The SVI implied volatility model and its calibration

ALEXANDER AURELL

Master of Science Thesis
Stockholm, Sweden 2014

The SVI implied volatility model and its calibration

A L E X A N D E R A U R E L L

Master's Thesis in Mathematical Statistics (30 ECTS credits)
Master Programme in Mathematics (120 credits)
Royal Institute of Technology year 2014
Supervisor at ORC were Jonas Hägg and Pierre Bäcklund
Supervisor at KTH was Boualem Djehiche
Examiner was Boualem Djehiche

TRITA-MAT-E 2014:53
ISRN-KTH/MAT/E--14/53--SE

Royal Institute of Technology
School of Engineering Sciences

KTH SCI
SE-100 44 Stockholm, Sweden

URL: www.kth.se/sci

KUNGLIGA TEKNISKA HÖGSKOLAN

Abstract

Division of Mathematical Statistics
School of Engineering Sciences

Master of Science

The SVI implied volatility model and its calibration

by Alexander Aurell

The SVI implied volatility model is a parametric model for stochastic implied volatility. The SVI is interesting because of the possibility to state explicit conditions on its parameters so that the model does not generate prices where static arbitrage opportunities can occur. Calibration of the SVI model to real market data requires non-linear optimization algorithms and can be quite time consuming. In recent years, methods to calibrate the SVI model that use its inherent structure to reduce the dimensions of the optimization problem have been invented in order to speed up the calibration.

The first aim of this thesis is to justify the use of the model and the no static arbitrage conditions from a theoretic point of view. Important theorems by Kellerer and Lee and their proofs are discussed in detail and the conditions are carefully derived. The second aim is to implement the model so that it can be calibrated to real market implied volatility data. A calibration method is presented and the outcome of two numerical experiments validate it.

The performance of the calibration method introduced in this thesis is measured in how big a fraction of the total market volume the method manages to fit within the market spread. Tests show that the model manages to fit most of the market volume inside the spread, even for options with short time to maturity.

Further tests show that the model is capable to recalibrate an SVI parameter set that allows for static arbitrage opportunities into an SVI parameter set that does not.

Key words: SVI, stochastic implied volatility, static arbitrage, parameter calibration, Kellerer's theorem, Lee's moment formula.

Acknowledgements

I would like to thank my supervisours at ORC, Jonas Hägg and Pierre Bäcklund, for introducing me to the topic. Throughout the project you have actively discussed my progress and results and you have often inspired me to come up with new ideas and approaches. Your constant support has been invaluable to me.
I would like to thank my advisor at KTH, Professor Boualem Djehiche, for great feedback, academic guidance and patience.
Thank you, Ty Lewis at ORC, for your many valuable inputs about the world of finance and thank you, Tove Odland at KTH, for sharing your research and opinions on the Nelder-Mead method with me.
Furthermore I would like to thank Fredrik Lannsjö and Helena E. Menzing for proofreading my thesis and commenting on the content.
Finally, I am most grateful for the support from Helena and my family during my entire studies.

Alexander Aurell
Stockholm, September 2014

Contents

List of Figures

Chapter 1

Introduction

This introduction will briefly state the background, purpose and delimitations of the work done in this thesis. Also a summary of all the chapters is given.

1.1 Background

When pricing financial contracts such as options it is common practice to use the Black-Scholes framework. Black-Scholes assumes that options with all parameters equal, except the strike price, are to be priced with the same implied volatility parameter value. This however stands in contradiction with the real world where market prices imply that the volatility depends on the strike price. One way that practioners handle this problem is to create implied volatility surfaces. An implied volatility surface is a function,

$$(\text{Time to maturity}, \text{Strike}) \mapsto \text{Implied volatility}(\text{Time to maturity}, \text{Strike}).$$

The work flow is basically as follows:

(1) Provided by option prices from the market for a range of strikes and maturities one gets the corresponding implied volatilities by inverting the Black-Scholes price function.

(2) Use the result from (1) to create an implied volatility surface for all (Strike, Maturity)-points.

(3) When pricing an option with a given strike and maturity get the implied volatility to use from the surface.

There are several popular models that are used for the surface construction in (2). The stochastic volatility inspired, or SVI, model of the implied volatility surface was originally created at Merrill Lynch in 1999 and was introduced to the public in the presentation [1]. The model has two key properties that are often stated in the literature that followed [1] as reasons for its popularity amongst practitioners. It satisfies Lee's Moment Formula, a model free result that specifies the asymptotics for implied volatility. Therefore, the SVI model is valid for extrapolation far outside the avaliable data. Furthermore, it is stated that the SVI model is relatively easy to calibrate to market data so that the corresponding implied volatility surface is free of calendar spread arbitrage. The recent development of the SVI model has been towards conditions guaranteing the abscence of butterfly arbitrage. In [2] this problem is solved by restricting the parameters in the SVI model.

1.2 Purpose of the thesis

The purpose of this thesis is to motivate the usage of the SVI model from a theoretical point of view, and implement the SVI model so that a parametrized implied volatility surface can be fitted to market data. Furthermore, the model should be able to detect static arbitrage and eliminate it by a recalibration. The thesis aims to give thorough explanations of the underlying theoretical results, and do a complete derivation of the no static arbitrage conditions. It also aims to in detail present a calibration method for the SVI parameters to real market implied volatility data and evaluate its accuracy.

1.3 Outline of the thesis

This thesis is divided into five chapters. Chapter 1 introduces the topic of this thesis, states the purpose of it, summarizes it and states the delimitations that were done. In Chapter 2, two underlying theoretical results are presented. Conditions on the call prices that guarantees absence of static arbitrage is derived using Kellerer's theorem and these conditions are translated into conditions on the implied volatility surface. Lee's Moment Formula is recalled and its implications are discussed. In Chapter 3, the SVI parameterization for implied volatility and variations of it are presented. A concrete method for eliminating static arbitrage in the implied volatility smile is constructed. In Chapter 4, the calibration method used to fit the SVI model to market data is described in detail. The final chapter, Chapter 5, numerical results of the calibration method are presented.

1.4 Delimitations

Some proofs are omitted and the reason for this is either their extensive length or that they are unrelevant to the surrounding context. If a proof is omitted, a reference is given to a complete proof.
Calendar spread arbitrage is only treated in theory. Sufficient conditions for the elimination of calendar spread arbitrage are derived together with the sufficient conditions for elimination of butterfly arbitrage, but no numerical experiments are done trying to eliminate calendar spread arbitrage from SVI surfaces fitted to real market data.
Numerous delimitations are made in the last chapter, mainly because of the lack of time. The speed of the calibration is not investigated. Only one set of weights are used in the optimization and only one type of options, call and put options on the S&P 500, are used.
The error of the fit is quantified in two ways, the fraction of the total market that the SVI fitted inside the price spread and the distance from the SVI fitted prices to the price spread. These results are not compared to the results of other calibration methods.

Chapter 2

Kellerer's Theorem and Lee's Moment Formula

In this chapter, a theoretical approach to the elimination of static arbitrage in an implied volatility surface will be presented. It begins with setting up the model framework used in the rest of the thesis and then gives a short introduction to stochastic implied volatility. After that, sufficient conditions for the absence of static arbitrage on the surface of call prices,

$$(\text{Time to maturity}, \text{Strike}) \mapsto \text{Call price}(\text{Time to maturity}, \text{Strike}),$$

are derived through an application of Kellerer's Theorem. At every time to maturity the density of the underlying's price process martingale is matched with the density of a martingale that has the Markov property. The conditions on the call surface are then translated into conditions for the implied volatility surface.

Additionally, Lee's Moment Formula is examined. It is a model free result that states conditions on the asymptotes of implied volatility smiles. Iimplied volatility smile is the name in finance for a time to maturity-section of the implied volatility surface. For a fixed, positive t,

$$\text{Implied volatility smile}(\text{Strike}) = \text{Implied volatility}(\text{Time to maturity} = t, \text{Strike})$$

All models that extrapolate implied volatility generated from market data in the strike direction should satisfy Lee's Moment Formula.

2.1 Stochastic implied volatility

The model is set up in a probability space $(\Omega, (\mathcal{F}_t, t \geq 0), \mathbb{Q})$. The filtration $(\mathcal{F}_t, t \geq 0)$, indexed by the time t, is generated by a 2-dimensional Brownian motion (B^0, B^1) and $\mathbb{Q}$ is the measure under which the underlying's discounted price process is a martingale. There are $M + 2$ traded objects in the model, the underlying with price process $S(t) = S_t$, a set of M European call options with strike prices and times to maturity (K, T) written on S_t and a risk-free investment with constant, positive interest rate r.

The underlying price process is assumed to be represented by the dynamics

$$\mathrm{d}S_t = rS_t\mathrm{d}t + \sigma S_t \mathrm{d}B_t^0, \tag{2.1}$$

where σ is stochastic. The Black-Scholes price of a call option is given by the Black-Scholes formula stated in Equation (2.2) and the implied volatility. The implied volatility will be denoted by σ_{imp}. The name and notation emphesizes that it is implied from the Black-Scholes pricing formula. Hence, there is a difference between the volatility of the underlying's price process, σ from Equation (2.1), and the implied volatility, σ_{imp} from Equation (2.2). The model is not set up in a Black-Scholes world since σ is not assumed to be constant but depends on both strike price and time to maturity. Therefore, the Black-Scholes formula merely serves as a convenient tool to describe option prices. Note that σ_{imp} does not necessarily equal σ. An important note to bear in mind as motivation for using implied volatility, is that it is generally easier to observe the implied volatility on the market than it is to observe the volatility of the underlying's price process.

The formal definition of implied volatility of an option is the parameter σ_{imp} that gives the observed option price, C, when inserted into the Black-Scholes call price formula,

$$C^{\text{BS}}(\tau, K, \tau\sigma_{\text{imp}}^2; S, r, t) = S_t\mathcal{N}(d_1) - e^{-r\tau}K\mathcal{N}(d_2). \tag{2.2}$$

Here $\mathcal{N}(x)$ is the cumulative distribution function of a standard normally distributed random variable, t is the current time, T is the maturity time for the option, $\tau = T - t$ is the time to maturity for the option and d_1 and d_2 are the Black-Scholes auxiliary functions,

$$\begin{aligned} d_1(\tau, K, \tau\sigma_{\text{imp}}^2; S, r, t) &= \frac{\log(S_t/K) + \tau r + \tau\sigma_{\text{imp}}^2/2}{\sqrt{\tau\sigma_{\text{imp}}^2}}, \\ d_2(\tau, K, \tau\sigma_{\text{imp}}^2; S, r, t) &= \frac{\log(S_t/K) + \tau r - \tau\sigma_{\text{imp}}^2/2}{\sqrt{\tau\sigma_{\text{imp}}^2}}. \end{aligned} \tag{2.3}$$

The dependece on the variables S, r, and T is denoted behind a semicolon, separating them from τ, K and σ. This is done to clarify that in this thesis, they are of lesser interest than the variables τ, K and σ_{imp}. More than often, the dependence of the variables behind the semicolon will be surpressed. This paragraph is summarized in the following definition.

Definition 1. (Implied volatility) Let a call option be written on the underlying S at time t with strike price K and expiry time T. Let the observed market price for this option be C. The implied volatility of the option is the unique value of σ_{imp} that solves

$$C = C^{\text{BS}}(\tau, K, \tau\sigma^2_{\text{imp}}; S, r, t).$$

An alternative, but equivalent, definition of implied volatility can be stated by replacing the underlying's price process with the forward price.

Definition 2. The forward price process of the underlying S is

$$F(t, \tau) = F_{[t,t+\tau]} = e^{r\tau} S_t.$$

If using $F_{[t,t+\tau]}$ instead of S_t, the implied volatility of an option is the parameter σ_{imp} that gives the observed compounded option price, $Ce^{r\tau}$, when substituted into what is called the Black call price formula,

$$C^{\text{B}}(\tau, K, \tau\sigma^2_{\text{imp}}; F, r, t) = F_{[t,t+\tau]}\mathcal{N}(d_1) - K\mathcal{N}(d_2). \tag{2.4}$$

By using the forward price instead of the underlying's price process, the auxiliary functions in Equation (2.3) simplify to

$$\begin{aligned} d_1(\tau, K, \tau\sigma^2_{\text{imp}}; F, r, t) &= \frac{\log\left(F_{[t,t+\tau]}/K\right) + \frac{1}{2}\sigma^2_{\text{imp}}\tau}{\sqrt{\tau\sigma^2_{\text{imp}}}}, \\ d_2(\tau, K, \tau\sigma^2_{\text{imp}}; F, r, t) &= \frac{\log\left(F_{[t,t+\tau]}/K\right) - \frac{1}{2}\sigma^2_{\text{imp}}\tau}{\sqrt{\tau\sigma^2_{\text{imp}}}}. \end{aligned} \tag{2.5}$$

The Black call price formuka will be used instead of the Black-Scholes call price formula especially in Appendix C.

So far, all we have said about σ is that it is random. If this is the case, σ_{imp} should also be random. This is in accordance with observations of the real market, where volatility across strike and across time is not constant but behaves in a stochastic manner. Hence let the implied variance, σ^2_{imp}, for an option written on S with maturity T and strike K

have the following dynamics,

$$\mathrm{d}\sigma^2_{\mathrm{imp}}(\tau, K, \tau\sigma^2_{\mathrm{imp}}; S, r, t) = \alpha(\tau, K, \tau\sigma^2_{\mathrm{imp}}; S, r, t)\,\mathrm{d}t + \eta\beta(\tau, K, \tau\sigma^2_{\mathrm{imp}}; S, r, t)\,\mathrm{d}B^1_t, \quad (2.6)$$

where $\langle \mathrm{d}B^0_t, \mathrm{d}B^1_t \rangle = \rho\,\mathrm{d}t, \rho \in [-1, 1]$. In Chapter 8.6 in [3], it is derived that the Black-Scholes model with non-constant but deterministic implied volatility is retrieved in the limit $\eta \to 0$. This is not necessarily an advantage for the model performance, however it is almost a practical requirement as the Black-Scholes model is the core of the intuition of practitioners. The choice of specifying the dynamics of the implied variance instead of the implied volatility is made to follow the line of literature in [3] and [4].

Being able to model correlation between the underlying's price process and the corresponding implied volatility is necessary, as it can be observed on the real market. In Figure 2.1, historical data for the S&P 500 index and the VIX index are presented. The S&P 500 is a stock market index based on the performance of the 500 largest companies which are listed at the New York Stock Exchange or the NASDAQ, while the VIX index is a measure of the implied volatility of the S&P 500 index. In this figure it can be seen that when the S&P 500 suffers a severe drop, such as during the Russian crisis of 1998, the end of the IT-bubble in 2002, the U.S. housing bubble in 2008 and the European sovereign debt crisis in 2011, the VIX index rose dramatically. Heuristically, a negative correlation between the S&P 500 and the VIX can be established.

2.2 Static arbitrage

This section introduces the concept of static arbitrage and how it differs from dynamic arbitrage, the kind of arbitrage that is treated in The Fundamental Theorem of Asset Pricing.

Definition 3 (Dynamic arbitrage opportunity)**.** A dynamic arbitrage opportunity is a costless trading strategy that gives a positive future profit with positive probability and has no probability of a loss.

The problem with this definition is that the opportunity depends on a too big set of data than is desired or even available in practical situations. For example, in continuous time the definition depends on the path properties of underlying's price processes. In practice only past prices at discrete times are observable. Working with static arbitrage, which is defined in Definition 4 below, suits this situation.

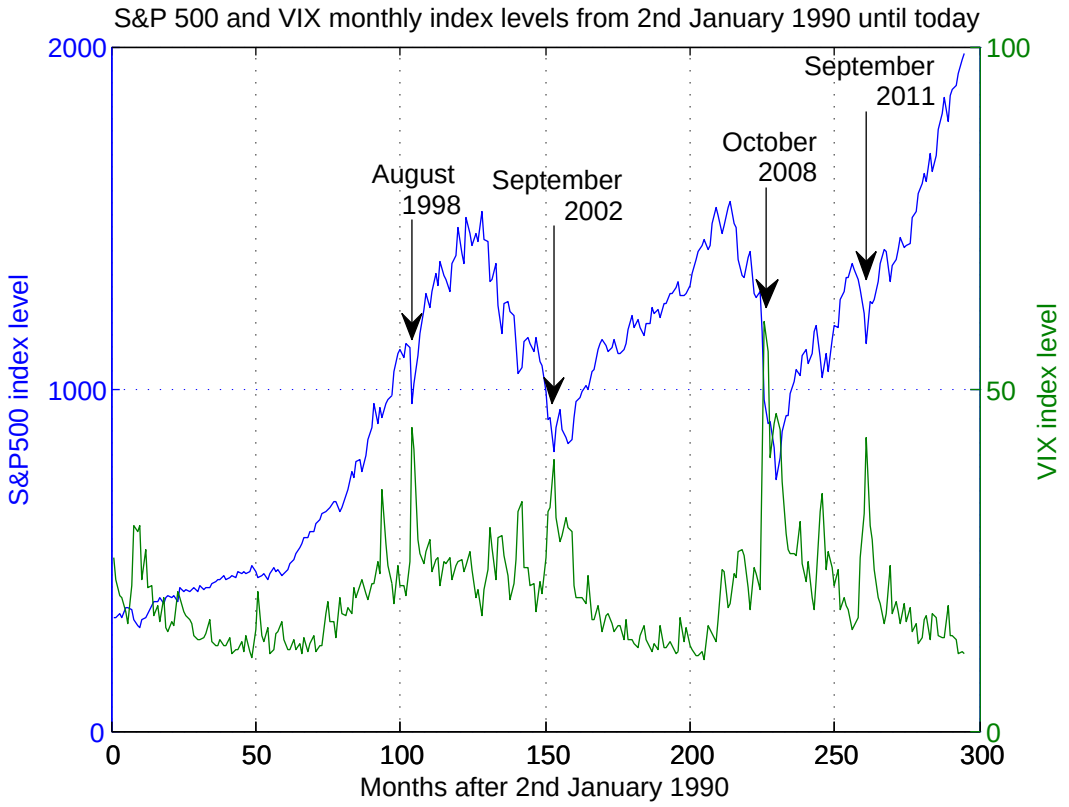

FIGURE 2.1: S&P 500 (left vertical axis) and VIX (right vertical axis) monthly index values plotted from January 1990 until today. Some important months are highlighted to visualize the negative correlation between the underlying's price process and the underlying's volatility. Data gathered from *finance.yahoo.com* on 7/8/2014.

Definition 4 (Static arbitrage opportunity)**.** A static arbitrage opportunity is a dynamic arbitrage opportunity where positions in the underlying at a particular time only can depend on time and actual corresponding price.

The Fundamental Theorem of Asset Pricing tells us that no dynamic arbitrage is equivalent to the existence of an equivalent martingale measure. From Definition 4, the following relaxed connection for static arbitrage was established in [5] and [6]. Instead of starting with a complete probability space and seeking martingales via a change of measure, as is the case in the elimination of dynamic arbitrage, the authors of [5] and [6] start with a family of densities, $\{q(X,t), t > 0\}$, of random variables, $\{X_t, t > 0\}$, indexed by t. These densities can be interpreted as measures, which will be done in Section 2.3. The authors proceed to show that if there exist some probability space on which it is possible to define a martingale $M(t)$ with the Markov property so that the law of $M(t)$ is $q(M,t)$ for each t, then the process $X = (X_t, t > 0)$ does not admit static arbitrage. We will refer to these laws, or densities, as t-marginals and we will call two proceeses that agree on their t-marginals for all t associated processes.

Definition 5 (Associated process)**.** If $(X_t; t \geq 0)$ and $(Y_t; t \geq 0)$ are two stochastic processes indexed by t, they are said to be associated if they have the same t-marginals for all t.

The theory that tell us whether an underlying asset, observable through call option prices on the market, and a process that is a martingale are associated or not is based on a theorem by Kellerer [7], which will be recalled in Section 2.3. The Markov property for stochastic processes is defined in Definition 6, which is the definition of Durrett [8].

Definition 6 (Markov property for stochastic processes)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_s, s > 0)$ and let $(S, \mathcal{S})$ be a measurable space. An $\mathcal{F}_s$-adapted process $X = (X_t, t > 0) : \Omega \mapsto S$ is said to have the Markov property if for each $s \in \mathcal{S}$ and each $s, t > 0$ with $s < t$,

$$\mathbb{P}\left(X_t \in s | \mathcal{F}_s\right) = \mathbb{P}\left(X_t \in s | X_s\right).$$

Equivalently, the process has the Markov property if for all $t \geq s \geq 0$ and for all bounded and measurable $f : S \mapsto \mathbb{R}$,

$$\mathbb{E}\left[f(X_t) | \mathcal{F}_s\right] = \mathbb{E}\left[f(X_t) | X_s\right].$$

An easy application of the tower property of conditional expectations shows that no dynamic arbitrage implies no static arbitrage, since the information set available when trading under no static arbitrage is a subset of that used when trading under no dynamic arbitrage. On the other hand, no static arbitrage does not imply no dynamic arbitrage, and this is best illustrated through a reproduction of an example in [6]. Let a process be defined on a grid with two levels, as in Figure 2.2. Call the process on the grid in Figure 2.2 for M_t. The procees M_t can not be a martingale since $\mathbb{E}[M_1 | M_{0.5}] \neq M_{0.5}$. But on the other hand, M_t has exactly the same t-marginals as the process in Figure 2.3 that goes up or down one price tick at each node with equal probability, and this process is a martingale!

