sigma\sigma^*(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R}) \leq \gamma_t I_2; \tag{3.3.3}$$

$$\mathbb{E}\Big[\sum_{j=1,2} |\sigma_j(s, X_s, V_s) - \sigma_j(t, X_t, V_t)|^2 1_{\{\tau_R \geq s\}}\Big] \leq L_t^2(s-t) \tag{3.3.4}$$

which correspond to hypothesis **(H)** in [8]. Then, according to Theorem 1 in [8], the estimate

$$\mathbb{P}(|(X_t,V_t) - (x_t,v_t)| \leq R_t, 0 \leq t \leq T) \geq \exp\Big(-Q(\mu)\Big(1 + \int_0^T F_{x,v,R}(t)dt\Big)\Big) \tag{3.3.5}$$

holds with the *rate function*

$$F_{x,v,R}(t) = \frac{1}{h} + \frac{(x_t')^2 + (v_t')^2}{\lambda_t} + 2(c_t^2 + L_t^2)\Big(\frac{1}{\lambda_t} + \frac{1}{R_t^2}\Big). \tag{3.3.6}$$

and the constant $Q(\mu)$ given by

$$Q(\mu) = \frac{q_\mu}{\phi_{\lambda,\gamma}^2} \ln \frac{q_\mu}{\phi_{\lambda,\gamma}}, \tag{3.3.7}$$

where

$$\phi_{\lambda,\gamma} = \inf_{t\leq T} \frac{\lambda_t}{\gamma_t}; \qquad q_\mu = 8^{12} e^2 \mu^{73}. \tag{3.3.8}$$

(We actually denote $\phi_{\lambda,\gamma}$ the constant ρ in [8]). The following proposition is the starting point to prove Theorem 3.2.1.

Proposition 3.3.1. *Assume conditions* (R) *and* (G)*. Let x_t, v_t, R_t satisfy (3.3.1) and consider a process $(X_t, V_t) = (X_t, V_t; 0 \leq t \leq T)$ satisfying (3.2.1)-(3.2.2). Let moreover*

$$R_t \leq Rv_t, \quad t \in [0,T] \tag{3.3.9}$$

hold for a fixed $R \in (0,1)$. Then, setting $\Theta = (K, \overline{\eta}, \overline{\sigma}, R, V_0)$, there exist strictly positive constants $c = c_\Theta; L = L_\Theta; \gamma = \gamma_\Theta; \lambda = \lambda_{\Theta,\underline{\eta},\underline{\sigma}}$ such that for every $0 \leq t < s \leq T$ the *conditions (3.3.2)-(3.3.3)-(3.3.4) are fulfilled by the curves*

$$c_t = cv_t; \qquad L_t^2 = L_T v_t;$$

$$\gamma_t = \gamma v_t; \qquad \lambda_t = \rho_\perp^2 \lambda v_t.$$

L_T is given by $L_T = Le^{C_2T^2}$, where C_2 is the constant appearing in Lemma (3.6.1). The curves $c_t, L_t, \gamma_t, \lambda_t$ belong respectively to $L(\mu, h), L(\sqrt{\mu}, h), L(\mu, h), L(\mu, h)$.

Proof. In what follows we shall repeatedly apply the inequality $\sqrt{v} \le 1 + v$, $v > 0$.
(3.3.2): We notice that for every $t, x, v \in [0, T] \times \mathbb{R} \times [0, \infty)$,

$$\begin{aligned}
|b(t, v, x)| + \sum_{j=1,2} |\sigma_j(t, v, x)| \\
&\le \frac{1}{2}\overline{\eta}^2 v + K(1 + v) + (\rho + \rho_\perp)\overline{\eta}\sqrt{v} + \overline{\sigma}\sqrt{v} \\
&\le K + (\rho + \rho_\perp)\overline{\eta} + \overline{\sigma} + \Big(\frac{1}{2}\overline{\eta}^2 + K + (\rho + \rho_\perp)\overline{\eta} + \overline{\sigma}\Big)v \\
&\le c(1 + v)
\end{aligned}$$

where the last holds with $c = \frac{1}{2}\overline{\eta}^2 + K + 2\overline{\eta} + \overline{\sigma}$. Then, employing the condition (3.3.9) on the radius and the fact that $v_t \ge V_0$ for any $t \in [0, T]$ by (3.3.1),

$$\begin{aligned}
|b(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R})| + \sum_{j=1,2} |\sigma_j(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R})| \\
&\le c(1 + V_{t\wedge\tau_R}) \le c(1 + (v_t + R_t)) \\
&\le c(1 + R)(1 + v_t) \le 2c(1 + R)\frac{1 \vee V_0}{V_0} v_t
\end{aligned} \tag{3.3.10}$$

and the last inequality holds since $(1 + v) \le 2\frac{V_0 \vee 1}{V_0} v$ for any $v > V_0$.
(3.3.3): Let $\sigma\sigma^*_{i,j}(t, x, v) = \sum_{k=1,2} \sigma_k^i(t, x, v)\sigma_k^j(t, x, v)$, $i, j = 1, 2$. The condition

$$\lambda_t I_2 \le \sigma\sigma^*(t, X_{t\wedge\tau_R}, V_{t\wedge\tau_R}) \le \gamma_t I_2 \tag{3.3.11}$$

for the given λ_t, γ_t will follow from the computation of the eigenvalues of $\sigma\sigma^*$. Denoting $\eta := \eta(t, x)$ and $\sigma := \sigma(t, v)$ for simplicity of notation, we have

$$\sigma\sigma^*(t, x, v) = \begin{pmatrix} \eta^2 v & \rho\eta\sigma v \\ \rho\eta\sigma v & \sigma^2 v \end{pmatrix}$$

hence the smallest, respectively the largest, eigenvalue satisfy

$$\begin{aligned}\overline{\lambda}_t(x,v) &= \frac{1}{2}\Big(\eta^2 v + \sigma^2 v - \sqrt{\big(\eta^2 v + \sigma^2 v\big)^2 - 4\eta^2\sigma^2 v^2 \rho_\perp}\Big)\\ &\geq \rho_\perp^2 \frac{\eta^2\sigma^2 v^2}{\eta^2 v + \sigma^2 v} \geq \rho_\perp^2 \frac{\underline{\eta}^2\underline{\sigma}^2}{2(\overline{\eta}^2 + \overline{\sigma}^2)} v\\ \overline{\gamma}_t(x,v) &= \frac{1}{2}\Big(\eta^2 v + \sigma^2 v + \sqrt{\big(\eta^2 v + \sigma^2 v\big)^2 + 4\eta^2\sigma^2 v^2 \rho_\perp}\Big)\\ &\leq \eta^2 v + \sigma^2 v \leq (\overline{\eta}^2 + \overline{\sigma}^2) v\end{aligned} \tag{3.3.12}$$

Proceeding as before we have

$$\begin{aligned}\overline{\lambda}_t(X_{t\wedge\tau_R}, V_{t\wedge\tau_R}) &\geq \rho_\perp^2 \frac{2(\underline{\eta}^2\underline{\sigma}^2)}{\overline{\eta}^2 + \overline{\sigma}^2}\frac{1-R}{1+R} v_t = \rho_\perp^2 \lambda(\underline{\eta}, \underline{\sigma}, \overline{\eta}, \overline{\sigma}, R) v_t;\\ \overline{\gamma}_t(X_{t\wedge\tau_R}, V_{t\wedge\tau_R}) &\leq (\overline{\eta}^2 + \overline{\sigma}^2)(1+R) v_t = \gamma(\overline{\eta}, \overline{\sigma}, R) v_t;\end{aligned}$$

Then (3.3.11) follows with λ_t, γ_t as in the statement of the proposition.
(3.3.4): Because of assumption (R), for every $s,t \in [0,T]\times[0,T]$, every $x,y \in \mathbb{R}\times\mathbb{R}$ and every $v,u \in [0,\infty)\times[0,\infty)$ we have

$$\begin{aligned}|\eta(s,x)\sqrt{v} - \eta(t,y)\sqrt{u}| &\leq \sqrt{u}K(|x-y| + |s-t|) + \overline{\eta}|\sqrt{v} - \sqrt{u}|\\ &\leq \sqrt{u}K(|x-y| + |s-t|) + \frac{\overline{\eta}}{2\min(\sqrt{v},\sqrt{u})}|v-u|\end{aligned} \tag{3.3.13}$$

and

$$\begin{aligned}|\sigma(s,v)\sqrt{v} - \sigma(t,u)\sqrt{u}| &\leq \sqrt{u}K(|x-y| + |s-t|) + \overline{\sigma}|\sqrt{v} - \sqrt{u}|\\ &\leq \sqrt{u}K(|x-y| + |s-t|) + \frac{\overline{\sigma}}{2\min(\sqrt{v},\sqrt{u})}|v-u|.\end{aligned} \tag{3.3.14}$$

It follows, for every $t \leq s \leq T$,

$$\begin{aligned}&\mathbb{E}\Big[|\eta(s,X_s)\sqrt{V_s} - \eta(t,X_t)\sqrt{V_t}|^2 1_{\{\tau_R \geq s\}}\Big]\\ &\leq 4K^2\mathbb{E}\big[V_t\big(|X_s - X_t|^2 + (s-t)^2\big)1_{\{\tau_R\geq s\}}\big] + \mathbb{E}\big[\frac{\overline{\eta}^2}{2\min(V_t,V_s)}|V_s - V_t|^2 1_{\{\tau_R\geq s\}}\big]\\ &\leq 4K^2\Big((1+R)v_t C_2 e^{C_2T^2}(s-t) + T(s-t)\Big) + \frac{\overline{\eta}^2}{2(1-R)v_t} C_2 e^{C_2T^2}(s-t)\\ &\leq C_2\Big(8(1+R)K^2\frac{V_0\vee 1}{V_0} + \frac{\overline{\eta}^2}{2(1-R)V_0^2}\Big)e^{C_2T^2} v_t (s-t)\end{aligned}$$

where C_2 is the constant considered in Lemma 3.6.1. Analogously,

$$\begin{aligned}&\mathbb{E}\Big[|\sigma(s,V_s)\sqrt{V_s}-\sigma(t,V_t)\sqrt{V_t}|^2 1_{\{\tau_R\geq s\}}\Big]\\&\leq 4K^2\Big((1+R)v_t C_2 e^{C_2T^2}(s-t)+T(s-t)\Big)+\frac{\overline{\sigma}^2}{2(1-R)v_t}C_2e^{C_2T^2}(s-t)\\&\leq C_2\Big(8(1+R)K^2\frac{V_0\vee 1}{V_0}+\frac{\overline{\sigma}^2}{(1-R)V_0^2}\Big)e^{C_2T^2}v_t(s-t).\end{aligned}$$

Estimate (3.3.4) then follows from the two previous inequalities and the expression of σ_1, σ_2. The last statement on the curves $c_t, L_t, \gamma_t, \lambda_t$ follows from the fact that the function af^p belongs to $L(\mu^p, h)$ if f belongs to $L(\mu, h)$, $p > 0$ and a is a positive constant. □

Basically, what Theorem 3.2.1 does is to compute the right hand side of (3.3.5) on a particular curve satisfying conditions (3.3.2)-(3.3.3)-(3.3.4), so that (3.3.5) translates into the explicit lower bound (3.2.6). The choice of the deterministic curve (x_t, v_t) considered in Theorem 3.2.1 is of course motivated by the form of the rate function (3.3.6). More precisely, consider any x_t, v_t, R_t that satisfy (3.3.1). Then, by Proposition 3.3.1, the estimate (3.3.5) holds with

$$F_{x,v,R}(t)=\frac{1}{h}+\frac{(x_t')^2+(v_t')^2}{\rho_\perp^2\lambda v_t}+2(c^2v_t^2+L_Tv_t)\Big(\frac{1}{\rho_\perp^2\lambda v_t}+\frac{1}{R_t^2}\Big), \tag{3.3.15}$$

$$\phi_{\lambda,\gamma}=\inf_{t\leq T}\frac{\lambda_t}{\gamma_t}=\frac{\rho_\perp^2\lambda}{\gamma}\inf_{t\leq T}\frac{v_t}{v_t}=\frac{\rho_\perp^2\lambda}{\gamma} \tag{3.3.16}$$

and $Q(\mu)$ as given in (3.3.7).

3.3.1 A Lagrangian minimization problem

We start from the simple observation that maximizing the lower bound in (3.3.5) is equivalent to minimizing the exponent $Q(\mu)(1+\int_0^T F_{x,v,R}(t)dt)$. Due to the presence of the competing terms $\frac{1}{v_t}+\frac{1}{R_t^2}$ and $(x_t')^2+(v_t')^2$ in $F_{x,v,R}$, we make the choice

$$R_t=\frac{1}{2}\sqrt{V_0v_t} \tag{3.3.17}$$

so that R_t^2 is proportional to v_t, and consider curves x_t, v_t such that $|x_t'|=|v_t'|$, precisely

$$x_t=sign(y)(v_t-V_0),\quad t\in[0,T]. \tag{3.3.18}$$

Equations (3.3.17) and (3.3.18) define R_t and x_t given v_t, as happens for (3.1.5). We remark that the radius in (3.3.17) satisfies the requirement $R_t\leq\frac{1}{2}v_t$ of Proposition 3.3.1. Moreover, if the arrival point $x_T=y$ of the curve x_t is given, the same will be for v_t. We define the "shifted" arrival point $\overline{y} = v_T$ setting

$$\overline{y} = |y| + V_0. \tag{3.3.19}$$

After (3.3.17)-(3.3.18), the rate function (3.3.15) reduces to

$$\overline{F}_v(t) = \frac{1}{h} + \frac{2}{\rho_\perp^2 \lambda v_t}(v_t')^2 + 2(c^2 v_t^2 + L_T v_t)\Big(\frac{1}{\rho_\perp^2 \lambda} + 1\Big)\frac{1}{v_t} \tag{3.3.20}$$

which is a function of the curve v_t only. Since we want to upper bound $\overline{F}_v$, we can get rid of all the constants and just keep the explicit dependence with respect to the curve v_t: defining

$$\Gamma_T = 1 \vee 2\frac{c^2 + L_T}{(V_0 \wedge 1)\rho_\perp^2 \lambda}, \tag{3.3.21}$$

we have

$$\overline{F}_v \leq \frac{1}{h} + \Gamma_T\Big(\frac{(v_t')^2}{v_t} + v_t\Big)$$

and the constant Γ carries the explicit dependence w.r.t the model parameter $\rho_\perp$. The strategy we shall follow is to consider

$$\mathcal{L}(v_t, v_t') = \frac{(v_t')^2}{v_t} + v_t \tag{3.3.22}$$

and to look for the solution of the minimization problem

$$\min_v \int_0^T \mathcal{L}(v_t, v_t') \tag{3.3.23}$$

with the constraints

$$v_0 = V_0; \qquad v_T = \overline{y}. \tag{3.3.24}$$

The problem (3.3.23)-(3.3.24) is the classical minimization problem in the Calculus of Variations for Lagrangian systems: a stationary point for the integral functional in (3.3.23) is the solution of the Euler-Lagrange equation

$$\frac{d}{dt}\frac{d\mathcal{L}}{dv'}(v_t, v_t') - \frac{d\mathcal{L}}{dv}(v_t, v_t') = 0$$

under the constraints (3.3.24). The Euler-Lagrange equation associated to the Lagrangian (3.3.22) is easily obtained to be:

$$\frac{v_t''}{v_t'} = \frac{v_t'}{2v_t} + \frac{v_t}{2v_t'}. \tag{3.3.25}$$

A closer look to Eq. (3.3.25) reveals that it can be turned into a linear second order ODE with the change of variables

$$u_t = \left(\frac{v_t}{V_0}\right)^{\frac{1}{2}}, \tag{3.3.26}$$

which indeed converts (3.3.25) into

$$u''_t - \frac{1}{4}u_t = 0, \tag{3.3.27}$$

now with the constraints

$$u_0 = 1; \qquad u_T = \left(\frac{\overline{y}}{V_0}\right)^{\frac{1}{2}}. \tag{3.3.28}$$

The explicit solution to (3.3.27)-(3.3.28) is easily found to be

$$u_t = \left(\frac{\overline{y}}{V_0}\right)^{\frac{1}{2}} \frac{\sinh(t/2)}{\sinh(T/2)} - e^{-T/2}\frac{\sinh(t/2)}{\sinh(T/2)} + e^{-t/2}. \tag{3.3.29}$$

The curve $\tilde{v}_t$ defined in (3.1.5) corresponds to the one given by (3.3.26) and (3.3.29). What Theorem 3.2.1 does, then, is to pick up this particular curve, to check for which values of μ, h and y the curve $\tilde{v}'_t$ belongs to $L(\mu, h)$ and satisfies $\tilde{v}'_t > 0$ from (3.3.1), hence to estimate the integral functional in (3.3.23).

Proof of Theorem 3.2.1. Step 1. We show that $\tilde{v}'_t > 0$, $t \in [0,T]$ and $\tilde{v}' \in L(4,h)$ with $h = \left(\frac{\overline{y}}{V_0}\right)^{\frac{1}{2}} \tanh(T/2)$, if y satisfies (3.2.5). Taking advantage of the notation introduced in (3.3.26), we have

$$\tilde{v}'_t = 2V_0 u_t u'_t.$$

and a simple calculation yields

$$u'_t = \left(\left(\frac{\overline{y}}{V_0}\right)^{\frac{1}{2}} - e^{-T/2}\right)\frac{\cosh(t/2)}{2\sinh(T/2)} - \frac{1}{2}e^{-t/2}. \tag{3.3.30}$$

We remark that $u_t > 0$ for every $t \in [0,T]$ as soon as $y > X_0$, hence by (3.3.27) $u''_t > 0$, too, and consequently u'_t is an increasing function. Using the expression of u'_0 given by (3.3.30), it is easy to verify that (3.2.5) implies $u'_t \geq u'_0 \geq \frac{1}{4} > 0$. Now, we simply observe that $u \in L(2, ||u'||_\infty^{-1})$: indeed, for every $s, t \in [0,T]$ such that $|s-t| < ||u'||_\infty^{-1}$,

$$u_s \leq u_t + ||u'||_\infty |s-t| \leq u_t + 1 \leq 2u_t$$

where the last holds because $u_t \geq u_0 = 1$. Analogously, $u' \in L(2, ||u||_\infty^{-1})$ because

$$u'_s \leq u'_t + ||u''||_\infty |s-t| \leq u'_t + (1-a)^2 ||u||_\infty |s-t| \leq 2u'_t$$

holds if $|s-t| < ||u||_\infty^{-1}$, employing in the last step the fact that $u'_t \geq u'_0 \geq \frac{1}{4}$. Because u_t and

u'_t are increasing, we have

$$||u||_\infty = u_T = \Big(\frac{\overline{y}}{V_0}\Big)^{\frac{1}{2}}; \quad ||u'||_\infty = u'_T \leq \Big(\frac{\overline{y}}{V_0}\Big)^{\frac{1}{2}} \frac{1}{2\tanh(T/2)}$$

Observing that $\tanh(T/2) < 1$, we conclude that both u and u' belong to the class $L(2,h)$ with $h = \Big(\frac{V_0}{\overline{y}}\Big)^{\frac{1}{2}} \tanh(T/2)$. The fact that $\tilde{v}'_t \in L(4,h)$ for the same h now follows from the property $cfg \in L(\mu_f \times \mu_g, h_f \wedge h_g)$ if $f \in L(\mu_f, h_f)$, $g \in L(\mu_g, h_g)$ and c is a constant.

Step2. We estimate the integral functional at the right hand side of (3.3.5).
By Proposition 3.3.1 and the computations at the beginning of the current section, we know that the rate function $F_{\tilde{x},\tilde{v},\tilde{R}}$ is upper bounded by $\overline{F}_{\tilde{v}}$ defined in (3.3.20), more precisely $F_{\tilde{x},\tilde{v},\tilde{R}} \leq \frac{1}{h} + \Gamma_T\Big(\frac{(\tilde{v}'_t)^2}{\tilde{v}_t} + \tilde{v}_t\Big)$. Making once again use of u defined in (3.3.26), we have $(\tilde{v}'_t)^2 = 4V_0^2 u_t^2 (u'_t)^2$, hence

$$\begin{aligned}
\int_0^T F_{\tilde{x},\tilde{v},\tilde{R}}(t)dt &\leq \int_0^T \Big(\frac{1}{h} + \Gamma_T\Big(\frac{(\tilde{v}'_t)^2}{\tilde{v}_t} + \tilde{v}_t\Big)\Big)dt \\
&\leq \frac{T}{h} + \Gamma_T \int_0^T \Big(4V_0 \frac{u_t^2(u'_t)^2}{u_t^2} + V_0 u_t^2\Big)dt \\
&\leq \frac{T}{h} + 4\Gamma_T V_0 \int_0^T ((u'_t)^2 + u_t^2)dt \\
&\leq \Big(\frac{\overline{y}}{V_0}\Big)^{\frac{1}{2}} \frac{T}{\tanh(T/2)} + 4\Gamma_T V_0 \int_0^T ((u'_t)^2 + u_t^2)dt
\end{aligned}$$

and we just have to integrate the expressions for u_t and u'_t over $[0,T]$. Since we are interested in an upper bound for the integral, we simplify the computations using

$$u_t \leq \Big(\frac{\overline{y}}{V_0}\Big)^{\frac{1}{2}} \Big(\frac{\sinh(t/2)}{\sinh(T/2)} + 1\Big), \qquad u'_t \leq \frac{1}{2}\Big(\frac{\overline{y}}{V_0}\Big)^{\frac{1}{2}} \frac{\cosh(t/2)}{\sinh(T/2)}.$$

Hence, setting

$$\begin{aligned}
&c_T^{(1)} = \int_0^T \Big(\frac{\sinh(t/2)}{\sinh(T/2)} + 1\Big)^2 dt; \quad c_T^{(2)} = \frac{1}{4}\frac{1}{\sinh(T/2)^2} \int_0^T \cosh(t/2)^2 dt \\
&\tilde{c}_T = 2\Big(\frac{T}{\tanh(T/2)} + 4V_0(c_T^{(1)} + c_T^{(2)})\Big)
\end{aligned}$$

we obtain that, if y satisfies (3.2.5) so that in particular $(\frac{\overline{y}}{V_0})^{1/2} < \frac{\overline{y}}{V_0} < 2|y|$,

$$\int_0^T F_{\tilde{x},\tilde{v},\tilde{R}}(t)dt \leq \tilde{c}_T \Gamma_T |y| = \tilde{c}_T \Gamma_T |y|.$$

We remark that have $c_T^{(1)} \leq \int_0^T 4dt \leq 4T$, $c_T^{(2)} \leq \frac{1}{4}\frac{T}{\tanh(\frac{T}{2})^2} \leq \frac{1}{T} + T$ and $\frac{T}{\tanh(T/2)} \leq T+1$, hence $\tilde{c}_T \leq 4(20V_0+1)(\frac{1}{T}+T)$ for a positive constant c depending on V_0. On the other hand, recalling the expression of Γ_T from (3.3.21), we have $\Gamma_T \leq \frac{\Gamma}{\rho_\perp^2} e^{C_2 T^2}$ for a positive constant $\Gamma \geq 1$ depending on $V_0, k, \underline{\eta}, \underline{\sigma}, \overline{\eta}, \overline{\sigma}$ but not on $\rho_\perp$ or T, hence $\tilde{c}_T \Gamma_T \leq 4(20V_0+1)\frac{\Gamma}{\rho_\perp^2}(\frac{1}{T}+T)e^{C_2T^2} \leq \frac{c^*}{\rho_\perp^2}(\frac{1}{T}+1)e^{c^*T^2}$ with $c^* \geq 1$. Now, by (3.3.16), the constant $Q(\mu)$ in (3.3.7) is given by

$$Q(\mu) = \frac{\gamma^2 q}{\rho_\perp^4 \lambda^2} \ln \frac{\gamma q}{\rho_\perp^2 \lambda}$$

with $q = 8^{12} e^2 4^{73}$. Eventually multiplying the constant c^* by $\frac{\gamma^2 q}{\lambda^2} \ln \frac{\gamma q}{\lambda}$, we conclude that

$$\exp\Big(-Q(\mu)\Big(1+\int_0^T F_{\tilde{x},\tilde{v},\tilde{R}}(t)dt\Big)\Big) \leq \exp\Big(-2c^*\Big(\frac{1}{T}+1\Big)e^{c^*T^2}\frac{1}{\rho_\perp^6}\Big(\ln\Big(\frac{1}{\rho_\perp}\Big)+1\Big)\times|y|\Big) \quad (3.3.31)$$

for every y satisfying (3.2.5). Using (3.3.5) and the definition of ψ in (3.2.4), the proof is completed. □

We now prove Corollary 2.

Proof of Corollary 2. We consider $y \in R^*$ with $|y| > (1-V_0)/2$ and we now define $\overline{y} = 2|y|+V_0$ and consider $\tilde{x}_t, \tilde{v}_t, \tilde{R}_t$ as in (3.1.5). We remark that $\tilde{R}_T = \frac{1}{2}\sqrt{V_0 \wedge 1}\sqrt{2|y|+V_0} \leq \frac{y}{2}$ if $|y|$ is larger that the larger root of $|y|^2V_0 - 2|y| - V_0$. This holds in particular if $|y| > 2(V_0 \vee 1)(1+V_0)$ as in (3.2.8). If $y > 0$, we write

$$\mathbb{P}(X_T > y) \geq \mathbb{P}\Big(|X_T - 2y| \leq \frac{y}{2}\Big) \geq \mathbb{P}(|X_T - 2y| \leq \tilde{R}_T) \geq \mathbb{P}(|(X_T, V_T) - (2y, \overline{y})| \leq \tilde{R}_T)$$

and

$$\begin{aligned}\mathbb{P}(X_T < -y) \geq \mathbb{P}\Big(|X_T - (-2y)| \leq \frac{|y|}{2}\Big) &\geq \mathbb{P}(|X_T - (-2y)| \leq \tilde{R}_T) \\ &\geq \mathbb{P}(|(X_T, V_T) - (-2y, \overline{y})| \leq \tilde{R}_T)\end{aligned}$$

and in both cases the last term is larger than $\mathbb{P}(|(X_t, V_t) - (\tilde{x}_t, \tilde{v}_t)| \leq \tilde{R}_t, t \in [0,T])$. Theorem 3.2.1 then yields estimate (3.2.9). To prove (3.2.10), we shall first show that if $\mathbb{E}[e^{pX_T}] < \infty$, then $p \leq c_T\psi(\rho_\perp)$. Indeed, it is sufficient to observe that if $\mathbb{E}[e^{pX_T}] = C < \infty$, $p > 0$, then $\mathbb{P}(X_T > y) \leq Ce^{-p\times y}$ for all $y > 0$ by Markov' inequality:

$$\mathbb{P}(X_T > y) = \mathbb{P}(e^{pX_T} > e^{py}) \leq e^{-p\times y}\mathbb{E}[e^{pX_T}]. \quad (3.3.32)$$

Since (3.2.9) and (3.3.32) hold simultaneously for all y from a certain range on, clearly this implies $p \leq c_T\psi(\rho_\perp)$. With the same argument and using the estimate for $\mathbb{P}(X_T < -y)$, one shows that if $\mathbb{E}[e^{-qX_T}] < \infty$, $q > 0$, then $q \leq c_T\psi(\rho_\perp)$. Finally, the estimate (3.2.11) on the implied volatility is a direct consequence of moment formula (3.1.2) and of (3.2.10), recalling that the function φ is decreasing. $\square$

3.4 Proof of results in 3.2.3

We introduce some compact notation that will be used throughout this section. For t, s with $0 \le t < s \le T$ and $x_1, v_1 \in \mathbb{R} \times [0, \infty)$ we denote $(X_u^{t,x_1}, V_u^{t,v_1}; t \le u \le s)$ the solution of (3.2.1)-(3.2.2) on $[t, s]$ with initial conditions $X_t = x_1$ and $V_t = v_1$. We denote Y_u^{t,x_1,v_1} the couple $(X_u^{x_1}, V_u^{v_1})$ and

$$x_u^{x_1,x_2} = x_1 + \frac{x_2 - x_1}{s-t}(u-t), \quad u \in [t, s]$$
$$v_u^{v_1,v_2} = v_1 + \frac{v_2 - v_1}{s-t}(u-t), \quad u \in [t, s] \tag{3.4.1}$$

the line segments between (t, x_1),(s, x_2) and (t, v_1),(s, v_2) respectively. For $y \neq 0$ and a couple of radii R_1, R_2 with $0 < R_2 \le R_1 \le |y|$, we define

$$A_{t,s}^{x_1,v_1}(y, R_2) := \big\{|Y_u^{t,x_1,v_1} - (x_u^{x_1,y}, v_u^{v_1,y+|V_0|})| \le R_2, u \in [t, s]\big\};$$

$$p_{t,s}(y, R_1, R_2) = \inf_{(x_1,v_1)\in B_{R_1}(y,|y|+V_0)} \mathbb{P}\big(A_{t,s}^{x_1,v_1}(y, R_2)\big).$$

Moreover, we set

$$\epsilon_0 = \frac{\rho_\perp \underline{\eta}\, \underline{\sigma}}{4\sqrt{2}\rho\, \overline{\eta}} \wedge 1 \tag{3.4.2}$$

with $\epsilon_0 = 1$ if $\rho = 0$, and

$$\delta_0 = \frac{\epsilon_0^2 q}{160K^2} \wedge \frac{T}{2}; \qquad q = \mathbb{P}\Big(\sup_{u\le 1} |b_u| \le \frac{\epsilon_0}{4\sqrt{2}\,\overline{\sigma}}\Big) \tag{3.4.3}$$

where $(b_u, u \ge 0)$ is a standard Brownian motion under $\mathbb{P}$. The following lemma provides some estimates that will be used in the proof of Proposition 3.2.1 and Theorem 3.2.2.

Lemma 3.4.1. *Let $y \in \mathbb{R}$ with $|y| > 16$ and R_1, R_2 with $0 < R_2 \le R_1 \le \sqrt{|y|}$. Assume (R'). Then, for any $0 \le t < s \le T$,*

$$p_{t,s}(y, R_1, R_2) \ge \exp\Big(-c_T\psi(\rho_\perp)\Big(\frac{R_1^2}{(s-t)y} + \frac{y^2}{R_2^2}(s-t)\Big)\Big) \tag{3.4.4}$$

where c_T is the constant defined in (3.2.7). Moreover, if $y > 0$, for any $t > 0$ and any $0 < \delta < \frac{\delta_0}{y} \wedge t$ we have

$$\inf_{v\in B_{\epsilon_0\sqrt{\delta y}/2}(y)} \mathbb{P}\Big(|V_s^{t,v} - y| < \epsilon_0\sqrt{\delta y}, t-\delta \le s \le t;$$

$$\Big|\int_{t-\delta}^{t}(\sigma(u,V_u^{t,v})-\sigma(t-\delta,V_{t-\delta}^{t,v}))\sqrt{V_u^{t,v}}dW_u^1\Big|\leq\epsilon_0\sqrt{\delta y}\Big)\geq\frac{1}{2}q. \quad (3.4.5)$$

The proof of this lemma is not particularly enlightening for the rest of our study, hence we postpone it to Appendix 3.6.2. Here we give the proof of Proposition 3.2.1.