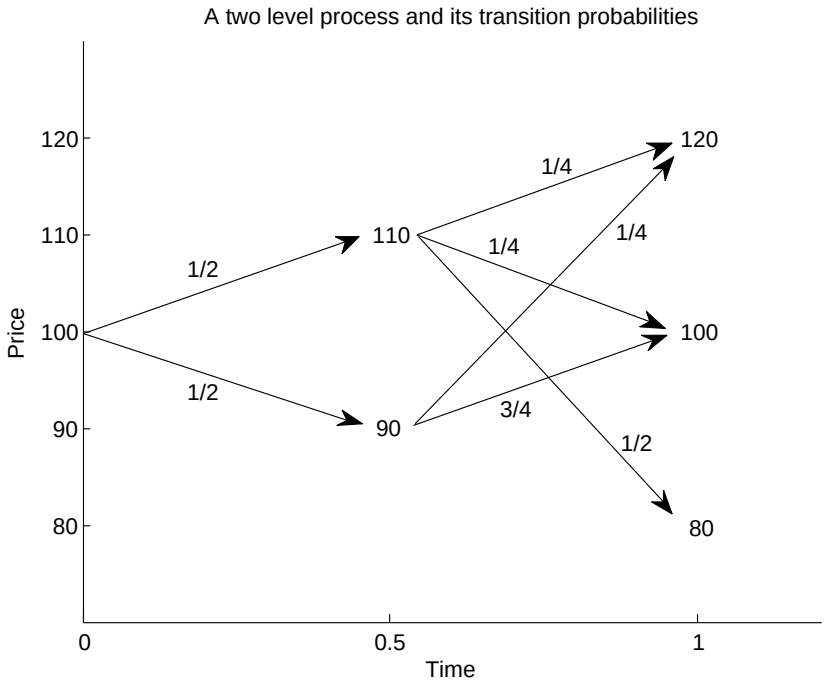

FIGURE 2.2: An example of a process that is not a martingale but for which there exists a martingale with the same distribution at each discrete step.

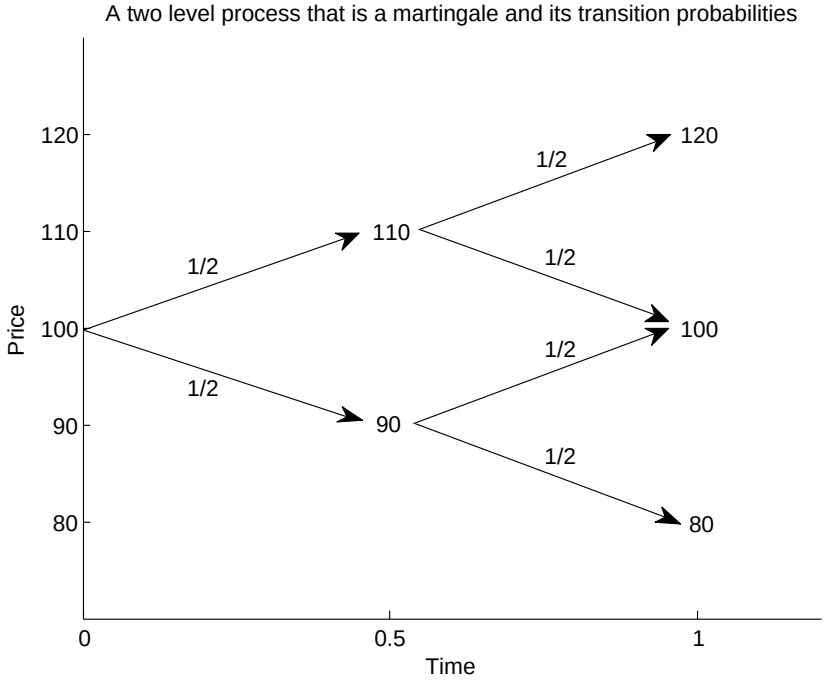

FIGURE 2.3: A process that is a martingale and has the same distribution at each discrete step as the process in Figure 2.2.

2.3 Kellerer's theorem: derivation

In this section, conditions for no static arbitrage on a call surface for European options are derived using Kellerer's theorem, which will be proved following the work of Hirsch, Roynette and Yor in [9] and [10]. We begin with some definitions.

Definition 7 ($\mathcal{M}_f$)**.** $\mathcal{M}_f$ is the set of all probability measures μ on $\mathbb{R}$ such that

$$\int |x|\mu(\mathrm{d}x) < \infty.$$

Definition 8 (Call function)**.** For $\mu \in \mathcal{M}_f$ and $x \in \mathbb{R}$, the corresponding call function is defined as

$$C_\mu(x) = \int_{\mathbb{R}} (y-x)^+ \mu(\mathrm{d}y).$$

From these definitions, we can derive three properties of the call function C_μ.

Proposition 1. C_μ is non-negative and convex.

Proof. Since $(y-x)^+ \geq 0$ for all $(x,y) \in \mathbb{R}^2$ and $\mu \in \mathcal{M}_f$, C_μ is non-negative. Convexity follows from the non-negativity of μ together with the following calculation,

$$\begin{aligned} \frac{\partial^2 C_\mu}{\partial x^2}(x) &= \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\int_{\mathbb{R}} (y-x)^+ \mu(\mathrm{d}y)\right) \\ &= \frac{\partial}{\partial x}\int_x^\infty -\mu(\mathrm{d}y) \\ &= \mu(x). \end{aligned}$$

□

Proposition 2. C_μ satisfies

$$\lim_{x\to\infty} C_\mu = 0.$$

Proof. Let $f_n = (y-n)^+$ for $n \in \mathbb{Z}$, and let $g_n = f_0 - f_n$. Then $\{g_n\}$ is an increasing, non-negative sequence of measurable functions that converges pointwise to f_0 and monotone convergence applies to g_n,

$$\lim_{n\to\infty}\int_{\mathbb{R}} g_n(y)\mu(\mathrm{d}y) = \int_{\mathbb{R}} \lim_{n\to\infty} g_n(y)\mu(\mathrm{d}y).$$

Thus we have

$$\begin{aligned} \lim_{n\to\infty}\left(\int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y) - \int_{\mathbb{R}} f_n(y)\mu(\mathrm{d}y)\right) &= \int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y) - \int_{\mathbb{R}} \lim_{n\to\infty} f_n(y)\mu(\mathrm{d}y) \\ &= \int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y), \end{aligned}$$

and

$$\begin{aligned}\lim_{x\to\infty} C_\mu(x) &= \lim_{n\to\infty}\int_{\mathbb{R}} f_n(y)\mu(\mathrm{d}y)\\ &= 0.\end{aligned}$$

□

Proposition 3. There exists a real number a so that

$$\lim_{x\to-\infty} C_\mu(x) + x = a.$$

Proof. Note first that $(y-x)^+ = \sup\{y,x\} - x$. Hence, since $\mu \in \mathcal{M}_f$,

$$\begin{aligned}C_\mu(x) &= \int_{\mathbb{R}} (y-x)^+\mu(\mathrm{d}y)\\ &= \int_{\mathbb{R}} \sup\{y,x\}\mu(\mathrm{d}y) - x\int_{\mathbb{R}}\mu(\mathrm{d}y)\\ &= \int_{\mathbb{R}} \sup\{y,x\}\mu(\mathrm{d}y) - x,\end{aligned}$$

so we get

$$\lim_{x\to-\infty} C_\mu(x) + x = \lim_{x\to-\infty}\int_{\mathbb{R}} \sup\{y,x\}\mu(\mathrm{d}y).$$

Now let $f_n(y) = \sup\{y,-n\}$ for $n \in \mathbb{Z}$, and let $g_n(y) = f_0(y) - f_n(y)$. Then $\{g_n\}$ is an increasing, non-negative sequence of measurable functions that converges pointwise to $f_0 - y$ and monotone convergence applies,

$$\lim_{n\to\infty}\int_{\mathbb{R}} g_n(y)\mu(\mathrm{d}y) = \int_{\mathbb{R}}\lim_{n\to\infty} g_n(y)\mu(\mathrm{d}y).$$

Thus, we get

$$\begin{aligned}\lim_{n\to\infty}\left(\int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y) - \int_{\mathbb{R}} f_n(y)\mu(\mathrm{d}y)\right) &= \int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y) - \int_{\mathbb{R}}\lim_{n\to\infty} f_n(y)\mu(\mathrm{d}y)\\ &= \int_{\mathbb{R}} f_0(y)\mu(\mathrm{d}y) - \int_{\mathbb{R}} y\mu(\mathrm{d}y),\end{aligned}$$

and the reuslt follows,

$$\begin{aligned}\lim_{x\to-\infty} C_\mu(x) + x &= \lim_{n\to\infty}\int_{\mathbb{R}} f_n(y)\mu(\mathrm{d}y)\\ &= \int_{\mathbb{R}} y\mu(\mathrm{d}y) < \infty.\end{aligned} \tag{2.7}$$

□

These three properties can also serve as a characterization of a probability measure, as the next proposition will tell us. This is a useful direction for our purpose.

Proposition 4. If $C : \mathbb{R} \mapsto \mathbb{R}$ has the properties in Proposition 1, Proposition 2 and Proposition 3,

P1. C is non-negative and convex,

P2. $C(x) \to 0$ as $x \to \infty$,

P3. There exists a real number a such that $C(x) + x \to a$ as $x \to -\infty$,

then there exists a unique $\mu \in \mathcal{M}_f$ such that $C = C_\mu$. Furthermore, this μ is the second derivative of C in the sense of distributions.

Proof. The proof of this direction can be found either in Proposition 2.1 of [10] or Lemma 7.23 of [11]. A slightly different version of the theorem is proven in Theorem 2.1 of [12]. The proof is omited here. □

Additionally, three useful properties of the call function can be derived.

Proposition 5. If $\mu \in \mathcal{M}_f$ then

i) for all $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$ such that $x_1 \leq x_2$,

$$0 \leq C_\mu(x_1) - C_\mu(x_2) \leq x_2 - x_1,$$

ii) for all $x \in \mathbb{R}$

$$C_\mu(x) + x - \int_{\mathbb{R}} x\mu(\mathrm{d}y) = \int_{\mathbb{R}} (x - y)^+ \mu(\mathrm{d}y),$$

iii) $\lim_{x \to -\infty} C_\mu(x) + x = \int_{\mathbb{R}} y\mu(\mathrm{d}y).$

Proof of i). Using the rewriting of the integrand in the call function from the proof of Proposition 3,

$$(y - x)^+ = \sup\{x, y\} - x,$$

the upper inequality can be derived,

$$\begin{aligned} C_\mu(x_1) - C_\mu(x_2) &= \int_{\mathbb{R}} \sup\{y, x_1\}\mu(\mathrm{d}y) - x_1 - \int_{\mathbb{R}} \sup\{y, x_2\}\mu(\mathrm{d}y) + x_2 \\ &= x_2 - x_1 + \underbrace{\int_{\mathbb{R}} \sup\{y, x_1\} - \sup\{y, x_2\}\mu(\mathrm{d}y)}_{\leq 0} \\ &\leq x_2 - x_1. \end{aligned}$$

The lower inequality follows from the fact that the call function is non-increasing in x, which was shown in the proof of Proposition 4.

Proof of ii). This is another application of the rewriting of the integrand in the call function from the proof of Proposition 3,

$$\begin{aligned} C_\mu(x) + x - \int_{\mathbb{R}} x\mu(\mathrm{d}y) &= \int_{\mathbb{R}} (y-x)^+\mu(\mathrm{d}y) + x - \int_{\mathbb{R}} y\mu(\mathrm{d}y) \\ &= \int_{\mathbb{R}} (y-x)^+\mu(\mathrm{d}y) + \int_{\mathbb{R}} x\mu(\mathrm{d}y) - \int_{\mathbb{R}} y\mu(\mathrm{d}y) \\ &= \int_{\mathbb{R}} \sup\{y,x\} - x + x - y\mu(\mathrm{d}y) \\ &= \int_{\mathbb{R}} (x-y)^+\mu(\mathrm{d}y). \end{aligned}$$

Proof of iii). This is an application of monotone convergence to the rewriting

$$\lim_{x\to-\infty} \int_{\mathbb{R}} (x-y)^+\mu(\mathrm{d}y) = \lim_{n\to\infty} \int_{\mathbb{R}} (-n-y)^+\mu(\mathrm{d}y).$$

□

The following special subset of $\mathcal{M}_f$ is useful for us.

Definition 9 (Uniformly integrable subset)**.** A subset $\mathcal{H}$ of $\mathcal{M}_f$ is said to be uniformly integrable if

$$\lim_{c\to\infty} \sup_{\mu\in\mathcal{H}} \int_{|x|\geq c} |x|\mu(\mathrm{d}x) = 0.$$

Note that if $\mathcal{H}$ is uniformly integrable then

$$\sup_{\mu\in\mathcal{H}} \left\{ \int |x|\mu(\mathrm{d}x) \right\} < \infty.$$

The two next propositions treat peacocks. A peacock is a family of stochastic processes or measures that have special convex properties.

Definition 10 (Peacock, measure version)**.** Let $(\mu_t; t \geq 0)$ be a family of probability measures on $\mathbb{R}$ indexed by t. Then $(\mu_t; t \geq 0)$ is a peacock if

i) for all $t \geq 0$, $\int |x|\mu_t(\mathrm{d}x) < \infty$,

ii) for all convex $\Psi : \mathbb{R} \to \mathbb{R}$, the map

$$\begin{aligned} g : [0,\infty) &\to (-\infty,\infty], \\ t &\mapsto \int \Psi(x)\mu_t(\mathrm{d}x), \end{aligned}$$

is increasing.

Proposition 6. Let the family of measures $(\mu_t; t \geq 0)$ be in $\mathcal{M}_f$. Let furthermore $\int_{\mathbb{R}} x\mu_t(\mathrm{d}y)$ be independent of t. Then $(\mu_t; t \geq 0)$ is a peacock if and only if for all $x \in \mathbb{R}$, the map $t \mapsto C(t,x)$, where $C(t,x) = C_{\mu_t}(x)$, is increasing.

Proof. The assumption that $(\mu_t; t \geq 0)$ is in $\mathcal{M}_f$ makes the family of measures satisfy condition i) of Definition 10. Thus let us focus on condition ii) of Definition 10.

Assume that $(\mu_t; t \geq 0)$ is a peacock, and hence satisfies condition ii) of Definition 10. Let $\Psi(x) = (x-c)^+, c \in \mathbb{R}$. Then $C(t,x) = \int \Psi(x)\mu_t(\mathrm{d}x)$, and the map $t \to C(t,x)$ is increasing. The proof of the other direction is long. It can be derived from Corollary 2.62 together with Theorem 2.58 in [11] and it is omitted here. □

Proposition 7. Assume that $(\mu_t; t \geq 0)$ is a peacock. Assume furthermore that $\int_{\mathbb{R}} x\mu_t(\mathrm{d}y)$ is independent of t. Then

1. the set $\{\mu_t; 0 \leq t \leq T\}$ is uniformly integrable,
2. $\lim\limits_{|x|\to\infty} \sup\{C(t,x) - C(s,x) \ : \ 0 \leq s \leq t \leq T\} = 0.$

Proof of 1. Note that if $c \geq 0$,

$$|y| I_{|y|\geq c} \leq (2|y| - c)^+ .$$

Then, since $(2|y| - c)^+$ is convex and $(\mu_t; t \geq 0)$ is a peacock,

$$\sup_{t\in[0,T]} \int_{|y|\geq c} |y|\mu_t(\mathrm{d}y) \leq \int_{\mathbb{R}} (2|y| - c)^+ \mu_T(\mathrm{d}y).$$

Let $f_n(y) = (2|y| - n)^+$ for $n \in \mathbb{N}$, then $\{f_n\}$ is a sequence of measurable functions with the pointwise limit 0. Also, $|f_n(y)| = f_n(y) \leq 2|y|$ for all $n \in \mathbb{N}$, and under the law of μ_T,

$$2\int_{\mathbb{R}} |y|\mu_T(\mathrm{d}y) < \infty,$$

since $\mu_T \in \mathcal{M}_f$. Hence by dominated convergence

$$\begin{aligned} \lim_{c\to\infty} \sup_{t\in[0,T]} \left\{ \int_{|y|\geq c} |y|\mu_t(\mathrm{d}y) \right\} &\leq \lim_{c\to\infty} \int_{\mathbb{R}} (2|y| - c)^+ \mu_T(\mathrm{d}y) \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} f_n(y)\mu_T(\mathrm{d}y) = 0, \end{aligned}$$

which proves that $\{\mu_t; 0 \leq t \leq T\}$ is uniformly integrable.

Proof of 2. Since $(y-x)^+$ is a convex function in x and $(\mu_t; t \geq 0)$ is a peacock, condition *ii*) in Definition 10 tells us that the call function is an increasing function in t and we get the inequality

$$\sup\{C(t,x) - C(s,x) : 0 \leq s \leq t \leq T\} \leq C(T,x) - C(0,x).$$

By Proposition 1, the call function is non-negative, so $C(0,x) \geq 0$ and

$$\sup\{C(t,x) - C(s,x) : 0 \leq s \leq t \leq T\} \leq C(T,x).$$

By Proposition 2,

$$\lim_{x\to\infty} \sup\{C(t,x) - C(s,x) : 0 \leq s \leq t \leq T\} = 0,$$

and we have proven the statement for large positive limit of x.. Now we have to prove it in the large negative limit of x. Since $\int_{\mathbb{R}} x\mu_t(\mathrm{d}y)$ is independent of t,

$$C(t,x) - C(s,x) = \left(C(t,x) + x - \int_{\mathbb{R}} x\mu_t(\mathrm{d}y)\right) - \left(C(s,x) + x - \int_{\mathbb{R}} x\mu_s(\mathrm{d}y)\right).$$

By the same arguments as above, $C(t,x)$ is increasing in t, so

$$\sup\{C(t,x) - C(s,x) : 0 \leq s \leq t \leq T\} \leq C(T,x) + x - \int_{\mathbb{R}} x\mu_T(\mathrm{d}y).$$

Taking the large negative limit and using Proposition 3 completes the proof,

$$\lim_{x\to-\infty} \sup\{C(t,x) - C(s,x) : 0 \leq s \leq t \leq T\} \leq \lim_{x\to-\infty} \left(C(T,x) + x - \int_{\mathbb{R}} x\mu_T(\mathrm{d}y)\right) = 0.$$

□

In the end of this section we will connect two processes via association, defined in Definition 5. The connection will be made through an application of the following uniqueness theorem for solutions of the Fokker-Planck equation.

Theorem 1 (M. Pierre's Uniqueness Theorem for the Fokker-Planck Equation)**.** Let the map

$$\begin{aligned} a: \mathbb{R}_+ \times \mathbb{R} &\to \mathbb{R}_+ \\ (t,x) &\mapsto a(t,x) \end{aligned}$$

be continuous such that $a(t,x) > 0$ for all $(t,x) \in (0,\infty) \times \mathbb{R}$, and let $\mu \in \mathcal{M}_f$. Then there exists at most one family of probability measures $(p(t,\mathrm{d}x); t \geq 0)$ such that

(FP1) $t \mapsto p(t, \mathrm{d}x)$ is weakly continuous,

(FP2) $p(0, \mathrm{d}x) = \mu(\mathrm{d}x)$ and

$$\frac{\partial p}{\partial t} - \frac{\partial^2}{\partial x^2}(ap) = 0, \quad \text{in } S'\left((0,\infty) \times \mathbb{R}\right),$$

where S' is the space of Schwartz distributions.

Proof. The proof is omitted for two reasons. Most importantly, it is not relevant to the understanding of Kellerer's theorem. Secondly, it is very long. A thorough version of the proof can be found in [13], Chapter 6.1. □

When the underlying's price process and the martingale have been connected, we would like to establish the Markov property for the martingale. A stronger property than the Markov property will be established for a certain class of stochastic processes in the next theorem, but first we need some new notation. As was done in [10], Definition 4.1 from [14] is used; if $(X_t; t \geq 0)$ is an $\mathbb{R}$-valued stochastic process then $\mathcal{F}^X$ is the filtration generated by X,

$$\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}, \ \forall\ t \geq 0.$$

For a Lipschitz continuous function $f : \mathbb{R} \to \mathbb{R}$, let $L(f)$ denote its Lipschitz constant. Let X be an $\mathbb{R}$-valued process. We say X has the Lipschitz-Markov property if there exists a Lipschitz continuous $f : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant $L(f) < 1$ such that for all bounded and continuous functions $g : \mathbb{R} \to \mathbb{R}$ with $L(g) \leq 1$ and all $s \in [0, t]$,

$$f(X_s) = \mathbb{E}\left[g(X_t) \,|\, \mathcal{F}_s^X\right].$$

The Lipschitz-Markov property implies the Markov property defined in Definition 6.

The following theorem tells us that a certain kind of process has the Lipschitz-Markov property.

Theorem 2. Let the map

$$\begin{aligned} \sigma : \mathbb{R}_+ \times \mathbb{R} &\to \mathbb{R}, \\ (t, x) &\mapsto \sigma(t, x), \end{aligned}$$

be continuous and such that $\partial_x \sigma$ exists and is continuous. Let furthermore X_0 be an integrable random variable and $(B_t; t \geq 0)$ be a standard Brownian motion independent of X_0. Then

$$X_t = X_0 + \int_0^t \sigma(s, X_s)\, \mathrm{d}B_s$$

has a unique solution with the Lipschitz-Markov property.

Proof. The proof is omitted here for the same reason as the proof of Theorem 1. The proof can be found in [10]. □

Finally we arrive to the theorem that establishes the connection.

Theorem 3 (Kellerer's Theorem)**.** Let $(X_t; t \geq 0)$ be an $\mathbb{R}$-valued integrable stochastic process indexed by t with t-marginals $(p(t,x); t \geq 0)$. Let furthermore $\int_{\mathbb{R}} xp(t, \mathrm{d}x)$ be independent of t, and let

$$\begin{aligned} C : \mathbb{R}_+ \times \mathbb{R} &\to \mathbb{R}_+, \\ (t,x) &\mapsto \mathbb{E}\left[(X_t - x)^+\right]. \end{aligned}$$

Asssume the following,

1. $C \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R})$ *and*

$$p(t,x) = \frac{\partial^2 C}{\partial x^2}(t,x), \ \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R},$$

2. p is positive on $\mathbb{R}_+ \times \mathbb{R}$, $\partial_t C$ is positive on $(0,\infty) \times \mathbb{R}$ and

$$\sigma(t,x) = \sqrt{\frac{2}{p}\frac{\partial C}{\partial t}}, \ \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}.$$

Then

$$Z_t = Z_0 + \int_0^t \sigma(s, Z_s)\, \mathrm{d}B_s$$

has a unique strong solution $(Y_t; t \geq 0)$ which is a martingale associated with $(X_t; t \geq 0)$ satisfying the Lipschitz-Markov property. This in turn implied absence of static arbitrage on the call surface C by the discussion in Chapter 2.2.

Proof. The proof will be divided into three steps. First, using Theorem 1, we prove that X and Y are associated processes. In the second step we prove that Y is a martingale. In the third step we show that Y has the Lipschitz-Markov property, using Theorem 2.

Step 1. We start with investigating the process X. Recall that the elements of the family $(p(t,x); t \geq 0)$ are the t-marginals of $(X_t; t \geq 0)$. Let $a(t,x) = \sigma^2(t,x)/2$, then

$$\frac{\partial^2}{\partial x^2}(ap) = \frac{\partial^3}{\partial x^2 \partial t} C = \frac{\partial}{\partial t} p.$$

Also $t \mapsto p(t,x)$ is a continuous function since $\partial_{xx}C$ is continuous by assumption 1. in the statement of the theorem. Hence $(p(t,x); t \geq 0)$ satisfies the Fokker-Planck equation. Furthermore, since $a(t,x)$ is positive everywhere, Theorem 1 tells us that $(p(t,x); t \geq 0)$ is the unique family of measures that does this.

We now investigate the process Y. Let φ be a real-valued function on the real line that is twice differentiable and has compact support within $Y(\mathbb{R})$, the image of Y. Itō's fomula says,

$$\mathrm{d}\varphi(Y_t) = \frac{\partial \varphi}{\partial x}(Y_t)\,\mathrm{d}Y_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial x^2}(Y_t)\,\mathrm{d}\langle Y\rangle_t.$$

The dynamics of Y_t is by Itō's formula,

$$\begin{aligned} \mathrm{d}Y_t &= 0\,\mathrm{d}t + \sigma(t,Y_t)\,\mathrm{d}B_t + 0\,\mathrm{d}t, \\ \mathrm{d}\langle Y\rangle_t &= \sigma^2(t,Y_t)\,\mathrm{d}t. \end{aligned}$$

Hence, we get

$$\mathrm{d}\varphi(Y_t) = \sigma(t,Y_t)\frac{\partial \varphi}{\partial x}(Y_t)\,\mathrm{d}B_t + \frac{1}{2}\sigma^2(t,Y_t)\frac{\partial^2 \varphi}{\partial x^2}(Y_t)\,\mathrm{d}t.$$

Integrating from 0 to t, with $a(t,x) = \sigma^2(t,x)/2$, yields,

$$\varphi(Y_t) - \varphi(Y_0) = \int_0^t \sigma(s,Y_s)\frac{\partial \varphi}{\partial x}(Y_s)\,\mathrm{d}B_s + \int_0^t a(s,Y_s)\frac{\partial^2 \varphi}{\partial x^2}(Y_s)\,\mathrm{d}s.$$

Under expectation conditioned on $Y_0 = y_0$, the first integral vanishes since $\varphi \in C^2(Y(\mathbb{R}))$. With $q(t,Y)$ as the law of Y_t, we get

$$\int_{Y_t(\mathbb{R})} \varphi(x) q(t,\mathrm{d}x) + y_0 = \int_{Y_t(\mathbb{R})} \int_0^t a(s,x)\frac{\partial^2 \varphi}{\partial x^2}(x) q(\mathrm{d}s,\mathrm{d}x).$$

The last step is to take the derivative with respect to time. In the proceeding calculations, the fact that $\varphi(Y_t)$ has compact support in the interior of $Y_t(\mathbb{R})$ is used in the partial integration. The left hand side becomes,

$$\frac{\partial}{\partial t}\int_{Y_t(\mathbb{R})} \varphi(x) q(t,\mathrm{d}x) + y = \int_{Y_t(\mathbb{R})} \varphi(x)\frac{\partial q}{\partial t}(t,\mathrm{d}x),$$

and the right hand side becomes,

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{Y_t(\mathbb{R})}\int_0^t a(s,x)\frac{\partial^2\varphi}{\partial x^2}q(\mathrm{d}s,\mathrm{d}x) &= \int_{Y_t(\mathbb{R})} a(t,x)\frac{\partial^2\varphi}{\partial x^2}q(t,\mathrm{d}x)\\
&= \left[a(t,x)\frac{\partial\varphi}{\partial x}q(t,x)\right]_{\partial Y_t(\mathbb{R})}\\
&\quad -\int_{Y_t(\mathbb{R})}\frac{\partial\varphi}{\partial x}\frac{\partial}{\partial x}\left(a(t,x)q(t,\mathrm{d}x)\right)\\
&= -\left[\varphi(x)\frac{\partial}{\partial x}\left(a(t,x)q(t,x)\right)\right]_{\partial Y_t(\mathbb{R})}\\
&\quad +\int_{Y_t(\mathbb{R})}\varphi(x)\frac{\partial^2}{\partial x^2}\left(a(t,x)q(t,\mathrm{d}x)\right)\\
&= \int_{Y_t(\mathbb{R})}\varphi(x)\frac{\partial^2}{\partial x^2}\left(a(t,x)q(t,\mathrm{d}x)\right).
\end{aligned}$$

To summarize, we have shown that

$$\int_{Y_t(\mathbb{R})}\varphi(x)\frac{\partial q}{\partial t}(t,\mathrm{d}x) = \int_{Y_t(\mathbb{R})}\varphi(x)\frac{\partial^2}{\partial x^2}\left(a(t,x)q(t,\mathrm{d}x)\right),$$

but this is nothing else than (FP2) in Theorem 1, in the sense of distributions. Thus the law of the t-marginals of $(Y_t; t \geq 0)$ satisfies the Fokker-Planck equation with the same $a(t,x)$ as the law of the t-marginals of $(X_t; t \geq 0)$. Therefore, by Theorem 1, the laws are the same and by defintion, X and Y are associated.

Step 2. Note one thing about condition expectations: by standard definition, $\mathbb{E}[X \mid Y]$ is the unique random variable such that for all bounded and measurable random variables Z,

$$\mathbb{E}[\mathbb{E}[X \mid Y]Z] = \mathbb{E}[XZ].$$

Let ϕ be a real-valued function on the real line that is twice differentiable and such that

$$\begin{aligned}
\phi(x) &= 1, \quad |x| \leq 1,\\
\phi(x) &= 0, \quad |x| \geq 2,\\
0 \leq \phi(x) &\leq 1, \quad \forall x \in \mathbb{R}.
\end{aligned}$$

Let for all $k > 0$, $\phi_k(x) = x\phi(x/k)$, let $h : \mathbb{R}^n \to \mathbb{R}$ be an arbitrary bounded and continuous function and let $0 \leq s_1 \leq \cdots \leq s_n \leq s \leq t$ be an arbitrary partition of the interval $[0,t]$. Set

$$\gamma_k = \mathbb{E}\left[h(Y_{s_1},\ldots,Y_{s_n})\phi_k(Y_t)\right] - \mathbb{E}\left[h(Y_{s_1},\ldots,Y_{s_n})\phi_k(Y_s)\right],$$

and $m = \sup_{x\in\mathbb{R}^n}\{|h(x)|\}$. Now $h(Y_{s_1},\ldots,Y_{s_n})\phi_k(Y_s)$ is a measurable functions that has the pointwise limit $h(Y_{s_1},\ldots,Y_{s_n})Y_s$. Furthermore $|h(Y_{s_1},\ldots,Y_{s_n})\phi_k(Y_s)| \leq m|Y_s|$ and $m|Y_s|$ is integrable since the density of Y_s is in $\mathcal{M}_f$. So by dominated convergence,

$$\lim_{k\to\infty}\gamma_k = \mathbb{E}\left[h(Y_{s_1},\ldots,Y_{s_n})Y_t\right] - \mathbb{E}\left[h(Y_{s_1},\ldots,Y_{s_n})Y_s\right].$$

If we can show that this limit is equal to zero, then by the previous comment on conditional expectation and the fact that h and the partition used were arbitrary, Y_s will be a martingale. Let us start with performing an Itō differentiation on ϕ_k. The dynamics of Y_t are from Step 1 known to be

$$\begin{aligned} \mathrm{d}Y_t &= \sigma(t,Y_t), \\ \mathrm{d}\langle Y\rangle_t &= \sigma^2(t,Y_t). \end{aligned}$$

Hence, we get

$$\mathrm{d}\phi_k = \sigma(t,Y_t)\frac{\partial\phi_k}{\partial x}(Y_t)\,\mathrm{d}B_t + \frac{1}{2}\sigma^2(t,Y_t)\frac{\partial^2\phi_k}{\partial x^2}(Y_t)\,\mathrm{d}t. \tag{2.8}$$

Integration from s to t of Equation 2.8 yields

$$\phi_k(Y_t) - \phi_k(Y_s) = \int_s^t \sigma(u,Y_u)\frac{\partial\phi_k}{\partial x}(Y_u)\,\mathrm{d}B_u + \int_s^t \frac{1}{2}\sigma^2(u,Y_u)\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\,\mathrm{d}u. \tag{2.9}$$

Multiplying both sides of Equation 2.9 with $h(Y_{s_1},\ldots,Y_{s_n})$ and taking the absolute value of them yields,

$$\begin{aligned} |h(Y_{s_1},\ldots,Y_{s_n})\left(\phi_k(Y_t)-\phi_k(Y_s)\right)| &= \left|h(Y_{s_1},\ldots,Y_{s_n})\left\{\int_s^t \sigma(u,Y_u)\frac{\partial\phi_k}{\partial x}(Y_u)\,\mathrm{d}B_u\right.\right. \\ &\quad \left.\left. + \int_s^t \frac{1}{2}\sigma^2(u,Y_u)\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\,\mathrm{d}u\right\}\right| \qquad (2.10) \\ &\leq m\left|\int_s^t \sigma(u,Y_u)\frac{\partial\phi_k}{\partial x}(Y_u)\,\mathrm{d}B_u\right. \\ &\quad \left. + \int_s^t \frac{1}{2}\sigma^2(u,Y_u)\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\,\mathrm{d}u\right|. \end{aligned}$$

Taking the expected value of both sides of Equation 2.10 and rearranging gives us, starting from γ_k,

$$\begin{aligned}
|\gamma_k| &= |\mathbb{E}\left[h(Y_{s_1},\dots,Y_{s_n})(\phi_k(Y_t)-\phi_k(Y_s))\right]| \\
&\leq \mathbb{E}\left[|h(Y_{s_1},\dots,Y_{s_n})(\phi_k(Y_t)-\phi_k(Y_s))|\right] \\
&\leq m\mathbb{E}\left[\left|\int_s^t \sigma(u,Y_u)\frac{\partial\phi_k}{\partial x}(Y_u)\,\mathrm{d}B_u + \int_s^t \frac{1}{2}\sigma^2(u,Y_u)\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\,\mathrm{d}u\right|\right] \\
&\leq m\mathbb{E}\left[\int_s^t \left|\sigma(u,Y_u)\frac{\partial\phi_k}{\partial x}(Y_u)\right|\,\mathrm{d}B_u + \int_s^t \left|\frac{1}{2}\sigma^2(u,Y_u)\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\right|\,\mathrm{d}u\right] \\
&= m\mathbb{E}\left[\int_s^t \frac{1}{2}\sigma^2(u,Y_u)\left|\frac{\partial^2\phi_k}{\partial x^2}(Y_u)\right|\,\mathrm{d}u\right].
\end{aligned}$$

Rewriting the expected value as an integral, where $p(t,x)$ is the density of Y_t, and by the assumption that $\sigma^2(t,x)/2 = \partial_t C(t,x)/p(t,x)$ we get

$$\begin{aligned}
|\gamma_k| &\leq m\int_{\mathbb{R}}\int_s^t \frac{1}{2}\sigma^2(u,x)\left|\frac{\partial^2\phi_k}{\partial x^2}(x)\right| p(u,x)\,\mathrm{d}u\,\mathrm{d}x \\
&= m\int_{\mathbb{R}}\int_s^t \frac{\partial C}{\partial u}(u,x)\left|\frac{\partial^2\phi_k}{\partial x^2}(x)\right|\,\mathrm{d}u\,\mathrm{d}x.
\end{aligned} \tag{2.11}$$

The derivative of ϕ_k in Equation 2.11 has no explicit t-dependence, so it can be moved outside the inner integral. The derivative of C can then be integrated to is antiderivative,

$$|\gamma_k| \leq m\int_{\mathbb{R}} (C(t,x)-C(s,x))\left|\frac{\partial^2\phi_k}{\partial x^2}(x)\right|\,\mathrm{d}x.$$

Note that ϕ_k is constant on the set $|x| \in \mathbb{R}\backslash[k,2k]$, so $\partial_{xx}\phi_k(x) = 0$ on $\mathbb{R}\backslash[k,2k]$. Furthermore, by definition of ϕ_k and the chain rule of differentiation,

$$\begin{aligned}
\int_k^{2k}\left|\frac{\partial^2\phi_k}{\partial x^2}(x)\right|\,\mathrm{d}x &= \int_k^{2k}\left|\frac{\partial}{\partial x}\left(\phi\left(\frac{x}{k}\right)+\frac{x}{k}\frac{\partial\phi}{\partial x}\left(\frac{x}{k}\right)\right)\right|\,\mathrm{d}x \\
&= \int_k^{2k}\left|\frac{1}{k}\left(2\frac{\partial\phi}{\partial x}\left(\frac{x}{k}\right)+\frac{x}{k}\frac{\partial^2\phi}{\partial x^2}\left(\frac{x}{k}\right)\right)\right|\,\mathrm{d}x \\
\{\text{let } x = ky\} &= \int_1^2\left|2\frac{\partial\phi}{\partial y}(y)+y\frac{\partial^2\phi}{\partial y^2}(y)\right|\,\mathrm{d}y.
\end{aligned}$$

Observe that the absolute value in the last integral is a continuous function and that the integration interval is a compact set. Hence the absolute value will have a maximum over the integration interval, and the left hand side is thus bounded by some $n \in \mathbb{R}$. For our "main inequality", this implies that

$$|\gamma_k| \leq mn\sup\{C(t,x)-C(s,x) : k \leq |x| \leq 2k\}.$$

By the assumption that $\mathbb{E}[Y_t]$ is independent of t, Proposition 6 tells us that the family $(p(t, \mathrm{d}x); t \geq 0)$ is a peacock. Then by Proposition 7,

$$\lim_{k\to\infty} |\gamma_k| \leq mn \lim_{|x|\to\infty} \sup\{C(t,x) - C(s,x)\} = 0$$

By the note on conditional expectation in the beginning of this step of the proof, Y_t is a martingale.

Step 3. It follows straight forward from Theorem 2 that Y_t has the Lipschitz-Markov property. □

2.4 Kellerer's theorem: implications on implied volatility

In this chapter, we will translate the restrictions that Kellerer's theorem enforce on the call surface in order for it to be free of static arbitrage into restrictions for the implied volatility surface. This will be done mainly by working with the Black formula, Equation (2.4). We begin by defining some new variables that will be used throughout the rest of this thesis.

A very useful variable when using expressions and formulas from the Black-Scholes model is the log-moneyness.

Definition 11 (Forward log-moneyness)**.** For a fixed time t, let $F_{[t,t+\tau]}$ be the forward price of an underlying S at time $t+\tau$, and let K be the strike price of some call option written on S at time t with expiry $t+\tau$. The forward log-moneyness x is defined by

$$x = \log\left(K/F_{[t,t+\tau]}\right).$$

When we are interested in a volatility smile, the time dependencies will always be surpressed since we will in these cases treat the time and time to maturity as constants. The usefulness of the forward log-moneyness lies not only in tiding up messy expressions, but also in its interpretation as the relative position of the option with respect to the forward price of the underlying. Other moneynesses can be defined, such as the underlying log-moneyness, $\log(K/S_t)$. If nothing else is specified, $\log(K/F)$ will be refered to only as the moneyness.

It will also be useful to introduce three variables as a complement to σ_{imp}. These three variables intermingle with σ_{imp} in the literature. In the defintions below, and the rest of this section, let τ be the time to maturity for an option.

Definition 12 (Total implied volatility)**.** The total implied volatility, θ_{imp} of an option with implied volatility σ_{imp}, is defined as

$$\theta_{\text{imp}} = \sqrt{\tau}\sigma_{\text{imp}}$$

Definition 13 (Implied variance)**.** The implied varaince, v_{imp}, of an option with implied volatility σ_{imp}, is defined as

$$v_{\text{imp}} = \sigma_{\text{imp}}^2$$

Definition 14 (Total implied variance)**.** The total implied variance, w_{imp}, of an option with implied volatility σ_{imp}, is defined as

$$w_{\text{imp}} = \tau\sigma_{\text{imp}}^2.$$

Using the total implied volatility together with moneyness has the advantage of simplifying the Black-Scholes auxiliary functions defined in Equation (2.5),

$$\begin{aligned} d_1(\tau, K, \tau\sigma_{\text{imp}}^2; F, r, t) &= \frac{\log(F_{[t,t+\tau]}/K) + \frac{1}{2}\tau\sigma_{\text{imp}}^2}{\sqrt{\tau\sigma_{\text{imp}}^2}} \\ &= -\frac{x}{\theta_{\text{imp}}} + \frac{\theta_{\text{imp}}}{2}, \\ d_2(\tau, K, \tau\sigma_{\text{imp}}^2; F, r, t) &= -\frac{x}{\theta_{\text{imp}}} - \frac{\theta_{\text{imp}}}{2}. \end{aligned}$$

We are now facing a conflict in notation. So far, we have used x as the argument of a call function $C_{\mu_t}(x)$. In the call function, x has the financial interpretation as the strike price. It is not wise to use x as a notation both for the strike price and for the moneyness. Also the variable t in the call function has the financial interpretation of time to maturity, and therefore this should be changed to τ. Therefore it is nessecary to make a change in the notation:

$$\begin{aligned} \text{Strike:} \quad x &\to K, \\ \text{Time to maturity:} \quad t &\to \tau. \end{aligned}$$

This notation is more in line with what is considered to be standard notation in mathematical finance. The variable x is henceforth reserved for the moneyness and the variable t is henceforth reserved for the current time.

So what restrictions have to be made on the implied volatility surface in order for the call surface it defines to be free of static arbitrage? It turns out that it is most convenient to state the definitions in terms of the total implied volatility, instead of the

implied volatility. Let us make a slight rewriting of the sufficient conditions implied by Kellerer's theorem, to ease a translation into conditions on implied volatility. The rewriting is stated as a theorem, Theorem 4 below, for clarity.

Theorem 4. An observed surface of call option prices written on some underlying S expiring at time T,

$$\begin{aligned} C : (0,\infty) \times \mathbb{R} &\to (0,\infty), \\ (\tau, K) &\mapsto \mathbb{E}\left[(S_{T-\tau} - K)^+\right], \end{aligned}$$

that is in $C^{2,2}$ is free of static arbitrage if the following five conditions hold.

(1) $\partial_\tau C > 0$.

(2) $\lim_{K\to\infty} C(\tau, K) = 0$.

(3) $\lim_{K\to-\infty} C(\tau, K) + K = a,\ a \in \mathbb{R}$.

(4) $C(\tau, K)$ is convex in K.

(5) $C(\tau, K)$ in non-negative.

Proof. The conditions (1)-(5) arise from the assumptions in Kellerer's theorem.

Condition (1) is stated as a condition on the call surface in Kellerer's theorem and will not be changed.

Condition (2)-(5) imply the existence and uniqueness of a positive p that satisfies $p = \partial_{KK} C$ through Proposition 4. These are the remaining conditions on the call surface in Kellerer's theorem. □

We would now like to translate conditions (1)-(5) in Theorem 4 into conditions on implied volatility. For this, we use the identity $C^{\mathrm{B}}(\tau, K, \tau\sigma^2_{\mathrm{imp}}) = e^{r\tau} C(\tau, K)$ that was introduced in Definition 1. Note that some dependencies have been dropped, since they are not of any interest here.

Theorem 5. The conditions (1)-(5) on call prices in Theorem 4 are implied by the following conditions on the implied volatility surface:

(A) $\partial_\tau w_{\text{imp}} = \partial_\tau \theta^2_{\text{imp}} > 0$.

(B) $\lim_{K\to\infty} d_1 = -\infty$.

(C) $\theta_{\text{imp}} \geq 0$.

(D) $\left(1 - \frac{x}{\theta_{\text{imp}}}\partial_x\left(\theta_{\text{imp}}\right)\right)^2 - \frac{\theta^2_{\text{imp}}}{4}\left(\partial_x\left(\theta_{\text{imp}}\right)\right)^2 + \theta_{\text{imp}}\partial_{xx}\left(\theta_{\text{imp}}\right) \geq 0$.

The inequality (D) is sometimes refered to as Durrleman's condition in the literature since it first appeared in [15]. It will be called Durrleman's condition in this thesis.