Proof (of Proposition 3.2.1). Step 1. We consider R_1, R_2 with $0 < R_1 < R_2 \leq \sqrt{|y|}$ and $\frac{T}{2} \leq t < s \leq T$. We have $\{Y_s \in B_{R_2}(y,|y|+V_0)\} \supset \{Y_t \in B_{R_1}(y,|y|+V_0)\} \cap \{Y_u \in B_{R_2}(x_u^{X_t,y}, v_u^{V_t,|y|+V_0}), t < u \leq s\}$. Hence, applying Markov property for the process Y

$$\begin{aligned}\mathbb{P}(Y_s\in B_{R_2}(y,|y|+V_0)) &\geq \mathbb{P}(\{Y_t\in B_{R_1}(y,|y|+V_0)\}\cap\{Y_u\in B_{R_2}(x_u^{X_t,u},v_u^{V_t,|y|+V_0}),t<u\leq s\})\\ &=\mathbb{E}\Big[1_{\{Y_t\in B_{R_1}(y,|y|+V_0)\}}\mathbb{E}\Big[1_{\{Y_u\in B_{R_2}(x_u^{X_t,y},v_u^{V_t,|y|+V_0}),t<u\leq s\}}|\mathcal{F}_t\Big]\Big]\\ &=\mathbb{E}\Big[1_{\{Y_t\in B_{R_1}(y,|y|+V_0)\}}\mathbb{E}\Big[1_{A_{t,s}^{x_1,v_1}(y,R_2)}|Y_t=(x_1,v_1)\Big]\Big]\\ &\geq \mathbb{P}(Y_t\in B_{R_1}(y,|y|+V_0))\times p_{t,s}(y,R_1,R_2).\end{aligned} \quad (3.4.6)$$

Step 2. We define the time step

$$\delta_j=\delta_j(y)=\frac{T}{2|y|^j},\quad j\geq 1.$$

Applying Lemma (3.4.1), for any $j \geq 1$ we have

$$\begin{aligned}\inf_{\frac{T}{2}\leq t\leq T-\delta_j} p_{t,t+\delta_j}(y,R^{(j-1)}(y),R^{(j)}(y)) &\geq \exp\Big(-c_T\psi(\rho_\perp)\Big(\frac{(R^{(j-1)}(y))^2}{\delta_j(y)|y|}+\frac{y^2}{(R^{(j)}(y))^2}\delta_j(y)\Big)\Big)\\ &=\exp\Big(-c_T\psi(\rho_\perp)\Big(2\frac{y^{2-j}}{T|y|^{1-j}}+\frac{y^2}{|y|^{1-j}}\frac{T}{2|y|^j}\Big)\Big)\\ &=\exp\Big(-2c_T\psi(\rho_\perp)\Big(\frac{1}{T}+T\Big)|y|\Big).\end{aligned} \quad (3.4.7)$$

On the other hand, $\tilde{R} = \frac{1}{2}\sqrt{(V_0\wedge 1)(|y|+V_0)} \leq \frac{1}{2}\sqrt{|y|+V_0} \leq \sqrt{\frac{|y|}{2}} \leq R^{(0)}(y)$. Applying Theorem (3.2.1) on the interval $[0,t]$, we have

$$\begin{aligned}\inf_{\frac{T}{2}\leq t\leq T}\mathbb{P}(Y_t\in B_{R^{(0)}(y)}(y,|y|+V_0)) &\geq \inf_{\frac{T}{2}\leq t\leq T}\mathbb{P}(Y_t\in B_{\tilde{R}}(y,|y|+V_0))\\ &\geq \inf_{\frac{T}{2}\leq t\leq T}\exp\Big(-c_t\Big(\frac{1}{t}+1\Big)\psi(\rho_\perp)|y|\Big) \quad (3.4.8)\\ &\geq \exp\Big(-2c_T\Big(\frac{1}{T}+1\Big)\psi(\rho_\perp)|y|\Big).\end{aligned}$$

Step 3. We fix $j \in \mathbb{N}^*$ and define

$$t_k^j = T - \sum_{h=1}^{k} \delta_{j-h+1}, \quad 0 \le k \le j,$$

so that $t_0 = T$ and $t_{k-1}^j - t_k^j = \delta_{j-k+1}$ for $1 \le k \le j$. Moreover, since $\sum_{k=1}^{\infty} \delta_k = \frac{T}{2} \sum_{k=1}^{\infty} \frac{1}{y^j} \le \frac{T}{2} \frac{1}{y-1} \le \frac{T}{2}$, we have $t_k^j \ge \frac{T}{2}$ for all $j \in \mathbb{N}^*, 1 \le k \le j$. Repeatedly applying (3.4.6) and (3.4.7), we get

$$\begin{aligned}
\mathbb{P}(Y_T \in B_{R^{(j)}}(y, |y| + V_0)) &= \mathbb{P}(Y_{t_0} \in B_{R^{(j)}}(y, |y| + V_0)) \\
&\ge \mathbb{P}(Y_{t_j} \in B_{R^{(0)}}(y, |y| + V_0)) \prod_{k=1}^{j} p_{t_k^j, t_{k-1}^j}(y, R^{(j-k)}(y), R^{(j-k+1)}(y)) \\
&\ge \mathbb{P}(Y_{t_j} \in B_{R^{(0)}}(y, |y| + V_0)) \times \exp\Big(-2jc_T\Big(\frac{1}{T} + T\Big)\psi(\rho_\perp)|y|\Big) \\
&\ge \exp\Big(-2(j+1)c_T\Big(\frac{1}{T} + T \vee 1\Big)\psi(\rho_\perp)|y|\Big)
\end{aligned}$$

and in the last step we have applied (3.4.8). Using the expression for the the constant c_T given in (3.2.7), we have $c_T\Big(\frac{1}{T} + T \vee 1\Big) \le c^*\Big(\frac{1}{T} + 1\Big)e^{c^*T^2}\Big(\frac{1}{T} + T \vee 1\Big) \le 2c^*\Big(\frac{1}{T^2} + 1\Big)e^{(c^*+1)T^2}$ and (3.2.12) is proved. □

Let us go back Theorem 3.2.2. To lower bound the density of X_T we follow the approach of [6], section 5. The idea is to treat X_T as a random variable of the form

$$F = x + G + R,$$

where $x \in \mathbb{R}$, $R \in \mathbb{D}^{2,\infty}$ and G is a Wiener integral $G = \sum_{j=1,2} \int_0^T h_j(t) dW_t^j$, with $h_j : [0, \infty) \to \mathbb{R}$ deterministic. Here $\mathbb{D}^{2,\infty}$ denotes the space of the random variables which are two times Malliavin differentiable in L^p for every $p \ge 2$. Remark that G is a centered Gaussian random variable with variance $\Delta = \sum_j \int_0^T h_j(t)^2 dt > 0$. Let $g_\Delta(y) = \frac{1}{\sqrt{2\pi\Delta}} \exp(-\frac{y^2}{2\Delta})$ denote the density of G and $||R||_{2,p}$ the stochastic Sobolev norm of R of order two. Our starting point is the following result due to Bally and Caramellino in [6], which we restate here in a form suitable for our purposes.

Proposition 3.4.1 (Proposition 8 in [6]). *If the law of F has a density p_F, then for any $f \in C_b(\mathbb{R})$ one has*

$$\int_R f(y) p_F(y) \ge \int_R f(y)\big(g_\Delta(y - x) - \epsilon(\Delta, R)\big) dy, \tag{3.4.9}$$

with

$$\epsilon(\Delta, R) = \frac{C^*}{\sqrt{\Delta}}\big(1 + ||R_\Delta||_{2,q^*}\big)^{l^*}||R_\Delta||_{2,q^*}$$

where $R_\Delta = R/\sqrt{\Delta}$ *and* C^*, q^*, l^* *are universal constants.*

Proof. Using point i) of Proposition 8 in [6], we know that there exists a probability measure $\overline{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ such that $\frac{d\overline{\mathbb{P}}}{d\mathbb{P}} \leq 1$ and the law of F under $\overline{\mathbb{P}}$ is absolutely continuous with respect to the Lebesgue measure. Again according to [6], the associated density $\overline{p}_F$ satisfies

$$\sup_{y\in\mathbb{R}} |\overline{p}_F(y) - g_\Delta(y - x)| \leq \epsilon(\Delta, R)$$

for the given $\epsilon(\Delta, R)$. (We refer to [6] for the explicit construction of the probability $\overline{\mathbb{P}}$). Then, for any $f \in C_b(\mathbb{R})$ we have

$$\int_R f(y)p_F(y) = \mathbb{E}[f(F)] \geq \mathbb{E}\Big[f(F)\frac{d\overline{\mathbb{P}}}{d\mathbb{P}}\Big] \geq \int_R f(y)\big(g_\Delta(y-x) - \epsilon(\Delta, R)\big)dy$$

which proves (3.4.9).

Remark 3.4.1. If the density p_F is continuous, then (3.4.9) implies $p_F(y) \geq g_\Delta(y - x) - \epsilon(\Delta, R)$ for all $y \in \mathbb{R}$.

Remark 3.4.2. We shall use conditional calculus in order to prove Theorem 3.2.2 : in particular, we will work with Malliavin derivatives only with respect to the Brownian noise $W_t, t \in [T-\delta, T]$, and consider conditional expectations with respect to $\mathcal{F}_{T-\delta}$, for a $\delta < T$. As in the proof of Theorem 2.2.4, this allows us to gain a free parameter δ in (3.4.9) that we can eventually optimize, and this feature turns out to be crucial in our analysis (cf. Propositions 3.4.2 and 3.4.3 hereafter). The use of conditional Malliavin calculus in order to derive lower bounds for the density of a random variable is not new and has been employed by, among others, [42], [4] and [32]. In our framework, we face some supplementary difficulties. Let us point them out: first, to estimate the marginal density of X_T we have to separately estimate the whole path of the stochastic volatility V up to time T. This was the motivation of estimate (3.4.5). Second, in order to manipulate the Sobolev norms of R we need all the involved random variables to be smooth in Malliavin sense, but this is not guaranteed in our framework due to the presence of the non-Lipschitz square-root coefficients in (3.2.1) and (3.2.2). This is why we introduce a regularization of the coefficients of the SDE, as we do hereafter.

Let us implement what stated in Remark 3.4.2. We consider the case of positive y in Theorem 3.2.2: the case of negative y is proven in the analogous manner. We assume $y > 2$

and introduce two parameters $\delta > 0$ and $l \in \mathbb{N}$ such that:

$$\delta < \frac{\delta_0}{y^2}; \qquad \frac{1}{y^l} < \frac{1}{2}\epsilon_0 \rho_\perp \underline{\eta} \sqrt{y\delta} \tag{3.4.10}$$

and ϵ_0, δ_0 as defined in (3.4.2) and (3.4.3). We remark that for such a value of δ we have $\epsilon_0\sqrt{\delta y} < \epsilon_0\sqrt{\delta_0} < 1$. Then, we consider a truncation function $\psi \in C_b^\infty(\mathbb{R},\mathbb{R})$ such that $\psi(x) = x$ for $|x-y| \le 1$, $\psi(x) = y - \frac{3}{2}$ for $x \le y-2$ and $\psi(x) = y + \frac{3}{2}$ for $x \ge y+2$. ψ can be defined in such a way that $|\psi|_0 \le y + \frac{3}{2} \le 2y$ and $\sum_{j=1}^k \sup_{x\in\mathbb{R}} |\psi^{(j)}(x)| \le 2^{\frac{k(k-1)}{2}}$. We define the sets

$$A_{\delta,l}(X,V) = \{|X_{T-\delta} - y| < \frac{1}{y^l}, |V_{T-\delta} - (y+V_0)| < \frac{\epsilon_0}{2}\sqrt{y\delta}\}$$

and

$$\overline{A}_\delta(V) = \Big\{|V_s - y| < \epsilon_0\sqrt{\delta y}, T-\delta < s \le T; \\ \Big|\int_{T-\delta}^T (\sigma(u,V_u) - \sigma(T-\delta, V_{T-\delta}))\sqrt{V_u}dW_u^1\Big| \le \epsilon_0\sqrt{\delta y}\Big\},$$

and denote

$$A_{\delta,l} = A_{\delta,l}(X,V) := A_{\delta,l}(X,V) \cap \overline{A}_\delta(V).$$

Finally, we consider $(\overline{X}_t, \overline{V}_t; T-\delta \le t \le T)$ the (unique strong) solution to the equation

$$\begin{aligned} \overline{X}_t &= X_{T-\delta} - \frac{1}{2}\int_{T-\delta}^t \eta(s,\overline{X}_s)^2\psi(\overline{V}_s)ds + \int_{T-\delta}^t \eta(s,\overline{X}_s)\sqrt{\psi(\overline{V}_s)}(\rho dW_s^1 + \rho_\perp dW_s^2), \\ \overline{V}_t &= V_{T-\delta} + \int_{T-\delta}^t \beta(s,\overline{V}_s)ds + \int_{T-\delta}^t \sigma(s,\overline{V}_s)\sqrt{\psi(\overline{V}_s)}dW_s^1. \end{aligned} \tag{3.4.11}$$

We remark that on the set $A_{\delta,l}$, $\psi(V_t) = V_t$ for all $t \in [T-\delta, T]$. Hence, since pathwise uniqueness holds for (3.4.11), we have $(X_t,V_t)(\omega) = (\overline{X}_t, \overline{V}_t)(\omega)$ for $(t,\omega) \in [T-\delta,T] \times A_{\delta,l}$ and in particular $A_{\delta,l} = A_{\delta,l}(X,V) \subset A_{\delta,l}(\overline{X},\overline{V})$. Under hypothesis (R'), $\overline{X}_t$ and $\overline{V}_t$ belong to the space $\mathbb{D}^{2,p}$ associated to $(W_t^1, W_t^2), t \in [T-\delta,T]$, for all $p > 1$.
We decompose the random variable $\overline{X}_T$ in the following way:

$$\overline{X}_T = G_0 + G + R,$$

where

$$\begin{aligned}
G_0 &= X_{T-\delta} + \rho\eta(T-\delta, X_{T-\delta}) \int_{T-\delta}^{T} \sqrt{\psi(\overline{V}_t)}dW_t^1 \\
G &= \rho_\perp \eta(T-\delta, X_{T-\delta}) \int_{T-\delta}^{T} \sqrt{\psi(\overline{V}_t)}dW_t^2 \\
R &= -\frac{1}{2}\int_{T-\delta}^{T} \eta(t, \overline{X}_t)^2 \psi(\overline{V}_t)dt \\
&\qquad + \int_{T-\delta}^{T} (\eta(t,\overline{X}_t) - \eta(T-\delta, X_{T-\delta}))\sqrt{\psi(\overline{V}_t)}(\rho dW_t^1 + \rho_\perp dW_t^2).
\end{aligned} \tag{3.4.12}$$

Conditional to $\mathcal{F}_{T-\delta} \vee \mathcal{F}_T^1$, the random variable G is a centered Gaussian with variance $I = \rho_\perp^2 \eta(T-\delta, X_{T-\delta})^2 \int_{T-\delta}^{T} \psi(\overline{V}_t)dt$. By the definition of ψ, we have $I \geq \rho_\perp^2 \underline{\eta}^2 \int_{T-\delta}^{T} (y - \frac{3}{2})dt \geq \frac{1}{2}\rho_\perp^2 \underline{\eta}^2 y\delta$. Similarly, we can see that an upper bound for I is given by $2\rho_\perp^2 \overline{\eta}^2 y\delta$, hence

$$\Delta \leq I = Var(G|\mathcal{F}_{T-\delta} \vee \mathcal{F}_T^1) \leq a\Delta \tag{3.4.13}$$

with

$$\Delta = \frac{1}{2}\rho_\perp^2 \underline{\eta}^2 y\delta, \qquad a = 4\frac{\overline{\eta}^2}{\underline{\eta}^2}.$$

Using (3.4.13) and Lemma (3.4.1), we can prove the following statement (which is the analogous of Lemma 5 in [4]):

Proposition 3.4.2. *Let $g(\cdot|F_{T-\delta} \vee \mathcal{F}_T^1)$ denote the density of G conditional to $F_{T-\delta} \vee \mathcal{F}_T^1$. Then, for any y, δ, l satisfying (3.4.10),*

$$g(y - G_0|F_{T-\delta} \vee \mathcal{F}_T^1) \geq \frac{1}{\rho_\perp \overline{\eta} e \sqrt{4\pi\delta y}} \qquad \text{on the set } A_{\delta,l}. \tag{3.4.14}$$

Proof. Recall that $I = \rho_\perp^2 \eta(T-\delta, X_{T-\delta})^2 \int_{T-\delta}^{T} \psi(\overline{V}_t)dt$. Moreover, let us set $J = \rho\eta(T - \delta, X_{T-\delta}) \int_{T-\delta}^{T} \sqrt{\psi(\overline{V}_t)}dW_t^1$ (so that $G_0 = X_{T-\delta} + J$). Then

$$g(y - G_0|F_{T-\delta} \vee \mathcal{F}_T^1) = \frac{1}{\sqrt{2\pi I}} \exp\Big(-\frac{1}{2I}(y - (X_{T-\delta} + J))^2\Big). \tag{3.4.15}$$

Since $I \geq \Delta$ and $|y - X_{T-\delta}| \leq \frac{1}{y^l}$ on $A_{\delta,l}$, on this set we have

$$\frac{|y - X_{T-\delta}|}{\sqrt{I}} \leq \frac{\frac{1}{y^l}}{\sqrt{\Delta}} \leq 1$$

where the last inequality holds because of (3.4.10). Now, using equation (3.4.11) for $\overline{V}$,

$$
\begin{aligned}
\sigma(T-\delta,\overline{V}_{T-\delta})\int_{T-\delta}^{T}\sqrt{\psi(\overline{V}_t)}dW^1_t = \overline{V}_T - \overline{V}_{T-\delta} - \int_{T-\delta}^{T}\beta(t,\overline{V}_t)dt \\
- \int_{T-\delta}^{T}(\sigma(t,\overline{V}_t)-\sigma(T-\delta,\overline{V}_{T-\delta}))\sqrt{\psi(\overline{V}_t)}dW^1_t,
\end{aligned}
$$

hence, on the set $A_{\delta,l}$

$$
\begin{aligned}
\Big|\int_{T-\delta}^{T}\sqrt{\psi(\overline{V}_t)}dW^1_t\Big| &\le \frac{1}{\underline{\sigma}}\Big(|\overline{V}_T-\overline{V}_{T-\delta}|+\int_{T-\delta}^{T}|\beta(t,\overline{V}_t)|dt \\
&\qquad + \Big|\int_{T-\delta}^{T}(\sigma(t,\overline{V}_t)-\sigma(T-\delta,\overline{V}_{T-\delta}))\sqrt{\psi(\overline{V}_t)}dW^1_t\Big|\Big) \\
&\le \frac{1}{\underline{\sigma}}\Big(2\epsilon_0\sqrt{y\delta}+K(1+y+\epsilon_0\sqrt{y\delta})\delta+\epsilon_0\sqrt{y\delta}\Big) \\
&\le \frac{1}{\underline{\sigma}}(2\epsilon_0+2K\sqrt{y\delta}+\epsilon_0)\sqrt{y\delta} \le \frac{4\epsilon_0}{\underline{\sigma}}\sqrt{y\delta}
\end{aligned}
$$

and the two last inequality are obtained using $K(1+y+\epsilon_0\sqrt{y\delta})\delta < K(1+y+1)\delta < 2Ky\delta$, then $2K\sqrt{y\delta}\le 2K\sqrt{\delta_0}\le\epsilon_0$ after (3.4.3). The previous estimate yields $|J|\le\frac{4\epsilon_0\rho\overline{\eta}}{\underline{\sigma}}\sqrt{y\delta}$, hence

$$
\frac{|J|}{\sqrt{I}} \le \frac{|J|}{\sqrt{\Delta}} \le \epsilon_0\frac{4\sqrt{2}\rho\overline{\eta}}{\rho_\perp\underline{\eta}\underline{\sigma}} \le 1
$$

and the last inequality holds after (3.4.2). Finally, for the exponential term in (3.4.15) we have

$$
\exp\Big(-\frac{1}{2I}\Big(y-(X_{T-\delta}-J)\Big)^2\Big) \ge e^{-\frac{1}{2}(1+1)^2} \ge e^{-2} \tag{3.4.16}
$$

on the set $A_{\delta,l}$. Since $I\le a\Delta = 2\rho_\perp^2\overline{\eta}y\delta$, (3.4.16) yields (3.4.14). □

The second result we need in order to prove Theorem 3.2.2 is an estimation of the reminder R. Let $R_\Delta := R/\sqrt{\Delta}$ as in Proposition 3.4.1.

Proposition 3.4.3. *Let y,δ,l satisfy (3.4.10). Then, for every $p>1$ there exists a positive constant c_p such that*

$$
||R_\Delta||_{T-\delta,\delta,2,p} \le \frac{c_p}{\rho_\perp\underline{\eta}}\sqrt{\delta y^{31}}\times e^{c_pT^p} \qquad \text{on the set } A_{\delta,l}. \tag{3.4.17}
$$

The constant c_p depends also on K but not on the other model parameters.

Remark 3.4.3. Similar estimates (with different powers of δ and y) could be obtained for $||R_\Delta||_{T-\delta,\delta,k,p}$, $k>2$, under the corresponding regularity assumptions on the coefficients η,β and σ.

We can now prove Theorem 3.2.2.

Proof (of Theorem 3.2.2). We make an explicit choice of δ and l:

$$\delta = \delta_0 \Theta_T^2 \times y^{-31}; \qquad l = 16; \qquad \text{with} \quad \Theta_T = \frac{\rho_\perp \underline{\eta}^2 2^{-l^*-5/2}}{\sqrt{\pi} e \overline{\eta} C^* c_p} e^{-2c_p T^p} \tag{3.4.18}$$

where C^* and c_p are the constant appearing in, respectively, (3.4.9) and (3.4.17). Conditions (3.4.10) are satisfied as soon as

$$y > M_T \quad \text{with} \quad M_T = 2/(\epsilon_0 \rho_\perp \underline{\eta} \sqrt{\delta_0} \Theta_T).$$

We now apply Proposition 3.4.1 to the law of $\overline{X}_T$ conditional to $\overline{\mathcal{F}} = \mathcal{F}_{T-\delta} \vee \mathcal{F}_T^1$. Let $p_{\overline{X}_T}(\cdot|\overline{\mathcal{F}})$ (resp. $g(\cdot|\overline{\mathcal{F}})$) denote the conditional density of $\overline{X}_T$ (resp. G), by Prop. 3.4.1 we have

$$\begin{aligned} \int_{y>M_T} f(y) p_{\overline{X}_T}(y|\overline{\mathcal{F}}) dy &\geq \int_{y>M_T} f(y) \big(g(y - G_0|\overline{\mathcal{F}}) - \epsilon(\Delta, R) \big) dy \\ &\geq \int_{y>M_T} f(y) \Big(\frac{1}{\rho_\perp \overline{\eta} e \sqrt{4\pi\delta y}} - \epsilon(\Delta, R) \Big) dy \end{aligned}$$

with

$$\begin{aligned} \epsilon(\Delta, R) &= \frac{C^*}{\sqrt{\Delta}} \big(1 + \mathbb{E}[||R_\Delta||_{2,q^*}|\overline{\mathcal{F}}] \big)^{l^*} \mathbb{E}[||R_\Delta||_{2,q^*}|\overline{\mathcal{F}}] \\ &= \frac{C^*}{\sqrt{\Delta}} \big(1 + \mathbb{E}[||R_\Delta||_{T-\delta,\delta,2,q^*}|\mathcal{F}_T^1] \big)^{l^*} \mathbb{E}[||R_\Delta||_{T-\delta,\delta,}|\mathcal{F}_T^1]. \end{aligned}$$

For the given value of δ, $||R_\Delta||_{T-\delta,\delta,2,q^*} < 1$ on the set $A_{\delta,l}$, hence by Prop. 3.4.3 $\epsilon(\Delta, R)$ is bounded by

$$\epsilon(\Delta, R) \leq \frac{C^* c_p}{\rho_\perp^2 \underline{\eta}^2} \times 2^{l*+1/2} \times \sqrt{\delta y^{31}} e^{C_p T^p}$$

on the set $A_{\delta,l}$. The value of δ in (3.4.18) is chosen in such a way that the right hand side in this last estimate is smaller than $\frac{1}{2} \times \frac{1}{\rho_\perp \overline{\eta} e \sqrt{4\pi\delta y}}$, hence

$$\int_{y>M_T} f(y) p_{\overline{X}_T}(y|\overline{\mathcal{F}}) dy \geq \int_{y>M_T} \frac{f(y)}{\rho_\perp \overline{\eta} e \sqrt{4\pi\delta y}} dy = \int_{y>M_T} \frac{f(y)}{\rho_\perp \overline{\eta} e \sqrt{4\pi\delta_0} \Theta_T} y^{15} dy$$

on the set $A_{\delta,l}$.

Let us now estimate the probability of the set $A_{\delta,l}$. Since $\frac{1}{y^l} < \frac{\epsilon_0}{2} \sqrt{y\delta}$ by condition (3.4.10), then $A_{T-\delta} \supset \{(X_{T-\delta}, V_{T-\delta}) \in B_{\frac{1}{y^l}}(y, y + V_0)\}$. Hence, we apply Prop. 3.2.1 for $j = 2l + 1$ and obtain

$$\mathbb{P}(A_{T-\delta}) \geq \mathbb{P}\big((X_{T-\delta}, V_{T-\delta}) \in B_{\frac{1}{y^l}}(y, y + V_0) \big) \geq \exp\Big(-(2l+2) d_{T-\delta} \psi(\rho_\perp) y \Big)$$

Since $2l + 2 = 34$ and $\frac{1}{(T-\delta)^2} < \frac{4}{T^2}$, then $(2l+2) d_{T-\delta} \leq 136 c^* \Big(\frac{1}{T^2} + 1 \Big) e^{(c^*+1)T} := e_T$ and $\mathbb{P}(A_{T-\delta}) \geq \exp(-e_T \psi(\rho_\perp) y)$. Applying Markov property for the process V and (3.4.5) in

Lemma 3.4.1, it is easy to see that

$$\mathbb{P}(A_{\delta,l}) \geq \frac{q}{2}\mathbb{P}(A_{T-\delta}) \geq \frac{q}{2}\exp\Big(-e_T\psi(\rho_\perp)y\Big). \tag{3.4.19}$$

Finally, let us denote $p_{X_T}(\cdot|\overline{\mathcal{F}})$ the density of X_T conditional to $\overline{\mathcal{F}}$. We have

$$\begin{aligned}
\int_{\mathbb{R}} f(y)p_{X_T}(y)dy &= \int_{\mathbb{R}} f(y)\mathbb{E}[p_{X_T}(y|\overline{\mathcal{F}})]dy \\
&\geq \int_{\mathbb{R}} f(y)\mathbb{E}[p_{X_T}(y|\overline{\mathcal{F}})1_{A_{\delta,l}}]dy \\
&= \int_{\mathbb{R}} f(y)\mathbb{E}[p_{\overline{X}_T}(y|\overline{\mathcal{F}})1_{A_{\delta,l}}]dy \\
&\geq \int_{\mathbb{R}} f(y)\frac{1}{\rho_\perp\overline{\eta}e\sqrt{4\pi\delta_0}\Theta_T}y^{15}\times\mathbb{P}(A_{\delta,l})dy.
\end{aligned}$$

Using estimate (3.4.19), we obtain (3.2.13). □

3.5 Conclusions

We have shown that the left and right tails of the distribution of the log price X decay no faster than exponentials in local stochastic volatility models driven by square root diffusions - namely, in the model class (3.1.3)-(3.1.4) - no matter how the (possibly time-dependent) skew function η, the volatility drift β and volatility of variance σ are chosen, provided they satisfy some reasonable boundedness and linear-growth conditions - namely, conditions (R) and (G) in section 3.2. Together with the elementary observation that e^X is an integrable supermartingale, this yields the “sandwich” estimate $e^{-c_1(t)y} \leq \mathbb{P}(X_t > y) \leq e^{-c_2(t)y}$ for large values of y. From the point of view of the financial modelisation, our estimate has an impact on moment explosion and, by Lee's moment formula, it translates into lower bounds on the asymptotic slopes of the implied volatility.

Our result is not limited to fixed-time marginal laws: we have shown that the exponential lower bound actually holds for the probability that the whole trajectory of the couple (X, V) remains in a “tube” of given deterministic radius around a given deterministic curve for all the times up to a given maturity. This means that our main estimate can also be applied to the two-dimensional joint distribution of X and V and to study the law of suprema of the components of the solution to (3.1.3)-(3.1.4). Back to the financial level, this can eventually lead to bounds on the prices of barrier and exotic options. We have also shown how one can apply density estimation techniques for locally-elliptic random variables on the Wiener space and Malliavin calculus tools to prove that a lower bound in the same asymptotic range holds the for the density of X as well.

3.6 Technical proofs

3.6.1 Preliminary estimates

Lemma 3.6.1. *Assume* (G) *and let* $(X_t, V_t; 0 \le t \le T)$ *be two processes satisfying (3.2.1)-(3.2.2). Then for every* $0 \le t \le s \le T$ *and every* $p \ge 1$ *there exist a positive constant* C_p *such that*

$$\mathbb{E}\Big[\sup_{t\le r\le s}\Big(|X_r - X_t|^{2p} + |V_r - V_t|^{2p}\Big)\Big|\mathcal{F}_t\Big] \le C_p(s-t)^p \exp(C_p s^{2p}). \tag{3.6.1}$$

C_p *also depends on the parameters* $K, \overline{\eta}, \overline{\sigma}$ *given in* (G) *and on* V_0.

Proof. Observing that both the functions $v \to \beta(t,v)$ and $v \to \sigma(t,v)\sqrt{v}$ have sub-linear growth under (G), (3.6.1) follows from the application of Burkholder's inequality and Gronwall's Lemma to the process $(V_t; 0 \le t \le T)$ satisfying (3.2.1), then to $(X_t; 0 \le t \le T)$ satisfying (3.2.2).

3.6.2 Proof of Lemma 3.4.1

Proof. Estimate (3.4.4): Consider $(x_1, v_1) \in B_{R_1}(y, |y| + V_0)$. On the set $A^{x_1,v_1}_{t,s}(x^{x_1,y}_\cdot, v^{v_1,|y|+V_0}_\cdot, R_\cdot)$, with $R_u = R_2, u \in [t,s]$, we have $V_u > |y| - R_1 - R_2 > |y| - 2\sqrt{|y|}$ and $V_u < |y| + R_1 + R_2 < |y| + 2\sqrt{|y|}$ for all $u \in [t,s]$. Therefore, $\frac{1}{2}|y| < V_u < 2|y|$ for all $u \in [t,s]$, if $|y| > 16$. Using estimates (3.3.10), (3.3.12), (3.3.13) and (3.3.14) it is easy to show that conditions (3.3.2),(3.3.3) and (3.3.4) are satisfied for the process $(X^{t,x_1}_u, V^{t,v_1}_u; t \le u \le s)$, the curves $x^{x_1,y}_\cdot, v^{v_1,|y|+V_0}_\cdot$, the radius $R_\cdot$ and the *constant* curves

$$\begin{aligned} c_u &= c|y|; \qquad & L_u^2 &= L_T|y|; \\ \gamma_u &= \gamma|y|; \qquad & \lambda_u &= \rho_\perp^2\lambda|y|, \end{aligned} \tag{3.6.2}$$

defined for $u \in [t,s]$, where c, L_T, γ, λ are the same as in Proposition 3.3.1. The derivatives $x'_\cdot, v'_\cdot$ and the radius $R_\cdot$ being constant, all the involved curves belong to $L(1,\infty)$. The factor $\phi_{\lambda,\gamma}$ in (3.3.7) is still given by $\phi_{\lambda,\gamma} = \frac{\rho_\perp^2\lambda}{\gamma}$, hence $Q_{\lambda,\gamma} = \frac{\gamma^2 q}{\rho_\perp^4\lambda^2}\ln\frac{\gamma q}{\rho_\perp^2\lambda}$ with $q = 8^{12}e^2$. On the other hand, using the constant Γ_T defined in (3.3.21), we have

$$\begin{aligned} \int_t^s F_{x,v,R}(u)du &= \int_t^s \Big(\frac{(x'_u)^2 + (v'_u)^2}{\lambda_u} + 2(c_u^2 + L_u^2)\Big(\frac{1}{\lambda_u} + \frac{1}{R_u^2}\Big)\Big)du \\ &\le 4\Gamma_T \int_t^s \Big(\frac{R_1^2}{(s-t)^2 y} + y^2\Big(\frac{1}{y} + \frac{1}{R_2^2}\Big)\Big)du \\ &\le 8\Gamma_T\Big(\frac{R_1^2}{(s-t)y} + \frac{y^2}{R_2^2}\Big)(s-t)\Big). \end{aligned}$$

Estimate (3.3.5) then tells that (3.4.4) holds with the same constant c_T defined in (3.2.7).