Proof. (A) *implies* (1)
In [16], the author provides a nice proof of this. Let us, without loss of generality, observe the market at time 0 and look at two contracts with the same moneyness, but with different expiry time, t_1 and t_2, $t_1 < t_2$, written on the same underlying, S. Since we are at time 0, the expiry time is also the time to maturity. If we want to keep moneyness constant, we need to require that the two options are written on different strikes, K_1 and K_2. From the definition of moneyness, Definition 11, together with the definition of the forward price, Definition 2, K_1 and K_2 are related in the following way,

$$\begin{aligned} \log\left(\frac{K_1}{F_{[0,t_1]}}\right) &= \log\left(\frac{K_2}{F_{[0,t_2]}}\right), \\ \Leftrightarrow \quad \frac{K_1}{S_0 e^{rt_1}} &= \frac{K_2}{S_0 e^{rt_2}}, \\ \Leftrightarrow \quad K_1 &= K_2 e^{-r(t_2-t_1)}. \end{aligned} \tag{2.12}$$

Thus, when differentiating the call price with respect to time to maturity, we do not have to care about changes in the forward price process. We want to achieve

$$C^{\text{BS}}(t_2, K_2, t_2\sigma^2_{\text{imp}}) > C^{\text{BS}}(t_1, K_1, t_1\sigma^2_{\text{imp}}). \tag{2.13}$$

If we multiply both sides of Equation 2.13 by $K_2^{-1}e^{rt_2}$, we get

$$\begin{aligned} \frac{e^{rt_2}C^{\text{BS}}(t_2, K_2, t_2\sigma^2_{\text{imp}})}{K_2} &> \frac{e^{rt_2}C^{\text{BS}}(t_1, K_1, t_1\sigma^2_{\text{imp}})}{K_1 e^{r(t_2-t_1)}} \\ &= \frac{e^{rt_1}C^{\text{BS}}(t_1, K_1, t_1\sigma^2_{\text{imp}})}{K_1}. \end{aligned}$$

Let the moneyness be constant. This implies that $F_{[0,t]}/K = e^{-x}$ is also constant. Then by Equation C.3, the function

$$\begin{aligned} f(w_{\text{imp}}) &= \frac{e^{rt}C^{\text{BS}}(t,K,w_{\text{imp}})}{K} \\ &= \frac{F_{[0,t]}}{K}\mathcal{N}(d_1) - \mathcal{N}(d_2) \end{aligned}$$

is an increasing function in total implied variance, w_{imp}. Hence, if we assume $\partial_\tau w_{\text{imp}} > 0$, then the call price C^{BS} will be an increasing function in time to maturity.

(B) *and* (C) *implies* (2)

Since (2) follows if the Black call price goes to zero as the strike price goes to infinity, we can examine the limit of C^{B} instead of the limit of C^{BS}. For the first term of C^{B}, that is $F\mathcal{N}(d_1)$, note that only if we have condition (B),

$$\lim_{K\to\infty} d_1(\tau,K,\tau\sigma_{\text{imp}}^2) = \lim_{x\to\infty} -\frac{x}{\theta_{\text{imp}}} + \frac{\theta_{\text{imp}}}{2} = -\infty,$$

do we have $\mathcal{N}(d_1(\tau,K,\tau\sigma_{\text{imp}}^2)) \to 0$ as $K \to \infty$. For the second term of C^{B}, that is $K\mathcal{N}(d_2)$, note that

$$d_2(\tau,K,\tau\sigma_{\text{imp}}^2) = -\frac{x}{\theta_{\text{imp}}} - \frac{\theta_{\text{imp}}}{2} = -\frac{1}{2}\left(\frac{2x}{\theta_{\text{imp}}} + \theta_{\text{imp}}\right). \tag{2.14}$$

Recall the inequality of arithmetic and geometric means.

Lemma 1. (Arithmetic-Geometric inequality) For any set of n non-negative real numbers $x_1,\ldots,x_n$,

$$\frac{\sum_{i=1}^n x_i}{n} \geq \sqrt[n]{\prod_{i=1}^n x_i}. \tag{2.15}$$

Since we want to examine the limit when K tends to infinity, K can be assumed to be positive in the following calculation. The forward price F is a price so it is always positive and finite. These two facts together imply that for a large enough K, x will be positive. If we assume that $\theta_{\text{imp}}(\tau,K,\tau\sigma_{\text{imp}}^2) \geq 0$, Equation (2.15) applies to the right hand side in Equation (2.14) and gives us

$$\begin{aligned} d_2(\tau,K,\tau\sigma_{\text{imp}}^2) &= -\frac{1}{2}\left(\frac{2x}{\theta_{\text{imp}}} + \theta_{\text{imp}}\right) \\ &\leq -\sqrt{\frac{2x}{\theta_{\text{imp}}}\theta_{\text{imp}}} \\ &= -\sqrt{2x}. \end{aligned}$$

Note that since $\mathcal{N}$ is a probability distribution, it is an increasing function and therefore

$$0 \leq e^x \mathcal{N}\left(d_2(\tau, K, \tau\sigma_{\text{imp}}^2)\right) \leq e^x \mathcal{N}\left(-\sqrt{2x}\right).$$

The right hand term tends to zero when K tends to infinity. The only condition that we made on θ_{imp} here is that is condition (C), that it should be non-negative.

(D) *implies* (4)
In Equation (C.1), it is derived that

$$\begin{aligned}
\partial_{KK} C^{\text{BS}} &= e^{-r\tau} \partial_{KK} C^{\text{B}} \\
&= \frac{Sn(d_1)}{K^2 \theta_{\text{imp}}} \left(\left(1 - \sqrt{\tau} \frac{x}{\theta_{\text{imp}}} \partial_x \left(\sigma_{\text{imp}}\right)\right)^2 - \tau \frac{\theta_{\text{imp}}^2}{4} \left(\partial_x \left(\sigma_{\text{imp}}\right)\right)^2 \right. \\
&\qquad \left. + \sqrt{\tau} \theta_{\text{imp}} \partial_{xx} \left(\sigma_{\text{imp}}\right) \right).
\end{aligned}$$

Since S, n and θ_{imp} are non-negative and since $\theta_{\text{imp}} = \sqrt{\tau}\sigma_{\text{imp}}$, the condition we need to impose to ensure convexity of C^{BS} in K is

$$\left(1 - \frac{x}{\theta_{\text{imp}}} \partial_x \left(\theta_{\text{imp}}\right)\right)^2 - \frac{\theta_{\text{imp}}^2}{4} \left(\partial_x \left(\theta_{\text{imp}}\right)\right)^2 + \theta_{\text{imp}} \partial_{xx} \left(\theta_{\text{imp}}\right) \geq 0, \tag{2.16}$$

in order to insure the convexity of C^{B} in K. Equation (2.16) can also be expressed in terms of the total implied variance. The following relation is taken from Equation (C.2),

$$\left(1 - \frac{x}{2w_{\text{imp}}} \partial_x \left(w_{\text{imp}}\right)\right)^2 - \frac{1}{4}\left(\frac{1}{w_{\text{imp}}} + \frac{1}{4}\right) \left(\partial_x \left(w_{\text{imp}}\right)\right)^2 + \frac{1}{2} \partial_{xx} \left(w_{\text{imp}}\right) \geq 0.$$

(B), (C) *and* (D) *imply* (3)
If (B), (C) and (D) hold then the call price is a convex, non-increasing function of K. Since the call price is assumed to be twice differentiable, the following limit exists

$$\lim_{h \to 0} \frac{C(\tau, K+h) - C(\tau, K)}{h}. \tag{2.17}$$

By Proposition 5 i), for all $h > 0$ the numerator satisfies

$$K - (K+h) \leq C(\tau, K+h) - C(\tau, K) \leq 0. \tag{2.18}$$

Applying Equation (2.18) and the definition of the partial derivative to Equation (2.17), we get

$$-1 \leq \partial_K C(\tau, K) \leq 0.$$

A similar argument to that which was made in Theorem 2.1 of [12] can be done for the previous inequality, leading to the limit

$$\lim_{K \to -\infty} \partial_K C(\tau, K) + 1 = 0. \tag{2.19}$$

Integrating Equation (2.19) with respect to K yields

$$\int \lim_{K \to -\infty} \partial_K C(\tau, K) + 1 \, \mathrm{d}K = a, \quad a \in \mathbb{R}.$$

Note that the explicit expression for the partial derivative is

$$\partial_K C(\tau, K) = \left((S_t - K)^+ + 1\right) \partial_K \mu_\tau(S_t).$$

Note furthermore that if we let f_n be defined by

$$f_n(K) = \left(\left(S_t - \frac{1}{n} - K\right)^+ + 1\right) \partial_K \mu_\tau(S_t),$$

then $\{f_n\}$ is a non-decreasing sequence of positive, measurable functions that converge pointwise to $\partial_K C(\tau, K)$. Hence the monotone convergence theorem is applicable and we can move the limit outside the integral to obtain the result,

$$\begin{aligned} a &= \int \lim_{K \to -\infty} \partial_K C(\tau, K) + 1 \, \mathrm{d}K \\ &= \lim_{K \to -\infty} \int \partial_K C(\tau, K) + 1 \, \mathrm{d}K \\ &= \lim_{K \to -\infty} C(\tau, K) + K. \end{aligned}$$

The Black-Scholes model implies (5)
The implied volatility is derived from the Black-Scholes model and therefore all assumptions of the Black-Scholes model will hold true for the implied volatility we calculate using the model. The Black-Scholes call price formula is derived as the unique solution to the Black-Scholes partial differential equation. This equation also has a call function of the form which was introduced in Defintion 8 as solution by the discounted Feynman-Kac theorem. Recall that a call function is an integral of a non-negative function with respect to a probability measure. Therefore, the Black-Scholes model implies that the call surface will be non-negative. □

2.5 Asymptotic bounds on the implied volatility smile

As was introduced in Definition 1, the implied volatility is the variable σ_{imp} that uniquely solves

$$e^{r\tau}C(\tau,K) = C^{\text{B}}(\tau,K,\tau\sigma_{\text{imp}}^2).$$

As before, let x denote the forward log-moneyness. Furthermore, let $g(x) \sim f(x)$ if $g(x)/f(x) \to 1$ as $x \to \infty$.

The intuition behind the result of this section is that it is crucial to match the asymptotics of $C^{\text{B}}(\tau,K,\tau\sigma_{\text{imp}}^2)$ with the asymptotics of $C(\tau,K)$, because if they are to agree for all K, we need to have $C \sim C^{\text{B}}$. This forced matching will have implications on the implied volatility. Let us investigate the limit behaviour of C^{B} and C as K goes to infinity and consolidate the intuition. As in many mathematical derivations, a special function appears that suits our needs well,

$$f_1(y) = \left(\frac{1}{\sqrt{y}} - \frac{\sqrt{y}}{2}\right)^2,$$
$$f_2(y) = \left(\frac{1}{\sqrt{y}} + \frac{\sqrt{y}}{2}\right)^2.$$

Note that d_1 and d_2 can be expressed in terms of the functions f_1 and f_2,

$$\begin{aligned} d_1(\tau,K,\tau\sigma_{\text{imp}}^2) &= -\left(\frac{x}{\theta_{\text{imp}}} - \frac{\theta_{\text{imp}}}{2}\right) \\ &= -\sqrt{x}\left(\frac{1}{\theta_{\text{imp}}/\sqrt{x}} - \frac{\theta_{\text{imp}}/\sqrt{x}}{2}\right) \\ &= -\sqrt{x}\left(\frac{1}{\sqrt{\theta_{\text{imp}}^2/x}} - \frac{\sqrt{\theta_{\text{imp}}^2/x}}{2}\right) \\ &= -\sqrt{xf_1(\theta_{\text{imp}}^2/x)}. \end{aligned}$$

An analogous calculation can be made to show that

$$d_2(\tau,K,\tau\sigma_{\text{imp}}^2) = -\sqrt{xf_2(\theta_{\text{imp}}^2/x)}.$$

With the functions f_1 and f_2 at hand, we get a nice expression of the Black call price when $\sigma_{\text{imp}} = \sqrt{\beta x}$ where β is a positive number,

$$C^{\text{B}}(\tau,K,\tau\beta x) = F\left(\mathcal{N}(-\sqrt{xf_1(\beta)}) - e^x\mathcal{N}(-\sqrt{xf_2(\beta)})\right). \tag{2.20}$$

Since we want to investigate the limit when K goes to positive infinity, we may without loss of generality assume that $\beta x > 0$ and the implied volatility in Equation (2.20) is therefore well defined. Equation (2.20) allows us to study the asymptotics of the Black call price when the implied variance is linear in moneyness. Through partial integration an asymptotic approximation can be derived for $\mathcal{N}(y)$. Using the fact that $\partial_y \mathcal{N}(y)$ is an even function,

$$\begin{aligned}
\mathcal{N}(-y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-y} e^{-t^2/2} \, \mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} e^{-t^2/2} \, \mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_{y^2/2}^{\infty} s^{-1/2} e^{-s} \, \mathrm{d}s \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{e^{-y^2/2}}{y} - \frac{1}{2} \int_{y^2/2}^{\infty} s^{-3/2} e^{-s} \, \mathrm{d}s \right).
\end{aligned}$$

Since both $s^{-3/2}$ and e^{-s} are decreasing functions on $\left[y^2/2, \infty\right)$, the integral on the right hand side can be bounded,

$$\begin{aligned}
\left| \int_{y^2/2}^{\infty} s^{-3/2} e^{-s} \, \mathrm{d}s \right| &\leq \frac{1}{y^3} \int_{x^2/2}^{\infty} e^{-s} \, \mathrm{d}s \\
&\leq \frac{e^{-y^2/2}}{y^3}.
\end{aligned}$$

Hence we have that

$$\mathcal{N}(-y) \sim \frac{e^{-y^2/2}}{y\sqrt{2\pi}}, \quad y \to \infty. \tag{2.21}$$

Using that $f_1(\beta) + 2 = f_2(\beta)$ together with the asymptotics from Equation (2.21), the asymptotics of C^{B} can be retrieved,

$$\begin{aligned}
C^{\mathrm{B}}(\tau, K, \tau\beta x) &= F \left(\mathcal{N}(-\sqrt{x f_1(\beta)}) - e^x \mathcal{N}(-\sqrt{x f_2(\beta)}) \right) \\
&\sim \frac{F}{\sqrt{2\pi}} \left(\frac{e^{-x f_1(\beta)/2}}{\sqrt{x f_1(\beta)}} - \frac{e^x e^{-x f_2(\beta)/2}}{\sqrt{x f_2(\beta)}} \right) \\
&= \frac{F}{\sqrt{2\pi x}} \left(\frac{e^{-x f_1(\beta)/2}}{\sqrt{f_1(\beta)}} - \frac{e^x e^{-x(f_1(\beta)+2)/2}}{\sqrt{f_2(\beta)}} \right) \\
&= \frac{e^{-x f_1(\beta)/2}}{B(\beta)\sqrt{x}}.
\end{aligned}$$

Here B is a function depending only on β. Having established the asymptotic properties for C^{B} for large strikes, now we need to do the same for C. Recall from Section 2.3

$$C(\tau, K) = \mathbb{E}[(S_t - K)^+],$$

where S_t is the underlying's price process. Inspired by [17], Lee derives an upper bound for C in [18] which suits our needs better than the standard bound $C(\tau, K) \leq \mathbb{E}[S_t]$. Note that for each $p > 0$ and for all $s \geq 0$,

$$s - x \leq \frac{s^{p+1}}{p+1} \left(\frac{p}{p+1} \right)^p e^{-xp}, \quad \forall x > 0,$$

since both sides of the inequality, if viewed as functions of s, have equal values and first derivatives at $s = (p+1)x/p$, but the right hand side has a positive second derivative. Note furthermore that the right hand side is non-negative, so

$$(s - x)^+ \leq \frac{s^{p+1}}{p+1} \left(\frac{p}{p+1} \right)^p e^{-xp}.$$

Exchanging s for the underlying S_t and taking expectations yields

$$C(\tau, K) \leq \mathbb{E}[S_t^{p+1}] \frac{1}{p+1} \left(\frac{p}{p+1} \right)^p e^{-xp}. \tag{2.22}$$

Hence, if S_t has finite $p + 1^{\text{th}}$ moment, then $C(\tau, K) = \mathcal{O}(e^{-xp})$ as $x \to \infty$. Comparing the asymptotics of C^{B} and C, we see that they agree if $f_1(\beta)/2 = p$. This idea, that the tail behaviour of the implied volatility smile carries the same information as the tail behaviour of the the option prices was made rigorous by Lee in [18]. He uses the connection between option prices and the number of finite moments of the underlying. This connection surley sounds reasonable, since option prices are bounded by moments by (2.22) and, since power payoffs are mixtures of call and put payoffs across a continuum of strikes, moments are bounded by option prices.

Theorem 6 (Lee's Large Strike Moment Formula)**.** Let

$$\begin{aligned} \hat{p} &= \sup \left\{ p \in (0, \infty) : \mathbb{E}\left[S_t^{1+p} \right] < \infty \right\}, \\ \beta_{\text{large}} &= \limsup_{x \to \infty} \frac{\sigma_{\text{imp}}^2(K)}{|x|} \end{aligned}$$

Then $\beta_{\text{large}} \in [0, 2]$ and

$$\hat{p} = \frac{1}{2\beta_{\text{large}}} + \frac{\beta_{\text{large}}}{8} - \frac{1}{2},$$

where $1/0 := \infty$. Equivalenty, $p \in [0, \infty]$ and

$$\beta_{\text{large}} = 2 - 4 \left(\sqrt{\hat{p}^2 + \hat{p}} - \hat{p} \right).$$

Proof. The proof is divided into three steps. In the first step we prove that $\beta_{\text{large}} \in [0, 2]$, in the second step we show that $\hat{p} \leq f_1(\beta_{\text{large}})/2$ and in the third step we show the complementary inequality which together with the second step implies that $\hat{p} = f_1(\beta_{\text{large}})/2$.

Step 1. If there exists an $\hat{x} > 0$ such that for all $x > \hat{x}$,

$$\sigma_{\text{imp}} < \sqrt{2|x|},$$

then by the definition of β_{large} in the theorem statement, $\beta_{\text{large}} \in [0, 2]$. This is equivalent to

$$C^{\text{B}}(\tau, K, \tau\sigma^2_{\text{imp}}) < C^{\text{B}}(\tau, K, \tau 2|x|), \quad x > \hat{x}, \tag{2.23}$$

since C^{B} is strictly increasing in the first argument. We know from the definition of implied volatility that the left hand side of (2.23) is equal to $C(\tau, K) = \mathbb{E}\left[(S_t - K)^+\right]$. Now $\{(S_t - K)^+\}_{x>0}$ is a family of non-negative random variables that converge to 0 as x goes to infinity and are bounded from above by S_τ. Furthermore, $\mathbb{E}[S_t] < \infty$ since we have assumed that the call prices exist. Then, by dominated converge,

$$\lim_{x\to\infty} C(\tau, K) = \lim_{x\to\infty} \mathbb{E}\left[(S_\tau - K)^+\right] = 0.$$

For the right hand side of (2.23), note that

$$\begin{aligned} C^{\text{B}}(\tau, K, \tau 2|x|) &= F\left(\mathcal{N}(0) - e^x \mathcal{N}(-\sqrt{2|x|})\right) \\ &= F\left(\frac{1}{2} - e^x \mathcal{N}(-\sqrt{2|x|})\right). \end{aligned}$$

By l'Hopital's rule,

$$\lim_{x\to\infty} \frac{\mathcal{N}(-\sqrt{2|x|})}{e^{-x}} = \lim_{x\to\infty} \frac{2(2|x|)^{-1/2} e^{-(-\sqrt{2|x|})^2/2}}{e^{-x}} = \lim_{x\to\infty} \frac{\sqrt{2}e^{-x}}{\sqrt{|x|}e^{-x}} = 0,$$

so

$$\lim_{x\to\infty} C^{\text{B}}(\tau, K, \tau 2|x|) = \frac{F}{2},$$

and the first step of the proof is finished.

Step 2. In the this and the third step, we need a special limit. For $\beta \in (0,2)$ and a constant c,

$$\begin{aligned}
\lim_{x\to\infty} \frac{e^{-cx}}{C^{\mathrm{B}}(\tau,K,\tau\beta|x|)} &= \lim_{x\to\infty} \frac{e^{-cx}}{F\left(\mathcal{N}(-\sqrt{xf_1(\beta)}) - e^x\mathcal{N}(-\sqrt{xf_2(\beta)})\right)} \\
&= \lim_{x\to\infty} \frac{ce^{-cx}}{F\left(n\left(-\sqrt{xf_1(\beta)}\right)\sqrt{\frac{f_1(\beta)}{x}} - e^x n\left(-\sqrt{xf_2(\beta)}\right)\sqrt{\frac{f_2(\beta)}{x}}\right)} \\
&= \lim_{x\to\infty} \frac{ce^{-cx}}{F\left(e^{-xf_1(\beta)/2}\sqrt{\frac{f_1(\beta)}{x}} - e^x e^{-xf_2(\beta)/2}\sqrt{\frac{f_2(\beta)}{x}}\right)} \\
&= \lim_{x\to\infty} \frac{ce^{-cx}}{F\left(e^{-xf_1(\beta)/2}\left(\sqrt{\frac{f_1(\beta)}{x}} - \sqrt{\frac{f_2(\beta)}{x}}\right)\right)} \\
&= \lim_{x\to\infty} \frac{c}{F}\left(\frac{\sqrt{x}}{\sqrt{f_1(\beta)}-\sqrt{f_2(\beta)}}\right)e^{x(f_1(\beta)/2-c)} \\
&= \begin{cases} 0, & c > f_1(\beta)/2, \\ \infty, & c \le f_1(\beta)/2. \end{cases}
\end{aligned}$$

Let $\beta \in (0,2)$ and $p \in (f_1(\beta)/2, \hat{p})$ where $\hat{p}$ is defined as in the theorem statement. By (2.22) and the previous limit we have that when $x \to \infty$,

$$\frac{C^{\mathrm{B}}(\tau,K,\tau\sigma^2_{\mathrm{imp}})}{C^{\mathrm{B}}(\tau,K,\tau\beta|x|)} = \frac{\mathcal{O}(e^{-px})}{C^{\mathrm{B}}(\tau,K,\tau\beta|x|)} \to 0. \tag{2.24}$$

Note now that $f_1(\beta)$ is strictly decreasing when $\beta \in (0,2)$. This implies that for any $\beta \in (0,2)$ with $f_1(\beta)/2 < \hat{p}$, we have $\beta_{\mathrm{large}} \le \beta$ and hence we need to have $\hat{p} \le f_1(\beta_{\mathrm{large}})/2$ in order for the limit (2.24) to be a constant.