(3.4.5): We fix $t > 0$, $v \in B_{\epsilon_0\sqrt{\delta y}/2}(y)$ and write $V = V^{t,v}$ for simplicity. We set $B = \{|V_s -$

$y| < \epsilon_0\sqrt{\delta y}, t-\delta \le s \le t\}$ and $C = \Big\{\Big|\int_{t-\delta}^t (\sigma(u,V_u)-\sigma(t-\delta,V_{t-\delta}))\sqrt{V_u}dW_u^1\Big| \le \epsilon_0\sqrt{\delta y}\Big\}$. Since $\epsilon_0\sqrt{\delta y} < \epsilon_0\sqrt{\delta_0} < 1$, on the set B we have $V_s < 1+y < 2y$ and $|\beta(s,V_s)| < K(2+y) < 2Ky$ for all $s \in [t-\delta, t]$. Hence, on the set B

$$|V_s - v| \le 2Ky\delta + \Big|\int_{t-\delta}^s \sigma(u,V_u)\sqrt{V_u \wedge 2y}dW_u^1\Big|$$

and $2Ky\delta < \frac{\epsilon_0}{4}\sqrt{\delta y}$ because $\delta < \delta_0/y$. Therefore

$$\mathbb{P}(B) \ge \mathbb{P}\Big(\sup_{t-\delta\le s\le t}|V_s - v| \le \frac{\epsilon_0}{2}\sqrt{\delta y}\Big) \ge \mathbb{P}\Big(\sup_{t-\delta\le s\le t}\Big|\int_{t-\delta}^s \sigma(u,V_u)\sqrt{V_u\wedge 2y}dW_u^1\Big| \le \frac{\epsilon_0}{4}\sqrt{\delta y}\Big).$$

We time-change the stochastic integral into b_{A_s}, where $(b_s; s\ge 0)$ is a standard Brownian motion and we denote $A_s = \int_{t-\delta}^s \sigma(u,V_u)^2(V_u\wedge 2y)du$ the quadratic variation (Dubins & Schwartz theorem, cf. Th. 3.4.6 in [40]). Since A_s, $s\in[t-\delta,t]$, is uniformly bounded by $\overline{\sigma}^2 2y\delta$, we have $\sup_{t-\delta\le s\le t}|b_{A_s}| \le \sup_{0\le s\le 2\overline{\sigma}^2 y\delta}|b_s|$. Using the scaling propertyof the Brownian motion $(b_{cs}; s\ge 0) \sim (\sqrt{c}b_s; s\ge 0)$ we obtain

$$\mathbb{P}(B) \ge \mathbb{P}\Big(\sup_{0\le s\le 2\overline{\sigma}^2 y\delta}|b_s| \le \frac{\epsilon_0}{4}\sqrt{\delta y}\Big) = \mathbb{P}\Big(\sup_{0\le s\le 1}|b_s| \le \frac{\epsilon_0}{4\sqrt{2\overline{\sigma}}}\Big) = q.$$

The same arguments lead to

$$\begin{aligned}\mathbb{P}(B\cap C^c) &\le \mathbb{P}\Big(\Big|\int_{t-\delta}^t (\sigma(u,V_u)-\sigma(t-\delta,V_{t-\delta}))\sqrt{V_u}dW_u^1\Big| > \epsilon_0\sqrt{\delta y}; V_u < 2y, t-\delta\le u\le t\Big)\\ &\le \mathbb{P}\Big(\Big|\int_{t-\delta}^t (\sigma(u,V_u)-\sigma(t-\delta,V_{t-\delta}))\sqrt{V_u\wedge 2y}dW_u^1\Big| > \epsilon_0\sqrt{\delta y}\Big)\\ &= \mathbb{P}(|\tilde{b}_{B_t}| > \epsilon_0\sqrt{\delta y})\end{aligned}$$

where $B_t = \int_{t-\delta}^t (\sigma(u,V_u)-\sigma(t-\delta,V_{t-\delta}))^2(V_u\wedge 2y)du$ and $(\tilde{b}_s; s\ge 0)$ is a standard Brownian motion. Using (R), we have $B_t \le 2K^2(\delta^2 + |V_u - V_{t-\delta}|^2)\cdot 2y\cdot\delta \le 4K^2(\delta^2 + 4\epsilon_0^2\delta y)\cdot y\delta \le 20K^2\epsilon_0^2(\delta y)^2$, hence

$$\begin{aligned}\mathbb{P}(B\cap C^c) &\le \mathbb{P}\Big(\sup_{s\le 20K^2\epsilon_0^2(\delta y)^2}|\tilde{b}_s| > \epsilon_0\sqrt{\delta y}\Big)\\ &= \mathbb{P}\Big(\sup_{s\le 1}|\tilde{b}_s| > \frac{1}{\sqrt{20K^2\delta y}}\Big)\\ &\le 20K^2\delta y\,\mathbb{E}\Big[\sup_{s\le 1}|\tilde{b}_s|^2\Big] \le 80K^2\delta_0 \le \frac{1}{2}q\end{aligned}$$

and we have used Doob's inequality and the value of δ_0 to get the two last inequalities. We conclude that

$$\mathbb{P}(B\cap C) = \mathbb{P}(B) - \mathbb{P}(C^c\cap B) \ge q - \frac{1}{2}q = \frac{1}{2}q.$$

$\square$

3.6.3 Tools of Conditional Malliavin calculus

We briefly introduce the main elements of conditional Malliavin calculus. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $(\mathcal{F}_t, t \geq 0)$, and a d-dimensional Brownian motion $(W_t, t \geq 0)$ with respect to this filtration. We fix some $t > 0$ and $\delta > 0$. Conditional Malliavin calculus amounts to consider the standard Malliavin derivative operators, but focusing on the derivatives with respect to $W_t, t \in [t, t+\delta]$ on the one hand, and to replace expectations with conditional expectations with respect to $\mathcal{F}_t$ on the other hand.

Let us introduce some compact notation, denoting $\Theta_k = \{1, \dots, d\}^k$ the set of multi-indexes of length k with components in $\{1, \cdots, d\}$, and $\mathbb{R}^{\Theta_k} := \{(x_\alpha)_{\alpha \in \Theta_k} : x_\alpha \in \mathbb{R}\}$. For $F \in \mathbb{D}^{k,p}$ and $\alpha \in \Theta_k$, we now denote $D^{k,\alpha} F$ the derivative of F of order k and index α, and $D^k_{s_1,\dots,s_k} F := (D^{k,\alpha}_{s_1,\dots,s_k} F)_{\alpha \in \Theta_k}$. The conditional Malliavin calculus is based on the use of the following scalar product and norm: for every fixed $t, \delta > 0$ and $U, V : [t, t+\delta) \to R^{\Theta_k}$ we define

$$\langle U, V \rangle_{t,\delta,k} := \int_{[t,t+\delta)^k} \sum_{\alpha \in \Theta_k} V^\alpha(s_1, \dots, s_k) U^\alpha(s_1, \dots, s_k) ds_1 \cdots ds_k,$$

$$|V|^2_{t,\delta,k} := \langle V, V \rangle^2_{t,\delta,k} = \int_{[t,t+\delta)^k} \sum_{\alpha \in \Theta_k} |V^\alpha(s_1, \dots, s_k)|^2 ds_1 \cdots ds_k.$$

For $F \in \mathbb{D}^{k,p}$, we define the following Sobolev norms:

$$||F||^2_{t,\delta,k} := \sum_{i=0}^{k} |D^i F|^2_{t,\delta,i} = \sum_{i=0}^{k} \sum_{\alpha \in \Theta_i} \int_{[t,t+\delta)^i} |D^{i,\alpha}_{s_1,\cdots,s_i} F|^2 ds_1 \cdots ds_i,$$

$$||F||^p_{t,\delta,k,p} := \mathbb{E}[||F||^p_{t,\delta,k} | \mathcal{F}_t]; \quad |||F|||^p_{t,\delta,k,p} := ||F||^p_{t,\delta,k,p} - \mathbb{E}[|F|^p | \mathcal{F}_t].$$

Notice that $||F||_{t,\delta,k,p}$ is not a constant (as happens in the standard Malliavin calulus) but an $\mathcal{F}_t$-measurable random variable. The standard norm $||F||_{k,p}$ corresponds to $||F||_{0,\infty,k,p}$. Remark that by the definition of $|||F|||_{t,\delta,k,p}$, using the elementary inequality $\frac{b^{p/2}}{2} \leq (a^2 + b)^{p/2} \leq 2^{p/2} b^{p/2}$ for positive a, b and $p \geq 2$, we have

$$\frac{1}{2} f \leq |||F|||^p_{t,\delta,k,p} \leq 2^{p/2} f, \tag{3.6.3}$$

with $f = \mathbb{E}\Big[\Big(\sum_{i=1}^k \sum_{\alpha \in \Theta_i} \int_{[t,t+\delta)^i} |D^{i,\alpha}_{s_1,\dots,s_i} F|^2 ds_1 \cdots ds_i\Big)^{p/2}\Big]$.
We will make use of the two following inequalities: first, let $F, G \in \cap_{p \geq 1} \mathbb{D}^{k,p}$. Then, for every $p \geq 1$,

$$||FG||_{t,\delta,k,p} \leq k! 2^k ||F||_{t,\delta,k,2p} ||G||_{t,\delta,k,2p}. \tag{3.6.4}$$

In addition, for every $k \geq 1$ there exists a constant $\mu(k)$ such that for every $\phi \in C^k_b$, every

$p > 1$ and every $F \in \mathbb{D}^{k,p}$, one has

$$|||\phi(F)|||_{t,\delta,k,p} \leq \mu(k)|\phi|_k |||F|||_{t,\delta,k,2^k p}, \tag{3.6.5}$$

where $|\phi|_k = \sum_{i=0}^k \sup_{x\in\mathbb{R}} |\phi^{(i)}(x)|$. Inequality (3.6.5) is a consequence of the chain rule and of (3.6.4). The proof of (3.6.4) is based on some rather standard (but cumbersome) computations and can be found in the Appendix of [4], Lemma 2.5.

Let us now consider diffusion processes. We consider some $T > 0$ and $0 < \delta < 1 \wedge T$ and $(Y_t; t \in [T-\delta, T])$ the unique strong solution to the equation

$$Y_t = Y_{T-\delta} + \int_{T-\delta}^t B(s, Y_s)ds + \sum_{j=1}^d \int_{T-\delta}^t A_j(s, Y_s)dW_s^j, \quad T-\delta \leq t \leq T,$$

where $Y_{T-\delta} \in L^2(\Omega, \mathcal{F}_{T-\delta}; \mathbb{R}^n)$ and $B, A_j \in Lip([T-\delta, T] \times \mathbb{R}^n; \mathbb{R}^n) \cap \mathcal{C}_b^\infty(\mathbb{R}^n; \mathbb{R}^n)$ for all $j = 1, \ldots, d$. Recall the function $e_p(t) = e^{t^{p/2}(t^{1/2}|B|_1 + |A|_1)^p}$ from (2.2.12) in Chapter (2) and let us define

$$N_k(A, B) := |A|_{k-1}^k \Big(|B|_k + |A|_k\Big)^{(k+1)^2}.$$

The following proposition gives the conditional version of the estimates in Chapter 2, Lemma 2.2.1 and Corollary 1.

Proposition 3.6.1. *For any $k \geq 1$ and any $p > 1$ there exists a positive constant $d_{k,p}$ depending on k, p but not on the bounds on B and A and their derivatives such that, for every $T - \delta \leq t \leq T$ and for every $l = 1, \ldots, m$,*

$$\sup_{\alpha\in\Theta_k} \sup_{s_1,\ldots,s_k \in [T-\delta,T)^k} \mathbb{E}\Big[\Big|D^{k,\alpha}_{s_1,\ldots,s_k} Y_t^l\Big|^p \Big| \mathcal{F}_{T-\delta}\Big]^{1/p} \leq d_{k,p} N_k(A, B) e_p(\delta)^{d_{k,p}} \tag{3.6.6}$$

$$|||Y_t^l|||_{T-\delta,\delta,k,p} \leq 2kd^k d_{k,p} \times \sqrt{\delta} \times N_k(A, B) e_p(\delta)^{d_{k,p}}. \tag{3.6.7}$$

Proof. Inequality (3.6.6) relies on the same proof as Lemma Lemma 2.2.1 in Chapter 2. Estimate (3.6.7) is a consequence of (3.6.6): for any $i = 1, \ldots, k$ and $l = 1, \ldots, n$, we have

$$\Big(\int_{[T-\delta,T)^i} |D^{i,\alpha}_{s_1,\ldots,s_i} Y_t^l|^2 ds_1 \cdots ds_i\Big)^{p/2} \leq \delta^{i(\frac{p}{2}-1)} \int_{[T-\delta,T)^i} |D^{i,\alpha}_{s_1,\ldots,s_i} Y_t^l|^p ds_1 \cdots ds_i,$$

hence, using (3.6.3) and $\sum_{i=1}^k card(\Theta_i) = \sum_{i=1}^k d^i \leq kd^k$, we obtain

$$\begin{aligned} |||Y_t^l|||^p_{T-\delta,\delta,k,p} &\leq 2^{p/2}(kd^k)^p \max_{i=1,\ldots,k} \delta^{i(\frac{p}{2}-1)} \max_{\alpha\in\Theta_i}\Big(\int_{[T-\delta,T)^i} \mathbb{E}[|D^{i,\alpha}_{s_1,\ldots,s_i} Y_t^l|^p | \mathcal{F}_{T-\delta}] ds_1 \cdots ds_i\Big) \\ &\leq 2^{p/2}(kd^k)^p \times \delta^{p/2} \times d_{k,p} N_k(A, B)^p e_p(\delta)^{pd_{k,p}}, \end{aligned}$$

which proves (3.6.7). □

3.6.4 Proof of Proposition 3.4.3

Proof. We assume without further mention that we are on the set $A_{\delta,l}$. Since $T-\delta$ and δ are fixed, we drop them from the notation and write $||F||_{k,p}$ instead of $||F||_{T-\delta,\delta,k,p}$ and so on.

We first show that estimate (3.4.17) holds for $\mathbb{E}[|R_\Delta|^p|\mathcal{F}_{T-\delta}]^{1/p}$. Using $|\eta(t,x)-\eta(s,y)|\sqrt{\psi(v)} \leq K\sqrt{y}(|s-t|+|y-x|)$ and applying Burkholder's inequality we obtain

$$\begin{aligned}\mathbb{E}[|R|^p|\mathcal{F}_{T-\delta}] &\leq c_p\Big(\delta^{p-1}\int_{T-\delta}^T \mathbb{E}[|\psi(\overline{V}_t)|^p|\mathcal{F}_{T-\delta}]dt \\ &\quad + \delta^{p/2-1}\int_{T-\delta}^T \mathbb{E}[(\eta(t,\overline{X}_t)-\eta(T-\delta,X_{T-\delta}))^p\psi(\overline{V}_t)^{p/2}|\mathcal{F}_{T-\delta}]dt\Big) \\ &\leq c_p\Big((\delta y)^p + \delta^{p/2-1}K^p y^{p/2}\int_{T-\delta}^T(\delta^p + ||\overline{X}_t - X_{T-\delta}||_{0,p}^p)dt\Big) \\ &\leq c_p(\delta y)^{p/2}\Big((\delta y)^{p/2} + K^p\delta^p + K^pC_p\delta^{p/2}e^{C_pT^p}\Big) \\ &\leq c_p(\delta y)^p e^{C_pT^p},\end{aligned}$$

where we have used $\delta < \sqrt{\delta} < \sqrt{\delta y}$ and Lemma 3.6.1 to estimate $||\overline{X}_t - X_{T-\delta}||_{0,p}^p = \mathbb{E}[|\overline{X}_t - X_{T-\delta}|^p|\mathcal{F}_{T-\delta}]$. Then we have

$$||R_\Delta||_{0,p} \leq c_p\frac{1}{\rho_\perp\underline{\eta}}\sqrt{\delta y}e^{C_pT^p}. \tag{3.6.8}$$

We now estimate the Sobolev norms of $\overline{X}$ and $\overline{V}$. Notice that by the definition of ψ we have $|\psi|_0 \leq y+\frac{3}{2} \leq 2y$ hence $|\psi(\cdot)|_k \leq c_k^{(1)}y$ for all $k\geq 1$, for some constant $c_k^{(1)}$. Similarly, $|\sqrt{\psi}|_0 \leq \sqrt{y+\frac{3}{2}} \leq 2\sqrt{y}$ and $\frac{d}{dv}\sqrt{\psi(v)} = \frac{\psi'(v)}{2\sqrt{\psi(v)}}$, hence

$$\Big|\frac{d}{dv}\sqrt{\psi(\cdot)}\Big|_0 \leq \frac{|\psi'|_0}{2\sqrt{y-\frac{3}{2}}} \leq \frac{|\psi'|_0}{\sqrt{2y}}$$

and it can be seen easily that $|\sqrt{\psi}(\cdot)|_k \leq c_k^{(1)}\sqrt{y}$ for an eventually different constant $c_k^{(1)}$. With a slight abuse of notation, we write $|\sigma\sqrt{\psi}|_k$ (resp. $|\eta\psi|_k$) for the $|\cdot|_k$-norm of the function $(t,v)\to\sigma(t,v)\sqrt{\psi(v)}$ (resp. $(t,x,v)\to\eta(t,x)\sqrt{\psi(v)}$). Then, using (3.6.7) with $k=2$, for any $t\in[T-\delta,T]$ we have

$$\begin{aligned}|||\overline{V}_t|||_{2,8p} &\leq 4\times d_{2,8p}\times\sqrt{\delta}\times|\sigma\sqrt{\psi}|_1^2(|\beta|_2+|\sigma\sqrt{\psi}|_2)^9 \\ &\quad\times\exp(d_{2,8p}\,\delta^{4p}(|\beta|_1+|\sigma\sqrt{\psi}|_1)^{8p}) \\ &\leq c_p^{(2)}K^9\times\sqrt{\delta}\times|\sqrt{\psi}|_2^9|\sqrt{\psi}|_1^2\times\exp(d_{2,8p}\,\delta^{4p}K^{16p}|\sqrt{\psi}|_1^{8p}) \\ &\leq c_p^{(2)}\times\sqrt{\delta y^{11}}\times\exp(c_p^{(2)}\times(\delta y)^{4p}).\end{aligned}$$

Hence, using $\delta y < 1$ we get $|||\overline{V}_t|||_{2,8p} \leq c_p^{(2)}\times\sqrt{\delta y^{11}}$, for a (eventually different) constant

$c_p^{(2)}$. An analogous estimate holds for $|||\overline{X}_t|||_{2,8p}$: observing that the only difference is in the contribution of the drift term $|\eta\psi|_2 \le Kc_3^{(1)}y$, we have

$$\begin{aligned}|||\overline{X}_t|||_{2,8p} &\le c_p^{(2)}\sqrt{\delta} \times |\psi|_2^9 \times |\sqrt{\psi}|_1^2 \times \exp(c_p^{(2)} \times (\sqrt{\delta}|\psi|_1)^{8p}) \\ &\le c_p^{(2)}\sqrt{\delta y^{20}}\end{aligned}$$

since $\sqrt{\delta}y < 1$, too. Now using (3.6.5) and denoting $\mu = \mu(2)$, we have $||\eta(t,\overline{X}_t)^2||_{2,2p} \le \mu|\eta|_2^2|||\overline{X}_t|||_{2,8p}$ and $||\psi(\overline{V}_t)||_{2,2p} \le \mu|\psi|_2|||\overline{V}_t|||_{2,8p} \le \mu c_3^{(1)}y|||\overline{V}_t|||_{2,8p}$. Hence, using (3.6.4), we get

$$\sup_{T-\delta\le t\le T} |||\eta(t,\overline{X}_t)^2\psi(\overline{V}_t)|||_{2,p} \le 2!2^2\mu^2K^2c_3^{(1)}y \times |||\overline{X}_t|||_{2,8p}|||\overline{V}_t|||_{2,8p} \le c_p\sqrt{\delta^2y^{31}}, \quad (3.6.9)$$

where the constant c_p also depends on K. We denote $R = -\frac{1}{2}I + J$, setting $I = \int_{T-\delta}^T \eta(t,\overline{X}_t)^2\psi(\overline{V}_t)dt$ and $J = \int_{T-\delta}^T(\eta_t - \eta_{T-\delta})\sqrt{\psi(\overline{V}_t)}(\rho dW_t^1 + \rho_\perp dW_t^2)$. Then, using (3.6.3)

$$\begin{aligned}|||I|||_{2,p}^p &\le 2^{p/2}\mathbb{E}\Big[\delta^{p-1}\int_{T-\delta}^T\Big(\sum_{k=1}^2\sum_{\alpha\in\Theta_k}\int_{[T-\delta,T)^k}|D_s^{k,\alpha}\eta(t,\overline{X}_t)^2\psi(\overline{V}_t)|^2\Big)^{p/2}dt\Big|F_{T-\delta}\Big] \\ &\le 2^{p/2+1}\delta^{p-1}\int_{T-\delta}^T |||\eta(t,\overline{X}_t)^2\psi(\overline{V}_t)|||_{2,p}^p dt \\ &\le c_p(\delta^4y^{31})^{p/2}.\end{aligned}$$

We now estimate the first Sobolev norm of J. For the ease of notation, we write η_t for $\eta(t,\overline{X}_t)$. For $T-\delta \le s \le t \le T$, we have $D_s^{1,j}((\eta_t - \eta_{T-\delta})\sqrt{\psi(\overline{V}_t)}) = D_s^{1,j}(\eta_t\sqrt{\psi(\overline{V}_t)})$, hence

$$\begin{aligned}D_s^{1,j}\int_{T-\delta}^T(\eta_t - \eta_{T-\delta})\sqrt{\psi(\overline{V}_t)}(\rho dW_t^1 + \rho_\perp dW_t^2) &= (\eta_s - \eta_{T-\delta})\sqrt{\psi(\overline{V}_s)})\rho_j \\ &\quad + \int_{T-\delta}^T D_s^{1,j}(\eta_t\sqrt{\psi(\overline{V}_t)})(\rho dW_t^1 + \rho_\perp dW_t^2) \quad (3.6.10)\end{aligned}$$

with $\rho_j = \rho 1_{j=1} + \rho_\perp 1_{j=2}$. Using bound (3.6.6) and proceeding as for (3.6.9) we obtain $\sup_{t\in[T-\delta,T]}\mathbb{E}\Big[\Big|D_s^{1,j}(\eta_t\sqrt{\psi(\overline{V}_t)})\Big|^p\Big|F_{T-\delta}\Big]^{1/p} \le c_p|\eta|_1|\sqrt{\psi}|_1\sqrt{y^{11+20}} \le c_p\sqrt{y^{32}}$, where c_p depends also on K. Then, starting from (3.6.3) and using Burkholder's inequalities

$$\begin{aligned}|||J|||_{1,p}^p &\le 2^p\delta^{p/2-1}\int_{T-\delta}^{T}||(\eta_s-\eta_{T-\delta})\sqrt{\psi(\overline{V}_s)}||_{0,p}^p ds\\ &\quad + \delta^{p-1}\sum_{j=1,2}\int_{T-\delta}^{T}\int_{T-\delta}^{T}\mathbb{E}\Big[|D_s^{1,j}(\eta_t\sqrt{\psi(\overline{V}_t)})|^p\Big|F_{T-\delta}\Big]dtds\\ &\le 2^pK^p\delta^{p/2-1}y^{p/2}\int_{T-\delta}^{T}(\delta+||\overline{X}_s-X_{T-\delta}||_{0,p})^p ds + c_p\delta^{p+1}y^{32p/2}\\ &\le c_p(\delta^2y^{32})^{p/2}e^{C_pT^p}.\end{aligned}$$

The Sobolev norms of higher order are estimated in a similar way, giving the bound $|||J|||_{2,p}\le c_p\sqrt{\delta^2y^{32}}e^{C_pT^p}$. Finally, $|||R|||_{2,p}\le|||I|||_{2,p}+|||J|||_{2,p}\le c_p\sqrt{\delta^2y^{32}}e^{C_pT^p}$ and this last estimate together with (3.6.8) yields $|||R_\Delta|||_{2,p}\le c_p\frac{1}{\rho_\perp\underline{\eta}}\sqrt{\delta y^{31}}e^{C_pT^p}$, which is (3.4.17). $\square$

Chapter 4

Tube estimates and general lower and upper bounds via time-change techniques

4.1 Introduction

The starting point for the study we develop in this chapter dates back to a discussion that Vlad Bally had with Emmanuel Gobet during a conference in Helsinky in 2008[1]. The tube estimate in [8], which is one of the core tools for the analysis carried in the previous chapter, is obtained by construction of an appropriate time-grid on the interval $[0,T]$, then exploiting the short time behaviour of Itô processes. E. Gobet suggested that martingale time-change techniques could be a powerful (yet simple) instrument to obtain similar results in this area. Here we take up this problem, and the results show that Gobet's intuition was correct: coupling elementary time-change techniques for martingales with the appropriate localization arguments, we manage to obtain similar results to [8], but the current estimates are sharper and the hypothesis are weaker. Last but not least, the machinery we have to settle here is considerably lighter, as a consequence the proofs of the main results are easier. Besides tube estimates, exploiting the same tools we derive upper bounds for suprema of Itô processes, namely we find functions $C(y)$ such that the estimate $\mathbb{P}\big(\sup_{t\leq T}|X_t|>y\big)\leq C(y)$ is significant in the asymptotic range $y\to\infty$. In Theorem 4.2.1 we formulate our basic estimate involving the time-change argument considering general Itô processes and tubes of simple geometry, which basically reduces to cylindric crowns of the form $\{(t,y)\in[0,T]\times\mathbb{R}^n : R_1\leq|y|\leq R_2\}$ (cf. hypothesis (H) and Remark 4.2.1 in section 4.2). The subsequent Propositions 4.3.2 and 4.3.3 specialise the result to diffusions and arbitrary deterministic tubes.

The first application of these results, in the spirit of the work of the previous chapters, is to work out tail estimates for fixed-time distribution functions and - when existing - densities. The lower bounds for the probability to stay inside a tube up to time T translate into lower bounds for the terminal distribution, using for example the techniques already settled in Chapter 3. In its turn, the upper bound on the supremum of the process over $[0,T]$ provides

[1]I am grateful to Prof. Vlad Bally for sharing with me these useful insights.

the corresponding upper bound on the law at time T: while the tails of the cdf can be trivially estimated by $\mathbb{P}(|X_T| > y) \le \mathbb{P}(\sup_{t\le T}|X_t| > y)$, to estimate the density we can apply Theorem 1.3.1 in Chapter 2. In particular, the current upper bound on $\sup_{t\le T}|X_t|$ provides a way to estimate, in a general framework, the factor P_t appearing in the upper bounds for the density and its derivatives in Theorem 1.3.1 (cf. Remark 4.3.1). We give applications of these results to the class of one-dimensional SDEs considered in Chapter 2, namely CIR/CEV-type processes with local coefficients as (1.3.1), showing that the tube estimates provided here are actually effective and allow to obtain significant tail estimates from above and below (cf. Propositions 4.3.1 and 4.3.4). The application to general diffusions is currently under study, and we pledge to fill this gap very quickly.

4.2 The basic estimate

Notation. *Matrices.* We denote $\mathcal{M}_n(\mathbb{R})$ the space of real $n\times n$ matrices and $\mathcal{S}_n^+ \subset \mathcal{M}_n(\mathbb{R})$ (resp. $\mathcal{S}_n^> \subset \mathcal{M}_n(\mathbb{R})$) the family of symmetric positive semidefinite (resp. positive definite) matrices. For $a \in \mathcal{M}_n(\mathbb{R})$, we denote $||a|| := \sup_{|y|=1}|ay|$ the operator norm and $\lambda(a) := \inf_{|y|=1}|ay|$. If $a\in\mathcal{S}_n^+$, then $||a||$, respectively $\lambda(a)$, is the larger, respectively the smaller, eigenvalue of a. $|\cdot|$ will still denote the absolute value for real numbers as well as the Euclidean norm for vectors, i.e. $|x| = \sqrt{\sum_i^n x_i^2}$ if $x\in\mathbb{R}^n$. We recall that for every $a\in\mathcal{S}_n^+$ it is uniquely defined in $\mathcal{S}_n^+$ a matrix $a^{1/2}$ (the square root of a) such that $a^{1/2}a^{1/2} = a$. Let us remark that $||a^{1/2}||^2 = ||a||$ and $\lambda(a^{1/2})^2 = \lambda(a)$. Moreover, if $a\in\mathcal{S}_n^>$, a is invertible, the inverse a^{-1} belongs to $\mathcal{S}_n^>$ and $(a^{-1})^{1/2} = (a^{1/2})^{-1}$.

In this section we give our main tube estimate. Let $(\Omega, \mathcal{F} := (\mathcal{F}_t)_{t\le T}, \mathcal{F}_T)$, $T>0$, be a filtered probability space and $W = (W_t^1,\dots,W_t^d; t\le T)$ a d-dimensional $\mathcal{F}$-Brownian motion. Let us consider an adapted n-dimensional Itô process, starting from zero:

$$Y_t = \int_0^t b(s,\omega,Y_s)ds + \sum_{j=1}^d \int_0^t \sigma_j(s,\omega,Y_s)dW_s^j, \quad t\le T, \tag{4.2.1}$$

where $b, \sigma_j : [0,T]\times\Omega\times\mathbb{R}^n \to \mathbb{R}^n$ are measurable functions such that the above integrals make sense. We assume that b and σ are adapted, in the sense that: for every $t\in[0,T]$ and $y\in\mathbb{R}^n$, $\omega\mapsto b(s,\omega,y)$ and $\omega\mapsto\sigma(t,\omega,y)$ are $\mathcal{F}_t$-measurable functions. Let moreover $R>0$ and $\tau_R = \inf\{t\ge 0 : |Y_t| > R\}$.

Remark 4.2.1. We may of course set $\bar{b}(t,\omega) = b(t,\omega,Y_t(\omega))$ and $\bar{\sigma}(t,\omega) = \sigma(t,\omega,Y_t(\omega))$ and write Y as a general Itô process. The parameterization of the coefficients in (4.2.1) allows us to "localize" our assumptions, in the spirit of the previous chapter. We will indeed ask the functions $b(t,\omega,y)$ and $\sigma(t,\omega,y)$ to satisfy some conditions only for y such that $\frac14 R \le |y| \le \frac34 R$ (cf. hypothesis (H) below).

We denote $a = \sigma\sigma^*$, that is $a^{ij} = \sum_{k=1}^d \sigma_k^i \sigma_k^j$, and we have $a \in \mathcal{S}_n^+$, $Tr(a) = \sum_{j=1}^d \sum_{i=1}^n |\sigma_j^i|^2 = \sum_{j=1}^d |\sigma_j|^2$. Let I be the set where a is invertible, $I = \{(t,\omega,y)\mathbb{R}_+ \times \Omega \times \mathbb{R}^n \setminus \{0\} : a(t,\omega,y) \in \mathcal{S}_n^+\}$. We define the auxiliary functions $q, h : I \to \mathbb{R}$ and $v : I \to \mathbb{R}^n$ by

$$q(t,\omega,y) = Tr(a(t,\omega,y)) - \frac{\langle a(t,\omega,y)y, y\rangle}{|y|^2},$$

$$h(t,\omega,y) = \frac{q^2(t,\omega,y)}{\langle a(t,\omega,y)y, y\rangle}$$

and

$$v(t,\omega,y) = q(t,\omega,y)\frac{a(t,\omega,y)y}{\langle a(t,\omega,y)y, y\rangle}.$$

These functions will come into play in the proof of Theorem 4.2.1 below. Notice that v is defined so that for every $y \in \mathbb{R}^n$ we have

$$\langle v(t,\omega,y), y\rangle = q(t,\omega,y)$$

and

$$\langle a^{-1}(t,\omega,y)v(t,\omega,y), v(t,\omega,y)\rangle = h(t,\omega,y).$$

Moreover, remark that when $n = 1$, q, h and v are identically zero.

We now state our assumptions. For the ease of notation, let us denote $a_{t,y} := a(t,\omega,y)$ and $b_{t,y} := b(t,\omega,y)$. In general, we drop the dependence with respect to $\omega \in \Omega$ when this can be done without generating confusion.