In the case when this last step is vacuously true, that is if there exists no $\beta \in (0,2)$ such that $f_1(\beta)/2 < \hat{p}$, we have by the definition in the statement of the theorem that $\beta_{\mathrm{large}} = 0$ and $\hat{p} = f_1(\beta_{\mathrm{large}})/2 = \infty$.

Step 3. In this step we will prove the complementary inequality, $\hat{p} \geq f_1(\beta_{\text{large}})/2$. From the defintion of $\hat{p}$, we see that it is enough to show that for any $p \in (0, f_1(\beta_{\text{large}})/2)$, $\mathbb{E}\left[S_\tau^{1+p}\right]$ is finite. To show this, we pick β such that $f_1(\beta)/2 \in (p, f_1(\beta_{\text{large}}))$. Then, as earlier, for large enough x,

$$\frac{C(\tau,K)}{e^{-xf_1(\beta)/2}} \leq \frac{C^{\text{B}}(\tau,K,\tau\beta|x|)}{e^{-xf_1(\beta)/2}} \to 0, \text{ as } x \to \infty.$$

Thus there exists a K^* so that for $K > K^*$, $C(\tau,K) < K^{-f_1(\beta)/2}$. Using the spanning relation from Appendix B with $k = 0$, we have

$$\begin{aligned}\mathbb{E}\left[S_t^{p+1}\right] &= \mathbb{E}\left[\int_0^\infty (p+1)pK^{p-1}(S_t-K)^+\mathrm{d}K\right] \\ &\leq (p+1)p\left[\int_0^{K^*} K^{p-1}C(\tau,K)\mathrm{d}K + \int_{K^*}^\infty K^{p-1-f_1(\beta)/2}\mathrm{d}K\right] < \infty.\end{aligned}$$

□

There is a corresponding theorem for small strikes. An analogous proof as the one for the large strike formula can be done, but a shorter one was presented in [19] that builds on what we already know from the large strike formula. Since the proofs are similar to a great extent, the proof is omitted.

Theorem 7 (Lee's Small Strike Moment Formula)**.** Let

$$\begin{aligned}\hat{q} &= \sup\left\{q \in (0,\infty) : \mathbb{E}\left[S_t^{-q}\right] < \infty\right\}, \\ \beta_{\text{small}} &= \limsup_{x\to-\infty} \frac{\sigma_{\text{imp}}^2(K)}{|x|t}.\end{aligned}$$

Then $\beta_{\text{small}} \in [0,2]$ and

$$\hat{q} = \frac{1}{2\beta_{\text{small}}} + \frac{\beta_{\text{small}}}{8} - \frac{1}{2},$$

where $1/0 := \infty$. Equivalently, $\hat{q} \in [0,\infty]$ and

$$\beta_{\text{small}} = 2 - 4\left(\sqrt{\hat{q}^2+\hat{q}} - \hat{q}\right).$$

The implications on the characteristics of implied volatility from Theorem 6 and Theorem 7 are important. The theorems determine that the implied volatility cannot grow faster than $\sqrt{|x|}$. That is, for large enough $|x|$, σ_{imp} has to be smaller or equal to $\sqrt{\beta|x|}$. Furthermore, unless S_t has finite moments of all orders which corresponds to the case where $\beta = 0$, the implied volatility cannot grow slower than $\sqrt{|x|}$.

Chapter 3

Parameterization of the implied volatility

A parametric model of the implied volatility comes with certain advantages. Observed implied volatilities, and hence call prices, can be inter- and extrapolated. Therefore a parametric implied volatility model can be used to price new contracts for which there are no quotes on the market. The implied volatility in a parametric model is function of strike and maturity with an explicit analytical expression. If the implied volatility is modeled as a smooth function it will also admit analytical explicit expression for its derivatives of all orders possibly saving computational time. A parametric model have to satisfy the conditions derived in Chapter 2 to be considered as feasible.

There exist several popular models for stochastic implied volatility, with the most popular being Stochastic Volatility Inspired (SVI) parameterization [1], the Stochastic alpha, beta, rho (SABR) parameterization [20] and Vanna-Volga (VV) model. We are concerned with the SVI, but it could be of some interest to mention some properties and limitations of the other models.

This chapter starts with a short examination of the three mentioned models. After this follows a summary of the different variations of the SVI parameterization that were introduced in [2] together with their interpretation. Finally, this chapter ends with a summary of the work in [2] that treats conditions on the SVI parameters that guarantee the absence of static abitrage in the implied volatility they define.

3.1 Popular stochastic volatility models

Stochastic volatility inspired (SVI)

The SVI parameterization of the total implied variance for a fixed time to maturity reads,

$$w_{\text{imp}}^{\text{SVI}}(x) = a + b\left(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\right), \tag{3.1}$$

where x is moneyness and $\{a, b, \sigma, \rho, m\}$ is the parameter set. The SVI parameter σ is not to be confused with the volatility of the underlying's price process, which is also denoted by σ! The first strength of the SVI is demonstrated in the following proposition.

Proposition 8. The SVI parameterization in Equation (3.1) satisfies Lee's large and small strike formulas.

Proof. The right asymptote is by [1]

$$\left(w_{\text{imp}}^{\text{SVI}}\right)_r(x) = a + b(1-\rho)(x-m).$$

The left asymptote is by [1]

$$\left(w_{\text{imp}}^{\text{SVI}}\right)_l(x) = a - b(1+\rho)(x-m).$$

They are both linear in moneyness hence satisfy Lee's formula. □

Note that these asymptotes imply through Lee's large and small strike formulas that the distribution of the underlying's price process has finite moments of all orders. This is a model limitation of the SVI, since by [19] the implied volatility may grow slower than $\sqrt{x}$ when the distribution on the underlying's price process does not have finite moments of all orders, because of for example fat tails. Work such as [21] and [22] tries to solve this by introducing more parameters into Equation (3.1), but these two models and their implementation is outside the area of interest for this thesis.

The second strength of the SVI model was established in [23]. It was shown that the implied volatility in the Heston model converges to the SVI in the long maturity limit. The Heston model assumes the same dynamics for the implied variance as was done in Chapter 2.1, but assigns the coefficients in Equation (2.6). The implied variance in the Heston model follows the dynamics

$$\mathrm{d}v_{\text{imp}} = \theta(\omega - v_{\text{imp}})\mathrm{d}t + \eta\sqrt{v_{\text{imp}}}\mathrm{d}B_t^1. \tag{3.2}$$

Here, ω is a long-time mean value of the variance and θ is the rate at which the variance reverts towards ω. The volatility of the variance process is η. This choice of dynamics is built on three assumptions, namely that the variance of S is a random process that

1. has a tendency to revert towards its long-term mean at some constant rate,
2. has a volatility that is proportional to the square root of its level,
3. has a source of randomness that is correlated with the randomness of the underlying's price process.

The assumption about mean reversion is often deemed as being the cause of mean reversion. If traders believe in the assumption of mean reversion, then traders will sell contracts when they are worth more than the long-term mean and buy when the contracts are worth less than the long term mean. Through their actions on the market, the traders will cause mean reversion.

The second assumption can be motivated by the observation of volatility clustering. In Figure 3.1 daily returns for the S&P 500 index are plotted. It is visible that larger movements cluster together and are interupted by periods of smaller movements. Heuristically, we see that the size of the movements depends on the size of the movements.

The third assumption was motivated in Chapter 2.1.

Stochastic alpha, beta, rho (SABR)

The underlying's price process is in the SABR modeled with the following dynamics,

$$\begin{aligned} \mathrm{d}S_t &= \sigma_t S_t^{\beta} \mathrm{d}B_t^0, \\ \mathrm{d}\sigma_t &= \alpha \sigma_t \mathrm{d}B_t^1, \\ \langle \mathrm{d}B_t^0, \mathrm{d}B_t^1 \rangle &= \rho \mathrm{d}t, \end{aligned}$$

where B^0 and B^1 are standard Brownian motions, $\beta \in [0, 1]$ is a skewness parameter, $\alpha \geq 0$ is the volatility of volatility and $\rho \in [-1, 1]$. A flaw in the SABR model is the lack of mean reversion, which makes it suitable for options with short time to maturity only. A strength of the SABR model is that it yeilds an explicit formula in the short time to maturity limit, which makes it possible to fit the parameters β, α and ρ to market data. The retrieved function with the fitted parameters can then be expanded using Taylor's formula to yield approximate implied volatilities to options with non-zero, but short, times to maturity. This is done for example in [1]. The SABR model is especially popular in the interest rate derivative markets.

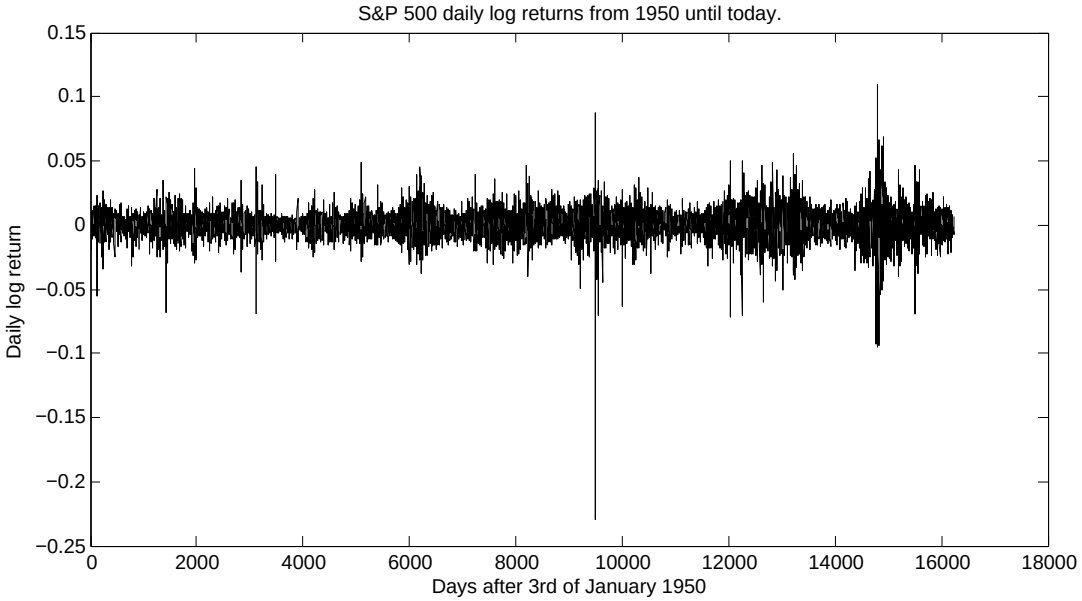

FIGURE 3.1: The daily log returns of S&P 500 from January 3rd 1950 until today. A slight clustering is visible. Data from finance.yahoo.com/ on 7/14/2014.

Vanna-Volga (VV)

The VV model is not based on assumptions on the underlying's price process but on a hedge of three big risks associated with the volatility of an option. The three risks are called vega, vanna and volga in the finance lingo. The vega is the sensitivity of the option value towards a change in implied volatility, the vanna is the sensitivity of the vega with respect to a change in the underlying's price process and finally the volga is the sensitivity of the vega with respect to changes in the implied volatility. A portfolio which hedges these three risks is created, and the Black-Scholes value of the option is adjusted by the value of this portfolio. The implied volatility can then be found by an inversion. The procedure is thoroughly explained in [24]. In [24], it is mentioned that the VV model does not guarantee convexity of the call prices, thus exposing the model to static arbitrage. The VV model is popular in the foreign exchange market.

3.2 SVI parametrizations and their interpretation

In [2], three equivalent versions of the SVI parameterization in Equation 3.1 are introduced; the raw, the natural and the jump-wing.

Natural parametrization

The natural parametrization of the SVI total implied variance reads

$$w_{\text{imp}}^{\text{SVI}}(x) = \Delta + \frac{\omega}{2}\left(1 + \eta\rho(x-\mu) + \sqrt{(\eta(x-\mu)+\rho)^2 + (1-\rho^2)}\right).$$

The natural parameterization is the functional form that appears as the limit of the Heston model. Unaesthetic expressions of ω and η in terms of the Heston parameters from Equation 3.2 can be found in [23], while ρ is simply the correlation between the Brownian motions driving the underlying's price process and the variance process and Δ and μ are vertical and horizontal shifts respectively. The natural parameterization is not as useful implementationwise as the two subsequent parameterizations.

Raw parametrization

The raw parametrization of SVI total implied variance reads

$$w_{\text{imp}}^{\text{SVI}}(x) = a + b\left(\rho(x-m) + \sqrt{(x-m)^2 + \sigma^2}\right).$$

This is the most frequently used version of the SVI parameterization in the litterature and it is also the original parameterization, introduced first in [1]. The effects on the volatility smile of the raw parameters are listed below:

- Increasing a results in a vertical translation of the smile in the positive direction.
- Increasing b decreases the angle between the put and call wing, i.e. tightens the smile.
- Increasing ρ results in a counter-clockwise rotation of the smile.
- Increasing m results in a horizontal translation of the smile in the positive direction.
- Increasing σ reduces the at-the-money curvature of the smile.

At-the-money is the financial term for the point $x = 0$, the point where the strike price is equal to the forward price. These properties of the raw parameters have been visualized in many articles, for example in [1].

Having bounds on the feasible region of these parameters is very useful when implementing the SVI model, since we will eventually have to do an optimization to find the, in some sense, best parameter set for an input data set. From the properties listed above, some obvious bounds for the raw parameters can be deduced. First of all, it is an empirical fact that volatility smiles have a positive at-the-money curvature. Hence we impose

$$\sigma > 0. \tag{3.3}$$

This bound has also computational advatages that will become clear to the reader in Chapter 4.3. The same fact also forces us to impose

$$b \geq 0. \tag{3.4}$$

Furthermore, we dont want an optimal parameter set to give a total implied variance curve that is systematically greater than the largest observed total variance. If $\{w_i\}$ are the observed total variances we want to fit our SVI to, we impose

$$a \leq \max_i\{w_i\}. \tag{3.5}$$

By the same reasoning, if $\{x_i\}$ are the moneynesses corresponding to the observed total variances, we impose

$$m \geq 2\min_i\{x_i\}, \tag{3.6}$$

$$m \leq 2\max_i\{x_i\}. \tag{3.7}$$

Finally, since ρ is the correlation between the Brownian motions driving the underlying's price process and the underlying's variance process we need to have

$$\rho \in [-1, 1]. \tag{3.8}$$

More bounds will be derived in Chapter 4.3, but from a more analytical point of view. A final note about the raw parameters is that they relate to the natural parameters in the following way,

$$\begin{aligned}(a, b, \rho, m, \sigma) &= \left(\Delta + \frac{\omega}{2}(1-\rho^2), \frac{\omega\eta}{2}, \rho, \mu - \frac{\rho}{\eta}, \frac{\sqrt{1-\rho^2}}{\eta}\right), \\ (\Delta, \mu, \rho, \omega, \eta) &= \left(a - \frac{\omega}{2}(1-\rho^2), m + \frac{\rho\sigma}{\sqrt{1-\rho^2}}, \rho, \frac{2b\sigma}{\sqrt{1-\rho^2}}, \frac{\sqrt{1-\rho^2}}{\sigma}\right).\end{aligned}$$

These relations give the possibility of an interpretation of the raw parameters in terms of the Heston model parameters.

Jump-Wing (JW) parameterization

The JW parameterization is defined in terms of the raw parameters,

$$\begin{aligned} v_\tau &= \frac{a + b\left(-\rho m + \sqrt{m^2+\sigma^2}\right)}{\tau}, \\ \psi_\tau &= \frac{1}{\sqrt{w_\tau}}\frac{b}{2}\left(\rho - \frac{m}{\sqrt{m^2+\sigma^2}}\right), \\ p_\tau &= \frac{1}{\sqrt{w_\tau}} b(1-\rho), \\ c_\tau &= \frac{1}{\sqrt{w_\tau}} bt(1+\rho), \\ \hat{v}_\tau &= \frac{1}{\tau}\left(a + b\sigma\sqrt{1-\rho^2}\right), \end{aligned} \tag{3.9}$$

where $w_\tau = v_\tau \tau$. Note that this parameterization depends explicitly on the time to maturity, τ. In [2], an inversion formula for Equation (3.9) is presented. The parameters are more tractable for traders than the two previous parameterization. More importantly for the aim of this thesis, there is literature on how to find conditions on them to prevent static arbitrage in the implied volatility surface. That will be the topic of Chapter 3.3. The tractability comes from the interpretations of the parameters, which are presented below.

- v_τ is the at-the-money implied variance. This is easily seen, since in the raw parameterization $v_\tau = w^{\text{SVI}}_{\text{imp}}(x)/\tau\Big|_{x=0}$.
- ψ_τ is the at-the-money implied volatility skew. This requires some calculation,

$$\begin{aligned} \partial_x \sqrt{\frac{w^{\text{SVI}}_{\text{imp}}(x)}{\tau}}\Bigg|_{x=0} &= \frac{1}{2\sqrt{w^{\text{SVI}}_{\text{imp}}(x)\tau}}\Bigg|_{x=0} \partial_x w(x)|_{x=0} \\ &= \frac{1}{2\sqrt{v_\tau \tau}} b\left(\rho + \frac{2x-2m}{2\sqrt{(x-m)^2+\sigma^2}}\right)\Bigg|_{x=0} \\ &= \frac{1}{\sqrt{w_\tau}}\frac{b}{2}\left(\rho - \frac{m}{\sqrt{m^2+\sigma^2}}\right) = \psi_\tau. \end{aligned}$$

- p_τ is the slope of the left wing, the part of the smile corresponding to out-of-the-money put options, scaled with the at-the-money total implied volatility,

$$\begin{aligned} \lim_{x\to-\infty} \frac{w^{\text{SVI}}_{\text{imp}}(x)}{x} &= \lim_{x\to-\infty} \frac{1}{x}\left(a + b\left(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\right)\right) \\ &= \lim_{x\to-\infty} \frac{a}{x} + b\left(\rho\left(1+\frac{m}{x}\right) + \sqrt{\left(1+\frac{m}{x}\right)^2 + \left(\frac{\sigma}{x}\right)^2}\right) \\ &= b(\rho+1) = p_\tau\sqrt{w_\tau} \end{aligned}$$

- c_τ is the slope of the right wing, the part of the smile corresponding to out-of-the-money call options, scaled with the at-the-money total implied volatility. The calculation is analoguous to that for p_τ.
- $\hat{v}_\tau$ is the minimum level implied variance. This was seen in the examination of the raw parameters.

3.3 The restriction: SSVI

We begin this section by introducing two standard definitions from the literature. Recall the conditions from Theorem 5.

Definition 15 (Butterfly arbitrage)**.** For a fixed and positive real τ, the implied volatility smile $\sigma_{\text{imp}}(\tau, K)|_{\tau=\tau_0}$ is free of butterfly arbitrage if and only if conditions (B) and (D) from Theorem 5 are satisfied.

Definition 16 (Calendar spread arbitrage)**.** An implied volatility surface $\sigma_{\text{imp}}(\tau, K)$ is free of calendar spread arbitrage if and only if condition (A) from Theorem 5 is satisfied.

Together with condition (C) from Theorem 5, absence of butterfly and calendar spread arbitrage is equivalent to absence of static arbitrage. As was discussed in the proof of Theorem 5, condition (C) will always be satisfied since we are using the implied volatility implied by the Black-Scholes call price formula.

With a parameterization, we could now calibrate the parameters to fit input volatility data, but we would like to do this in such a way that the output call prices are free of static arbitrage. In literature such as [2] and [25], the authors express hopelessness in front of the task to state conditions for the SVI parameters in order to guarantee no butterfly arbitrage. On the other hand, the easiness of eliminating calendar spread arbitrage is pronounced as a great strenth of the SVI model. The problem with butterfly arbitrage was partly solved in [2] where the authors propose a new parameterization for the total implied variance which they name Surface SVI (SSVI),

$$w_{\text{imp}}^{\text{SSVI}}(\tau, x) = \frac{\theta_\tau}{2}\left(1 + \rho x\phi(\theta_\tau) + \sqrt{(x\phi(\theta_\tau) + \rho)^2 + (1 - \rho^2)}\right).$$

The requirements on the new parameters θ_t and ϕ are that θ_t has to be positive and ϕ has to be a smooth function from $(0, \infty)$ to $[0, \infty)$. As earlier, ρ can take values in $[-1, 1]$. Note that the SSVI parameterization is the natural parameterization with its parameters restricted as $(\Delta, \mu, \rho, \omega, \eta) = (0, 0, \rho, \theta_t, \phi(\theta_t))$. Hence, the number of

variables that the parameterization depends on has gone down from five to three. Note also that θ_t is the at-the-money total implied variance.