(H) There exist some functions $c_b, c_\sigma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and $\overline{a} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathcal{S}_n^+$ such that for every $t \in [0,T]$ and every $y \in \mathbb{R}^n$ such that $\frac{1}{4}R \leq |y| \leq \frac{3}{4}R$,

$$a_{t,y} > 0 \quad \text{and} \quad \langle a_{t,y}\xi, \xi\rangle \leq \langle \overline{a}(R,t)\xi, \xi\rangle, \qquad \forall \xi \in \mathbb{R}^n, \ \mathbb{P}\text{-a.s.}$$

$$|a_{t,y}^{-1/2} b_{t,y}| \leq c_b(R,t), \qquad \mathbb{P}\text{-a.s.}$$

$$h(t,y) \leq c_\sigma(R,t), \qquad \mathbb{P}\text{-a.s.}$$

$$\int_0^T \big(||\overline{a}(R,s)|| + c_b^2(R,s) + c_\sigma^2(R,s)\big)ds < \infty.$$

Theorem 4.2.1. *Under (H),*

$$\mathbb{P}\Big(\sup_{t\leq T} |Y_t| \leq R\Big) \geq \exp\big(-(\sqrt{\overline{\alpha}_T(R)} + \sqrt{\beta_T(R)})^2\big) \tag{4.2.2}$$

with

$$\alpha_T(R) = \int_0^T ||\overline{a}(R,s)||ds, \quad \overline{\alpha}_T(R) = 2\log\Big(\frac{\pi}{2}\Big) + \frac{2\pi^2}{R^2}\alpha_T(R)$$

$$\beta_T(R) = \int_0^T \big(c_b^2(R,s) + 1_{\{n>1\}}c_\sigma(R,s)\big)ds.$$

Moreover, if $16\alpha_T(R)(\beta_R(T)+\ln(2)) < R^2$*,*

$$\mathbb{P}\Big(\sup_{t\le T}|Y_t| \ge R\Big) \le \exp\big(-\big(\sqrt{\underline{\alpha}_T(R)} - \sqrt{\beta_T(R)}\big)^2\big) \tag{4.2.3}$$

with

$$\underline{\alpha}_T(R) = \frac{R^2}{16\alpha_T(R)} - \ln(2).$$

Comment 4.2.1. Estimate (4.2.2) will be used for small values of R, and estimate (4.2.3) for large values of R.

Remark 4.2.2. The hypotheses under which the tube estimate 4.2.2 holds are actually weaker with respect to the ones in [8]. In particular, we demand no Lipschitz assumptions on σ.

Proof. Estimate (4.2.2). We consider the stopping times $\tau_R' := \inf\{t : |Y_t| \ge \frac{1}{2}R\}$ and $\tau_R'' := \inf\{t \ge \tau_R' : \big||Y_t| - \frac{R}{2}\big| \ge \frac{1}{4}R\}$, with $\inf\emptyset = T$. It is clear that $\tau_R'' \le \tau_R$ so that $P(\tau_R \ge T) \ge P(\tau_R'' \ge T)$. We will look for a lower bound on this last probability.

Step 1. Let $A_R(t) := \{\tau_R' \le t \le \tau_R''\}$. Notice that on the set $A_R(t)$ we have $0 < \frac{R}{4} \le |Y_t| \le \frac{3R}{4}$, hence $v(t,Y_t)$ is well defined and $a(t,Y_t)$ is invertible. Let us define the d-dimensional process $\theta = (\theta(t), t\in[0,T])$ by

$$\theta(t) := 1_{A_R}(t)\sigma^*(t,Y_t)a^{-1}(t,Y_t)\big(b(t,Y_t) + v(t,Y_t)\big).$$

We remark that θ_t is zero outside $A_R(t)$ and that we have, for every $t\in[0,T]$:

$$\begin{aligned}
|\theta(t)|^2 &= 1_{A_R(t)}\langle a^{-1}(t,Y_t)(b(t,Y_t)+v(t,Y_t)), b(t,Y_t)+v(t,Y_t)\rangle \\
&\le 1_{A_R(t)} \times 2\Big(\langle a^{-1}(t,Y_t)b(t,Y_t), b(t,Y_t)\rangle + \langle a^{-1}(t,Y_t)v(t,Y_t), v(t,Y_t)\rangle\Big) \\
&\le 1_{A_R(t)} \times 2\Big(|a^{-1/2}(t,Y_t)b(t,Y_t)|^2 + h(t,Y_t)\Big) \\
&\le 1_{A_R(t)} \times 2\big(c_b^2(R,t) + 1_{\{n>1\}}c_\sigma(R,t)\big),
\end{aligned} \tag{4.2.4}$$

where the indicator function comes from the fact that h is identically zero when $n=1$.

Step 2 (Girsanov drift-change). Let us define

$$\overline{W}_t = W_t + \int_0^t \theta(s)ds, \quad e_t = \exp\Big(-\int_0^t \langle\theta(s), dW_s\rangle - \frac{1}{2}\int_0^t |\theta(s)|^2ds\Big), \quad t\le T.$$

By (4.2.4) and assumption (H), the process $(e_t; t\le T)$ is a martingale under $\mathbb{P}$. By Girsanov's theorem, the process $(\overline{W}_t; t \leq T)$ is a d-dimensional Brownian motion under the probability measure $\widetilde{\mathbb{P}}$ defined by $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = e_T$. Now, since $\sigma(t, Y_t)\theta(t) = b(t, Y_t) + v(t, Y_t)$ on $A_R(t)$, for $\tau'_R \leq t \leq \tau''_R$ we have

$$\begin{aligned} Y_t &= Y_{\tau'_R} + \int_{\tau'_R}^t \Big(b(t, Y_t) - \sum_{j=1}^d \sigma_j(s, Y_s)\theta^j(s)\Big)ds + \sum_{j=1}^d \int_{\tau'_R}^t \sigma_j(s, Y_s)(dW_s^j + \theta^j(s)ds) \\ &= Y_{\tau'_R} - \int_{\tau'_R}^t v(s, Y_s)ds + \sum_{j=1}^d \int_{\tau'_R}^t \sigma_j(s, Y_s)d\overline{W}_s^j. \end{aligned}$$

We now apply Itô's formula to $|Y_t|$. Using the relations $\partial_{x_i}|x| = \frac{x_i}{|x|}$ and $\partial^2_{x_i x_j}|x| = \frac{1}{|x|}(\delta_{ij} - \frac{x_i x_j}{|x|^2})$, we have

$$\begin{aligned} |Y_t| = |Y_{\tau'_R}| &+ \int_{\tau'_R}^t \frac{1}{|Y_s|}\Big(-\langle v(s, Y_s), Y_s\rangle + Tr(a(s, Y_s)) - \frac{\langle a(s, Y_s)Y_s, Y_s\rangle}{|Y_s|^2}\Big)ds \\ &+ \sum_{j=1}^d \int_{\tau'_R}^t \frac{\langle \sigma_j(s, Y_s), Y_s\rangle}{|Y_s|} d\overline{W}_s^j. \end{aligned}$$

By the definition of v, the drift term in the previous equation cancels and, since $|Y_{\tau'_R}| = \frac{R}{2}$, we have

$$|Y_t| = \frac{R}{2} + \sum_{j=1}^d \int_0^t 1_{\{\tau'_R \leq s \leq \tau''_R\}} \frac{\langle \sigma_j(s, Y_s), Y_s\rangle}{|Y_s|} d\overline{W}_s^j =: \frac{R}{2} + m_t,$$

for $\tau'_R \leq t \leq \tau''_R$, where $m_t := \sum_{j=1}^d \int_0^t 1_{A_R(s)} \frac{\langle \sigma_j(s, Y_s), Y_s\rangle}{|Y_s|} d\overline{W}_s^j$.

Step 3 (Time-change). The process m_t is a martingale with quadratic variation

$$\langle m\rangle_t = \int_0^t 1_{A_R(s)} \frac{\langle a(s, Y_s)Y_s, Y_s\rangle}{|Y_s|^2} ds \leq \int_0^t \frac{\langle \overline{a}(R, s)Y_s, Y_s\rangle}{|Y_s|^2} ds \leq \alpha_T(R), \quad \mathbb{P}\text{-a.s.}$$

In particular, by Dubins & Schwarz Theorem there exists a one-dimensional $\widetilde{\mathbb{P}}$-Brownian motion $b = (b(t); t \leq T)$ such that $b(\langle m\rangle_t) = m_t$.

Step 4. Conclusion. We have

$$\begin{aligned}
\widetilde{\mathbb{P}}(\tau_R' \le T \le \tau_R'') &= \widetilde{\mathbb{P}}\Big(\tau_R' \le T, \sup_{\tau_R' \le t \le T}\Big||Y_t| - \frac{R}{2}\Big| \le \frac{R}{4}\Big) \\
&= \widetilde{\mathbb{P}}\Big(\tau_R' \le T, \sup_{\tau_R' \le t \le T}|b(\langle m\rangle_t)| \le \frac{R}{4}\Big) \\
&\ge \widetilde{\mathbb{E}}\Big[1_{\{\tau_R' \le T\}}\widetilde{\mathbb{E}}\big[1_{\{\sup_{\tau_R' \le t \le T}|b(\langle m\rangle_t)| \le \frac{R}{4}\}}|\mathcal{F}_{\tau_R'}\big]\Big] \\
&\ge \widetilde{\mathbb{P}}(\tau_R' \le T) \times \sup_{\tau \le T}\widetilde{\mathbb{P}}\Big(\sup_{\tau \le t \le T}|b(\langle m\rangle_t)| \le \frac{R}{4}\Big) \\
&\ge \widetilde{\mathbb{P}}(\tau_R' \le T) \times \widetilde{\mathbb{P}}\Big(\sup_{0 \le t \le \alpha_T(R)}|b(t)| \le \frac{R}{4}\Big) \\
&\ge \frac{2}{\pi}\,\widetilde{\mathbb{P}}(\tau_R' \le T)\exp\Big(-\frac{2\pi^2}{R^2}\alpha_T(R)\Big)
\end{aligned}$$

where we have applied estimate (4.4.2) of Lemma 4.4.1, Section 4.4, in the last step. It follows that

$$\widetilde{\mathbb{P}}(\tau_R'' \ge T) = \widetilde{\mathbb{P}}(\tau_R' > T) + \widetilde{\mathbb{P}}(\tau_R' \le T \le \tau_R'') \ge \frac{2}{\pi}\exp\Big(-\frac{2\pi^2}{R^2}\alpha_T(R)\Big) = \exp(-\overline{\alpha}_T(R)).$$

Now, by Holder's inequality, for every $p > 1$ we have

$$\widetilde{\mathbb{P}}(\tau_R'' \ge T) = \mathbb{E}[1_{\{\tau_R'' \ge T\}}e_T] \le \mathbb{P}(\tau_R'' \ge T)^{(p-1)/p}\mathbb{E}[e_T^p]^{1/p}. \tag{4.2.5}$$

Since, by the martingale property,

$$\begin{aligned}
\mathbb{E}[e_T^p] &= \mathbb{E}\Big[\exp\Big(-\int_0^T\langle p\theta(s), dW_s\rangle - \frac{1}{2}\int_0^T|p\theta(s)|^2ds + \frac{p(p-1)}{2}\int_0^T|\theta(s)|^2ds\Big)\Big] \\
&\le \exp\Big(p(p-1)\int_0^T(c_b^2(R,s) + c_\sigma(R,s))ds\Big) = \exp\Big(p(p-1)\beta_T(R)\Big),
\end{aligned}$$

from this last inequality and (4.2.5) we obtain, for all $p > 1$,

$$\begin{aligned}
\mathbb{P}\Big(\sup_{t \le T}|Y_t| \le R\Big) \ge \mathbb{P}(\tau_R'' \ge T) &\ge \frac{\widetilde{\mathbb{P}}(\tau_R'' \ge T)^{p/(p-1)}}{\mathbb{E}[e_T^p]^{1/(p-1)}} \\
&\ge \exp\Big(-\frac{p}{p-1}\overline{\alpha}_T(R) - p\beta_T(R)\Big). \qquad (4.2.6)\\
&= \exp\Big(-\frac{p}{p-1}\overline{\alpha}_T(R) - p\beta_T(R)\Big).
\end{aligned}$$

Finally, we optimize the right hand side expression over p : it is easy to see (and will be proven later on) that

$$\max_{p>1}\Big\{\exp\Big(-\frac{p}{p-1}\overline{\alpha}_T(R) - p\beta_T(R)\Big)\Big\} = \exp\big(-\big(\sqrt{\overline{\alpha}_T(R)} + \sqrt{\beta_T(R)}\big)^2\big) \tag{4.2.7}$$

Estimates (4.2.6) and (4.2.7) yield (4.2.2).

Estimate 4.2.3. We have $\mathbb{P}(\tau_R \le T) \le \mathbb{P}(\tau''_R \le T)$, and we look for an upper bound on this last probability. We proceed in the analogous way: we remark that we have

$$\begin{aligned}
\widetilde{\mathbb{P}}(\tau''_R \le T) &= \widetilde{\mathbb{P}}\Big(\tau'_R \le \tau''_R \le T, \sup_{\tau'_R \le t \le \tau''_R} \Big||Y_t| - \frac{R}{2}\Big| = \frac{R}{4}\Big) \\
&= \widetilde{\mathbb{P}}\Big(\tau'_R \le \tau''_R \le T, \sup_{\tau'_R \le t \le \tau''_R} |b(\langle m \rangle_t)| = \frac{R}{4}\Big) \\
&\le \widetilde{\mathbb{P}}\Big(\sup_{0 \le t \le \alpha_T(R)} |b(t)| \ge \frac{R}{4}\Big),
\end{aligned}$$

and hence, applying estimate (4.4.3) of Lemma 4.4.1,

$$\widetilde{\mathbb{P}}(\tau''_R \le T) \le 2\exp\Big(-\frac{R^2}{16\alpha_T(R)}\Big) = \exp\Big(-\Big(\frac{R^2}{16\alpha_T(R)} - \ln(2)\Big)\Big) = \exp(\underline{\alpha}_T(R)).$$

On the other hand, for every $p > 1$ we have

$$\begin{aligned}
\tilde{\mathbb{E}}[e_T^{-p}] &= \mathbb{E}[e_T^{-(p-1)}] \\
&= \mathbb{E}\Big[\exp\Big(\int_0^T \langle (p-1)\theta(s), dW_s\rangle - \frac{1}{2}\int_0^T |(p-1)\theta(s)|^2 ds + \frac{p(p-1)}{2}\int_0^T |\theta(s)|^2 ds\Big)\Big] \\
&\le \exp\Big(p(p-1)\int_0^T (c_b^2(R,s) + c_\sigma(R,s))ds\Big) = \exp\big(p(p-1)\beta_T(R)\big).
\end{aligned}$$

Once again by Holder's inequality, we have

$$\begin{aligned}
\mathbb{P}\Big(\sup_{t \le T} |Y_t| \ge R\Big) = \mathbb{P}(\tau_R \le T) \le \mathbb{P}(\tau''_R \le T) &= \tilde{\mathbb{E}}\big[1_{\{\tau''_R \le T\}} e_T^{-1}\big] \\
&\le \widetilde{\mathbb{P}}(\tau''_R \le T)^{(p-1)/p} \times \tilde{\mathbb{E}}[e_T^{-p}]^{1/p} \\
&\le \exp\Big(-\frac{p-1}{p}\underline{\alpha}_T(R) + (p-1)\beta_T(R)\Big)
\end{aligned}$$

Now, if $\underline{\alpha}_T(R) > \beta_T(R)$, that is if

$$16\alpha_R(T)(\beta_R(T) + \ln(2)) < R^2, \tag{4.2.8}$$

the right hand expression has a minimum over $p > 1$, and it is easy to see that

$$\min_{p>1}\Big\{\exp\Big(-\frac{p-1}{p}\underline{\alpha}_T(R) + (p-1)\beta_T(R)\Big)\Big\} = \exp\Big(-\Big(\sqrt{\underline{\alpha}_T(R)} - \sqrt{\beta_T(R)}\Big)^2\Big). \tag{4.2.9}$$

Proof of (4.2.7) and (4.2.9): Maximizing $\exp(-f(p))$ amounts to minimize $f(p)$. In

(4.2.7), we have $f(p) = \frac{p}{p-1}\overline{\alpha}_T(R) + p\beta_T(R)$. Derivation yields

$$f'(p) = \frac{(p-1)^2\beta_T(R) - \alpha_T(R)}{(p-1)^2}$$

hence, noticing that $\beta_T(R) > 0$, $\alpha_T(R) > 0$, the minimum of f on $(1,\infty)$ is attained at $p^* = 1 + \sqrt{\frac{\alpha_T(R)}{\beta_T(R)}}$. Direct computation yields $f(p^*) = (\sqrt{\alpha_T(R)} + \sqrt{\beta_T(R)})^2$.

Similar computations show that $p \to g(p) := \frac{p-1}{p}\underline{\alpha}_T(R) + (1-p)\beta_T(R)$ has a maximum on $(1,\infty)$ if and only if $\underline{\alpha}_T(R) > \beta_T(R)$, that is if (4.2.8) holds. In this case we have $\max_{p>1} g(p) = (\sqrt{\underline{\alpha}_T(R)} - \sqrt{\beta_T(R)})^2 > 0$ and we obtain (4.2.9). We notice that if (4.2.8) does not hold, g is decreasing over $(1,\infty)$ and $\exp(-g(p)) \geq \exp(-g(1)) = 1$ and $\exp(-g(p))$is not significant as an upper bound on $\mathbb{P}(\tau''_R \leq T)$. □

4.3 Lower and upper estimates of distribution functions: application to diffusions

4.3.1 Upper Bounds

In this section we show how the "upper-bound" part of Theorem 4.2.1 (that is, estimate (4.2.3)) can be directly applied in order to obtain tail upper bounds for the laws of diffusions. We will work on an example, focusing on the class of one-dimensional diffusions considered in Chapter 2, section 3.

Exemple 1. Let

$$X_t = x + \int_0^t \tilde{b}(X_s)ds + \sigma \int_0^t X_s^\alpha dW_s, \quad t \leq T \tag{4.3.1}$$

where

- $0 < \alpha < 1$;
- There exists $\kappa > 0$ and $\beta \in [0,1]$ such that

$$\tilde{b}(x) \leq \kappa x^\beta \qquad \text{for every } x > 1. \tag{4.3.2}$$

In the spirit of the work of Chapter 3, we do not deal here with the issue of existence and uniqueness of solutions for (4.3.1). The result we give hereafter holds indeed for any continuous adapted process $X = (X_t, t \leq T)$ such that $\mathbb{P}(X_t \geq 0, t \leq T) = 1$ and X_t satisfies (4.3.1) for all t.

Proposition 4.3.1. If (4.3.2) holds with $\beta < 1$ and R is large enough, precisely

$$R > 4x \vee \left(2^5 \ln(2)\sigma^2 \times T \vee \frac{1}{\kappa^2 T}\right)^{1/(2(1-\alpha))} \vee \left(2^{8+4\alpha}\kappa^2 T^2\right)^{1/(2(1-\beta))} \tag{4.3.3}$$

then

$$\mathbb{P}\Big(\sup_{t\leq T}|X_t-x|>R\Big)\leq \exp\Big(-\frac{R^{2(1-\alpha)}}{2^7\sigma^2T}\Big). \tag{4.3.4}$$

In particular, for every $z>x+R$,

$$\mathbb{P}(X_T>z)\leq \exp\Big(-\frac{(z-x)^{2(1-\alpha)}}{2^7\sigma^2T}\Big). \tag{4.3.5}$$

If (4.3.2) holds with $\beta=1$ and (4.3.3) holds, then (4.3.4) and (4.3.5) hold for T small enough, precisely

$$T<2^{-(4+2\alpha)}\kappa^{-1}. \tag{4.3.6}$$

Remark 4.3.1. Let us come back to the estimate for the density of the law provided in Chapter 2. When $\alpha\in[1/2,1)$, under the appropriate regularity assumptions on the drift coefficient $\tilde b$, X_T admits a smooth density p_T on $(0,\infty)$, for every $T>0$. More precisely, if $\tilde b\in C^\infty(\mathbb{R})$ and $\tilde b(0)\geq 0$, then by the same arguments as in Lemma 2.4.1, eq. (4.3.1) admits a unique strong solution X and by Theorem 2.2.4, the law of X_T restricted to $(0,\infty)$ has a smooth density p_T. An upper bound on this density is given by

$$p_T(z)\leq \Big(1+\frac{1}{T^{3/2}}\Big)P_T(z)\Lambda(z),$$

where Λ is an explicitly known function with polynomial growth and $P_T(z)=\mathbb{P}\big(\inf_{s\in[T/2,T]}|X_s-z|\leq 3\big)$ (to be precise, we have set $R=1$ in Theorem 2.2.4 and we denote $\Lambda(z):=\Lambda_3(T\wedge 1,z)$. The polynomial growth of the function Λ follows from the same arguments as in Proposition 2.3.2). An issue is how to estimate the term $P_T(z)$ (a problem which is already tackled in Chapter 2, relying on some arguments specifically designed for equations such as (4.3.1)). We can now upper bound $P_T(z)$ by means of the previous proposition: we simply write

$$P_T(z)\leq \mathbb{P}\Big(\sup_{s\in[T/2,T]}X_s\geq z-3\Big)\leq \mathbb{P}\Big(\sup_{s\leq T}|X_s-x|\geq z-x-3\Big),$$

then use (4.3.4). Doing this we eventually obtain the same asymptotic estimate (up to multiplicative constants) as the one derived in Chapter 2, Proposition 2.3.3.

Proof. We consider the process

$$Y_t=X_t-x=\int_0^t\tilde b(x+Y_s)ds+\sigma\int_0^t(x+Y_s)^\alpha dW_s.$$

Let us define $a(y)=\sigma^2(x+y)^{2\alpha}$ and $b(y)=\tilde{b}(x+y)$. It is easy to see that, for $\frac{1}{4}R\le y\le\frac{3}{4}R$,

$$\begin{aligned}a(y)&\le\sigma^2(x+\frac{3}{4}R)^{2\alpha}\le\sigma^2R^{2\alpha}\\ \frac{|b(y)|}{\sqrt{a(y)}}&\le\frac{\kappa(x+\frac{3}{4}R)^\beta}{\sigma(x+\frac{1}{4}R)^\alpha}\le\frac{4^\alpha\kappa}{\sigma}\times R^{\beta-\alpha}\end{aligned}$$

where we have used $x<\frac{1}{4}R$ because of (4.3.3). Then the hypothesis (H) holds for the process Y and the constant functions $\overline{a}(R,t)=\sigma^2R^{2\alpha}$ and $c_b(R,t)=\frac{4^\alpha\kappa}{\sigma}R^{\beta-\alpha}$. We have $\alpha_T(R)=\int_0^T\overline{a}(R,t)dt=\sigma^2R^{2\alpha}T$ and $\beta_T(R)=\int_0^Tc_b(R,t)^2dt=\big(\frac{4^\alpha\kappa}{\sigma}\big)^2R^{2(\beta-\alpha)}T$. Moreover, $\beta_T(R)<\ln(2)$ by (4.3.3) hence

$$\begin{aligned}16\alpha_T(R)(\beta_T(R)+\ln(2))&\le16\sigma^2R^{2\alpha}T\times2\big(\frac{4^\alpha\kappa}{\sigma}\big)^2R^{2(\beta-\alpha)}T\\&=2^{5+4\alpha}\kappa^2T^2R^{2\beta}.\end{aligned}$$

Case $\beta<1$. When $\beta<1$, by (4.3.3) the previous inequality reads

$$16\alpha_T(R)(\beta_T(R)+\ln(2))\le R^2 \tag{4.3.7}$$

therefore all the conditions in order to apply estimate (4.2.3) are verified. Remark that by (4.3.3) we also have$\underline{\alpha}_T(R)=\frac{R^2}{16\sigma^2R^{2\alpha}T}-\ln(2)\ge R^{2(1-\alpha)}\frac{1}{32\sigma^2T}$, hence

$$\big(\sqrt{\underline{\alpha}_T(R)}-\sqrt{\beta_T(R)}\big)^2\ge\frac{R^{2(1-\alpha)}}{32\sigma^2T}\Big(1-4^{\alpha+3/2}\kappa\frac{T}{R^{1-\beta}}\Big)^2\ge\frac{R^{2(1-\alpha)}}{32\sigma^2T\times4} \tag{4.3.8}$$

since $4^{\alpha+3/2}\kappa\frac{T}{R^{1-\beta}}<\frac{1}{2}$ by (4.3.3). Estimate (4.2.3) in Theorem 4.2.1 yields (4.3.4). The inequality (4.3.5) follows obviously using (4.3.4).

Case $\beta=1$. Condition (4.3.7) holds if $2^{5+4\alpha}\kappa^2T^2$, which is in its turn assured by (4.3.6). Analogously, (4.3.8) holds if $4^{\alpha+3/2}\kappa T<\frac{1}{2}$, that is if condition (4.3.6) holds. □

4.3.2 Lower Bounds

We start by developing an "interface" allowing to apply estimate (4.2.2) in Theorem 4.2.1 to obtain tube estimates for diffusion processes. In particular, since Theorem 4.2.1 only considers constant radius R and no deterministic curve, we want to introduce a curve on which the tube is centered. This can be achieved simply by working on the shifted process X_t-x_t. On the other hand, instead of introducing a time-dependent radius R_t, we will rescale the process by multiplying it by an invertible matrix. This reasonings are implemented in the following two propositions.

Let: $R > 0$, $x \in \mathbb{R}^n$, $\phi \in C([0,T];\mathbb{R}^n)$,

$$X_t = x + \int_0^t \tilde{b}(s, X_s)ds + \sum_{j=1}^d \int_0^t \tilde{\sigma}_j(s, X_s)dW_s^j, \quad t \le T,$$

and

$$x_t = x + \int_0^t \phi_s ds, \quad t \le T.$$

Let us denote $\tilde{a}(t,x) = \tilde{\sigma}\tilde{\sigma}^*(t,x)$. We consider the following conditions:

(I) $\tilde{a}(t, x_t) > 0$;

(G) *(growth condition)* There exists $\mu(R) > 1$ and a function $c : \mathbb{R}_+ \to \mathbb{R}^n$ such that for every $t \in [0,T]$ and every y such that $|y - x_t| \le \frac{3}{4}R$,

$$\left|\tilde{a}(t,x_t)^{-1/2}\big(\tilde{b}(t,y) - \phi_t\big)\right| \le |\tilde{a}(t,x_t)^{-1/2}c(t)| \quad \mathbb{P}\text{-a.s.}$$

$$\frac{1}{\mu(R)}\langle \tilde{a}(t,x_t)\xi,\xi\rangle \le \langle \tilde{a}(t,y)\xi,\xi\rangle \le \mu(R)\langle \tilde{a}(t,x_t)\xi,\xi\rangle \quad \forall \xi \in \mathbb{R}^n, \mathbb{P}\text{-a.s.}$$

$$Tr(\tilde{a}(t,y)) \le \mu(R)Tr(\tilde{a}(t,x_t)) \quad \mathbb{P}\text{-a.s.}$$

Proposition 4.3.2. *Under (I) and (G),*

$$\mathbb{P}(|X_t - x_t| \le R, t \in [0,T]) \ge \exp\big(-(\sqrt{\overline{\alpha}_T(R)} + \sqrt{\beta_T(R)})^2\big) \tag{4.3.9}$$

with

$$\alpha_T(R) = \mu(R)\int_0^T ||\tilde{a}(s,x_s)||ds, \quad \overline{\alpha}_T(R) = 2\log\Big(\frac{\pi}{2}\Big) + \frac{2\pi^2}{R^2}\alpha_T(R),$$
$$\beta_T(R) = \mu(R)\int_0^T \Big(\big|\tilde{a}(s,x_s)^{-1/2}c(t)\big|^2 + 1_{\{n>1\}}\mu(R)^2\frac{9}{16}R^2\frac{Tr(\tilde{a}(s,x_s))^2}{\lambda(\tilde{a}(s,x_s))}\Big)ds.$$

Proof. We consider the process

$$Y_t = X_t - x_t = \int_0^t (\tilde{b}(s, x_s + Y_s) - \phi_s)ds + \sum_{j=1}^d \int_0^t \tilde{\sigma}_j(s, x_s + Y_s)dW_s^j, \quad t \le T.$$

For every $y \in \mathbb{R}^n$ such that $|y| \le \frac{3}{4}R$ and for every $\xi \in R^n$ we have:

$$\langle \tilde{a}(t, x_t + y)\xi,\xi\rangle \le \mu(R)\langle \tilde{a}(t,x_t)\xi,\xi\rangle$$

and

$$\begin{aligned}|a^{-1/2}(t,x_t+y)\big(\tilde b(t,x_t+y)-\phi_t\big)|^2 &= \langle \tilde a^{-1}(t,x_t+y)\big(\tilde b(t,x_t+y)-\phi_t\big),\big(\tilde b(t,x_t+y)-\phi_t\big)\rangle\\ &\le \mu(R)\langle \tilde a^{-1}(t,x_t)\big(\tilde b(t,x_t+y)-\phi_t\big),\big(\tilde b(t,x_t+y)-\phi_t\big)\rangle\\ &= \mu(R)|a^{-1/2}(t,x_t)c(t)|^2.\end{aligned}$$

We now claim that

$$\frac{\Big(Tr(\tilde a(t,x_t+y))-\frac{\langle \tilde a(t,x_t+y)y,y\rangle}{|y|^2}\Big)^2}{\langle \tilde a(t,x_t+y)y,y\rangle}\le \mu(R)^3\frac{9R^2}{16}\frac{Tr(\tilde a(t,x_t))^2}{\lambda(\tilde a(t,x_t))}. \tag{4.3.10}$$

Then, the hypothesis (H) is satisfied by the process $Y_t=X_t-x_t$ and the functions $\overline{a}(R,t)=\mu(R)\tilde a(t,x_t)$; $c_b(R,t)=\sqrt{\mu(R)}|\tilde a(t,x_t)^{-1/2}c(t)|$; $c_\sigma(R,t)=\mu(R)^3\frac{9}{16}R^2\frac{Tr(\tilde a(t,x_t))^2}{\lambda(\tilde a(t,x_t))}$. Estimate (4.2.2) in Theorem 4.2.1 yields the desired result.