Proposition 9. The JW parameters corresponding to the SSVI are

$$\begin{aligned} v_t &= \theta_t/t, \\ \psi_t &= \frac{1}{2}\rho\sqrt{\theta_t}\phi(\theta_t), \\ p_t &= \frac{1}{2}\sqrt{\theta_t}\phi(\theta_t)(1-\rho), \\ c_t &= \frac{1}{2}\sqrt{\theta_t}\phi(\theta_t)(1+\rho), \\ \hat{v}_t &= \frac{\theta_t}{t}\left(1-\rho^2\right). \end{aligned}$$

Proof. In the calculations below, text under the equality sign denotes change of parameterization. The leftmost term is always stated in the JW parameterization.

$$\begin{aligned} v_t \underset{\text{Raw}}{=}\ & \frac{1}{t}\left(a + b\left(-\rho m + \sqrt{m^2+\sigma^2}\right)\right) \\ \underset{\text{Natural}}{=}\ & \frac{1}{t}\left(\Delta + \frac{\omega}{2}\left(1-\rho^2\right) + \frac{\omega\eta}{2}\left(-\rho\left(\mu - \frac{\rho}{\eta}\right) + \sqrt{\left(\mu-\frac{\rho}{\eta}\right)^2 + \left(\frac{\sqrt{1+\rho^2}}{\eta}\right)^2}\right)\right) \\ \underset{\text{SSVI}}{=}\ & \frac{1}{t}\left(\frac{\theta_t}{2}\left(1-\rho^2\right) + \frac{\theta_t\phi(\theta_t)}{2}\left(\frac{\rho^2}{\phi(\theta_t)} + \sqrt{\frac{\rho^2}{\phi(\theta_t)^2} + \frac{1+\rho^2}{\phi(\theta_t)^2}}\right)\right) \\ =\ & \frac{1}{t}\left(\frac{\theta_t}{2} - \frac{\theta_t\rho^2}{2} + \frac{\theta_t\rho^2}{2} + \frac{\theta_t}{2}\right) = \frac{\theta_t}{t}, \end{aligned}$$

$$\begin{aligned} \psi_t \underset{\text{Raw}}{=}\ & \frac{1}{\sqrt{w_t}}\frac{b}{2}\left(\rho - \frac{m}{\sqrt{m^2+\sigma^2}}\right) \\ \underset{\text{Natural}}{=}\ & \frac{1}{\sqrt{v_t t}}\frac{\omega\eta}{4}\left(\rho - \frac{\mu - \frac{\rho}{\eta}}{\sqrt{\left(\mu-\frac{\rho}{\eta}\right)^2 + \left(\frac{\sqrt{1-\rho}}{\eta}\right)^2}}\right) \\ \underset{\text{SSVI}}{=}\ & \frac{1}{\sqrt{\theta_t}}\frac{\theta_t\phi(\theta_t)}{4}\left(\rho + \frac{\rho}{\phi(\theta_t)\sqrt{\frac{\rho^2}{\phi(\theta_t)^2} + \frac{1-\rho^2}{\phi(\theta_t)^2}}}\right) \\ =\ & \frac{1}{2}\rho\sqrt{\theta_t}\phi(\theta_t), \end{aligned}$$

$$
\begin{aligned}
p_t &\underset{\text{Raw}}{=} \frac{1}{\sqrt{w_t}} b(1-\rho) \\
&\underset{\text{Natural}}{=} \frac{1}{\sqrt{v_t t}} \frac{\omega\eta}{2}(1-\rho) \\
&\underset{\text{SSVI}}{=} \frac{1}{\sqrt{\theta_t}} \frac{\theta_t \phi(\theta_t)}{2}(1-\rho) \\
&= \frac{1}{2}\sqrt{\theta_t}\phi(\theta_t)(1-\rho),
\end{aligned}
$$

$$
\begin{aligned}
c_t &\underset{\text{Raw}}{=} \frac{1}{\sqrt{w_t}} b(1+\rho) \\
&\underset{\text{SSVI}}{=} \frac{1}{2}\sqrt{\theta_t}\phi(\theta_t)(1+\rho),
\end{aligned}
$$

$$
\begin{aligned}
\hat{v} &\underset{\text{Raw}}{=} \frac{1}{t}\left(a + b\sigma\sqrt{1-\rho^2}\right) \\
&\underset{\text{Natural}}{=} \frac{1}{t}\left(\Delta + \frac{\omega}{2}(1-\rho^2) + \frac{\omega\eta}{2}\frac{\sqrt{1-\rho^2}}{\eta}\sqrt{1-\rho^2}\right) \\
&= \frac{1}{t}\left(\Delta + \omega(1-\rho^2)\right) \\
&\underset{\text{SSVI}}{=} \frac{\theta_t}{t}(1-\rho^2).
\end{aligned}
$$

□

Using the SSVI, sufficient conditions for absence of static arbitrage are stated in [2].

Theorem 8. The SSVI is free of butterfly arbitrage if

(1) $\theta_\tau \phi(\theta_\tau)\,(1+|\rho|) < 4$,

(2) $\theta_\tau \phi(\theta_\tau)^2\,(1+|\rho|) \leq 4$.

Proof. The proof is omitted. It is based on a rewriting of condition (D) of Theorem 5, Durrleman's condition, and can be found in [2]. □

Theorem 9. The SSVI is free of calendar spread arbitrage if and only if

(1) $\partial_\tau \theta_\tau \geq 0$,

(2) $0 \leq \partial_{\theta_\tau}\left(\theta_\tau \phi(\theta_\tau)\right) \leq \frac{1}{\rho^2}\left(1+\sqrt{1-\rho^2}\right)\phi(\theta_\tau)$.

Proof. The proof is omitted. It is based on a rewriting of condition (A) of Theorem 5, $\partial_\tau w_{\text{imp}} \geq 0$, and can be found in [2]. □

Using Proposition 9, Theorem 8 can be translated into the JW parameterization.

Proposition 10. The SSVI is free of butterfly arbitrage if

(1) $\sqrt{v_\tau \tau}\max(p_\tau, c_\tau) < 2$,

(2) $(p_\tau + c_\tau)\max(p_\tau, c_\tau) \leq 2$.

Proof. Note that $(1 + |\rho|) < a$ if and only if $\max\big(1-\rho, 1+\rho\big) < a$. Using this, (1) is given by

$$\begin{aligned}\sqrt{v_\tau \tau}\max(p_\tau, c_\tau) &= \sqrt{\theta_\tau}\max\left(\frac{1}{2}\sqrt{\theta_\tau}\phi\left(\theta_\tau\right)(1-\rho), \frac{1}{2}\sqrt{\theta_\tau}\phi\left(\theta_\tau\right)(1+\rho)\right) \\ &= \frac{\theta_\tau \phi(\theta_\tau)}{2}\max\big(1-\rho, 1+\rho\big),\end{aligned}$$

and (2) is given by

$$\begin{aligned}(p_\tau + c_\tau)\max(p_\tau, c_\tau) &= \frac{1}{2}\sqrt{\theta_\tau}\phi(\theta_\tau)(1-\rho+1+\rho)\frac{\sqrt{\theta_\tau}\phi(\theta_\tau)}{2}\max\big(1-\rho, 1+\rho\big) \\ &= \frac{\theta_\tau \phi(\theta_\tau)^2}{2}\max(1-\rho, 1+\rho).\end{aligned}$$

□

Suppose that we have calibrated an SVI to some input data and know the calibrated JW parameters. Then [2] tells us that if we fix the JW parameters v_τ, ψ_τ and p_τ of this smile and the repick c_τ and $\hat{v}_\tau$ as

$$c_\tau = p_\tau + 2\psi_\tau, \tag{3.10}$$

$$\hat{v}_\tau = v_\tau \frac{4p_\tau c_\tau}{\left(p_\tau + c_\tau\right)^2}, \tag{3.11}$$

then by Proposition 10 we are guaranteed to have a smile that is free of butterfly arbitrage. Another way to do this is to fix v_τ, ψ_τ and c_τ and the repick p_τ and $\hat{v}_\tau$ using Equation (3.10) and Equation (3.11). This result will be a main component in the calibration algorithm, which is presented in the next chapter.

Chapter 4

Parameter calibration

In this chapter a method to calibrate the SVI parameters to market data will be presented. The line of work is close to, and inspired by, the work of [26] and [27]. First, the parameter bounds of Chapter 3.2 are complemented so that we have a lower and an upper bound for all the raw parameters. Then a new parameterization, called the Quasi-Explicit parameterization, is introduced. Finally the necessary optimization is reviewed and the calibration algorithm is summarized. Related to this chapter is Appendix A where the optimization algorithm that was used, the Nelder-Mead method, is described.

4.1 More parameter bounds

Recall the raw SVI parameterization of the total implied variance,

$$w_{\text{imp}}^{\text{SVI}}(x) = a + b\left(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\right).$$

If $\rho^2 \neq 1$, then $w_{\text{imp}}^{\text{SVI}}(x)$ has the unique minimum $a + b\sigma\sqrt{1-\rho^2}$. If $\rho = 1$, then $w_{\text{imp}}^{\text{SVI}}(x)$ is non-decreasing and if $\rho = -1$, then $w_{\text{imp}}^{\text{SVI}}(x)$ is non-increasing.
If σ goes to 0, then $w_{\text{imp}}^{\text{SVI}}(x)$ becomes a piecewise affine function of the form c_1x+c_2. This will cause the calibration to be an ill-posed problem. An example from [26] illuminates this. If we let σ go to 0 and let m be greater than the largest observed moneyness then

$$\begin{aligned} c_1 &= b(\rho-1), \\ c_2 &= a - bm(\rho-1), \end{aligned}$$

which leaves us with infinitely many choices of the triplet (a, b, ρ). Some bounds have to be imposed to prevent an optimization algorithm to get stuck. To make a good decision

we need to take a look at another case. Let σ go to infinity and let at the same time a go to negative infinity. Using a Taylor expansion and the $\sim$ notation from Chapter 2.5, we get for large σ and negative a that

$$\begin{aligned} w_{\text{imp}}^{\text{SVI}}(x) &= a + b\left(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\right) \\ &= -|a| + b\rho(x-m) + b\sigma\sqrt{1+\left(\frac{x-m}{\sigma}\right)} \\ &\sim -|a| + b\rho(x-m) + b\sigma\left(1+\frac{1}{2}\left(\frac{x-m}{\sigma}\right)\right). \end{aligned}$$

Letting $|a| = b\sigma$, we get

$$\lim_{\substack{\sigma\to\infty \\ a\to-\infty \\ |a|=b\sigma}} w_{\text{imp}}^{\text{SVI}}(x) = b\rho(x-m).$$

In this affine smile, we can find $b\rho$ and m, but not the triplet (b,ρ,m) uniquely. To solve these problems above, and gain stability in an optimization, we set a lower bound for a,

$$a \geq 0, \tag{4.1}$$

and reset the lower bound for σ from Equation 3.3,

$$\sigma \geq \sigma_{\min} > 0. \tag{4.2}$$

The value of $\sigma_{\min}$ is a user choice, but should be picked to be very small in comparison to typical parameter values.

No upper bound can be derived for σ, but a small integer always does the job. An implementation using for example

$$\sigma \leq 10 \tag{4.3}$$

would certainly cover all smiles appearing from market input data since such a large σ lifts the whole smile above 100% total implied variance, but if computer power is the limitation a smaller integer such as 1 is also a good choice for the upper bound.

In [28] the authors derive a necessary condition for the absence of dynamic arbitrage for the total implied variance that reads

$$|\partial_x w_{\text{imp}}(x)| \leq 4, \quad \forall x, \forall \tau.$$

This is a condition that is widely used in the SVI literature, because it yields an upper bound for b,

$$b \leq \frac{4}{\tau\,(1+|\rho|)}. \tag{4.4}$$

Let us end this section with a summary of all bounds that we have derived for the raw parameters:

$$\begin{aligned} 0 &\leq a \leq && \max_i\{w_i\}, \\ 0 &\leq b \leq && \frac{4}{\tau\,(1+|\rho|)}, \\ -1 &\leq \rho \leq && 1, \\ \min_i\{x_i\} &\leq m \leq && \max_i\{x_i\}, \\ 0 < \sigma_{\min} &\leq \sigma \leq && 10 \text{ (or another "large" number).} \end{aligned} \tag{4.5}$$

4.2 A new parameterization

The more parameters to calibrate, the more computing power is needed to do each optimization. In the SVI model there are 5 parameters and if a dimension reduction was possible it would speed up the calibration of the SVI. An observation in this line of though was made in [26]. Let

$$y(x) = \frac{x-m}{\sigma}.$$

Under this change of variables, the total implied variance in the raw parameterization reads,

$$\begin{aligned} w_{\text{imp}}^{\text{SVI}}(x) &= a + b\sigma\left(\rho y(x) + \sqrt{y(x)^2+1}\right) \\ &= \hat{a} + dy(x) + cz(x), \end{aligned} \tag{4.6}$$

where

$$\begin{aligned} \hat{a} &= a, & d &= \rho b\sigma, \\ c &= b\sigma, & z(x) &= \sqrt{y(x)^2+1}. \end{aligned}$$

Observe that given a value of (σ, m), $w_{\text{imp}}^{\text{SVI}}(x)$ in Equation (4.6) depends linearly on $\hat{a}, d, c$. This "split" of the parameter set, into a linear and a non-linear part, will be extremely useful in the optimization that is to follow. The parameterization in Equation (4.6) will henceforth be refered to as the Quasi-Explicit (QE) parameterization, inspired by the name of the paper where it was first defined.

The parameters $\hat{a}, d, c$ will from now on be refered to as the inner parameters and σ, m as the outer parameters.

Before we move on to the optimization, a word of caution to anyone who wants to implement this. In the article [26], the authors have incorrectly stated that $w_{\text{imp}}^{\text{SVI}}$ is the implied variance, instead of the *total* implied variance. Therefore, they have an extra τ factor throughout the paper.

4.3 Optimization

Translating the ten bounds in Equation 4.5 into bounds for the inner parameters $\hat{a}, d, c$ in the Quasi-Explicit parameterization, we see that the triplet $(\hat{a}, d, c)$ has a compact and convex domain in $\mathbb{R}^3$,

$$\mathcal{D} = \left\{ (\hat{a}, d, c) : \begin{cases} 0 \leq c \leq 4\sigma \\ |d| \leq c \\ |d| \leq 4\sigma - c \\ 0 \leq \hat{a} \leq \max\limits_i \{w_i\} \end{cases} \right\},$$

where $\{w_i\}$ is as in Equation (3.5) the input set of observed total variances. For a fixed pair (σ, m), we are facing the optimization

$$\min_{(\hat{a},d,c)\in\mathcal{D}} f_{x_i,w_i}(\hat{a}, d, c), \tag{4.7}$$

where f_{x_i,w_i} is the quadratic cost function,

$$f_{x_i,w_i}(\hat{a}, d, c) = \sum_{i=1}^{n} \left(\hat{a} + dy(x_i) + cz(x_i) - w_i\right)^2 .$$

Note that f_{x_i,w_i} is a smooth and convex function. Hence by a standad theorem from calculus, f has a unique minimum for every fixed (σ, m) on the convex and compact set $\mathcal{D}$. Furthermore, the gradient of f is linear in $(\hat{a}, d, c)$, so the solution can easily be computed analytically by calculating the inverse of a 3×3 equation system. This makes it possible to do an extremely fast optimization of the inner parameters given the outer parameters. A smart optimization algorithm should take advantage of this and involve an "outer optimization" that optimizes (σ, m) and an "inner optimization", that in every step of the outer optimization optimizes $(\hat{a}, d, c)$ over $\mathcal{D}$. Let $(\hat{a}^*, d^*, c^*)$ be the solution of Equation (4.7) and let (a^*, b^*, ρ^*) be the corresponding raw parameters.

Then the whole optimization can be stated as

$$\min_{(\sigma,m)\in\mathcal{E}} \sum_{i=1}^{n} \left(w_{(a^*,b^*,\sigma,\rho^*,m)} - w_i\right)^2, \tag{4.8}$$

where $\mathcal{E}$ is the domain of (σ, m). The situation for the outer optimization is not as bright as for the inner. First of all, f_{x_i,w_i} depends non-linearly on (σ, m). Secondly, f_{x_i,w_i} usually has multiple minima which makes gradient methods unreliable. An optimizer that can work around these concerns is the Nelder-Mead method with probabilistic restarts. It is treated in detail in Appendix A.

When we have arrived to a optimal set of raw parameters, we can check if this set gives an implied volatility smile that is free of static arbitrage by checking if it satisfies Durlemann's condition an condition (B) from Theorem 5. In [28], it is shown that (B) is equivalent to call prices tending to zero in the large strike limit, which is always the case in observed data. Therefore, it is sufficient to check Durrleman's condition. If the implied volatility smile satisfies it, the calibration is complete. If it does not a recalibration is needed. This is the second step of the optimization. By translating the calibrated parameters from raw to JW, we get a JW parameter set $(v_\tau, \psi_\tau, p_\tau, c_\tau, \hat{v}_\tau)$ that admits static arbitrage. We can then use Equation 3.10 and Equation 3.11 to repick c_τ and $\hat{v}_\tau$ and retrieve a new JW parameter set $(v_\tau, \psi_\tau, p_\tau, c'_\tau, \hat{v'}_\tau)$ which defines an implied volatility smile that is guaranteed to be free of static arbitrage. Now observe that Durrleman's condition is a continuous function of the parameters c_τ and $\hat{v}_\tau$. This guarantees us to find a solution to the optimization

$$\begin{aligned} \min_{(\bar{c}_\tau,\bar{v}_\tau)\in[c_\tau,c'_\tau]\times[\hat{v}_\tau,\hat{v'}_\tau]} \quad & \left(w^{\text{SVI}}_{\text{imp}}(x; c_\tau, \hat{v}_\tau) - w^{\text{SVI}}_{\text{imp}}(x; \bar{c}_\tau, \bar{v}_\tau)\right)^2, \\ \text{such that} \quad & \text{Durrleman's condition is satisfied.} \end{aligned} \tag{4.9}$$

The optimization (4.9) yields, in a sense, the SVI closest to the originally calibrated SVI that is free of static arbitrage. Since we may encounter multiple minima and non linearlity here aswell, the Nelder-Mead method can be implemented to solve this optimization.

The following flow chart summarizes the optimization.

1. Solve the optimization in Equation (4.8). Let the Nelder-Mead method work over $\mathcal{E}$ and for every pair (σ, m)

   i). Find the global minimum f^* of Equation (4.7). This minimum is guaranteed to exist and is attained at (a^*, b^*, ρ^*).

   ii). Check if (a^*, b^*, ρ^*) is in $\mathcal{D}$.

   Yes? Move to step 1.iv).

   No? Move to step 1.iii).

   iii) Find the minimum f^* of Equation (4.8) on the boundary $\partial\mathcal{D}$. This is guaranteed to exist and is attained at (a^*, b^*, ρ^*).

   iv) Store the solution parameter set (a^*, b^*, ρ^*) and the function minimum f^*.

2. Extract the parameter set corresponding to the smallest f^*. This is the solution $(a^*, b^*, \sigma^*, \rho^*, m^*)$ to the optimization Equation (4.8).

3. Check if the SVI generated from $(a^*, b^*, \sigma^*, \rho^*, m^*)$ satisfies Durrleman's condition.

   Yes? Optimization completed.

   No? Continue to step 4.

4. Generate the JW parameter set $(v_\tau^*, \psi_\tau^*, p_\tau^*, c_\tau^*, \hat{v}_\tau^*)$ corresponding to $(a^*, b^*, \sigma^*, \rho^*, m^*)$ through Equation (3.9).

5. Generate an SVI free of static arbitrage from $(v_\tau^*, \psi_\tau^*, p_\tau^*, c_\tau^*, \hat{v}_\tau^*)$ by repicking c_τ^* as c_τ' according to Equation (3.10) and repicking $\hat{v}_\tau^*$ as $\hat{v}_\tau'$ according to Equation (3.11).

6. Solve the optimization in Equation (4.9) with the Nelder-Mead method and arrive to a JW parameter set $(v_\tau^*, \psi_\tau^*, p_\tau^*, \bar{c}_\tau, \bar{v}_\tau)$.

7. The optimization is complete.

Chapter 5

Numerical experiments

5.1 The Vogt smile: the elimination of static arbitrage

The Vogt smile constitutes an example of an SVI parameter set that does not satisfy Durrleman's condition. The set is in the raw parameterization

$$(a, b, \sigma, \rho, m) = (-0.041, 0.1331, 0.4153, 0.306, 0.3586). \tag{5.1}$$

The constraint in Equation (4.1) implies that the parameter set in Equation (5.1) could never be an output from the calibration algorithm described in Chapter 4. It is however possible to generate an implied volatility smile and its corresponding Durrleman's condition, and these are presented in Figure 5.1. The right plot illustrates the violation of Durrleman's condition; the graph is not everywhere non-negative.

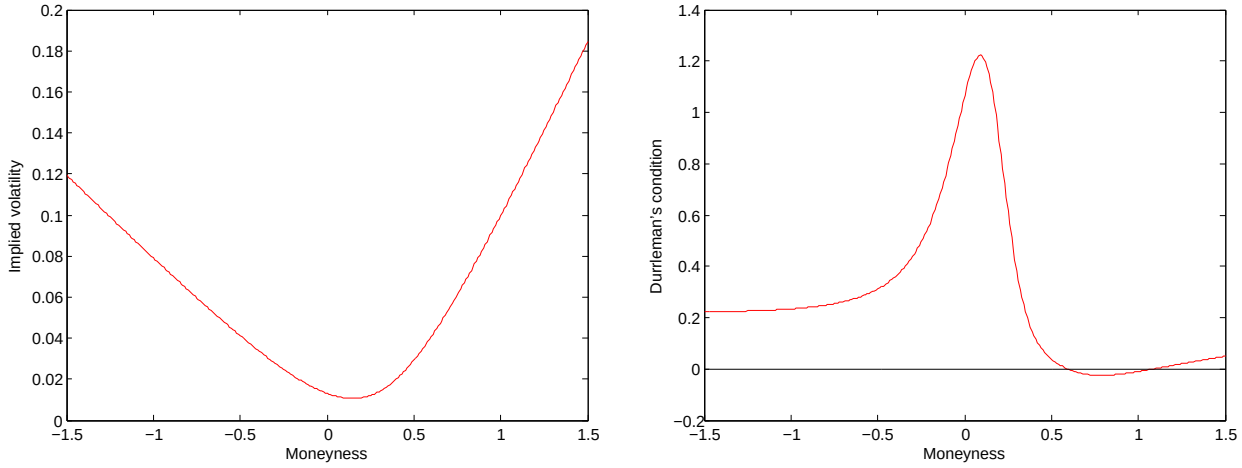

FIGURE 5.1: *Left plot:* The implied volatility corresponding to the parameters in Equation (5.1) is plotted against moneyness. *Right plot:* Durrleman's condition corresponding to the parameters in Equation (5.1) is plotted against moneyness.