Proof of (4.3.10). $Tr(\tilde a(t,x_t+y))$ equals the sum of the (positive) eigenvalues of $\tilde a(t,x_t+y)$, while $\frac{\langle \tilde a(t,x_t+y)y,y\rangle}{|y|^2}$ is clearly smaller than the larger eigenvalue. Hence $0<Tr(\tilde a(t,x_t+y))-\frac{\langle \tilde a(t,x_t+y)y,y\rangle}{|y|^2}<Tr(\tilde a(t,x_t+y))$ and

$$\begin{aligned}\frac{\Big(Tr(\tilde a(t,x_t+y))-\frac{\langle \tilde a(t,x_t+y)y,y\rangle}{|y|^2}\Big)^2}{\langle \tilde a(t,x_t+y)y,y\rangle}&\le Tr(\tilde a(t,x_t+y))^2\times\frac{|y|^2}{\inf_{|z|=1}\langle \tilde a(t,x_t+y)z,z\rangle}\\ &\le \mu(R)^2 Tr(\tilde a(t,x_t))^2\times\frac{9}{16}R^2\frac{\mu(R)}{\lambda(\tilde a(t,x_t))}.\end{aligned}$$

□

In the previous proposition we have applied Theorem 4.2.1 to the shifted process X_t-x_t. As previously stated, we can also rescale the process multiplying by an invertible matrix. If we consider $\tilde a(t,x_t)^{-1/2}(X_t-x_t)$, applying Itô's formula it is easy to see that the diffusion matrix of the resulting process is $\tilde a(t,x_t)^{-1}\tilde a(t,X_t)$. This is not the identity matrix but it is not "far" from it, since the process X_t remains in the tube around the curve x_t. In particular, we can ask for this matrix to be elliptic up to τ_R. Then, assume (I) and let

$$\Gamma_T(R)=\big\{(t,y)\in[0,T]\times\mathbb{R}^n:|\tilde a(t,x_t)^{-1/2}(y-x_t)|\le\frac{3}{4}R\big\}.$$

We consider the following conditions:

(D) $t\to\tilde a(t,x_t)^{ij}\in C^1([0,T])$;

(G') *(growth condition)* There exists $\mu(R)>1$ and a function $c:\mathbb{R}_+\to\mathbb{R}^n$ such that for every $(t,y)\in\Gamma_T(R)$,

$$|\tilde a(t,x_t)^{-1/2}(\tilde b(t,y)-\phi_t)|\le|\tilde a(t,x_t)^{-1/2}c(t)|\quad \mathbb{P}\text{-a.s.}$$

$$\frac{1}{\mu(R)}\langle\tilde a(t,x_t)\xi,\xi\rangle \le \langle\tilde a(t,y)\xi,\xi\rangle \le \mu(R)\langle\tilde a(t,x_t)\xi,\xi\rangle \quad \forall\xi\in\mathbb{R}^n, \mathbb{P}\text{-a.s.}$$

Proposition 4.3.3. *Under (I), (D) and (G'),*

$$\mathbb{P}\big(\big|\tilde a(t,x_t)^{-1/2}(X_t-x_t)\big|\le R, t\in[0,T]\big) \ge \exp\big(-(\sqrt{\overline{\alpha}_T(R)}+\sqrt{\beta_T(R)})^2\big) \tag{4.3.11}$$

with

$$\begin{aligned}\overline{\alpha}_T(R) &= 2\log\Big(\frac{\pi}{2}\Big)+\mu(R)\frac{2\pi^2}{R^2}T,\\ \beta_T(R) &= 1_{\{n>1\}}\mu(R)^3n^2\frac{9R^2}{16}T+2\mu(R)\int_0^T\Big(\big|\tilde a(s,x_s)^{-1/2}c(t)\big|^2\\ &\qquad +\frac{9}{64}R^2\|\tilde a(t,x_t)^{-1}\frac{d}{dt}(\tilde a(t,x_t))\|^2\Big)ds.\end{aligned}$$

Proof. We denote $\tilde a_t := \tilde a(t,x_t)$ for simplicity. Let us define $b,\sigma_j : B_0(\frac{3}{4}R)\to\mathbb{R}^n$, $j=1,\dots,n$ and $a : B_0(\frac{3}{4}R)\to S_+(n)$ by

$$\begin{aligned}b(t,y) &:= \tilde a_t^{-1/2}\big(\tilde b(t,x_t+\tilde a_t^{1/2}y)-\phi_t\big)+\frac{d}{dt}\big(\tilde a_t^{-1/2}\big)\tilde a_t^{1/2}y\\ \sigma_j(t,y) &:= \tilde a_t^{-1/2}\tilde\sigma_j(t,x_t+\tilde a_t^{1/2}y), \qquad j=1,\dots,d\\ a(t,y) &:= \sigma\sigma^*(t,y)=\tilde a_t^{-1/2}\tilde a(t,x_t+\tilde a_t^{1/2}y)\tilde a_t^{-1/2}.\end{aligned}$$

Then, an application of the Ito's formula allows one to show that the process $Y=(Y_t,t\le T)$, $Y_t=\tilde a_t^{-1/2}(X_t-x_t)$, satisfies

$$Y_t=\int_0^t b(s,Y_s)ds+\sum_{j=1}^d\int_0^t\sigma_j(s,Y_s)dW_s^j.$$

Let us consider $t\in[0,T]$, $y\in\mathbb{R}^n$ with $|y|\le\frac{3}{4}R$ and $\xi\in\mathbb{R}^n$. By (G'), we have

$$\begin{aligned}\langle a(t,y)\xi,\xi\rangle &= \langle\tilde a(t,x_t+\tilde a_t^{1/2}y)\tilde a_t^{-1/2}\xi,\tilde a_t^{-1/2}\xi\rangle\\ &\le \mu(R)\langle\tilde a(t,x_t)\tilde a(t,x_t)^{-1/2}\xi,\tilde a(t,x_t)^{-1/2}\xi\rangle=\mu(R)|\xi|^2\end{aligned} \tag{4.3.12}$$

and, with analogous reasoning,

$$\langle a(t,y)\xi,\xi\rangle \ge \frac{1}{\mu(R)}|\xi|^2. \tag{4.3.13}$$

As a consequence of (4.3.12), $Tr(a(t,y))\le\mu(R)n$. Hence, proceeding as in the proof of

(4.3.10) in Proposition 4.3.2, we have

$$\frac{\Big(Tr(a(t,y)) - \frac{\langle a(t,y)y,y\rangle}{|y|^2}\Big)^2}{\langle a(t,y)y,y\rangle} \leq Tr(a(t,y))^2\mu(R)|y|^2 \leq \mu(R)^3 n^2 \frac{9}{16}R^2.$$

Now, it is easy to see that

$$\frac{d}{dt}\big(\tilde{a}_t^{-1/2}\big) = -\frac{1}{2}\tilde{a}_t^{-1}\frac{d}{dt}(\tilde{a}_t)\tilde{a}_t^{-1/2}$$

so that $\frac{d}{dt}\big(\tilde{a}_t^{-1/2}\big)\tilde{a}_t^{1/2}y = -\frac{1}{2}\tilde{a}_t^{-1}\frac{d}{dt}(\tilde{a}_t)y$, hence

$$\begin{aligned}|b(t,y)|^2 &\leq 2\Big(\big|\tilde{a}_t^{-1/2}\big(\tilde{b}(t,x_t+\tilde{a}_t^{1/2}y)-\phi_t\big)\big|^2 + \frac{1}{4}\big|\tilde{a}_t^{-1}\frac{d}{dt}(\tilde{a}_t)y\big|^2\Big)\\ &\leq 2\Big(\big|\tilde{a}_t^{-1/2}c(t)\big|^2 + \frac{9}{64}R^2\|\tilde{a}_t^{-1}\frac{d}{dt}(\tilde{a}_t)\|^2\Big).\end{aligned}$$

On the other hand, by (4.3.13) we have $|a(t,y)^{-1/2}b(t,y)|^2 = \langle \tilde{a}(t,y)^{-1}b(t,y), b(t,y)\rangle \leq \mu(R)|b(t,y)|^2$. As a consequence of the inequalities previously obtained, the hypothesis (H) is satisfied by the process Y and the functions

$$\begin{aligned}&\overline{a}(R,t) = \mu(R)I_n, \quad t\leq T; \qquad c_\sigma(t) = \mu(R)^3n^2\frac{9}{16}R^2, \quad t\leq T\\ &c_b(R,t) = 2\mu(R)\big(|\tilde{a}(t,x_t)^{-1/2}c(t)|^2 + \frac{9}{64}R^2\|\tilde{a}(t,x_t)^{-1}\frac{d}{dt}(\tilde{a}(t,x_t))\|^2\big).\end{aligned}$$

Estimate (4.2.2) in Theorem 4.2.1 yields the desired result. □

We can now continue the example of the previous section. We give lower bounds for the tails of the distribution function of a process X_t satisfying (4.3.1) that are in the same range (on the log-scale) as the upper bound (4.3.5) in Prop. (4.3.1).

Proposition 4.3.4 (Example 1, continued)**.** Let $X = (X_t, t\leq T)$ be a continuous process satisfying (4.3.1). Then, there exist strictly positive constants c_T and d_T depending also on x and α, such that for every z large enough, precisely

$$z > x\big(1+2\sinh(T/2)\big)^{1/(1-\alpha)}, \tag{4.3.14}$$

we have

$$\mathbb{P}(X_t > z) \geq \exp\big(-d_T - c_T z^{2(1-\alpha)}\big). \tag{4.3.15}$$

Proof. We will give the proof in three steps, mainly following the scheme set up in Chapter 3, Theorem 3.2.1 and Corollary 2. Let us consider

$$R = \frac{2}{3\sigma}x^{1-\alpha}. \tag{4.3.16}$$

Step 1. We refer to the notation in the proof of Proposition (4.3.1). Let us consider a curve $x_\cdot \in C^1([0,T])$ such that

$$x_0 = x, \quad x_T = 2z, \quad x'_t > 0 \quad \forall t \in [0,T]. \tag{4.3.17}$$

Clearly $t \mapsto a(x_t) = \sigma^2 x_t^{2\alpha} \in C^1([0,T])$. For every $t \in [0,T]$ and every y such that $|y - x_t| \leq \frac{3}{4}Ra(x_t)^{1/2} = \frac{3}{4}R\sigma x_t^\alpha$, we have

$$a(y) \leq \sigma^2\big(x_t + \frac{3}{4}R\sigma x_t^\alpha\big)^{2\alpha} \leq 2^{2\alpha}\sigma^2 x_t^{2\alpha},$$

since $\frac{3}{4}R\sigma x_t^\alpha \leq x^{1-\alpha}x_t^\alpha \leq x_t$ by (4.3.16). Similar computations lead to $a(y) \geq \frac{\sigma^2}{2^{2\alpha}}x_t^{2\alpha}$; moreover, since $\beta \leq 1$,

$$|b(y) - x'_t| \leq \kappa y^\beta + x'_t \leq \kappa\big(x_t + \frac{3}{4}R\sigma x_t^\alpha\big)^\beta + x'_t \leq 2\kappa x_t + x'_t.$$

Then, the hypothesis (I), (D) and (G') are satisfied for the process X, the curve $x_\cdot$ and with $\mu(R) = 2^{2\alpha}$, $c(t) = 2\kappa x_t + x'_t$. Then we have

$$\begin{aligned}\beta_T(R) &= 2^{2\alpha+1}\int_0^T \Big(\frac{(2\kappa x_t + x'_t)^2}{\sigma^2 x_t^{2\alpha}} + \frac{9R^2}{64} \times 4\alpha^2\Big(\frac{x'_t}{x_t}\Big)^2\Big)dt \\ &\leq 2^{2\alpha+1}\int_0^T \Big(\frac{4\kappa^2}{\sigma^2}x_t^{2(1-\alpha)} + \frac{5}{4\sigma^2}\frac{(x'_t)^2}{x_t^{2\alpha}}\Big)dt\end{aligned}$$

and the last inequality comes from (4.3.16) and the fact that $x_t^{2\alpha} = x^{2\alpha}\big(\frac{x_t}{x}\big)^{2\alpha} \leq x^{2\alpha}\big(\frac{x_t}{x}\big)^2 = \frac{x_t^2}{x^{2(1-\alpha)}}$.

Step 2 (Lagrange optimization). Let us set $\Gamma(x) = \frac{2^{2\alpha+1}}{\sigma^2}\max(4\kappa^2, \frac{5}{4})$. We are led to the minimization of the functional

$$\Gamma(x)\int_0^T \mathcal{L}(x_t, x'_t)dt \tag{4.3.18}$$

under the constraints (4.3.17), where $\mathcal{L}(x,x') = x^{2(1-\alpha)} + \frac{(x')^2}{x^{2\alpha}}$. The Euler-Lagrange equation associated to $\mathcal{L}$ is easily found to be

$$\frac{x''_t}{x'_t} = \alpha\frac{x'_t}{x_t} + (1-\alpha)\frac{x_t}{x'_t}. \tag{4.3.19}$$

Eq. (4.3.19) can be explicitly solved by means of the change of variables $u_t = \big(\frac{x_t}{x}\big)^{1-\alpha}$, which converts (4.3.19) into the linear equation

$$u''_t - (1-\alpha)^2 u_t = 0, \quad u_0 = 1, \quad u_T = (2z/x)^{1-\alpha}. \tag{4.3.20}$$

The solution to (4.3.20) is given by

$$u_t = \Big(\frac{2z}{x}\Big)^{1-\alpha}\varphi(t) - e^{-(1-\alpha)T}\varphi(t) + e^{-(1-\alpha)t},$$

where $\varphi_t = \frac{\sinh((1-\alpha)t)}{\sinh((1-\alpha)T)}$. Let us set $\tilde{x}_t = x u_t^{1/(1-\alpha)}$. It is now a matter of computation to evaluate the functional in (4.3.18) on this particular curve. A direct computation shows that $u'_t > 0$, hence $\tilde{x}'_t > 0$, if z satisfies (4.3.14). Moreover, we have

$$\begin{aligned} & u_t \le \Big(\frac{2z}{x}\Big)^{1-\alpha}\varphi(t) + e^{-(1-\alpha)t}, \quad u'_t \le \Big(\frac{2z}{x}\Big)^{1-\alpha}\frac{\cosh((1-\alpha)t)}{\sinh((1-\alpha)T)}, \\ & \frac{(\tilde{x}'_t)^2}{\tilde{x}_t^{2\alpha}} = \frac{x^{2(1-\alpha)}}{(1-\alpha)^2}(u'_t)^2 \end{aligned}$$

hence

$$\begin{aligned} \int_0^T \mathcal{L}(\tilde{x}_t, \tilde{x}'_t)dt &\le x^{2(1-\alpha)} \int_0^T \Big(u_t^2 + \frac{(u'_t)^2}{(1-\alpha)^2}\Big)dt \\ &\le 2(2z)^{2(1-\alpha)} \int_0^T \Big(\varphi(t)^2 + \frac{\cosh^2((1-\alpha)t)}{(1-\alpha)^2\sinh^2((1-\alpha)T)}\Big)dt + x^{2(1-\alpha)}\frac{1-e^{-2(1-\alpha)T}}{1-\alpha} \end{aligned}$$

Setting

$$\begin{aligned} c_T &:= 2^{4-2\alpha}\Gamma(x)\int_0^T \Big(\varphi(t)^2 + \frac{\cosh^2((1-\alpha)t)}{(1-\alpha)^2\sinh^2((1-\alpha)T)}\Big)dt \\ d_T &:= 2\alpha_T(R) + 2\Gamma(x)x^{2(1-\alpha)}\frac{1-e^{-2(1-\alpha)T}}{1-\alpha} \\ &= 4\log\Big(\frac{\pi}{2}\Big) + 2^{2\alpha+2}\frac{9\sigma^2\pi^2}{4x^{2(1-\alpha)}}T + 2\Gamma(x)x^{2(1-\alpha)}\frac{1-e^{-2(1-\alpha)T}}{1-\alpha} \end{aligned}$$

we have $(\sqrt{\alpha_T(R)} + \sqrt{\beta_T(R)})^2 \le 2(\alpha_T(R) + \beta_T(R)) \le d_T + c_T z^{2(1-\alpha)}$. Estimate (4.3.11) in Proposition 4.3.3 then yields

$$\mathbb{P}\big(|X_t - \tilde{x}_t| \le \sigma R \tilde{x}_t^\alpha, t \in [0,T]\big) \ge \exp\big(-d_T - c_T z^{2(1-\alpha)}\big). \qquad (4.3.21)$$

Step 3. We conclude by observing that $z > R\sigma z^\alpha$ if (4.3.14) holds, hence

$$\mathbb{P}(X_T \ge z) \ge \mathbb{P}(|X_T - 2z| \le \sigma R z^\alpha) \ge \mathbb{P}\big(|X_t - \tilde{x}_t| \le \sigma R \tilde{x}_t^\alpha, t \in [0,T]\big)$$

and (4.3.21) yields (4.3.15). □

4.4 A lemma on suprema of Brownian motion

Here we prove a lemma about the law of suprema of the one-dimensional Brownian motion.

Lemma 4.4.1. *Let $(b_t; t \geq 0)$ be a standard one-dimensional Brownian motion. Then, for every $T > 0$ and every $\epsilon > 0$,*

$$\mathbb{P}\Big(\sup_{t\leq T} |b_t| < \epsilon\Big) = \sum_{n=2k+1, k\geq 0}^{\infty} (-1)^{\frac{n-1}{2}} \frac{4}{\pi n} e^{-\frac{\pi^2 n^2 T}{8\epsilon^2}}, \tag{4.4.1}$$

in particular

$$\mathbb{P}\Big(\sup_{t\leq T} |b_t| < \epsilon\Big) \geq \frac{2}{\pi} e^{-\frac{\pi^2 T}{8\epsilon^2}}. \tag{4.4.2}$$

We also have

$$\mathbb{P}\Big(\sup_{t\leq T} |b_t| > \epsilon\Big) \leq 2e^{-\epsilon^2/2T}. \tag{4.4.3}$$

Proof. Estimate (4.4.3) is proven in [57], pg. 60, relying on exponential brownian martingales. On the other hand, according to [39], pg. 520-521, we have

$$\mathbb{P}\Big(\sup_{t\leq T} |b(t)| < \epsilon\Big) = \sum_{n=1}^{\infty} \exp\Big(-\frac{\lambda_n T}{\epsilon^2}\Big) \phi_n(0) \int_{-1}^{1} \phi_n(y) dy, \tag{4.4.4}$$

where $0 < \lambda_1 \leq \lambda_2 \leq \ldots$ are the eigenvalues and $(\phi_n)_n$ the corresponding eigenfunctions of the eigenvalue problem

$$\begin{cases} \frac{1}{2}\phi''(x) + \lambda\phi(x) = 0, & x \in (-1, 1) \\ \phi(-1) = \phi(1) = 0. \end{cases} \tag{4.4.5}$$

Therefore, in order to prove (4.4.1) we are left with the explicit computation of the spectrum. It is easy to see that the solution to (4.4.5) is given by

$$\lambda_n = \frac{\pi^2 n^2}{8}, \quad \phi_n(x) = \sin\Big(\frac{\pi n}{2}(x-1)\Big).$$

Now, by direct computation

$$\phi_n(0) \int_{-1}^{1} \phi_n(y) dy = \begin{cases} 0 & \text{if } n \text{ is even} \\ (-1)^{\frac{n-1}{2}} \frac{4}{\pi n} & \text{if } n \text{ is uneven,} \end{cases}$$

hence, in the sum at the right hand side of (4.4.4) only the uneven terms give a contribution, and direct substitution gives (4.4.1). (4.4.2) follows from (4.4.1) □

Part II

Implied volatility and the SVI parametric model: calibration and time-dependent extensions

Chapter 5

Quasi-Explicit Calibration of Gatheral's SVI model

Abstract

We present a procedure - based on dimension reduction in the parameters space - providing a quasi-explicit calibration of J. Gatheral's SVI implied variance model. The resulting parameter identification is reliable, robust and stable.

Keywords: SVI model, Calibration.

Note

The study developed in this chapter is motivated by some smart remarks of Claude Martini on the properties of the SVI model. I implemented the numerical tests that are presented in section 5.4 relying on Zeliade Systems' code development framework. I would like to thank Zeliade Systems' team for the related technical support (in particular Steve Younan and Tarek Bouz for several insights on C$\#$) and for providing with the market data.

5.1 A simple model and a delicate calibration

In [29], Jim Gatheral proposes the following parametric form to model the fixed-maturity smile:

$$\sigma^2_{SVI}(x) = a + b\Big(\rho(x-m) + \sqrt{(x-m)^2+\gamma^2}\Big), \tag{5.1.1}$$

where σ^2_{SVI} is the implied variance (square of implied volatility) at fixed time-to-maturity T, x the log-moneyness $x := log(K/F_0)$ and a, b, ρ, m, γ are the model parameters. (5.1.1) is known as the Stochastic Volatility Inspired (SVI) model since the functional form of the curve has been inspired by the results on the large-time asymptotics of implied variance in the Heston model.

Parametric models (SVI, approximations of implied volatility in CEV or SABR models, others) are of common use in the treatment of the volatility surface. Apart from the extrapolation of smile points, they provide a smoothing of the market smile and the consequent facilities in the calibration of stochastic models for the underlying (including the reconstruction of a local volatility surface via Dupire's formula, for which interpolation in time is also needed). It is widely known that the SVI functional form (5.1.1) proves to have outstanding performances in the calibration to single-maturity slices of the implied smile on Equity indexes. Nevertheless, it is also common knowledge that the least square calibration of (5.1.1) is typically affected by the presence of several local minima. To our experience, even when SVI is calibrated to simulated data, i.e. a smile produced by SVI itself, local minima that are difficult to sort out (least square objective $\approx 10^{-8}$ for reasonable volatilities, $\sigma \approx 20\% - 40\%$) are found far away from the global one (objective $= 0$).
This unpleasant feature brings some difficulties, when one wants to design a reliable and robust parameter identifications strategy in SVI model. Because of the local minima, the solution yield by a least square optimizer usually displays a strong dependence on the input starting point. Of course, smart initial guesses of parameters can be done by looking to the "geometry" of the observed smile (asymptotic slopes, minimum). Then, one can run the calibration using more than one non-linear optimizer and/or restarting from several different initial guesses. Nonetheless, the strategy to find the initial point is not defect-free and requires attention, since the desired smile features are not available in all the cases and often not for all maturities (e.g. the wings are not both observed, smile has no clearly visible minimum). The calibration reset, though useful, does not guarantee to overcome all the local minima: the final choice of optimal parameters can still be ambiguous, since the same smile can be - remarkably well - calibrated with sets of parameters that are totally different one from the other. The big issue, then, becomes the stability of calibrated parameters through time-to-maturity. This feature comes into play in a significant way when one is trying to parameterize the whole volatility surface.

This chapter presents a procedure providing a trustworthy and robust calibration of the SVI parametric form (5.1.1), having the pleasant feature to be almost insensitive to the initial parameter guess. We rely on some simple observations on the symmetries of the functional form (5.1.1) to downsize the minimization problem from dimension 5 (the number of parameters in (5.1.1)) to dimension 2 (namely, m and γ), while the optimization over the remaining 3 is performed exactly (except for a few minimum searches in dimension 1, which can nevertheless be performed accurately and fast). Last but not least, the method yields an optimal parameter set which is automatically consistent with the arbitrage constraint on the slopes of the implied variance. The procedure is presented in section 5.3. We start by reviewing the main properties of the SVI parametric family in the next section.

5.2 Parameter constraints and limiting cases

The parameters a, b, ρ, m, γ in general depend on time-to-maturity T. We assume that b, γ, ρ satisfy

$$b > 0, \quad \gamma \geq 0, \quad \rho \in [-1, 1].$$

Further conditions on b and ρ follow from arbitrage conditions (cf. §5.2.2 hereafter). We will discuss some constraints on the parameters γ and a in §5.2.3 and §5.2.4. We recall that in this chapter we just look to the parameterization of time-slices of the implied variance.

5.2.1 Slopes and minimum

We review the main interesting properties of the parametric form (5.1.1). As pointed our in cf.[29], it is easy to see that the left and right asymptotes are respectively

$$\sigma_L^2(x) = a - b(1-\rho)(x-m),$$
$$\sigma_R^2(x) = a + b(1+\rho)(x-m).$$

The term adding to a in (5.1.1) is always positive and convex with respect to x. Simple computations show that σ_{SVI}^2 has a unique minimum point if $\rho^2 \neq 1$ and is in its "degenerate" form if $\rho^2 = 1$, precisely:

- if $\rho^2 \neq 1$, the minimum is $min\{\sigma_{SVI}^2\} = a + b\gamma\sqrt{1-\rho^2}$ attained at $x^* = m - \frac{\rho\gamma}{\sqrt{1-\rho^2}}$;
- if $\rho^2 = 1$, σ_{SVI}^2 is not-increasing for $\rho = -1$ and not-decreasing for $\rho = 1$ and
  - if $\gamma \neq 0$, σ_{SVI}^2 is strictly monotone and the minimum is never attained (nevertheless, $\sigma_{SVI}^2 \to a$ respectively for $x \to \infty$ or $x \to -\infty$);
  - if $\gamma = 0$, σ_{SVI}^2 has the shape of a Put or Call payoff of strike m (precisely, σ_{SVI}^2 is worth a for $x \geq m$ if $\rho = -1$ and for $x \leq m$ if $\rho = 1$).

5.2.2 Arbitrage constraints (b and ρ)

A necessary condition for the absence of arbitrage is given by a constraint on the maximal slopes of the total implied variance $T\sigma_{SVI}^2(x)$. As found in [58], this condition reads

$$\forall\, x, \forall\, T, \quad |T\partial_x \sigma_{SVI}^2(x)| \leq 4. \tag{5.2.1}$$

As stated in [29], this translates into the following equivalent condition on b and ρ :

$$b \leq \frac{4}{(1+|\rho|)T}. \tag{5.2.2}$$

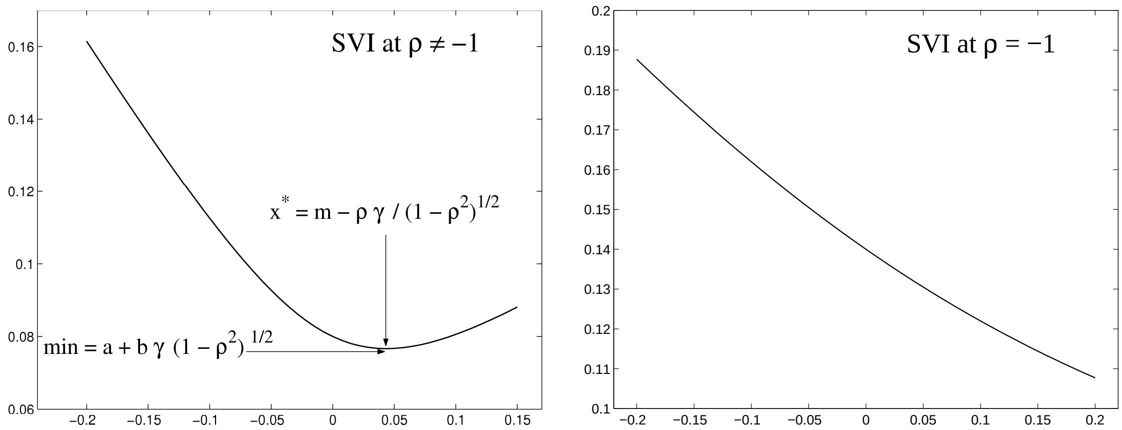

Figure 5.1: Examples of SVI smile shapes. SVI parameters are $a = 0.04, b = 0.4, \rho = -0.4, m = 0.05, \gamma = 0.1$ (left) and $a = 0.04, b = 0.2, \rho = -1, m = 0.1, \gamma = 0.5$ (right).

5.2.3 Limiting cases $\gamma \to 0$ and $\gamma \to \infty$ (almost-affine smiles)

As observed in §5.2.1 for the case $\rho^2 = 1$, letting $\gamma \to 0$ also makes the SVI parameterisation piecewise affine. In this case, the implied variance reads

$$\sigma^2_{SVI}(x) = a + b(\rho \mp 1)(x - m) \tag{5.2.3}$$

respectively in the two regions $x < m$ and $x > m$. Smiles which can be excellently fitted with a monotone affine parameterisation $\sigma^2(x) = px + q$ are not uncommon on Equity indexes, in particular for largest maturities: we will refer to them as to "almost-affine" smiles. Clearly, the calibration of SVI to an almost-affine smile is an ill-posed problem: thinking of a downward smile to fix ideas, it is sufficient to let $\gamma \to 0$, then m to be greater than the largest observed log-moneyness (in order to pick up the minus sign in (5.2.3)) and the matching of the two relevant quantities, i.e. smile slope p and intercept q, gives the two equations

$$\begin{aligned} b(\rho - 1) &= p, \\ a - bm(\rho - 1) &= q \end{aligned}$$

which correspond to infinitely many choices for the parameters a, b, ρ.

The same kind of limiting behaviour is attained in the limit $\gamma \to \infty$ and $a \to -\infty$, in the precise way that we specify as follows. Assume $a < 0$ and $\gamma >> 1$, then

$$\begin{aligned}\sigma^2_{SVI}(x) &= a + b\Big(\rho(x-m) + \sqrt{\gamma^2 + (x-m)^2}\Big) \\ &= -|a| + b\rho(x-m) + b\gamma\sqrt{1 + \frac{(x-m)^2}{\gamma^2}} \\ &\sim -|a| + b\rho(x-m) + b\gamma\Big(1 + \frac{(x-m)^2}{2\gamma^2}\Big) \\ &\sim_{|a|=b\gamma} \; b\rho(x-m) + b\frac{(x-m)^2}{2\gamma}.\end{aligned}$$

Hence

$$\lim_{\gamma\to\infty, a\to-\infty, |a|=b\gamma} \sigma^2_{SVI}(x) = b\rho(x-m)$$

for any value of x, and this correspond again to an affine smile whose slope and intercept will identify the product $b\rho$ and the parameter m, but not b, ρ and m separately.

Smiles tend to flatten with increasing time to maturity, and curved smiles can continuously deform into almost-affine ones. Since robustness and stability of the calibration are the features we have in mind, we would like the calibration strategy to avoid falling into the instable behaviour caused by the limiting SVI cases we have outlined above. Hence, we decide to choose a strictly positive lower bound for γ ($\gamma_{\min} = 0.005$ in our numerical tests) and state that if this threshold is reached, then an unambiguous calibration of SVI is not doable, in the sense that any precise choice of model parameters is arbitrary (of course one could decide, for example, to inherit one of the SVI parameters from the calibration to the previous time-slice - if any - but this goes back to user choices). In the same way, we want to avoid the $\gamma \to \infty$, $a \to -\infty$ limiting behaviour: to do this, we claim that a good choice is to constrain a to positive values. Setting an upper bound for γ would prevent it from assuming too high values, but this would not avoid the situation where γ sticks to $\gamma_{\max}$ and a becomes very negative.

5.2.4 Constraints on a

The constraints on parameter a have been partially discussed in the previous section. *A priori*, negative values of a could be allowed, since the positivity of the parameterisation is simply achieved by asking that the minimum of σ^2_{SVI} (when attained) be non-negative, i.e.

$$a \geq -b\gamma\sqrt{1-\rho^2}$$

(if the minimum is not attained, then $\rho^2 = 1$ and the previous condition becomes $a \geq 0$). In the previous section we have explained that, in order to avoid the phenomena of "coupling" of low values of a and high values of γ, we prefer to assume

$$a \geq 0. \tag{5.2.4}$$

We just add here that an obvious upper bound on a is:

$$a \leq \max_i \{\sigma_i^2\}, \tag{5.2.5}$$

where the σ_i^2's are the market variances at the given maturity. Condition (5.2.5) simply follows from a consistent vertical location of the graph of (5.1.1): clearly the curve σ_{SVI}^2 giving the optimal fit cannot be systematically greater than the largest observed variance.

5.3 A convex optimization problem with Linear Program

Let us now come to the calibration problem. As it stands, the calibration of the SVI parametric form, cast as a least square problem, yields an optimization in dimension 5. In thee following we show that, relying on some simple observations on the properties of the functional form (5.1.1), one can reformulate the problem reducing the main dimension from 5 to 2.

5.3.1 Dimension reduction: drawing out the linear objective

We focus hereafter on the total variance $\tilde{v} = T\sigma_{SVI}^2$ rather than on variance. The main ingredient of the method is the fact that, by means of the change of variables

$$y = \frac{x - m}{\gamma}$$

the SVI parametrization transforms into

$$\tilde{v}(y) = aT + b\gamma T(\rho y + \sqrt{y^2 + 1}).$$

This expression nicely shows how, for fixed values of m and γ, the support of the curve $T\sigma_i^2$ is fully determined by the factors a, ρ and the product $b\gamma$. Thus, most important, if we redefine the model parameters as

$$c = b\gamma T$$
$$d = \rho b\gamma T$$
$$\tilde{a} = aT,$$

then $\tilde{v}(y)$ turns out to depend *linearly* on c, d, a:

$$\tilde{v}(y) = \tilde{a} + dy + c\sqrt{y^2 + 1}.$$

Therefore, for fixed m and γ, we look for the solution of the problem:

$$(P_{m,\gamma}) \qquad \min_{(c,d,\tilde{a})\in D} f_{\{y_i,v_i\}}(c,d,\tilde{a})$$

where $f_{\{y_i,v_i\}}$ is the cost function

$$f_{\{y_i,v_i\}}(c,d,\tilde{a}) = f(c,d,\tilde{a}) = \sum_{i=1}^{n}\left(\tilde{a} + dy_i + c\sqrt{y_i^2+1} - \tilde{v}_i\right)^2,$$

with $\tilde{v}_i = T\sigma_i^2$, and D is the compact and convex domain (a parallelepipedon)

$$D = \begin{cases} 0 \le c \le 4\gamma \\ |d| \le c \text{ and } |d| \le 4\gamma - c \\ 0 \le \tilde{a} \le \max_i\{\tilde{v}_i\} \end{cases}$$

which is obtained from bounds (5.2.2) and (5.2.4)-(5.2.5) on parameters b, ρ and a. Letting $(c^*, d^*, \tilde{a}^*)$ denote the solution of $P_{m,\gamma}$ and (a^*, b^*, ρ^*) the corresponding triplet for a, b, ρ, then the complete calibration problem is restored as

$$(P) \qquad \min_{m,\gamma}\sum_{i=1}^{n}(\sigma_{SVI}^2(m,\gamma,a^*,b^*,\rho^*;x_i) - \sigma_i^2)^2.$$

Our goal is therefore to solve $P_{m,\gamma}$ in the fastest and most accurate way: once this is done, the only task left is to look for the solution of the 2-dim problem P.

5.3.2 Explicit solution of the reduced problem

Let us focus on $P_{m,\gamma}$ (the reduced problem). This is a convex optimization problem with linear program, and *all* the constraints defining the admissible domain D are linear. It is clearly seen, then, that this problem admits an explicit solution, and it becomes extremely easy to deal with.
The cost function f is convex and, if the target smile contains at least three points (an assumption that we are going to make for the following), the gradient of f is zero at just one point, and this is the unique global minimum of f. Therefore, only two scenarios are possible:

- the minimum of f over D is attained at the interior of D, and this is the global minimum of f;
- the minimum of f over D is attained on the boundary ∂D.

Then, this yields the simple recipe:

Step 1. find the global minimizer of f, solving the *linear* system $\nabla f = 0$. If the output belongs to D, then stop;

Step 2. if *Step 1* yields a global minimum outside D, then look for $\min_{\partial D} f$.