The time to maturity is needed to generate the JW parameters corresponding to the raw parameters in Equation (5.1), but its value is not of importance and can be arbitrarily chosen. For simplicity, and to follow Gatheral in [2] as close as possible, the time to maturity is set to $\tau = 1$. The JW parameters corresponding to the parameters in Equation (5.1) are then

$$(v_\tau, \psi_\tau, p_\tau, c_\tau, \hat{v}_\tau) = (0.0131, -0.1366, 0.7472, 1.5826, 0.0106). \tag{5.2}$$

If $v_\tau, \psi_\tau, p_\tau$ are fixed and $c_\tau, \hat{v}_\tau$ are repicked according to Equation (3.10) and Equation (3.11), the new set of JW parameters gives a volatility smile that is guaranteed to be free of butterfly arbitrage. The new parameter values are

$$\begin{aligned} c'_\tau &= p_\tau + 2\psi_\tau = 0.4740, \\ \hat{v}'_\tau &= v_\tau \frac{4p_\tau c'_\tau}{(p_\tau + c'_\tau)^2} = 0.0124, \end{aligned}$$

and the set of JW parameters that guarantees absence of butterfly arbitrage is thus

$$(v_\tau, \psi_\tau, p_\tau, c'_\tau, \hat{v}'_\tau) = (0.0131, -0.1366, 0.7472, 0.4740, 0.0124). \tag{5.3}$$

The implied volatility and Durrleman's condition corresponding to this set of parameters is presented in Figure 5.2 together with the arbitrageable volatility from Figure 5.1.

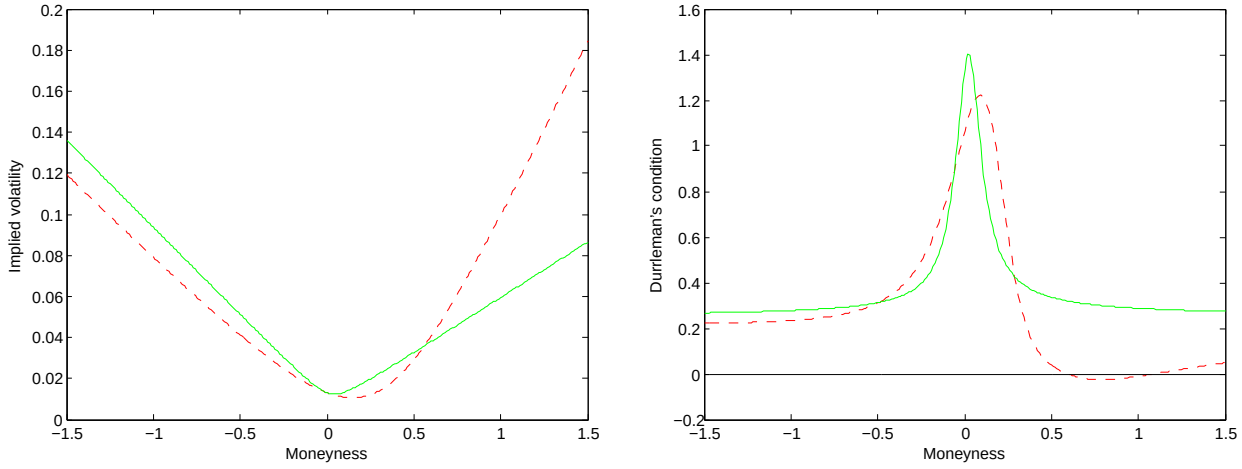

FIGURE 5.2: Solid lines correspond to the parameter set in Equation (5.3) and dashed lines correspond to the parameter set in Equation (5.1). *Left plot:* Implied volatility corresponding to the parameter sets is plotted against moneyness. *Right plot:* Durrleman's condition corresponding to the parameter sets is plotted against moneyness.

In the left plot of Figure 5.2, the effect of repicking c_τ is observable. Since the value of c_τ was lowered to about a thrid of its original value, the slope of the right wing of the smile is flatter. The right plot confirms that the SVI corresponding to the parameters

in Equation (5.3) do satisfy Durrleman's condition, since the green solid line is positive everywhere.

Recall the discussion in Chapter 4 leading to the optimization in Equation (4.9). The JW parameterization is continuous in all its parameters. Therefore, there exists a JW parameterization $(v_\tau, \psi_\tau, p_\tau, c_\tau^*, \hat{v}_\tau^*)$ between the one in Equation (5.2) and the one in Equation (5.3) that satisfies Durrleman's condition and is as close as possible, in some sense, to $(v_\tau, \psi_\tau, p_\tau, c'_\tau, \hat{v}'_\tau)$. This parameterization is found by optimizing $(c_\tau, \hat{v}_\tau)$ over the set $[0.4740, 1.5826] \times [0.0106, 0.0124]$ according to Equation (4.9). The Nelder-Mead method was used as optimization algorithm. The optimal result was found to be

$$\begin{aligned} c_\tau^* &= 1.3743, \\ \hat{v}_\tau^* &= 0.0110, \end{aligned}$$

and is illustrated in Figure 5.3.

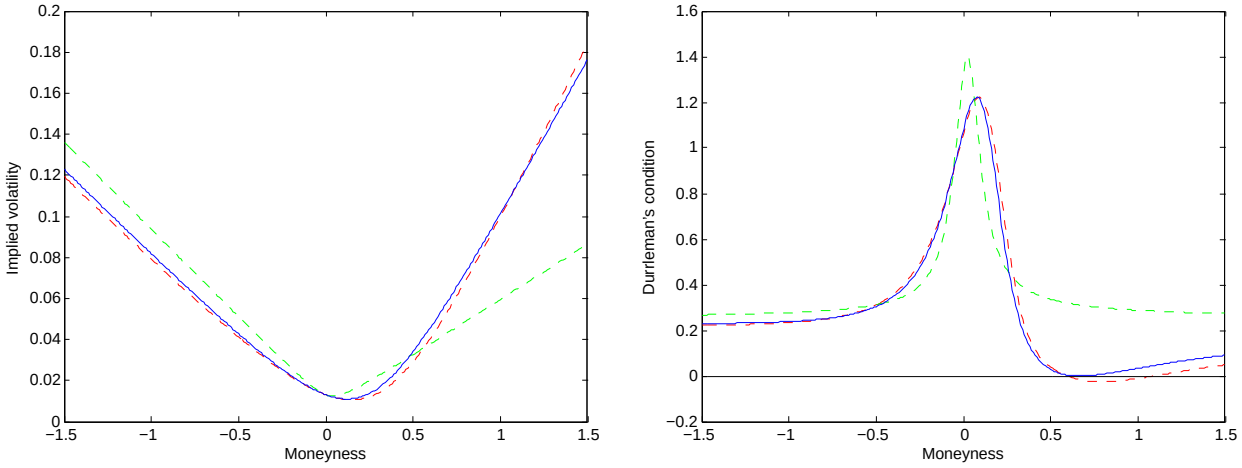

FIGURE 5.3: The solid lines correspond to the, according to Equation (4.9), optimal set of JW parameters, the dashed red lines that lie on top of the solid lines correspond to the original parameter set in Equation (5.1) and the deviant green dashed lines correspond to the parameter set in Equation (5.3). *Left plot:* Implied volatility corresponding to the parameter sets is plotted against moneyness. *Right plot:* Durrleman's condition corresponding to the parameter sets is plotted against moneyness.

5.2 Calibration to market data: the weighting of options

In this section, the SVI model will be fitted to raw market data. Before the calibration method is applicable, the market data has to be prepared. This preperation is treated in Section 5.2.1. One set of data on European put and call options written on the S&P 500 index, commonly refered to as SPX options. The S&P 500 index was discussed in Chapter 2.1.

5.2.1 Data preperation

The market data used consists of a set of strikes and for each strike, bid and ask prices for European call and put options written on that strike together with their corresponding traded volume. Also, the price of the underlying and the time to maturity is given. All prices are quoted in USD. Without loss of generality, let the current time be 0.

To apply the calibration, the forward price of the underlying and the interest rate of the market need to be known. The forward price of the underlying can be derived from the Put-Call parity,

$$C(\tau, K) - P(\tau, K) = S_0 - e^{-r\tau}K. \tag{5.4}$$

The forward price of the underlying, $F_{[0,\tau]} = S_0 e^{r\tau}$, is hence equal to the strike at which the call put and the put price are equal. This may happen in between data points. In that case the strike where the prices agree is found by a linear interpolation between the nearest data points. Note that since we now know the price of the underlying and the forward price of the underlying for some expiry time we also know the interest rate from today, $t = 0$, until that time.

All variables that are needed in the calibration are now at hand. It is time to look at the quality of the information each data point carries. The quality, and therefore importance, of the information of a data point will be judged by the trading volume of that option. If there are options on the market without volume, illiquid options, then these are removed from the data set. The illiquid options provide us with prices at which nobody is trading and therefore they carry no information about the real price of the option. In the next step, call options that have a strike price lower than the market price of the underlying and put options with a strike price that is higher than the market price of the underlying are discarded. Those that are kept, the out-of-the-money options, tend to be more precise in pricing since they are bearing higher risk than the ones we discarded, the in-the-money options. The out-of-the-money put option prices are then translated into in-the-money call option prices via Equation (5.4) to create a

discrete version of a call function. At each strike the latest ask and bid price for each option is avaliable. The mid price for the option is the mean of the ask price and the bid price. By inverting the Black-Scholes formula, implied volatilities corresponding to the mid, ask and bid call prices can be obtained. An example of a prepared data set is shown in Figure 5.4.

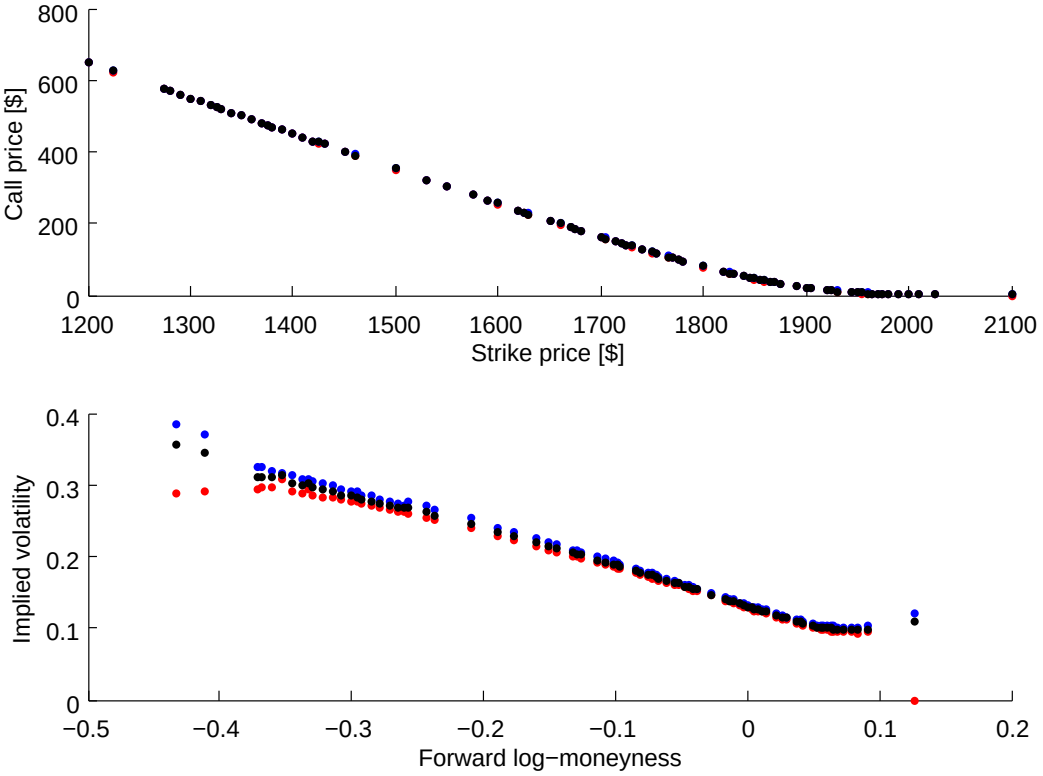

FIGURE 5.4: Market data from liquid, out-of-the-money options written on the SPX with underlying price $1857.6. *Upper plot:* Prices for call options plotted against strike prices. The ask, mid and bid prices are not distinguishable on this scale. *Lower plot:* Implied volatilities for the same option that was treated in the upper plot plotted against moneyness. Here we can distinguish between the bid implied volatility, which are plotted as red dots, mid implied volatility, black dots, and ask implied volatility, blue dots.

5.2.2 SPX options

This section will take the reader through a trail of thought of the author during a test of the calibration method on real market data. As the reader will notice, the first results are not very attractive. These results serve as motivation to further elaborate the calibration, by introducing option weights and caps. The gist of this section is that there are several ways to put weights on the data points, and that the best way to do it, using the calibration method from Chapter 4, depends on the data set.

The first choice that is faced is what implied volatility data to use in the calibration. Ask, bid and mid implied volatility data is avaliable, but all values in the interval between the bid and the ask, called the bid-ask spread, are feasible. The SVI in this calibration are calibrated to mid implied volatility. In Figure 5.5, the SVI fit together with bid (smaller) and ask (larger) implied volatilites are presented. Only every third ask and bid implied volatility is plotted, so that it is possible to distinguish between data point in the graphs.

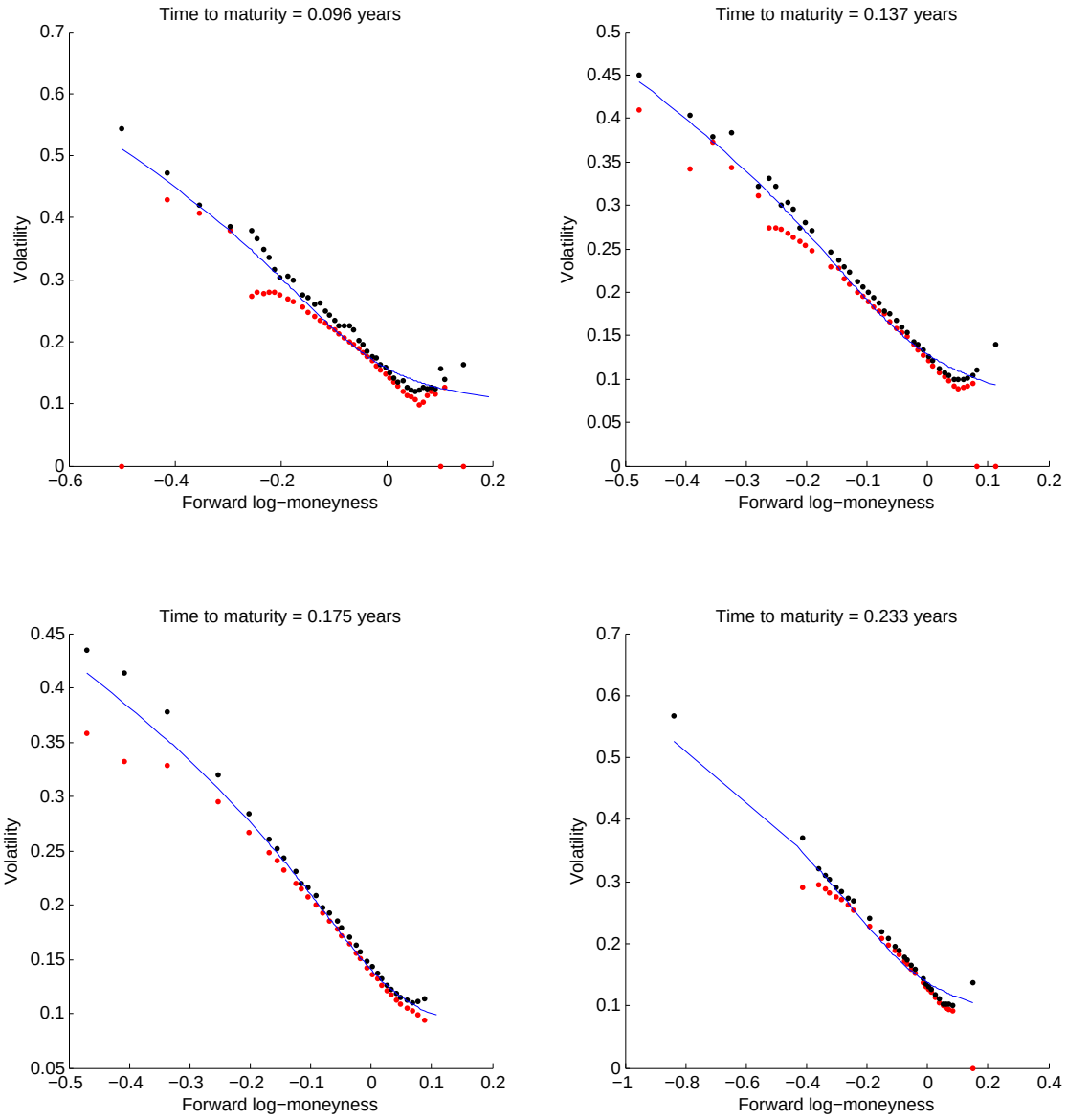

FIGURE 5.5: Implied volatility plotted against moneyness for four different times to maturity. The red dots are bid implied volatility, the blue line is the SVI fit to mid implied volatility and the black dots are ask implied volatility. Only every third ask and bid implied volatility is plotted.

The SVI implied volatility smile looks stiff, and fails to fit the implied volatility well around at-the-money for three of the four times to maturity. This calibration error is quantified in Figure 5.6 in terms of fractions of the total trading volume. The option volume traded differs across the strikes. The total trading volume at a specific time to maturity is the options volume that has been traded across all strikes. The volume at all the strikes where the SVI call prices lie within the bid-ask call market price spread are summed. This sum divided by the total trading volume is chosen as a measure of how well the SVI fits the market data. The fractions corresponding to the SVI implied volatility in Figure 5.5 are plotted in Figure 5.6, and for the three times to maturity where the SVI implied volatility did not manage to fit the at-the-money options, the fractions are small.

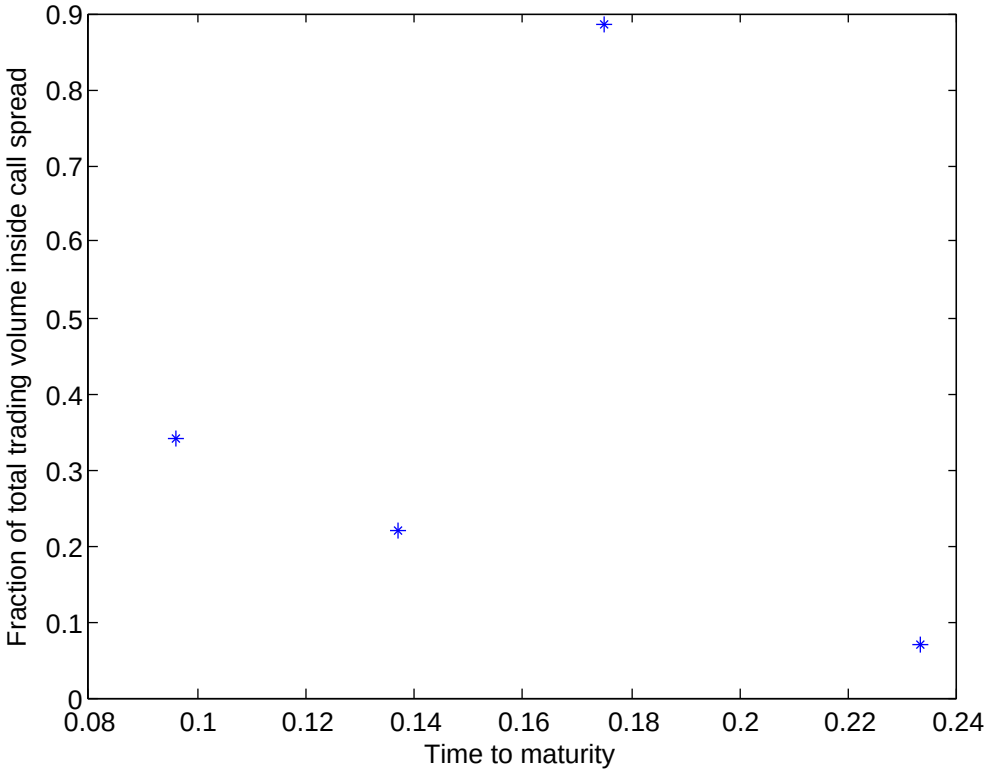

FIGURE 5.6: The fraction of the total trading volume that the SVI implied volatility manages to fit within the bid-ask call market price spread is plotted against time to maturity for the implied volatility smile that has been fitted.

The connection between the low values in Figure 5.6 and the bad fit around at-the-money is explained by the fact that most of the traded volume lies in the moneyness interval $[-0.1, 0.05]$. In Figure 5.7, the relative trading volume, that is the fraction of the total trading volume that is traded at a specific strike, is plotted against moneyness using data from the shortest time to maturity. Also, the accumulated relative trading

volume, that is all the volume traded at all strikes smaller or equal to a specific strike, is plotted against moneyness. The concentration of volume is clearly visible in the steep ascent of the thick red line, about 80% of the trading volume was lying in the moneyness interval $[-0.1, 0.05]$.