The gradient of the cost function f is the affine mapping

$$\begin{aligned}\frac{1}{2}\nabla f(c,d,\tilde{a}) &= A\begin{pmatrix} c \\ d \\ \tilde{a}\end{pmatrix} - b \\ &= \begin{pmatrix} (n+Y_2) & Y_4 & Y_3 \\ Y_4 & Y_2 & Y_1 \\ Y_3 & Y_1 & n \end{pmatrix}\begin{pmatrix} c \\ d \\ \tilde{a}\end{pmatrix} - \begin{pmatrix} vY_2 \\ vY \\ v \end{pmatrix}\end{aligned} \tag{5.3.1}$$

with

$$\begin{array}{ll} Y_1 = \sum_i y_i & Y_2 = \sum_i y_i^2 \\ Y_3 = \sum_i \sqrt{y_i^2+1} & Y_4 = \sum_i y_i\sqrt{y_i^2+1} \end{array}$$

and

$$\begin{aligned} vY_2 &= \textstyle\sum_i \tilde{v}_i\sqrt{y_i^2+1} \\ vY &= \textstyle\sum_i \tilde{v}_i y_i \\ v &= \textstyle\sum_i \tilde{v}_i. \end{aligned}$$

Then, *Step 1* corresponds to the solution of the 3×3 linear system $A \times (c,d,\tilde{a}) = b$, whose numerical treatment goes without saying. The constrained optimization problem addressed in *Step 2* can be solved applying a Lagrange multipliers method for each of the (flat) sides of the domain D. Then, *Step 2* involves only the solution of 3×3 linear systems, plus a few explicit one-dimensional minimizations along the perimeter of the sides. We accept to solve these one-dimensional minimum searches that do not have explicit solution (there are three of them) employing any one-dimensional optimizer (e.g., Brent's method), which is fast and reliable enough not to deteriorate global efficiency.
Once the two steps are accomplished, the solution of the reduced problem $(P_{m,\gamma})$ is achieved in explicit form. The calibration of (5.1.1) is then carried out solving the 2-dimensional problem (P) with some iterative optimizer, whose performances will be extremely enhanced with respect to the original problem in full dimension.

5.4 Numerical Results

In this section we display some numerical tests and calibration results showing the performances of the Quasi-Explicit method.

Table 5.1 compares the "direct" procedure, i.e. standard least square calibration in dimension 5, and the Quasi-Explicit method for a SVI model calibrated to simulated data, i.e. a smile generated by SVI itself, for fixed $T = 1$. The Root Mean Square Error RMSE is $\sqrt{\sum_i(\sigma^2_{SVI}(x_i) - \sigma^2_i)^2}$, hence when looking to RMSE values one must take into account that the natural scale is the one of a variance. For the Standard Least Square calibration, the input value for a is inferred from the minimum observed variance, and the calibration is restarted 10 times from 10 randomly chosen points ($b \in [0, 0.5], \rho \in [-1, 1], m \in [2\min(x_i) < 0, 2\max(x_i) > 0], \gamma \in [0, 1]$). We do not go to great effort here in identifying a smart initial guess for all the parameters since the randomized procedure works quite well anyway, and our intention is rather to display the performances of the Quasi-Explicit calibration. We recall that, for the latter, no inputs for a, b and γ are needed; moreover, we take the initial guesses for m and γ as simple as it might be, i.e. a randomly chosen point. Standard Least Square optimization is performed with truncated-Newton algorithm, while the optimization over m and γ for the Quasi-Explicit method employs Nelder-Mead simplex algorithm. As it is seen from Table 5.1, even if a classical calibration can work properly, the Quasi-Explicit technique brings the objective to extremely small values. Moreover, the calibration we have obtained in the case $\rho = -0.9$ finely shows how a downward SVI smile can be more-than-reasonably calibrated with a SVI smile reaching its minimum (and then pointing upwards) for large values of the log-moneyness, hence with a set of parameters which is far away from the true one.

Table 5.2 and Figure 5.2 display the result of the Quasi-Explicit calibration of SVI model to the market-implied smile on DAX and EuroStoxx 50 indexes, for two different dates (20 August and 22 September 2008, respectively). Concerning Table 5.2, the quality of the fit is excellent through all maturities, somehow worse just for the very shortest one, $T = 1$ month (but this has seemed quite normal to us, since the target market smile for this maturity presented some irregularities). Calibrated parameters show a good stability: obviously, a dependence w.r.t. time to maturity is expected - time-dependence is not taken into account by SVI model - but the important fact here is that the parameters do not show a noisy behaviour (smooth time-dependence is particularly seen for a and b). In particular, concerning the behaviour of ρ: ρ is different than -1 just for the first time-slice, which is the only non downward-pointing smile (cf. Figure 5.2, the smile on DAX on 20 August shows the same feature), and then it sticks to -1, because for all other maturities the smile is purely decreasing. The Quasi-Explicit method has indeed the tendency to fit decreasing smiles with anti-correlated SVI parameterizations.

5.5 Conclusions

Given the excellent performances of the Quasi-Explicit method - at least for what concerns calibration on Equity indexes - we claim that this methodology responds properly to the question of how to obtain an unambiguous identification of a time-slice of implied variance

Method	Parameters	a	b	ρ	m	γ	RMSE
	True value	0.04	0.1	-0.5	0.0	0.1	
Standard LS	start. pt	0.048	10 r.p.	10 r.p.	10 r.p.	10 r.p.	
	calibrated	0.041	0.098	-0.51	$-4e^{-4}$	0.095	$7.9e^{-8}$
Quasi-Expl.	start. pt	-	-	-	1 r.p.	1 r.p.	
	calibrated	0.040	0.10	-0.50	$-8e^{-7}$	0.100	$5.0e^{-14}$
	True value	0.1	0.06	-0.9	0.24	0.06	
Standard LS	start. pt	0.11	10 r.p.	10 r.p.	10 r.p.	10 r.p.	
	calibrated	0.004	0.11	-0.19	0.73	0.58	$8.3e^{-8}$
Quasi-Expl.	start. pt	-	-	-	1 r.p.	1 r.p.	
	calibrated	0.10	0.060	-0.90	0.240	0.060	$3.4e^{-17}$

Table 5.1: Calibration of SVI model to simulated data. The two calibration strategies, Standard Least Square ($dim = 5$) and the Quasi-Explicit method are compared, for $T = 1$. For Standard Least Square, the starting value for a is inferred from minimum variance.

T (Yrs)	a	b	ρ	m	γ	RMSE
0.082	0.027	0.234	0.068	0.100	0.028	$1.6e^{-6}$
0.16	0.030	0.125	-1.0	0.074	0.050	$2.8e^{-7}$
0.26	0.032	0.094	-1.0	0.093	0.041	$2.1e^{-7}$
0.33	0.028	0.105	-1.0	0.096	0.072	$1.3e^{-7}$
0.58	0.026	0.080	-1.0	0.127	0.098	$7.1e^{-8}$
0.83	0.026	0.066	-1.0	0.153	0.113	$1.8e^{-8}$
1.33	0.031	0.047	-1.0	0.171	0.065	$5.2e^{-8}$
1.83	0.037	0.039	-1.0	0.152	0.030	$9.1e^{-10}$
2.33	0.036	0.036	-1.0	0.200	0.083	$1.3e^{-9}$
2.82	0.038	0.036	-1.0	0.170	0.139	$2.4e^{-9}$
3.32	0.034	0.032	-1.0	0.246	0.199	$7.2e^{-10}$
4.34	0.044	0.028	-1.0	0.188	0.069	$2.6e^{-7}$

Table 5.2: Calibration of SVI model to the implied smile on the DAX Index on 20 August 2008. Each maturity is separately calibrated.

in terms of a set of SVI parameters. Once the high-quality fit is achieved, the SVI functional form can serve in many ways. Besides smile point extrapolation, one can recast the calibration of any stochastic model for the underlying as a calibration to the smooth objective (5.1.1). The matching of the geometry (levels, slopes and curvature) of the two model smiles may lead to explicit mappings of SVI parameters onto the ones of the chosen model, in the spirit (in another context) of the calibration methodology of [55]. Of course, this subject includes the issue of extracting a local volatility surface with Dupire's formula: since this needs interpolation in time, at this level the time-interpolation mechanism becomes as well a crucial point.

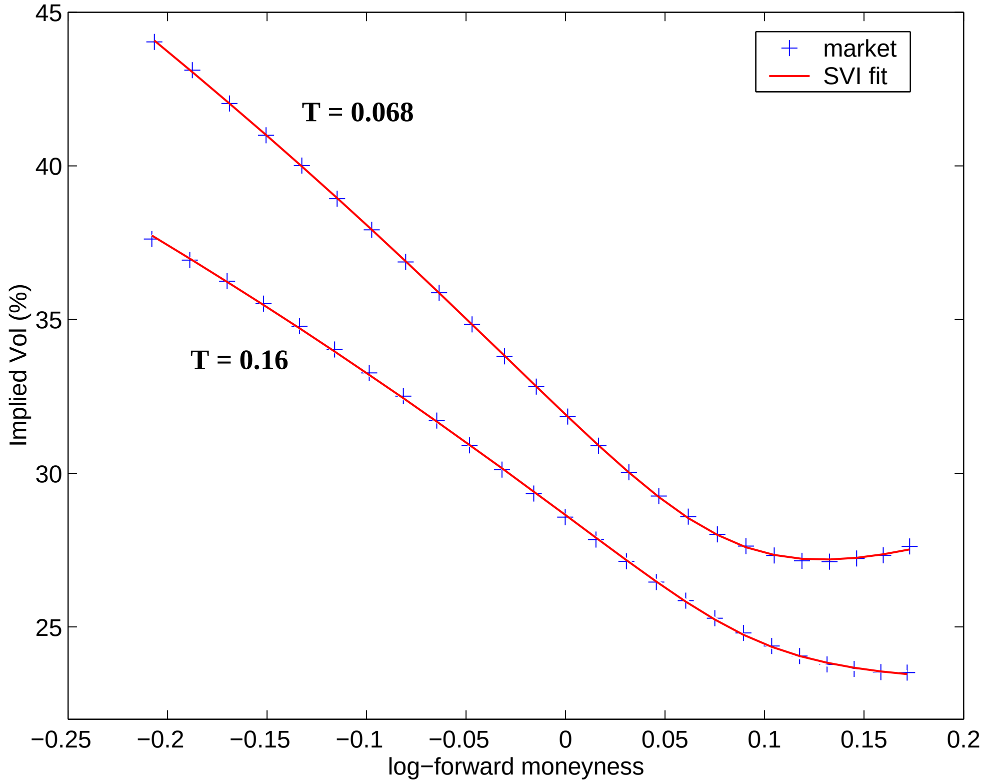

Figure 5.2: Calibration of SVI model to the implied smile on the EuroStoxx 50 Index on 22 September 2008, for the two shortest maturities.

Chapter 6

The Term Structure of Implied Volatility in Symmetric Models

Abstract

We study the term structure of the implied volatility in a situation where the smile is symmetric. Starting from the result by Tehranchi [62] that a symmetric smile generated by a continuous martingale necessarily comes from a mixture of normal distributions, we derive representation formulae for the at-the-money (ATM) implied volatility level and curvature in a general symmetric model. As a result, the ATM curve is directly related to the Laplace transform of the quadratic variation of the log price. To deal with the remaining part of the volatility surface, we build a time dependent SVI-type [29] approximation which matches the ATM and extreme moneyness structure. As an instance of a symmetric model, we consider uncorrelated Heston: in this framework, our representation of the ATM volatility takes semi-closed (and easy to implement) form and proves to be extremely accurate, at the same time the time-dependent SVI approximation displays considerable performances in a wide range of maturities and strikes. In addition, we show how to apply our results to a skewed smile by considering a displaced model. Finally, a noteworthy fact is that all along the paper we will deal only with Laplace transforms and not with Fourier transforms, thus avoiding any complex-valued function.

Keywords: Implied volatility, Term structure, Symmetric smile, SVI, Heston, real-valued functions.

Note

The results in this Chapter have been submitted for publication in the *International Journal of Theoretical and Applied Finance*, in a joint paper with C. Martini. The corresponding preprint can be found at [20]. I owe to Claude Martini brilliant insights on implied volatility modeling: besides him, I would like to thank Antoine Jacquier (Imperial College and

Zeliade Systems) and José da Fonseca (Auckland University of Technology, Ecole Supérieure d'Ingénieurs Léonard de Vinci and Zeliade Systems) for useful discussions on this subject and for reading an earlier version of the manuscript [20], and Pierre Cohort (Zeliade Systems) for insights on displaced models.

6.1 Introduction

The implied volatility is usually recovered from the prices of options by numerical inversion of the Black-Scholes formula. Looking at the implied volatility as a function of time to maturity and strike of the option, the non-linearity of the BS formula makes it difficult to work out the analytical properties of such a function. In recent years, several authors have studied the (static and dynamic) properties the implied volatility surface must exhibit in an arbitrage-free model, with a painstaking emphasis on its behaviour in limiting cases: extreme strikes, short and large maturities. Concerning the dependence with respect to strike, a few major theoretical results are known in a model-independent framework: in particular, the celebrated moment formula (1.2.8) by Lee [47] relates the extreme-strike slopes of the implied volatility to the critical moments of the underlying. Let us recall that the "right" part of the moment formula (i.e. for the right part of the smile) reads:

$$\limsup_{x\to+\infty} \frac{t\sigma(t,x)^2}{x} = \psi(u^*-1) \tag{6.1.1}$$

where $\sigma(t,x)$ denotes the implied volatility of a European Call option with maturity t and strike $K = F_0 e^x$, $\psi(u) = 2 - 4(\sqrt{u^2+u} - u)$ and $u^* = u^*(t) := \sup\{u \geq 1 : \mathbb{E}[F_t^u] < \infty\}$ is the critical moment of the underlying price $F = (F_t)_{t\geq 0}$. An analogous formula holds for the left part of the smile, i.e. for the $\limsup_{x\to-\infty}$. In order to apply this formula in fashionable models, some authors have concentrated their efforts on the computation of the critical moments $u^*(t)$, see Andersen & Piterbarg [2] and Keller-Ressel [41] in the framework of stochastic volatility. The study of short- (resp. long-) time asymptotics of the implied volatility is motivated by the research of efficient calibration strategies of the underlying stochastic model to the market smile at short (resp. long) maturities. Short time results have been obtained following a PDE approach, as in Medvedev and Scaillet [52], Berestycki et al. [10] and Lewis [48] for stochastic volatility models. Some recent works provide deep insights on the large-time behaviour, as done by Tehranchi [61] in a general setting and Keller-Ressel [41] for Affine Stochastic Volatility Models. Given its high tractability, a particular attention has been given to the Heston model [38]. Recall that under the risk-neutral measure, the

Heston model for the forward price F_t of an asset is given by the SDE

$$\begin{aligned} dF_t &= F_t\sqrt{V_t}dW_t \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dZ_t \\ d\langle W, Z\rangle_t &= \rho dt, \end{aligned} \tag{6.1.2}$$

with $F_0 > 0$ and deterministic initial variance $V_0 > 0$. When the two Brownian motions W and Z are independent, i.e. $\rho = 0$, we refer to (6.1.2) as to the uncorrelated Heston model. A survey of the recent advances on the Heston model (in the general framework $\rho \in [-1, 1]$) can be found in [60]. For Heston, sharp results are available on the behaviour of implied volatility at extreme strikes (cf. again [2] and [41] for the computation of the asymptotic slopes and the refinements from Friz et al. [28]) and at short and long maturities (cf. Forde & Jacquier [25], and Forde, Jacquier & Mijatovic [26]).

Parametric forms modeling the implied volatility surface must of course be consistent with the theoretical behaviour described above. So does Gatheral's SVI parameterisation [29]. Recall from the previous chapter that SVI reads:

$$\sigma_{SVI}^2(x)^2 = a + b\big(r(x-m) + \sqrt{(x-m)^2 + \gamma^2}\big), \tag{6.1.3}$$

where σ_{SVI}^2 models the implied variance as a function of the log-moneyness x *at fixed maturity* by means of the five factors a, b, r, m, γ. The parametric form at the rhs of (6.1.3) has been inspired by the known results on the large-time asymptotics of the ATM value and the wings of implied variance in the Heston model. Very recently, Gatheral & Jacquier [31] have shown that the large maturity ($t \to \infty$) Heston smile *does* belong to the SVI family (6.1.3), thus confirming Gatheral's original conjecture. On the other hand, still very recently Friz and other authors [28] have shown that at finite maturities ($t < \infty$) the Heston smile is *not* exactly described by a SVI, and that an additional term should be added to the SVI parameterisation to account for the fine behaviour of the wings. More precisely, relying on some sharp asymptotic estimates of the probability distribution of the Heston spot price, the authors show the validity of the following expansion for large log-moneyness x:

$$\sigma(x,t)^2 t \approx \Big(\beta_1(t)\sqrt{x} + \beta_2(t) + \beta_3(t)\frac{\log(x)}{\sqrt{x}}\Big)^2, \tag{6.1.4}$$

where the coefficients $\beta_1(t), \beta_2(t), \beta_3(t)$ are explicit functions of the maturity, the critical moment $u^*(t)$ and the model parameters.

Despite of all the aforementioned recent advances on the asymptotics of the implied volatility surface, in the general setting fewer results are available in order to describe the behaviour of the implied volatility *close to the money* and at *intermediate maturities*. We mention that Gatheral [30] proposes an heuristic derivation, and discusses the validity, of an approximation of the ATM implied variance in the Heston model, of the form $\sigma(t,0)^2 = (V_0 - \theta')\frac{1-e^{-k't}}{k't} + \theta'$, where θ', k' are functions of Heston parameters.

The main aim and contribution of this chapter is to tackle the problem of the implied volatility term structure in a general setting, at least in a situation where the smile is symmetric. Tehranchi [62] recently showed that a symmetric smile (symmetric meaning $\sigma(t,x) = \sigma(t,-x)$, for all x and t) generated by an underlying continuous martingale necessarily comes from a mixture of normal distributions (cf. eq. (6.2.2) in this chapter for the precise statement of this property). Using this result, together with some properly designed representations of the normal density and cumulative distribution, we derive representation formulae of the density and the cumulative distribution of the log-price in a general symmetric model (cf. Proposition 6.2.4). Then, basing on some simple equations relating the distribution of the underlying price directly to the ATM implied volatility curve, we exploit the previous result to work out the corresponding formulae for the ATM implied volatility level and curvature. As a result, the ATM implied volatility curve is directly related to the Laplace transform of the quadratic variation of the underlying log price: consequently, our representation formulae take a semi-closed form (i.e. the integral of an explicit function) as soon as this Laplace transform is known as a function of model parameters. To deal with the remaining part of the volatility surface, we build a time dependent SVI-type approximation which matches the ATM and "far-from-the-money" (extreme strikes) structure. This construction provides an almost closed-form expression of the ATM implied volatility in the uncorrelated Heston model, which is provided in section 6.3.1. In some numerical experiments we perform in section 6.5 we show that the resulting formula for the ATM implied volatility in the Heston model is extremely accurate (we compare the output to the standard pricers based on the inversion of Fourier transforms), and that the time-dependent SVI approximation displays considerable performances in a wide range of maturities and strikes. As an additional product of our approach, we will get a new quasi-explicit form of the density and of the cumulative distribution (for every strike) of the uncorrelated Heston log price (see Proposition 6.3.4). Even though it is clear that the symmetric smile assumption is not realistic in financial markets, our results can be straightforwardly applied to skewed smiles by considering a displaced model, as we do in section 6.4. Finally, a noteworthy fact is that all along the chapter we will deal only with Laplace transforms and not with Fourier transforms, thus avoiding complex numbers and the related oscillatory integrals.

6.2 Symmetric Models and Symmetric Smiles

We consider a positive continuous martingale $F = (F_t, t \geq 0)$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t, t \geq 0)$ satisfying the usual conditions. We denote $\mathbb{E}$ the expectation under $\mathbb{P}$. In the standard setting, F_t will be the forward price of an asset at time t and $\mathbb{P}$ the risk-neutral measure. If this is the case, since F_t is a forward price the interest rate, the possible dividends and the repo rates are already included in the price itself.

We will make use of some well known relations between the probability distribution of the log forward price $X_t = \log(F_t/F_0)$ and the Black-Scholes implied functions (volatility or total variance). Let us recall these results briefly. We refer to x as to the (log forward) moneyness of a Call option struck at $K = F_0 e^x$. Moreover, we denote $\omega(t,x) = t\sigma(t,x)^2$ the so-called total implied variance (or total variance for short) and $C_{BS}(x,\omega)$ the price of a Black-Scholes Call option maturing at t, with log moneyness x and total variance ω. By the definition of $\omega(x,t)$ we have

$$\mathbb{E}[(F_0 e^{X_t} - F_0 e^x)^+] = C_{BS}(x, \omega(x,t)).$$

Differentiating the above equation w.r.t. x yields[1]

$$-F_0 e^x \mathbb{P}(X_t > x) = \partial_x C_{BS}(x, \omega(x,t)) + \partial_\omega C_{BS}(x, \omega(x,t)) \partial_x \omega(x,t).$$

We use some well known identities for the derivatives of the Black-Scholes Call price,

$$\begin{aligned} \partial_x C_{BS}(x,\omega) &= -F_0 e^x N(d_2), \\ \partial_\omega C_{BS}(x,\omega) &= F_0 N'(d_2 + \sqrt{\omega}) \frac{1}{2\sqrt{\omega}} \end{aligned}$$

where $N(d) = \int_{-\infty}^d \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$ is the standard normal cdf and $d_2 = -\frac{x}{\sqrt{\omega}} - \frac{\sqrt{\omega}}{2}$. Together with the identity $e^{-x} N'(d_2 + \sqrt{\omega}) = N'(d_2)$, the previous relations yield

$$\mathbb{P}(X_t > x) = N(d_2) - \frac{N'(d_2)}{2\sqrt{\omega(t,x)}} \partial_x \omega(t,x) \tag{6.2.1}$$

where d_2 is computed on the total variance, i.e. $d_2 = -\frac{x}{\sqrt{\omega(t,x)}} - \frac{\sqrt{\omega(t,x)}}{2}$. eq. (6.2.1) holds in a model independent framework and relates the smile to the distributional properties of the underlying: the right hand side depends only on Black-Scholes implied quantities, while the left hand side is the (complementary) cdf of the log forward price X_t.

In the next section we will show how, in the presence of a symmetric smile, one can provide a representation formula for the left hand side of (6.2.1) on a very special curve in the (x,t) plane, the ATM curve $x = 0$. In the spirit of the chapter, this will be done focusing on Laplace rather than Fourier transforms, hence avoiding any complex integration. In section (6.2.4) we will show how to extend the same approach to the "away-from-the-money" case, that is to the computation of $\mathbb{P}(X_t > x)$ for any value of $x \in \mathbb{R}$.

[1] At this level, we assume that some sufficient conditions for the Call prices and the implied volatility to be differentiable with respect to x are satisfied. This is the case if the distribution of F_t is continuous (see for example [61]). In the situation we consider, it will be seen hereafter that the distribution of F_t admits a density for all $t > 0$.

We consider the following situation

Assumption 1 (Symmetric Smile)**.** *The smile is symmetric, in the sense that* $\sigma(t,x) = \sigma(t,-x)$ *holds for every* $t > 0$ *and* $x \in \mathbb{R}$.

Let us start by some remarks on the consequences of such an assumption. Several authors have investigated the relationship between a symmetric smile and the properties of the underlying stochastic dynamics. In the context of stochastic volatility (SV), Renault and Touzi [56] showed that an independent volatility process yields a symmetric smile. Carr and Lee [13] proved a converse of this result, showing that a symmetric smile is produced by an uncorrelated SV model. Carr and Lee's result actually holds for a much larger range of underlying dynamics: they prove that, as soon as the underlying price F is a càdlàg martingale, the smile is symmetric if and only if F is geometrically symmetric (cf. (6.2.24) in this paper for the exact formulation of this property). Tehranchi [62] characterizes the property of geometric symmetry in terms of the distributional properties of the log price. Precisely, Tehranchi shows that if F is a geometrically symmetric continuous martingale, then the distribution of $\log(F_t/F_0)$ is a mixture of normal distributions, in the precise sense that the distribution of $\log(F_t/F_0)$ conditional to its quadratic variation $\langle \log(F/F_0)\rangle_t$ is gaussian with variance $\langle \log(F/F_0)\rangle_t$ and mean $-\frac{1}{2}\langle \log(F/F_0)\rangle_t$, for any $t > 0$. Using the notation of the previous section $X_t = \log(F_t/F_0)$, this property reads

$$\mathbb{P}(X_t \in dy | \langle X\rangle_t) = \mathcal{N}\Big(-\frac{1}{2}\langle X\rangle_t, \langle X\rangle_t\Big)(dy). \tag{6.2.2}$$

Then, let us denote

$$p_t(y|\langle X\rangle_t) = \frac{1}{\sqrt{2\pi\langle X\rangle_t}} \exp\Big(-\frac{(2y+\langle X\rangle_t)^2}{8\langle X\rangle_t}\Big), \qquad t > 0, \quad y \in \mathbb{R}. \tag{6.2.3}$$

Eq. (6.2.2) tells that $p_t(y|\langle X\rangle_t)$ is the conditional density of X_t given $\langle X\rangle_t$. Apart for the trivial case where $\langle X\rangle_t = 0$ for some $t > 0$, we remark that (6.2.2) implies, in particular, that the law of X_t admits a density for all $t > 0$ and that this (unconditional) density is given by $p_t(y) = \mathbb{E}\big[p_t(y|\langle X\rangle_t)\big]$. Eqs (6.2.1) and (6.2.2)-(6.2.3) are our starting point to derive representation formulae for the implied volatility in the next two sections.

6.2.1 A representation formula for the ATM implied volatility

Under Assumption 1, $\partial_x \sigma(t,x)|_{x=0} = \partial_x \omega(t,x)|_{x=0} = 0$ for all $t > 0$. Hence, taking $x = 0$ in (6.2.1) and inverting the normal cdf,

$$\sigma(t,0) = -\frac{2}{\sqrt{t}}\, N^{-1}\Big(\mathbb{P}(X_t > 0)\Big), \quad t > 0. \tag{6.2.4}$$

We focus on the argument of N^{-1}, $\mathbb{P}(X_t > 0)$. Observing that (6.2.2) in particular implies $\mathbb{P}(X_t > 0|\langle X\rangle_t) = 1 - N(\frac{1}{2}\sqrt{\langle X\rangle_t}) = N(-\frac{1}{2}\sqrt{\langle X\rangle_t})$, we easily get

$$\mathbb{P}(X_t > 0) = \mathbb{E}\Big[N\Big(-\frac{1}{2}\sqrt{\langle X\rangle_t}\Big)\Big], \quad t > 0. \tag{6.2.5}$$

We now provide a "Laplace transform representation" of the normal cdf N (what we mean by this is made clearer in Lemma 6.2.1 hereafter). This kind of result will lead to a formula relying (6.2.5) - hence the ATM implied volatility $\sigma(t,0)$ - to the Laplace transform of the quadratic variation of X. The motivation for this approach to the implied volatility stems from the fact that a closed form expression for the Laplace transform of $\langle X\rangle_t$ is available in several financial models (e.g. Heston). As soon as this Laplace transform is known as a function of model parameters, our representation formula converts into a semi-explicit formula for the ATM implied volatility (and we can obtain much more, cf. sections 6.2.2 and 6.2.4). In section 6.3, we work out all the related computations in the case of the Heston model.

Lemma 6.2.1. [2] *Let N denote the standard normal cdf, $N(d) = \int_{-\infty}^{d} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}dz$. Then, for every $I \geq 0$*

$$N\Big(-\frac{1}{2}\sqrt{I}\Big) = \frac{1}{4\sqrt{2\pi}}\int_0^\infty \frac{e^{-(z+\frac{1}{8})I}}{(z+\frac{1}{8})\sqrt{z}}dz. \tag{6.2.6}$$

Proof. Eq. (6.2.6) holds for $I = 0$, since the rhs is worth $\frac{1}{4\sqrt{2\pi}}\int_0^\infty \frac{1}{(z+\frac{1}{8})\sqrt{z}}dz = \frac{1}{4\sqrt{2\pi}}4\sqrt{2}\arctan(2\sqrt{2}z)|_{z=0}^{z\to\infty} = \frac{1}{2}$. Then, the derivative of the rhs with respect to I is

$$-\frac{1}{4\sqrt{2\pi}}\int_0^\infty \frac{e^{-(z+\frac{1}{8})I}}{\sqrt{z}}dz = -\frac{e^{-\frac{I}{8}}}{4\sqrt{2\pi}}\int_0^\infty \frac{e^{-zI}}{\sqrt{z}}dz = -\frac{e^{-\frac{I}{8}}}{4\sqrt{2\pi I}}$$

which is integrable on $(0,\infty)$ and equal to $\partial_I N(-\frac{1}{2}\sqrt{I})$. □

Remark 6.2.1. It is easy to check that the function $I \to N(-\frac{1}{2}\sqrt{I})$ is completely monotonic (cf. [53] for the definition of complete monotonicity). Lemma 6.2.1 shows that this function is indeed the (shifted) Laplace transform of $\frac{1}{(z+\frac{1}{8})\sqrt{z}}$.

We plug (6.2.6) into the argument of N^{-1} in (6.2.5) and, exchanging the order of integrations by means of Fubini's theorem, we obtain

$$\mathbb{P}(X_t > 0) = \frac{1}{4\sqrt{2\pi}}\int_0^\infty \frac{\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]}{(z+\frac{1}{8})\sqrt{z}}dz, \quad t > 0. \tag{6.2.7}$$

Remark 6.2.2. Writing $\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big] = 1-(1-\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big])$ and using again $\frac{1}{4\sqrt{2\pi}}\int_0^\infty \frac{1}{(z+\frac{1}{8})\sqrt{z}}dz =$

[2]This useful representation of the normal distribution is due to C. Martini.

$\frac{1}{2}$, Eq (6.2.7) can be rewritten as

$$\mathbb{P}(X_t > 0) = \frac{1}{2} - \frac{1}{4\sqrt{2}\pi} \int_0^\infty \frac{1 - \mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]}{(z+\frac{1}{8})\sqrt{z}} dz,$$

showing that $\mathbb{P}(X_t > 0) < \frac{1}{2}$ for all $t > 0$. In turn, this implies $N^{-1}(\mathbb{P}(X_t > 0)) < 0$.

Eq (6.2.7) allows us to make our first statement on the term structure of implied volatility.

Proposition 6.2.1. *Under Assumption 1, the ATM implied volatility $\sigma(t, 0)$ satisfies*

$$\sigma(t,0) = -\frac{2}{\sqrt{t}} N^{-1}\Big(\frac{1}{4\sqrt{2}\pi} \int_0^\infty \frac{\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]}{(z+\frac{1}{8})\sqrt{z}} dz\Big), \quad t > 0. \tag{6.2.8}$$

Proof. Done above. □

6.2.2 A representation formula for the ATM probability density and the implied volatility curvature

There exists a well known relation between the second derivative of the implied volatility and the density of the log forward price X_t. In this section, we show how one can obtain a representation analogous to (6.2.8) for the ATM probability density $p_t(0)$: using this representation and the one obtained for the ATM implied volatility, we will deduce a similar formula for the ATM curvature $\partial_x^2 \sigma(t,x)^2|_{x=0}$ (resp. $\partial_x^2 \sigma(t,x)|_{x=0}$) of implied variance (resp. implied volatility).

We first recall the relationship between the density and the ATM curvature: differentiating with respect to x once more in (6.2.1) yields

$$-p_t(x) = N'(d_2)\big(\partial_x d_2 + \partial_\omega d_2 \cdot \partial_x \omega(t,x)\big) - \frac{N'(d_2)}{2\sqrt{\omega(t,x)}} \partial_x^2 \omega(t,x) - \frac{d}{dx}\Big(\frac{N'(d_2)}{2\sqrt{\omega(t,x)}}\Big)\partial_x \omega(t,x)$$

hence, taking $x = 0$ and using $\partial_x \omega(t,x)|_{x=0} = 0$,

$$p_t(0) = N'(d_2|_{x=0})\Big(-\partial_x d_2|_{x=0} + \frac{\partial_x^2 \omega(t,x)|_{x=0}}{2\sqrt{\omega(t,0)}}\Big).$$

Now using $d_2|_{x=0} = -\frac{\sqrt{\omega(t,0)}}{2}$, $\partial_x d_2|_{x=0} = -\frac{1}{\sqrt{\omega(t,0)}}$, $N'(d) = \frac{e^{-\frac{d^2}{2}}}{\sqrt{2\pi}}$, we get

$$p_t(0) = \frac{e^{-\frac{\omega(t,0)}{8}}}{\sqrt{2\pi\omega(t,0)}}\big(1 + \partial_x^2 \omega(t,x)|_{x=0}\big)$$

that is

$$\partial_x^2 \omega(t,x)|_{x=0} = 2\big(\sqrt{2\pi\omega(t,0)}e^{\frac{\omega(t,0)}{8}} p_t(0) - 1\big)$$

or equivalently, in terms of implied volatility $\sigma(t,x)$

$$\partial_x^2 \sigma(t,x)^2|_{x=0} = \frac{2}{t}\big(\sigma(t,0)\sqrt{2\pi t}e^{\frac{t\sigma^2(t,0)}{8}} p_t(0) - 1\big). \tag{6.2.9}$$

We remark that eq. (6.2.9) is actually written for the ATM curvature of implied variance $\sigma(t,x)^2$. Of course from (6.2.9) and (6.2.8) one can deduce the corresponding expression for the ATM curvature of implied volatility $\partial_x^2\sigma(t,x)|_{x=0} = \frac{\partial_x^2\sigma(t,x)^2|_{x=0}}{2\sigma(t,0)}$. Since in the next section we are mainly interested in building SVI-type approximations for the implied variance, we keep (6.2.9) and do not switch to implied volatility.

Let us now focus on $p_t(0)$. By (6.2.3) one has

$$p_t(0) = \frac{1}{\sqrt{2\pi}}\mathbb{E}\Big[\frac{1}{\sqrt{\langle X\rangle_t}} e^{-\frac{\langle X\rangle_t}{8}}\Big], \tag{6.2.10}$$

and the derivation of a formula analogous to (6.2.8) will be based on the same ingredient:

Lemma 6.2.2. *For every $I > 0$,*

$$\frac{1}{\sqrt{I}} = \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{e^{-zI}}{\sqrt{z}}dz. \tag{6.2.11}$$

Proof. By direct computation of the integral. □

Proposition 6.2.2. *Under Assumption 1, the density p_t of the log forward price X_t satisfies*

$$p_t(0) = \frac{1}{\pi\sqrt{2}}\int_0^\infty \frac{\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]}{\sqrt{z}}dz, \quad t > 0. \tag{6.2.12}$$

Moreover, the ATM curvature of the implied variance is given by

$$\partial_x^2 \sigma(t,x)^2|_{x=0} = \frac{2}{t}\Big(\sigma(t,0)\sqrt{2\pi t}e^{\frac{t\sigma^2(t,0)}{8}} p_t(0) - 1\Big). \tag{6.2.13}$$

Proof. For (6.2.12), plug (6.2.11) into (6.2.10) and exchange the order of integrations. Eq. (6.2.13) is (6.2.9). □

6.2.3 The optimal time-dependent SVI approximation

Recall Gatheral's SVI parameterisation of implied variance (6.1.3). In order for this parameterisation to fulfill Assumption 1, we must have $r = m = 0$, hence (6.1.3) simplifies to

$$\sigma_{SVI}(x)^2 = a + b\sqrt{x^2+\gamma^2}. \tag{6.2.14}$$

We recall that (6.2.14) holds for fixed maturity. Our main concern is how to account for term structure of implied variance in (6.2.14). Of course this can be done introducing time-dependent parameters,

$$\sigma_{SVI}(t,x)^2 = a(t) + b(t)\sqrt{x^2 + \gamma(t)^2}, \tag{6.2.15}$$

and the issue becomes how to choose $a(t), b(t), \gamma(t)$ so that the time-dependence is both smooth enough and suitable for the description of market features.

We answer the question of how to build a parameterisation within the class (6.2.15) that best matches the implied variance of a symmetric model. The construction of the time-dependent parameters will be based on the representation formulae (6.2.8) and (6.2.12). Since these formulae describe the ATM term structure (implied variance level and curvature), we have to add an ingredient that accounts for the behaviour far from the money. This is easily done observing that the parameter $b(t)$ in (6.2.15) gives the asymptotic slopes of $x \to \sigma_{SVI}(t,x)^2$, in the sense that

$$\lim_{x\to\pm\infty} \frac{\sigma_{SVI}(t,x)^2}{|x|} = b(t).$$

Hence, $b(t)$ can be related to the asymptotic slopes of total implied variance

$$\limsup_{x\to\pm\infty} \frac{t\,\sigma(t,x)^2}{|x|} = \beta(t) \tag{6.2.16}$$

which are provided by the moment formula (3.1.2) in a model-independent framework. Using these elements, we can make our SVI-type parameterisation of the implied variance explicit.

Proposition 6.2.3. *Let $\sigma(t,x)^2$ denote the implied variance under Assumption 1. Let $\beta(t)$ be given by (6.2.16). Then, setting*

$$a(t) = \sigma(t,0)^2 - \frac{\beta(t)^2}{t^2\partial_x^2\sigma(t,x)^2|_{x=0}}; \qquad b(t) = \frac{\beta(t)}{t}; \qquad \gamma(t) = \frac{\beta(t)}{t\partial_x^2\sigma(t,x)^2|_{x=0}},$$

the parametric form

$$\sigma_{SVI}(t,x)^2 = a(t) + b(t)\sqrt{x^2 + \gamma(t)^2} \tag{6.2.17}$$

satisfies the following "matching" properties:

$$\sigma_{SVI}(t,0)^2 = \sigma(t,0)^2; \tag{ATM level}$$

$$\partial_x^2\sigma_{SVI}(t,x)^2|_{x=0} = \partial_x^2\sigma(t,x)^2|_{x=0}; \tag{ATM curvature}$$

$$\lim_{x\to\pm\infty} \frac{t\,\sigma_{SVI}(t,x)^2}{|x|} = \lim_{x\to\pm\infty} \frac{t\,\sigma(t,x)^2}{|x|} = \beta(t). \tag{Asymptotic slopes}$$

Proof. The first and third equations are obvious. The second follows from the explicit computation of the second derivative of $x \to \sigma_{SVI}(t,x)^2$. $\square$

The performances of the parameterisation (6.2.17) will be illustrated in section 6.5 for the uncorrelated Heston model.

6.2.4 Extension to "away-from-the-money" probability distribution

Propositions 6.2.1 and 6.2.2 provide the time-dependence of the ATM implied volatility $\sigma(t,0)$ (equivalently, by (6.2.4), of $\mathbb{P}(X_t > 0)$) and of the density $p_t(0)$ on the very special ATM curve $x = 0$. It is actually possible to generalize the approach of the previous sections and to relate the density $p_t(x)$ and the cumulative distribution function $\mathbb{P}(X_t > x)$ at *any* point $x \in \mathbb{R}$ to the Laplace transform of $\langle X \rangle_t$, in such a way that the resulting representations avoid integration in the complex plane. We show hereafter how this can be done: as happens for Propositions 6.2.1 and 6.2.2, the fundamental ingredient is a Laplace-transform representation of the function giving the conditional density $p_t(\cdot|\langle X \rangle_t)$ in (6.2.3).

Lemma 6.2.3. *For every $I > 0$ and $x, y \in \mathbb{R}$, one has:*

$$\frac{1}{\sqrt{I}} e^{-\frac{(2y+I)^2}{8I}} = \frac{e^{-\frac{y}{2}}}{\sqrt{\pi}} \int_0^\infty \frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}} e^{-(z+\frac{1}{8})I} dz \tag{6.2.18}$$

and

$$\int_x^\infty \frac{1}{\sqrt{2\pi I}} e^{-\frac{(2y+I)^2}{8I}} dy = \frac{1}{4\sqrt{2}\pi} e^{-\frac{x}{2}} \int_0^\infty \frac{\cos(\sqrt{2z}x) - 2\sqrt{2z}\sin(\sqrt{2z}x)}{\sqrt{z}(z+\frac{1}{8})} e^{-(z+\frac{1}{8})I} dz. \tag{6.2.19}$$

Remark 6.2.3. Taking $y = 0$ in (6.2.18) and $x = 0$ in (6.2.19) one finds, respectively, (6.2.11) and (6.2.6).

Proof of Lemma 6.2.3. We start from the expression of Weber's integral ([46] p. 132, [22] Appendix)

$$\frac{b^\nu}{(2I)^{\nu+1}} e^{-\frac{b^2}{4I}} = \int_0^\infty e^{-u^2 I} J_\nu(bu) u^{\nu+1} du, \quad I > 0,\ b > 0,\ \nu > -1,$$

where J_ν is the ordinary Bessel function of the first kind of order ν. Taking $\nu = -\frac{1}{2}$ and using $J_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}}\cos(x), x \in \mathbb{R}$, we easily obtain

$$\begin{aligned} \frac{e^{-\frac{b^2}{4I}}}{\sqrt{I}} &= \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-u^2 I} \cos(bu) du \\ &= \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-zI} \frac{\cos(b\sqrt{z})}{\sqrt{z}} dz, \quad I > 0,\ b > 0. \end{aligned} \tag{6.2.20}$$

We now consider the lhs in (6.2.18) and, expanding the exponent $(2y+I)^2$ and using (6.2.20) with $b=\sqrt{2}y$, we get

$$\begin{aligned}\frac{1}{\sqrt{I}}e^{-\frac{(2y+I)^2}{8I}} &= e^{-\frac{y}{2}-\frac{I}{8}}\times\frac{e^{-\frac{y^2}{2I}}}{\sqrt{I}}\\ &= \frac{e^{-\frac{y}{2}}}{\sqrt{\pi}}\int_0^\infty \frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}}e^{-(z+\frac{1}{8})I}dz\end{aligned}$$

which is indeed(6.2.18).

Now integrating (6.2.18) with respect to y between x and ∞ and exchanging the integrations with respect to y and z (the integrability conditions are clearly satisfied) we get

$$\frac{1}{\sqrt{2\pi}}\int_x^\infty \frac{1}{\sqrt{I}}e^{-\frac{(2y+I)^2}{8I}}dy = \frac{1}{\pi\sqrt{2}}\int_0^\infty e^{-(z+\frac{1}{8})I}\int_x^\infty e^{-\frac{y}{2}}\frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}}dy\,dz. \tag{6.2.21}$$

By direct computation,

$$\int_x^\infty e^{-\frac{y}{2}}\frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}}dy = e^{-\frac{x}{2}}\frac{\cos(\sqrt{2z}x)-2\sqrt{2z}\sin(\sqrt{2z}x)}{4\sqrt{z}(z+\frac{1}{8})},$$

hence, making the substitution in (6.2.21), we finally get to

$$\frac{1}{\sqrt{2\pi}}\int_x^\infty \frac{1}{\sqrt{I}}e^{-\frac{(2y+I)^2}{8I}}dy = \frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}\int_0^\infty \frac{\cos(\sqrt{2z}x)-2\sqrt{2z}\sin(\sqrt{2z}x)}{\sqrt{z}(z+\frac{1}{8})}e^{-(z+\frac{1}{8})I}dz$$

which is (6.2.19). □

Let us go back to the the distribution of X_t. The consequence of Lemma 6.2.3 reads as follows.

Proposition 6.2.4. *Under Assumption 1, for every $t>0$ the density p_t and the cdf $\mathbb{P}(X_t>\cdot)$ of the log forward price X_t satisfy*

$$\begin{aligned} p_t(y) &= \frac{e^{-\frac{y}{2}}}{\pi\sqrt{2}}\int_0^\infty \frac{\cos\big(\sqrt{2z}y\big)}{\sqrt{z}}\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz,\\ \mathbb{P}(X_t>x) &= \frac{1}{4\sqrt{2\pi}}e^{-\frac{x}{2}}\int_0^\infty \frac{\cos(\sqrt{2z}x)-2\sqrt{2z}\sin(\sqrt{2z}x)}{\sqrt{z}(z+\frac{1}{8})}\mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz\end{aligned} \tag{6.2.22}$$

for every $x,y\in\mathbb{R}$.

Proof. Recall that $p_t(y)=\mathbb{E}[p_t(y|\langle X\rangle_t)]$ with $p_t(y|\langle X\rangle_t)$ given by (6.2.3). Using Lemma (6.2.3), eq. (6.2.18) we directly obtain the first claim in (6.2.22). For the cdf, repeatedly applying Fubini's theorem we have

$$\begin{aligned}
\mathbb{P}(X_t > x) &= \int_x^\infty p_t(y)dy = \int_x^\infty \mathbb{E}[p_t(y|\langle X\rangle_t)]dy \\
&= \mathbb{E}\Big[\int_x^\infty p_t(y|\langle X\rangle_t)\Big]dy \\
&= \mathbb{E}\Big[\int_x^\infty \frac{1}{\sqrt{2\pi\langle X\rangle_t}} e^{-\frac{(2y+\langle X\rangle_t)^2}{8\langle X\rangle_t}}\Big]dy \\
&= \frac{1}{4\sqrt{2}\pi} e^{-\frac{x}{2}} \int_0^\infty \frac{\cos(\sqrt{2z}x) - 2\sqrt{2z}\sin(\sqrt{2z}x)}{\sqrt{z}(z+\frac{1}{8})} \mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz
\end{aligned} \tag{6.2.23}$$

and we have applied Lemma 6.2.3, eq. (6.2.19), in the last step. □

6.2.5 A pricing formula for symmetric models

We end this section noticing that from Proposition 6.2.4 we can deduce a pricing formula for Vanillas in a general symmetric model.

Since F is a martingale, for any given $t > 0$ we can define a probability measure $\mathbb{P}^*$ (sometimes called the Share measure) setting $\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{F_t}{F_0}$. The property of geometric symmetry of F (see the discussion at the beginning of this section) reads

$$\mathbb{E}\Big[g\Big(\frac{F_t}{F_0}\Big)\Big] = \mathbb{E}\Big[\frac{F_t}{F_0} g\Big(\frac{F_0}{F_t}\Big)\Big] \tag{6.2.24}$$

for all bounded measurable g. Denoting $\mathbb{E}^*$ the expectation under $\mathbb{P}^*$, we have $\mathbb{E}\Big[g\Big(\frac{F_t}{F_0}\Big)\Big] = \mathbb{E}^*\Big[g\Big(\frac{F_0}{F_t}\Big)\Big]$ hence $\frac{F_0}{F_t}$ has the same law under $\mathbb{P}^*$ as $\frac{F_t}{F_0}$ under $\mathbb{P}$. In particular, for any $K > 0$, we have $\mathbb{E}[F_t 1_{\{F_t > K\}}] = F_0\mathbb{P}^*(F_t > K) = F_0\mathbb{P}^*(\frac{F_0}{F_t} < \frac{F_0}{K}) = F_0\mathbb{P}(\frac{F_t}{F_0} < \frac{F_0}{K}) = F_0\mathbb{P}(X_t < -x)$. Hence, denoting $C(t,x)$ the price of a European Call option maturing at t and struck at $K(x) = F_0 e^x$, we have

$$\begin{aligned}
C(t,x) &= \mathbb{E}\big[F_t 1_{\{F_t > K(x)\}}\big] - K(x)\mathbb{P}(F_t > K(x)) \\
&= F_0\mathbb{P}(X_t < -x) - K(x)\mathbb{P}(X_t > x) \\
&= F_0\big(1 - \mathbb{P}(X_t > -x) - e^x\mathbb{P}(X_t > x)\big).
\end{aligned} \tag{6.2.25}$$

This justifies the following

Proposition 6.2.5. *Under Assumption 1, the price of a European Call option maturing at t and struck at $K(x) = F_0 e^x$ is given by*

$$C(t,x) = F_0\Big(1 - \frac{1}{2\sqrt{2}\pi} e^{\frac{x}{2}} \int_0^\infty \frac{\cos(\sqrt{2z}x)}{\sqrt{z}(z+\frac{1}{8})} \mathbb{E}\big[e^{-(z+\frac{1}{8})\langle X\rangle_t}\big]dz.\Big) \tag{6.2.26}$$

Proof. Plug (6.2.22) in (6.2.25) and simplify. □

6.3 Uncorrelated Heston model

The uncorrelated Heston model (6.1.2) belongs to the class of symmetric models considered in section 6.2. In particular, according to (6.1.2) the process $X_t = \log(F_t/F_0)$ satisfies

$$X_t = -\frac{1}{2}\int_0^t V_s ds + \int_0^t \sqrt{V_s} dW_s, \quad t \geq 0,$$

so that $\langle X \rangle_t = \int_0^t V_s ds$ (the integrated instantaneous variance). The Laplace transform of the integrated variance $\int_0^t V_s ds$ is known in closed-form as a function of model parameters (and can be obtained relying on the affine properties of the variance process, cf. for example [57]). From now on we will assume $\kappa, \theta, \sigma > 0$ and will make use of the rescaling of model parameters considered in [60], setting

$$\overline{V}_0 = \frac{V_0}{\sigma}; \quad \alpha = \frac{\kappa}{\sigma}; \quad \psi = \frac{\kappa\theta}{\sigma^2}; \quad \bar{t} = \sigma t. \tag{6.3.1}$$

The next lemma recalls the well-known expression of the Laplace transform of $\int_0^t V_s ds$, taking advantage of the rescaling (6.3.1).

Lemma 6.3.1. *Let $(V_t; t \geq 0)$ be the unique strong solution to the second SDE in (6.1.2). Then, for every $t > 0$ and every $\lambda > 0$ one has*

$$\mathbb{E}\left[e^{-\lambda \int_0^t V_s ds}\right] = \Big(p_\alpha(a) + (1 - p_\alpha(a))e(t,a)\Big)^{-2\psi} \\ \times \exp\left(-\frac{\bar{t}\psi}{2}a - \frac{\overline{V}_0}{2}a\frac{p_\alpha(a)(1-\epsilon(t,a))}{p_\alpha(a) + (1-p_\alpha(a))\epsilon(t,a)}\right) \tag{6.3.2}$$

with

$$\epsilon(t,a) = \exp\Big(-\bar{t}\Big(\frac{a}{2} + \alpha\Big)\Big); \quad p_\alpha(a) = \frac{4\alpha + a}{4\alpha + 2a}; \quad a = \sqrt{8\lambda + 4\alpha^2} - 2\alpha.$$

Proof. Use the closed-form expression of $\mathbb{E}\left[e^{-\lambda \int_0^t V_s ds}\right]$ from [57] and compose with (6.3.1). □

All the results stated in section 6.2 apply to uncorrelated Heston: it is basically a matter of computation to employ (6.3.2) and of Propositions 6.2.1-6.2.5 to work out the corresponding semi-explicit formulae (i.e. formulae involving the integral of an explicit function) for the ATM volatility level and curvature. This is what we implement in the following subsection.

6.3.1 Semi-explicit formulae for implied volatility

Consider equations (6.2.8), (6.2.12) and (6.3.2). Having in mind the ease of implementation of the resulting formulae, we make the following subsequent changes of variables (recall that

$z > 0$):

$$a = \sqrt{8\Big(z + \frac{1}{8}\Big) + 4\alpha^2} - 2\alpha, \quad a \in [\sqrt{1+4\alpha^2} - 2\alpha, \infty) \tag{6.3.3}$$

$$v = \frac{a}{2\alpha} - v_0, \quad \text{with } v_0 = \sqrt{1 + \frac{1}{4\alpha^2}} - 1 > 0. \tag{6.3.4}$$

The substitution (6.3.3) comes into play naturally when considering (6.3.2), while (6.3.4) simplifies the dependence to the parameter α and shifts the integration back to $(0, \infty)$. After these substitutions and the proper simplifications (remark that we have $z = \frac{\alpha^2}{2} v(v + 2v_0 + 2)$), we get

$$\mathbb{E}[e^{-(z+\frac{1}{8})I_t}] = \tilde{h}(t, v), \tag{6.3.5}$$

with

$$\begin{aligned}
\tilde{h}(t, v) &= \tilde{P}(t, v)^{-2\psi} \exp\Big(-\alpha\psi\bar{t}(v_0 + v) - \alpha\overline{V}_0(v_0 + v)\frac{\tilde{P}(t, v) - \tilde{e}(t, v))}{\tilde{P}(t, v)}\Big);\\
\tilde{P}(t, v) &= \tilde{p}(v) + (1 - \tilde{p}(v))\tilde{e}(t, v);\\
\tilde{e}(t, v) &= \exp(-\alpha\bar{t}(v_0 + 1 + v));\\
\tilde{p}(v) &= \frac{v_0 + 2 + v}{2(v_0 + 1 + v)}; \qquad v_0 = \sqrt{1 + \frac{1}{4\alpha^2}} - 1.
\end{aligned} \tag{6.3.6}$$

Equations (6.2.7) and (6.2.12) become respectively

$$\mathbb{P}(X_t > 0) = \frac{1}{2\pi\alpha}\int_0^\infty \tilde{g}(v)\tilde{h}(t, v)dv$$

and

$$p_t(0) = \frac{\alpha}{\pi}\int_0^\infty \tilde{q}(v)h(t, v)dv,$$

with

$$\begin{aligned}
\tilde{g}(v) &= \frac{v_0 + 1 + v}{\sqrt{v}\sqrt{2(v_0 + 1) + v}\,(v_0 + v)(v_0 + 2 + v)};\\
\tilde{q}(v) &= \frac{v_0 + 1 + v}{\sqrt{v}\sqrt{2(v_0 + 1) + v}}.
\end{aligned} \tag{6.3.7}$$

Since $\frac{1}{2} < \tilde{p}(v) < 1$ and $0 < \tilde{e}(t, v) < 1$ for every $v, t > 0$, we observe that the function $\tilde{h}(t, v)$ is particularly well-behaved: in particular, $v \to \tilde{h}(t, v)$ is bounded for any value of t and exponentially decaying as $v \to \infty$. On the other hand, the functions $\tilde{g}$ and $\tilde{q}$ diverge as $\frac{1}{\sqrt{v}}$ at zero: to remove the singularity, we consider the additional change of variable

$$x = \frac{\sqrt{v}}{\sqrt{v}\ + 1}. \tag{6.3.8}$$

The substitution (6.3.8) allows at the same time to remove the singularities at zero and to squeeze the integration over the bounded interval $[0,1]$. Thus, we end up with the integral of *bounded* functions over the *fixed* interval $[0,1]$, a feature which is extremely convenient for numerical purposes. Here are our final formulae.

Proposition 6.3.1 (ATM formulae)**.** *In the uncorrelated Heston model (6.1.2), for all $t>0$ the ATM implied volatility $\sigma(t,0)$ and the ATM density $p_t(0)$ of the log forward price X_t satisfy*

$$\sigma(t,0) = -\frac{2}{\sqrt{t}}\, N^{-1}\Big(\frac{1}{\pi\alpha}\int_0^1 g(x)h(t,x)dx\Big), \tag{6.3.9}$$

respectively

$$p_t(0) = \frac{2\alpha}{\pi}\int_0^1 q(x)h(t,x)dx \tag{6.3.10}$$

with

$$\begin{aligned}
g(x) &= \frac{((v_0+1)\overline{x}^2+x^2)\overline{x}}{\sqrt{2(v_0+1)\overline{x}^2+x^2}\;(v_0\overline{x}^2+x^2)((v_0+2)\overline{x}^2+x^2)};\\
q(x) &= \frac{(v_0+1)\overline{x}^2+x^2}{\overline{x}^3\sqrt{2(v_0+1)\overline{x}^2+x^2}};\\
h(t,x) &= P(t,x)^{-2\psi}\exp\Big(-\alpha\psi\bar{t}\frac{v_0\overline{x}^2+x^2}{\overline{x}^2} - \alpha\overline{V}_0(v_0\overline{x}^2+x^2)\frac{P(t,x)-e(t,x)}{\overline{x}^2P(t,x)}\Big);\\
P(t,x) &= p(x)+(1-p(x))e(t,x); \qquad e(t,x) = \exp\Big(-\alpha\bar{t}\frac{(v_0+1)\overline{x}^2+x^2}{\overline{x}^2}\Big);\\
p(x) &= \frac{(v_0+2)\overline{x}^2+x^2}{2((v_0+1)\overline{x}^2+x^2)}; \qquad \overline{x}=(1-x); \qquad v_0 = \sqrt{1+\frac{1}{4\alpha^2}}-1.
\end{aligned}$$

The ATM curvature of the implied variance is given by

$$\partial_x^2\sigma(t,x)^2|_{x=0} = \frac{2}{t}\Big(\sigma(t,0)\sqrt{2\pi t}e^{\frac{t\sigma^2(t,0)}{8}}p_t(0)-1\Big). \tag{6.3.11}$$

Proof. Done above: the final claim is achieved performing the change of variable (6.3.8) in (6.3.6)-(6.3.7). □

Remark 6.3.1. Actually, with the change of variable (6.3.8) we have introduced a singularity in the function $q(x)$ as x tends to 1. But, unlike the singularity at zero, the divergence at $x=1$ is strongly dampened by the decreasing exponential in $h(t,x)$. Hence, as happens for (6.3.9), the integrand in (6.3.10) is smooth and bounded over $[0,1]$, too.

Prop 6.3.1 allows to compute the coefficients of the time-dependent SVI approximation given in Prop 6.2.3. Recall that the extreme log-moneyness slope of implied variance is given by formula (6.1.1) in terms of the critical moment u^*. In the Heston model, the critical moment can be computed as in Andersen and Piterbarg [2] or Keller-Ressel [41] (cf. also [60], Prop 11): these authors provide a closed-form formula for the explosion time

$t^*(u) = \sup\{t \geq 0 : \mathbb{E}[F_t^u] < \infty\}$, hence $u^*(t)$ is obtained by numerical inversion of the equation $t^*(u^*(t)) = t$. For completeness of the presentation, we restate here this well known result (in the uncorrelated case), proposing a small variant of the statements in [2] and [41].

Proposition 6.3.2. *Let $\beta(t) = \lim_{x\to\pm\infty} \frac{t\,\sigma(t,x)^2}{|x|}$, where $\sigma(t,x)$ is the implied volatility in the uncorrelated Heston model. Then*

$$\beta(t) = 4\alpha\Big(\sqrt{(1+\omega^*(t)^2) + \frac{1}{4\alpha^2}} - \sqrt{1+\omega^*(t)^2}\Big) \tag{6.3.12}$$

where $\omega^*(t)$ is the unique solution in $[\frac{\pi}{2t}, \frac{\pi}{t}]$ of $t^*(\omega^*(t)) = \frac{\alpha\sigma}{2}t$, with

$$t^*(\omega) = \frac{2\pi - \arccos\big(\frac{1-\omega^2}{1+\omega^2}\big)}{2\omega}. \tag{6.3.13}$$

This result can be proved without difficulty relying on affine principles as in [41] and performing an appropriate time rescaling of the involved Riccati equations. As it stands, the value of the asymptotic slopes given by (6.3.12) and (6.3.13) is not manifestly seen to coincide with Andersen and Piterbarg's and Keller-Ressel's one, but it can be checked that the values of $\beta(t)$ are actually the same. The advantage of the formulation in Prop (6.3.2) is to make the function t^* to be inversed independent of model parameters: as a consequence, the values of the root ω^* can be tabulated once for all, fastening the computations.

Let us come back to the SVI parameterisation. The functional form (6.2.17) we use to build our approximation of implied volatility clearly belongs the original SVI class [29]. As pointed out in the Introduction, very recently Friz and other authors [28] have shown that at finite maturities ($t < \infty$) the Heston smile is *not* exactly described by the SVI parameterisation and that an additional term should be added to SVI to account for the fine behaviour of the wings. Hence, to be consistent with Heston, we should take Friz et al.'s correction into account in our approximation: this would amount to an additional term of the form $2\beta_1(t)\beta_2(t)\sqrt{x}$. In our framework, it is not obvious how to implement such a correction term for *every* log-moneyness without crushing the ATM structure: a term proportional to $\sqrt{|x|}$ would of course not have the desired ATM convexity. Hence, even though the time-dependent SVI parameterisation would benefit from the introduction of such a term in the far wings, the close to the money structure would be affected and we retain the parametric form (6.2.17) as it stands. This works yet considerably well for reasonable ranges of model parameters, maturities and strikes (cf. the results in section 6.5).

Here we restate Prop 6.2.3 in the framework of uncorrelated Heston.

Proposition 6.3.3. *Let*

$$a(t) = \sigma(t,0)^2 - \frac{\beta(t)^2}{t^2\partial_x^2\sigma(t,x)^2|_{x=0}}; \qquad b(t) = \frac{\beta(t)}{t}; \qquad \gamma(t) = \frac{\beta(t)}{t\partial_x^2\sigma(t,x)^2|_{x=0}},$$

with $\sigma(t,0)$, $\partial_x^2\sigma(t,x)^2|_{x=0}$ *and* $\beta(t)$ *given by (6.3.9), (6.3.11) and (6.3.12) respectively. Then, the parametric form*

$$\sigma_{SVI}(t,x)^2 = a(t) + b(t)\sqrt{x^2+\gamma(t)^2} \tag{6.3.14}$$

has the same ATM level, ATM curvature and extreme log-moneyness slopes of the implied variance under the uncorrelated Heston model (6.1.2) (cf. the "matching" properties of Proposition 6.2.3).

The approximation (6.3.14) of the implied volatility in the uncorrelated Heston model is straightforward to implement and computationally cheap: once the ATM implied volatility and curvature and the asymptotic slopes of the smile have been computed for a given maturity t, the computation of the whole smile (i.e. $\sigma(t,x)$ for all the desired values of x) is instantaneous. The performances of this approximation for different strikes and maturities will be illustrated in section 6.5 by some numerical examples.

We close this section by making the statements of Proposition 6.2.4 specific to uncorrelated Heston.

Proposition 6.3.4. *In the uncorrelated Heston model (6.1.2), for all* $t>0$ *and* $k\in\mathbb{R}$ *the density* $p_t(k)$ *and the complementary cdf* $\mathbb{P}(X_t>k)$ *of the log forward price* X_t *are given by, respectively*

$$p_t(k) = \frac{2\alpha}{\pi}\int_0^1 q(x)c(k,x)h(t,x)dx, \tag{6.3.15}$$

$$\mathbb{P}(X_t>k) = \frac{e^{-k/2}}{\pi\alpha}\int_0^1 g(x)s(k,x)h(t,x)dx, \tag{6.3.16}$$

where

$$\begin{aligned}
c(k,x) &= \cos\Big(\frac{\alpha k x}{\overline{x}^2}r(x)\Big);\\
s(k,x) &= c(k,x) - 2\frac{\alpha x}{\overline{x}^2}r(x)\sin\Big(\frac{\alpha k x}{\overline{x}^2}r(x)\Big);\\
r(x) &= \sqrt{2(v_0+1)\overline{x}^2+x^2};\\
\overline{x} &= (1-x); \quad v_0 = \sqrt{1+\frac{1}{4\alpha^2}}-1
\end{aligned}$$

and the functions g,h,q *are the ones of Proposition 6.3.1.*

Proof. Use (6.3.5) and the changes of variables (6.3.3)-(6.3.4)-(6.3.8) in (6.2.22) and (6.2.26). □

As an immediate consequence (cf. Prop. 6.2.5), we obtain a formula for the price of a

European Call option in the uncorrelated Heston model:

$$C(t,k) = F_0 \mathbb{P}\Big(X_t < \log\frac{F_0}{K}\Big) - K\mathbb{P}\Big(X_t > \log\frac{K}{F_0}\Big)$$
$$= F_0\Big(1 - 2\frac{e^{\frac{k}{2}}}{\pi\alpha}\int_0^1 \frac{((v_0+1)\overline{x}^2 + x^2)\overline{x}\cos\big(\frac{\alpha k x}{\overline{x}^2} r(x)\big)}{\sqrt{2(v_0+1)\overline{x}^2+x^2}\,(v_0\overline{x}^2+x^2)((v_0+2)\overline{x}^2+x^2)} h(t,x)dx\Big), \tag{6.3.17}$$

where $C(t,k)$ is a Call maturing at t and struck at $K := F_0 e^k$ and the function $h(t,x)$ is given in Prop. 6.3.1.

Remark 6.3.2. It seems to us that the formulation of Call options prices in (6.3.17) is new, since setting $\rho = 0$ in the classical solutions for Call prices based on extended Fourier transforms (cf. [38], [48] and many others) is not sufficient to get rid of the complex-valued functions. Proposition 6.3.4 provides indeed a fully "real" (in the algebraic sense) approach to option pricing in the Heston model.

6.4 An application to a skewed smile

Since market smiles are skewed, the uncorrelated Heston model (6.1.2) is quite unlikely to fit a set of market option prices. Nevertheless, the SVI parameterisation (6.3.14) can be straightforwardly applied to a skewed smile, by considering a displaced model (cf. [59] and the subsequent literature). Let $(F_t, t \geq 0)$ be the forward price in a displaced Heston model:

$$\begin{cases} dF_t = (\beta F_t + (1-\beta)F_0)\sqrt{V_t}dW_t \\ \\ dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dZ_t, \end{cases} \tag{6.4.1}$$

where W, Z are independent Brownian motions and $\beta \in [0,1]$ is the displacement parameter introducing the asymmetry in the model ($\beta = 1$ corresponding to zero displacement). As is well known, option pricing under the model (6.4.1) boils down to option pricing in the standard Heston model (6.1.2). Indeed, let

$$\tilde{F}_t = \beta F_t + (1-\beta)F_0, \quad \tilde{V}_t = \beta^2 V_t, \tag{6.4.2}$$

then the couple $(\tilde{F}, \tilde{V})$ satisfies

$$\begin{cases} d\tilde{F}_t = \tilde{F}_t\sqrt{\tilde{V}_t}dW_t \\ \\ d\tilde{V}_t = \kappa(\tilde{\theta} - \tilde{V}_t)dt + \tilde{\sigma}\sqrt{\tilde{V}_t}dZ_t \\ \\ \tilde{V}_0 = \beta^2 V_0; \quad \tilde{\theta} = \beta^2\theta; \quad \tilde{\sigma} = \beta\sigma, \end{cases} \tag{6.4.3}$$

meaning that $(\tilde{F}, \tilde{V})$ is an uncorrelated Heston couple with "dislaced" parameters $\tilde{V}_0, \tilde{\theta}, \tilde{\sigma}$ (while F_0 and κ remain the same). The affine transformation of forward price and variance (6.4.2) translates into the following mapping for the price C^D of a Call option in the displaced model:

$$\begin{aligned} C^D(t,x) &:= \mathbb{E}[(F_t - F_0 e^x)^+] \\ &= \frac{1}{\beta}\mathbb{E}\big[\big(\tilde{F}_t - F_0(1-\beta+\beta e^x)\big)^+\big], \end{aligned} \tag{6.4.4}$$

hence the price of a Call on F_t is $1/\beta$ times the price of a Call on $\tilde{F}_t$, the new log-moneyness being $\log(1-\beta+\beta e^x)$.

We approximate the implied volatility of the "undisplaced" model (6.4.3) using the SVI parameterisation (6.3.14). Denoting $\sigma_{SVI}(t,x;V_0,k,\theta,\sigma)$ the SVI approximation of the implied variance in the uncorrelated Heston model with parameters $F_0, V_0 k, \theta, \sigma$ and (with obvious notation) $C_{BS}(t,x;\sigma)$ the price of a Black-Scholes Call, we have

$$C^D(t,x) \approx \frac{1}{\beta} C_{BS}\Big(t, \log(1-\beta+\beta e^x); \sigma_{SVI}\big(t, \log(1-\beta+\beta e^x); \beta^2 V_0, k, \beta^2\theta, \beta\sigma\big)\Big). \tag{6.4.5}$$

Equation (6.4.5) gives a fast and economical approach to the pricing of Vanillas in the *skewed* displaced model (6.4.1).

6.5 Numerical Results

In this section we show the validity and accuracy of the representation formulae obtained in section 6.3 for the implied volatility in uncorrelated Heston. At all the time, our benchmark Call option pricer is the usual semi-analytical pricer based on extended Fourier transforms (and the reference implied volatilities are obtained by numerical inversion of BS formula).

Table 6.1 compares the values of the implied volatility obtained by the standard pricing with the ones obtained with formula (6.3.9), for a set of Heston parameters. The integration in (6.3.9) is performed using a fixed-tolerance Gaussian quadrature. The two series are perfectly in line, proving the validity of formula (6.3.9). Table 6.2 analyses the performances of the time-dependent SVI parametric form (6.3.14): we refer to this model as to the "Optimal SVI" in the sense of the matching of the ATM and far from the money structure as stated in Proposition 6.3.3. Only the right half of the smile is considered, the left part being identical by symmetry. The errors on implied volatilities are considerably small for all maturities and log-moneynesses up to 0.15 and remain under reasonable limits for log-moneynesses between 0.15 and 0.25. Let us remark that the SVI approximation (6.3.14) performs particularly well when the maturity is large: this is consistent with the recent result by Gatheral & Jacquier [31] proving that the large time Heston smile does converges to a (properly parameterised) SVI model. For short maturities ($t = 1$ month) the approximation of the implied volatility is less precise, but we remark that the larger errors in Table 6.2 occur when the sensitivity of the option value with respect to the volatility is small and that in any case for short maturities one is most interested in log moneyness close to zero. The parameterisation (6.3.14) actually provides a very good approximation of the Heston implied volatility close to the money. Of course, the accuracy of formula (6.3.9) and of approximation (6.3.14) varies in the model parameter domain and the usual care is needed in applications; in particular, as the standard Heston pricers, the formula can be patched using Lewis' expansion for small volatility of volalitility [48].

Figure 6.1 shows some typical profiles of the integrands in (6.3.9) and (6.3.17). As already pointed out in section 6.3, the functions that are integrated in our semi-closed formulae are particularly well-behaved: in particular, they are bounded and smooth functions of the interval $[0, 1]$. Figure 6.2 is a sanity check of formula (6.3.17) (which we refer to as to the "Uncorrelated formula"): we compare with the Call option prices obtained by usual pricing, for three different maturities (1 month, $1y$, $3y$ from bottom to top). The uncorrelated formula (6.3.17) works as expected and shows perfect agreement with the standard pricing results: the advantage of this formulation with respect to the usual Fourier-based pricing is to contain real valued functions only. Finally, figure 6.3 displays some examples of skewed smiles produced by the displaced model (6.4.1) for different maturities (1 month, $1y$, $3y$ from top to bottom). The skewed smiles are obtained applying the mapping (6.4.4) to the Call option prices computed with the usual pricer ("semi-analytic" method) and using the SVI approximation of the implied variance as in (6.4.5).

	Maturity						
Pricer :	1 month	4 months	1y	2y	$3y$	5y	10y
Usual Fourier transforms	19.593	18.958	17.939	17.866	18.018	18.342	18.842
Semi-closed formula (6.3.9)	19.592	18.958	17.939	17.862	18.014	18.340	18.840
	Maturity						
Pricer :	1 month	4 months	1y	2y	$3y$	5y	10y
Usual Fourier transforms	27.448	26.002	22.657	21.076	20.441	20.045	20.197
Semi-closed formula (6.3.9)	27.446	26.001	22.652	21.071	20.436	20.044	20.195

Table 6.1: Values of the ATM implied volatility in the uncorrelated Heston model. The values obtained by standard pricing with Fourier transforms and the ones obtained with formula (6.3.9) are compared. Heston parameters are $V_0 = 0.04, \sigma = 0.5, \kappa = 1.0, \theta = 0.04$ (first series) and $V_0 = 0.08, \sigma = 0.8, \kappa = 0.5, \theta = 0.06$ (second series).

$t = 1$ **month**	log-moneyness						
Model:	0.0	0.025	0.05	0.10	0.15	0.20	0.25
FT pricer	2.2556	1.2472	0.6256	0.1225	0.0185	0.0023	0.0002
	19.59	19.68	19.92	20.71	21.68	22.70	23.72
Optimal SVI	2.2556	1.2470	0.6245	0.1194	0.0167	0.0023	0.0002
	19.59	19.68	19.91	20.60	21.41	22.24	23.06
$t = 1$ **year**	log- moneyness						
Model:	0.0	0.025	0.05	0.10	0.15	0.20	0.25
FT pricer	7.1461	6.0585	5.1162	3.6320	2.5830	1.8471	1.3286
	17.94	17.98	18.12	18.59	19.26	20.04	20.87
Optimal SVI	7.1461	6.0585	5.1152	3.6215	2.5519	1.7931	1.2570
	17.94	17.98	18.11	18.56	19.16	19.84	20.54
$t = 5$ **years**	log- moneyness						
Model:	0.0	0.025	0.05	0.10	0.15	0.20	0.25
FT pricer	16.2466	15.2217	14.2421	12.4230	10.8006	9.3480	8.0792
	18.34	18.35	18.37	18.43	18.56	18.70	18.89
Optimal SVI	16.2466	15.2217	14.2421	12.430	10.7930	9.3585	8.0802
	18.34	18.35	18.36	18.43	18.55	18.70	18.89

Table 6.2: Comparison of the values of the implied volatility in the uncorrelated Heston model obtained by usual pricing with Fourier transforms ("FT pricer") and with the time-dependent SVI approximation (6.3.14) ("Optimal SVI"). Heston parameters are $V_0 = 0.04, \sigma = 0.5, \kappa = 1.0, \theta = 0.04$. At each line, the first entry is the Call option price, the second entry the volatility.

6.6 Conclusions

We have obtained single-integral representation formulae for the ATM level and curvature of a symmetric smile generated by a continuous martingale. This result is based on the derivation of the corresponding representations of the law (cumulative distribution and density) of the log forward price X in terms of the Laplace transform of the quadratic variation of X. The resulting formulae take the form of the integral of a known *real-valued* function as soon as this Laplace transform in known in closed form, as e.g. in the uncorrelated Heston model. This result allows us to build time-dependent SVI parameterisations which match the ATM and extreme strike structure and are therefore optimal as SVI parameterisations of a symmetric volatility surface. This time-dependent SVI model provides a very good and computationally cheap approximation of the implied volatility in the uncorrelated Heston model close to the money. We have addressed how to apply this assembly to a skewed smile by considering a displaced model: a subject for future work is how to generalise this approach to parameterise the smiles generated by a correlated stochastic volatility model.

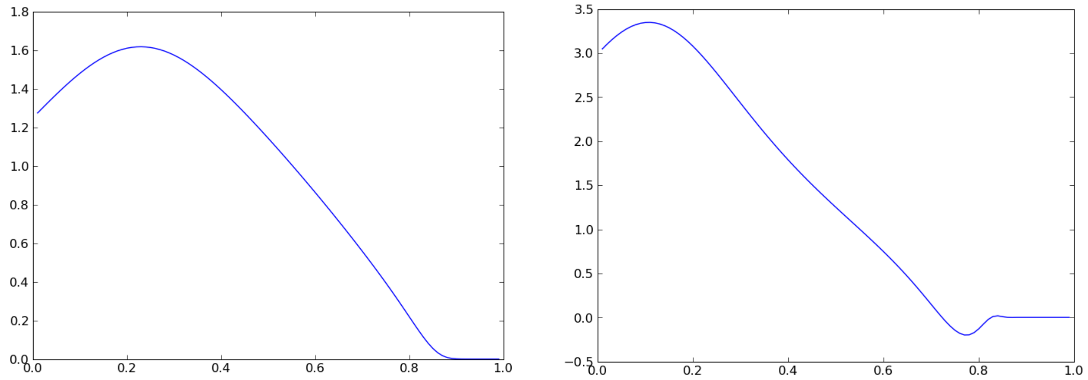

Figure 6.1: Profiles of the function to be integrated in formula (6.3.9) to obtain the 1 month ATM implied volatility (left) and in formula (6.3.17) to obtain the price of a 1 month, log-moneyness= 0.2 Call option (right) in the uncorrelated Heston model.

6.7 Proof of Proposition 6.3.2

The Heston model is an affine stochastic volatility model (ASVM), according to the definition of [41]. In particular, we have

$$\mathbb{E}\big[\exp(uX_t + vV_t)\big] = \exp(\phi(t,u,v) + uX_0 + \psi(t,u,v)V_0) \tag{6.7.1}$$

for every $(t,u,v) \in \mathbb{R}_+ \times \mathbb{R}^2$ (allowing for the case where both sides of the identity are infinite). The functions ϕ and ψ satisfy the system of Riccati equations

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

with

$$\begin{aligned} F(u,v) &= k\theta v \\ R(u,v) &= \frac{1}{2}u(u-1) + \frac{\sigma^2}{2}v^2 - kv + \rho\sigma uv. \end{aligned}$$

We aim here to compute the explosion time $t^*(u) = \sup\{t \geq 0 : \mathbb{E}[\exp(uX_t)] < \infty\}$, $u > 1$. Recalling that $X_0 = 0$, by (6.7.1) we have $\mathbb{E}[\exp(uX_t)] = \exp(\phi(t,u,0) + \psi(t,u,0)V_0)$. Since $\phi(t,u,v) = k\theta \int_0^t \psi(s,u,v)ds$, it follows that $t^*(u)$ is the explosion time of $\psi(\cdot,u,0)$, that is $t^*(u) = \sup\{t \geq 0 : \psi(t,u,0) < \infty\}$. Therefore we focus on the equation for ψ, in the uncorrelated case $\rho = 0$.

Proof of Proposition 6.3.2. Step 1 (Rescaling of the Riccati equation). Recalling the parameter rescaling (6.3.1), we have

$$R\big(u,\frac{v}{\sigma}\big) = \frac{1}{2}u(u-1) + v^2 - \alpha v = \frac{1}{2}(v-\alpha)^2 - \frac{1}{2}\big(\alpha^2 - u(u-1)\big)$$

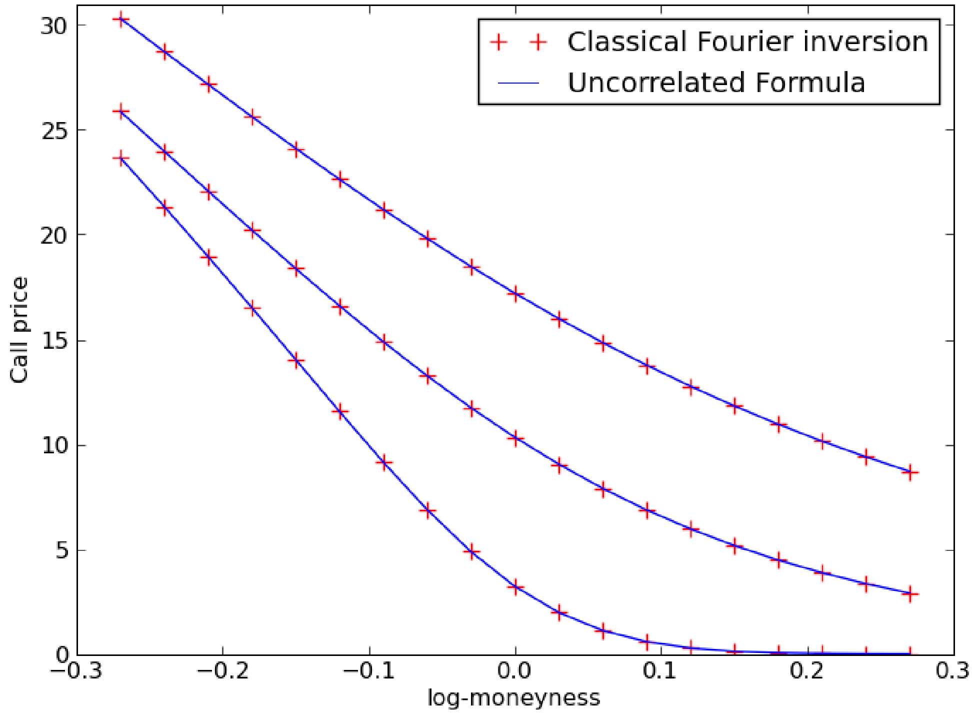

Figure 6.2: A "sanity check" of formula (6.3.17) (the "Uncorrelated formula") for Call option prices in the uncorrelated Heston model. Comparison with the usual pricing with Fourier transform is shown. From bottom curve to top curve, option maturities are 1 month, $1y$, $3y$. Heston parameters are $V_0 = 0.08, \sigma = 0.5, \kappa = 0.5, \theta = 0.08$.

Hence, defining

$$u(\omega) := \frac{1}{2} + \sqrt{\alpha^2(1+\omega^2) + \frac{1}{4}}, \quad \omega > 0; \qquad R_2(\omega, v) := \frac{2}{\alpha^2} R\big(u(\omega), \alpha\frac{v}{\sigma}\big)$$

we have $R_2(\omega, v) = (v-1)^2 + \omega^2$. It is easy to check that, for every $\omega > 0$, the solution to the Riccati equation

$$\partial_t \psi_2(t,\omega,v) = R_2(\omega, \psi_2(t,\omega,v)), \quad \psi_2(0,\omega,v) = \frac{\sigma}{\alpha} v \tag{6.7.2}$$

is such that

$$\psi(t, u(\omega), v) = \frac{\alpha}{\sigma}\psi_2\Big(\frac{\alpha\sigma t}{2}, \omega, v\Big) \tag{6.7.3}$$

for all t. We will show (Step 3) that for $u \le u(0)$, $\psi(t,u,0)$ is finite for all t. Hence it is sufficient to compute the explosion time of $\psi_2(\cdot,\omega,0)$: by (6.7.3), $\psi(\cdot,u(\omega),0)$ explodes at t iff $\psi_2(\cdot,\omega,0)$ explodes at $\frac{\alpha\sigma t}{2}$. Let us remark that the advantage of considering Eq (6.7.2) is the fact that ψ_2 depends on model parameters only through the initial condition.

Step 2 (Solution of the rescaled Riccati equation). For simplicity, we drop the dependence with respect to ω and write $\psi_2(t) := \psi_2(t,\omega,0)$. Let us denote $v^+ = 1 + i\omega$ and $v^- = 1 - i\omega$

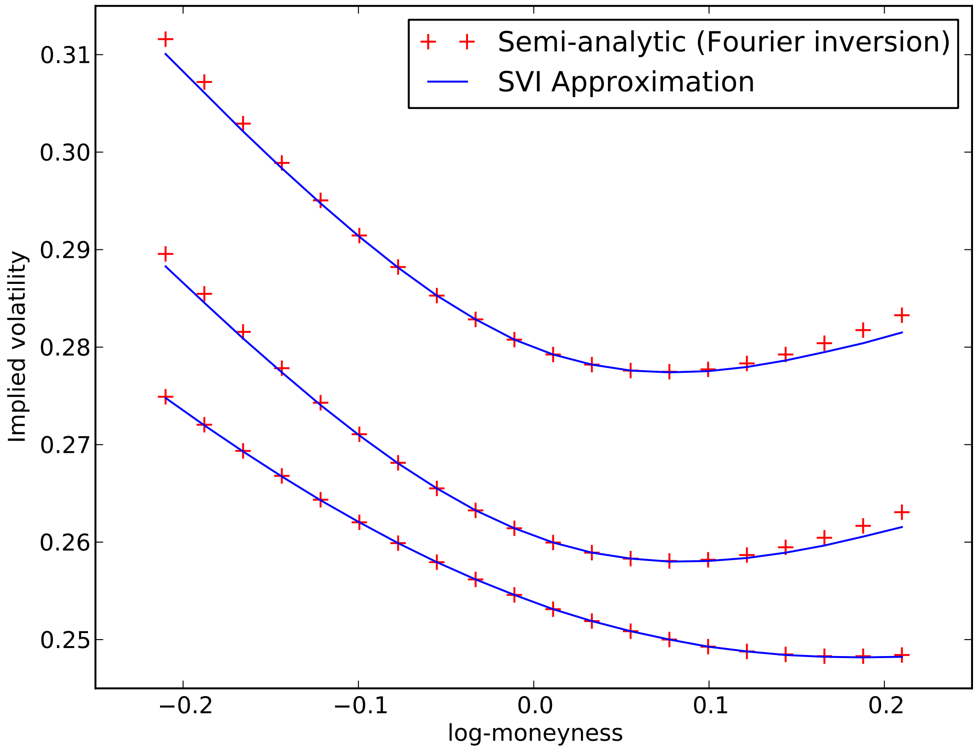

Figure 6.3: Skewed implied volatility in the displaced model (6.4.1): comparison of the values obtained by usual pricing with Fourier transforms ("Semi-analytic", plus sign) and the SVI approximation (6.3.14) (solid line). From top to bottom curve, option maturities are 1 month, $1y$, $3y$. Heston parameters are $V_0 = 0.08, \sigma = 0.5, \kappa = 0.5, \theta = 0.08$ and the displacement parameter is $\beta = 0.5$.

the two roots of $R_2(\omega, v)$, so that we have $R_2(\omega, v) = (v - v^+)(v - v^-)$. Let us also denote $\psi_2^{\pm}(t) = \psi_2(t) - v^{\pm}$. By (6.7.2), we have

$$\partial_t \psi_2^+(t) = \psi_2^+(t)\psi_2^-(t), \qquad \partial_t \psi_2^-(t) = \psi_2^-(t)\psi_2^+(t).$$

Then, introducing the initial condition $\psi_2^{\pm}(0) = \psi_2(0) - v^{\pm} = -v^{\pm}$,

$$\psi_2^+(t) = -v^+ \exp\Big(\int_0^t \psi_2^-(s)ds\Big)$$
$$\psi_2^-(t) = -v^- \exp\Big(\int_0^t \psi_2^+(s)ds\Big)$$

hence, taking ratios,

$$\frac{\psi_2(t) - v^+}{\psi_2(t) - v^-} = \frac{\psi_2^+(t)}{\psi_2^-(t)} = \frac{1 + i\omega}{1 - i\omega} \exp(2i\omega t).$$

It follows that $\psi_2(t)$ is finite valued until there is a t such that the lhs is worth 1, that is a t

such that

$$\exp(2i\omega t) = \frac{1 - i\omega}{1 + i\omega} = \frac{1 - \omega^2}{1 + \omega^2} - 2i\frac{\omega}{1 + \omega^2}. \tag{6.7.4}$$

As ω ranges from 0 to ∞, the lhs of this expression turns counterclockwise on the unit circle, while the rhs fills the lower half unit circle clockwise, from right to left. Hence, the smaller t which solves eq. (6.7.4) is such that $\pi < 2\omega t < 2\pi$ and that $\cos(2\omega t) = \frac{1-\omega^2}{1+\omega^2}$, that is

$$t^*(\omega) = \frac{2\pi - \arccos\left(\frac{1-\omega^2}{1+\omega^2}\right)}{2\omega}.$$

Recalling the time scaling (6.7.3), the critical exponent $u^*(t) = \inf\{u > 0 : \mathbb{E}[\exp(uX_t)] = \infty\} = \inf\{u > 0 : \psi(t, u, 0) = \infty\}$ is given by $u^*(t) = u(\omega^*(t))$, where $\omega^*(t)$ is the unique solution in $[\frac{\pi}{2t}, \frac{\pi}{t}]$ of $t^*(\omega^*(t)) = \frac{\alpha\sigma}{2}t$. Eq (6.3.12) now follows easily from the expression of $u(\omega)$ and from Lee's moment formula (6.1.1).

Step 3. Let us come back to $R(u, v)$: the two roots of $v \mapsto R(u, v)$ are

$$v^{\pm} = \frac{\alpha \pm \sqrt{\alpha^2 - u(u-1)}}{\sigma}.$$

If $0 < u < u(0) = \frac{1}{2} + \sqrt{\alpha^2 + \frac{1}{4}}$, then $\alpha^2 - u(u-1) > 0$ and v^+ and v^- are both real with $0 < v^- < v^+$. Then, the same arguments as in Step 2 yield

$$\frac{\psi(t, u, 0) - v^+}{\psi(t, u, 0) - v^-} = \frac{v^+}{v^-}\exp((v^+ - v^-)t)$$

and it is easy to see that this entails that $\psi(t, u, 0)$ is finite for all t. Finally, if $u = u(0)$, then $v^+ = v^- = \alpha/\sigma$ and $R(u(0), v) = \frac{\sigma^2}{2}\left(v - \frac{\alpha}{\sigma}\right)^2$. The equation $\partial_t\psi(t, u(0), 0) = \frac{\sigma^2}{2}\left(\psi(t, u(0), 0) - \frac{\alpha}{\sigma}\right)^2$ can be solved by separation of variables, giving

$$\psi(t, u(0), 0) = \frac{\alpha^2 t}{\alpha\sigma t + 2}$$

hence $\psi(t, u(0), 0)$ is finite for all t, too. $\square$

Bibliography

[1] E. Alos and C.-O. Ewald. Malliavin Differentiability of the Heston volatility and Applications to Option Pricing. *Advances in Applied Probability*, 40(1):144–162, 2008.

[2] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

[3] D.G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.

[4] V. Bally. Lower bounds for the density of locally elliptic Itô processes. *The Annals of Probability*, 34(6):2406–2440, 2006.

[5] V. Bally. Integration by parts formula for locally smooth laws and applications to equations with jumps I. Preprints Institut Mittag-Leffler, The Royal Swedish Academy of Sciences, 2007.

[6] V. Bally and L. Caramellino. Riesz transform and integration by parts formulas for random variables. Preprint arXiv:0911.2631, 2010.

[7] V. Bally and S. De Marco. Bounds on Stock Price probability distributions in general Local-Stochastic Volatility models. Submitted, available at arXiv:1006.3337v1, 2010.

[8] V. Bally, B. Fernández, and A. Meda. Estimates for the probability that Itô processes remain around a curve, and applications to finance. To appear in Stochastic Processes and Applications, 2009.

[9] S. Benaim and P. Friz. Regular variation and smile asymptotics. *Mathematical Finance*, 19(1):1–12, 2006.

[10] H. Berestycki, J. Busca, and I. Florent. Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics*, LVII:1352–1373, 2004.

[11] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 2007.

[12] M. Bossy and A. Diop. An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha, \alpha \in [1/2, 1)$. Technical report INRIA RR-5396, 2004.

[13] P. Carr and R. Lee. Put-call symmetry : Extensions and applications. *Mathematical Finance*, 19(4):523–560, 2009.

[14] P. Carr and D. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 3:463–520, 1999.

[15] J. Cox. Notes on option pricing I: Constant elasticity of variance diffusions. Working paper, Stanford University, reprinted in *Journal of Portfolio Management* 22:15-17, 1996.

[16] J. Cox, J. Ingersoll, and S. Ross. A Theory of the Term structure of Interest Rates. *Econometrica*, 53:385–407, 1985.

[17] R. C. Dalang and E. Nualart. Potential theory for hyperbolic spdes. *The Annals of Probability*, 32(3A):2099–2148, 2004.

[18] D. Davydov and V. Linetsky. The valuation and hedging of barrier and lookback options under the CEV process. *Management Science*, 47:949–965, 2001.

[19] S. De Marco. Smoothness and Asymptotic Estimates of densities for SDEs with locally smooth coefficients and Applications to square root-type diffusions, 2009. Forthcoming in The Annals of Applied Probability.

[20] S. De Marco and C. Martini. The term structure of implied volatility in symmetric models, with applications to Heston. Submitted, available at SSRN: http://ssrn.com/abstract=1622828, 2010.

[21] A. A. Dragulescu and V. M. Yakovenko. Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance*, 2:443–453, 2002.

[22] D. Dufresne. The integrated square-root process. Research Paper no. 90, Centre for Actuarial Studies, University of Melbourne, 2001.

[23] M. Forde. Tail asymptotics for diffusion processes, with applications to local volatility and CEV-Heston models. Working Paper. Available at arXiv:math/0608634v6, 2008.

[24] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under a general local-stochastic volatility model. Working Paper. Available at SSRN: http://ssrn.com/abstract=1514125, 2009.

[25] M. Forde and A. Jacquier. Small-time asymptotics for implied volatility under the Heston model. *International Journal of Theoretical and Applied Finance*, 2:861–876, 2009.

[26] M. Forde, A. Jacquier, and A. Mijatovic. Asymptotic formulae for implied volatility under the Heston model. Forthcoming in Proceedings of the Royal Society A, 2009.

[27] N. Fournier. Smoothness of the law of some one-dimensional jumping S.D.E.s with non-constant rate of jump. *Electronic Journal of Probability*, 13, 2008.

[28] P. Friz, S. Gerhold, A. Gulisashvili, and S. Sturm. On refined volatility smile expansion in the Heston model. To appear in Quantitative Finance, 2010.

[29] J. Gatheral. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. Presentation at Global Derivatives & Risk Management, Madrid, May 2004.

[30] J. Gatheral. The volatility surface: A practitioner's guide. *Wiley Finance*, 2006.

[31] J. Gatheral and A. Jacquier. Convergence of Heston to SVI. Working Paper. Available at SSRN: http://ssrn.com/abstract=1555251, 2010.

[32] H. Guérin, S. Méléard, and E. Nualart. Estimates for the density of a nonlinear Landau process. *Journal of Functional Analysis*, 238(2):649–677, 2006.

[33] A. Gulisashvili and E. M. Stein. Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics and Optimization*, 61(3):287–315, 2010.

[34] P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, 2002.

[35] P. Hagan, A. Lesniewski, and D. Woodward. Probability distribution in the SABR model of stochastic volatility. Working Paper. Available at http://lesniewski.us/papers/working/ProbDistrForSABR.pdf, 2005.

[36] P. Henry-Labordère. Analysis, geometry, and modeling in finance: Advanced methods in option pricing. Chapman & Hall/CRC, Financial Mathematics Series, 2008.

[37] P. Henry-Labordère. Calibration of local stochastic volatility models to market smiles. *Risk*, September 2009.

[38] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[39] N. Ikeda and S. Watanabe. Stochastic Differential Equations and Diffusion Processes. *North-Holland/Kodansha*, 1989.

[40] I. Karatzas and S.E. Shreve. Brownian Motion and Stochastic Calculus, second edition. *Springer-Verlag*, 1991.

[41] M. Keller-Ressel. Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21:73–98, 2011. Available at SSRN: http://ssrn.com/abstract=1732470.

[42] A. Kohatsu-Higa. Lower bounds for densities of uniformly elliptic random variables on Wiener space. *Probab. Theory and Related Fields*, 126:421–457, 2003.

[43] S. Kusuoka and D. Stroock. Application of the Malliavin calculus, Part I. *Stochastic analysis (Katata/Kyoto, 1982), North-Holland*, pages 271–306, 1984.

[44] S. Kusuoka and D. Stroock. Application of the Malliavin calculus, Part II. *Journal of the Faculty of Science Univ. Tokyo Sect IA Math*, pages 1–76, 1985.

[45] D. Lamberton and B. Lapeyre. Introduction to Stochastic Calculus Applied to Finance. *Chapman & Hall*, 1997.

[46] N. N. Lebedev. Special functions and their applications. *Dover, New York*, 1972.

[47] R. W. Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14:469–480, 2004.

[48] A. L. Lewis. Option valuation under stochastic volatility, with Mathematica code. *Finance press*, 2005.

[49] P.-L. Lions and M. Musiela. Correlations and bounds for stochastic volatility models. *Annales de l'Institut H. Poincaré*, 24:1–16, 2007.

[50] A. Lipton. The vol smile problem. *Risk*, February 2002.

[51] P. Malliavin. Stochastic Calculus of Variation and hypoelliptic operators. *Proc. Int. Symp on Stoch. Diff. Equations, Kyoto 1976*, pages 195–263, 1978.

[52] A. Medvedev and O. Scaillet. A simple calibration procedure of stochastic volatility models with jumps by short term asymptotic. FAME Research Paper Series rp93, International Center for Financial Asset Management and Engineering, 2004.

[53] K. S. Miller and S. G. Samko. Completely monotonic functions. *Integr. Transf. and Spec. Funct.*, 12(4):389–402, 2001.

[54] D. Nualart. The Malliavin calculus and Related Topics. *Springer-Verlag*, 1995.

[55] V. Piterbarg. A stochastic volatility forward LIBOR model with a term structure of volatility smiles. *Applied Mathematical Finance*, 12:147–185, 2003.

[56] E. Renault and Nizar Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, 6(3):279–302, 1996.

[57] D. Revuz and M. Yor. Continuous Martingales and Brownian motion, third edition. *Springer*, 1999.

[58] L. C. G. Rogers and M. R. Tehranchi. Can the implied volatility surface move by parallel shifts? *Finance and Stochastics*, 14(2):235–248.

[59] M. Rubinstein. Displaced diffusion option pricing. *Journal of Finance*, 38:125–144, 1983.

[60] Zeliade Systems. Heston 2009. Zeliade Systems White Paper. Available at http://www.zeliade.com/whitepapers/zwp-0004.pdf, 2009.

[61] M. R. Tehranchi. Asymptotics of implied volatility far from maturity. *Journal of Applied Probability*, 46(3):629–650, 2009.

[62] M. R. Tehranchi. Symmetric martingales and symmetric smiles. *Stochastic Processes and Their Applications*, 119(10):3785–3797, 2009.

[63] F. Viens and A. Vizcarra. Supremum concentration inequality and modulus of continuity for sub-nth chaos processes. *Journal of Functional Analysis*, 248:1–26, 2007.