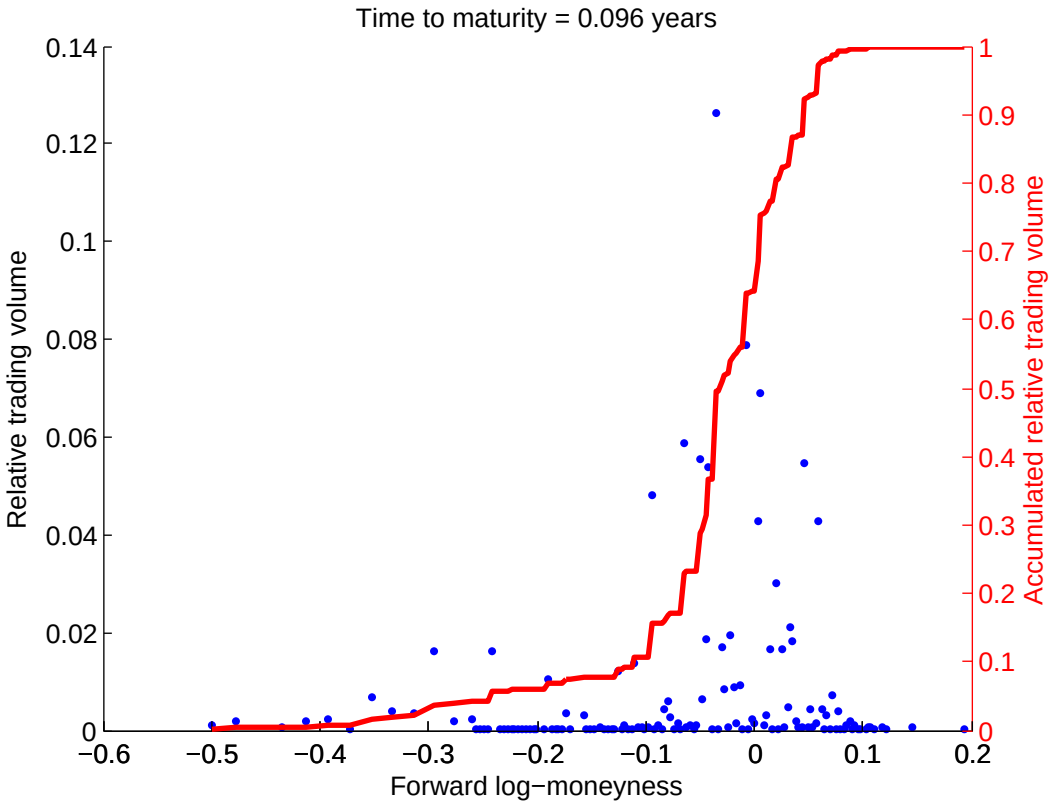

FIGURE 5.7: The concentration of trading volume around the origin.
Left axis: The relative trading volume, blue dots, is plotted against moneyness.
Right axis: The accumulated relative trading volume, red line, is plotted against moneyness.

The importance of the high volume options is not considered in the calibration method of Chapter 4 where all options were treated equally. Introducing weights on the options in the objective functions of the optimization reflecting their importance comes to mind as a natural solution to this. Therefore Equation (4.8) is replaced with the optimization

$$\min_{(\sigma,m)\in\mathcal{E}} \sum_{i=1}^{n} W_i \left(w_{(a^*,b^*,\sigma,\rho^*,m)} - w_i \right)^2, \tag{5.5}$$

where $W_1, \ldots, W_n$ are the weights. What weights to use is a choice of the implementor. The relative trading volume of the options, or some increasing function of it, is a strong candidate. The relative trading volume reflects each individual option's fraction of the total market volume. Another candidate is the vega of the option, or some increasing

function of it. The vega is a measurement of the sensitivity of the value of an option to changes in the implied volatility of the underlying. It is defined as the derivative of the option price with respect to implied volatility. In Figure 5.8, vega is plotted together with the relative trading volume against moneyness for the option data at the second shortest time to maturity.

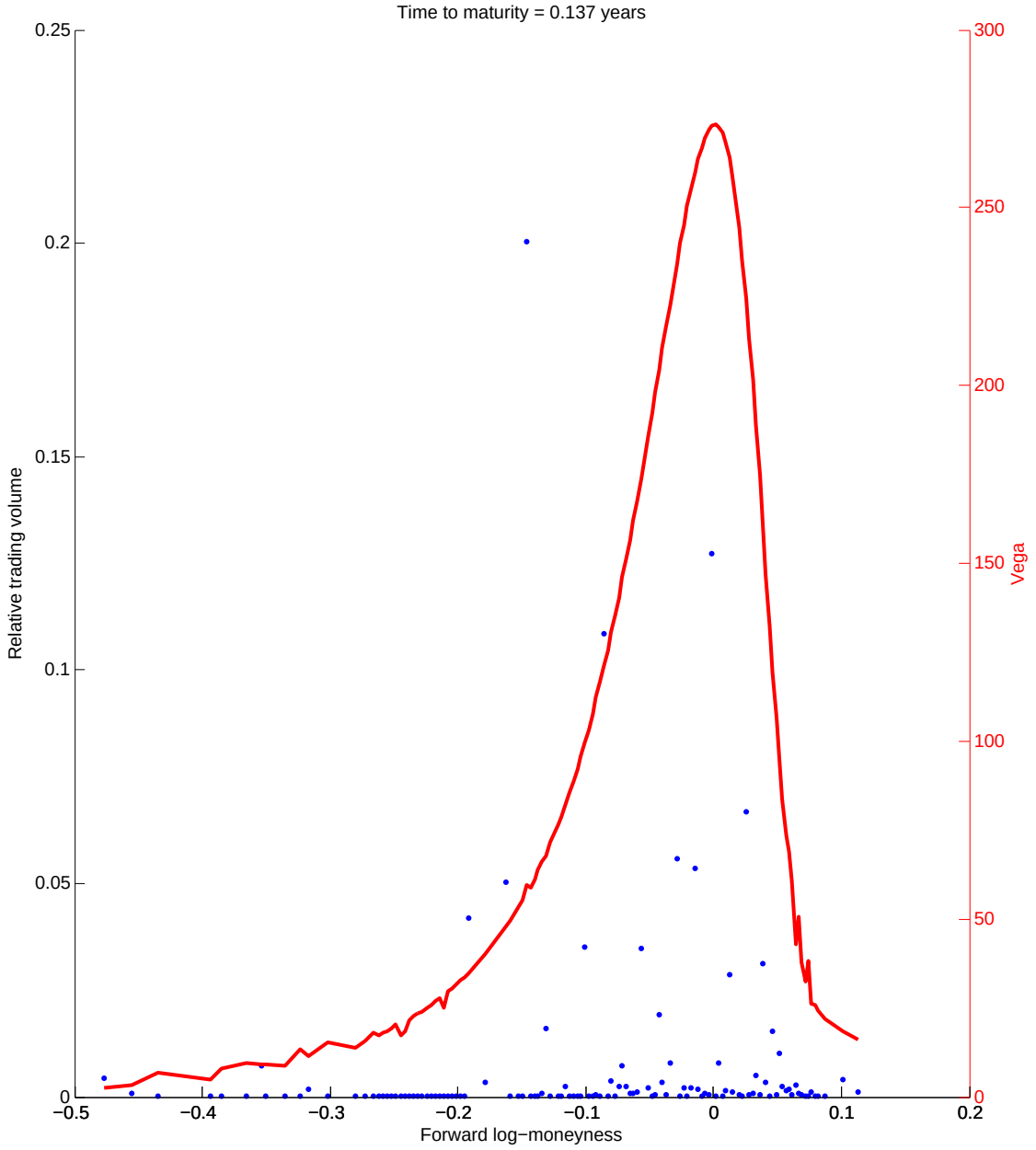

FIGURE 5.8: Two possible choices of option weights. *Left axis:* The relative trading volume is plotted against moneyness. Most of the volume is traded between -0.1 and 0.05, but some outliers can be seen on the negative part of the x-axis. *Right axis:* The vega of the option is plotted against moneyness. The sensitivity of the option price to changes in volatility is greatest around the origin and decays quickly in both directions.

The shape of the vega in Figure 5.8 is characteristic for the vega of a call option, hence prices of call options are not sensitive to large changes in volatility at strikes far from at-the-money. This implies that if the same volume is traded at some strike far from at-the-money and at some strike at-the-money, it is more important to fit the at-the-money option well. In this calibration the weights were chosen to be the product between the relative trading volume and the vega,

$$W_i = \text{Relative trading volume at strike number } i \times \text{Vega at strike number } i. \quad (5.6)$$

One more question needs to be settled. Since the SVI only has five parameters, it will never be able to fit the mid implied volatility for all options. If fewer options are used in the fit, the SVI will be able to fit these options better. Assume that the N heaviest options are used in the calibration. How does the SVI overall performance depend on N? Does the SVI nessecarily do better just because we calibrate it to more data points? In Figure 5.9, the fraction of the total trading volume that the SVI fitted inside the call spread is plotted against N, the number of options in heavy to light order, used in the calibration.

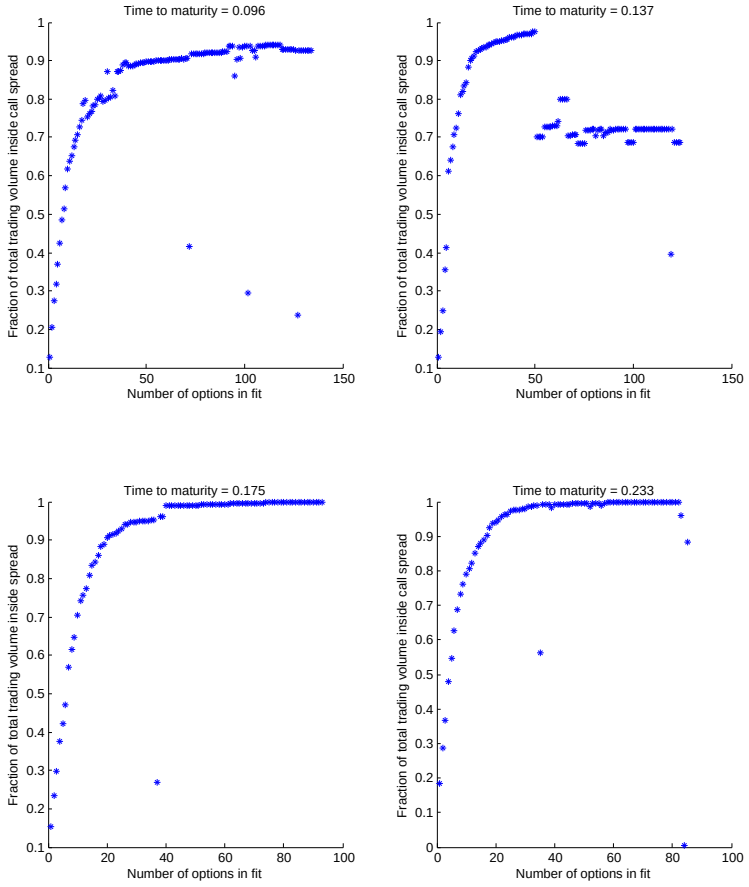

FIGURE 5.9: The fraction of the market which was fitted inside the call spread as a function of the number of options used in the calibration.

For three of the maturities, the general pattern is that the overall performance becomes better and better the more options we use. The top right corner plot breaks this pattern, and the shape of that plot is sometimes seen in market data. In the light of the plots in Figure 5.9 we chose to calibrate the SVI after the $N_{\max}$ heaviest options, where $N_{\max}$ is the number where the graphs in Figure 5.9 reach their maximum. To summarize,

- all options are weighted according to Equation (5.6),
- the options are ordered according to their weights,
- the $N_{\max}$ heaviest options are used in the calibration, where $N_{\max}$ is the number at which the experiment illustrated in Figure 5.9 attains its maximum.

The result of this new calibration is presented in Figure 5.10. Compared to the fits in Figure 5.5, the new calibration looks better. By looking at the fraction of total trading volume that the new calibration fit inside the bid-ask call price spread, this is confirmed. The fractions are presented in Figure 5.11. The fit is still not perfect for short times to maturity, but this is expected. Short lived options are generally much more difficult to fit.

Figure 5.11 shows that still not all the volume is hit by the SVI. It is interesting to know which options that were fitted outside the bid-ask spread and how far away from the spread these misfitted prices are. Where are the thieves lurking that the SVI misses, and how dangerous are they? This question is answered by Figure 5.12. Only for the shortest time to maturity does the SVI miss the bid-ask spread close to the at-the-money. Otherwise, the misses are far out-of-the-money.

Finally, no butterfly arbitrage was detected for the SVI implied volatilitiy in Figure 5.10. This was an expected result, since the SPX is one of the most liquid options on the market. In Figure 5.13, Durrleman's condition corresponding to the SVI fits is plotted. Is is everywhere positive, which implies absence of butterfly arbitrage.

We have now seen the calibration method in action. It is able to eliminate static arbitrage from an implied volatility smile given by a set of SVI parameters and it manages to produce good fits to real market data. Thus this research has generated a calibration model that fulfills the purpose of this thesis within the stated delimitations. We let the following plots conclude this chapter and the thesis.

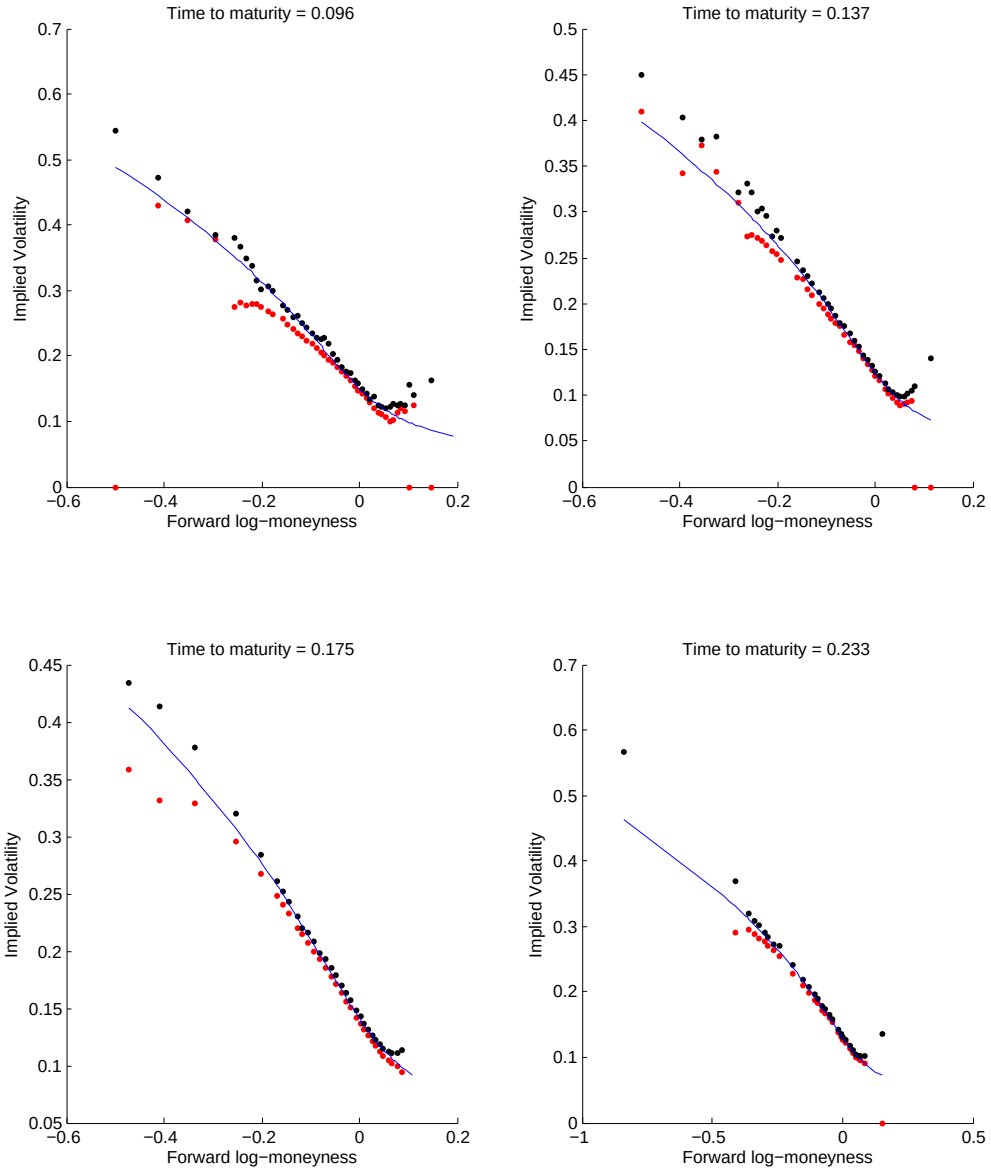

FIGURE 5.10: New SVI implied volatility fit using weights and caps in the calibration. The red dots are bid implied volatility, the blue line is the SVI fit to mid implied volatility and the black dots are ask implied volatility. Only every third ask and bid implied volatility is plotted.

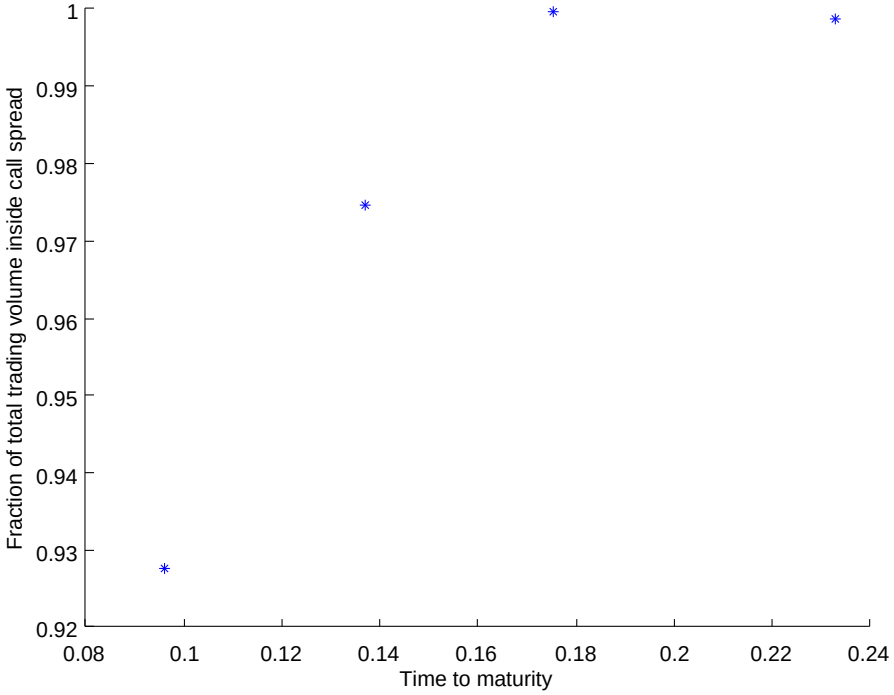

FIGURE 5.11: The fraction of the total trading volume that the new SVI implied volatility, calibrated with weights and caps, manages to fit within the bid-ask call price spread is plotted against time to maturity for the implied volatility smile that has been fitted.

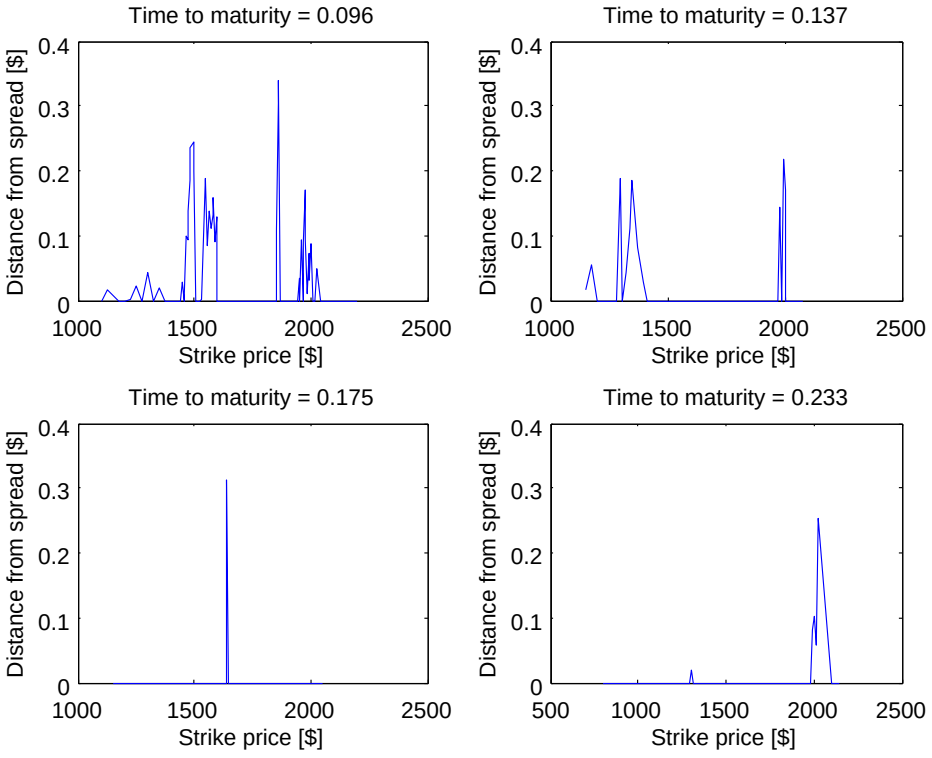

FIGURE 5.12: The amount that the SVI generated call prices miss the bid-ask call price spread at each strike plotted for the four times to maturity.

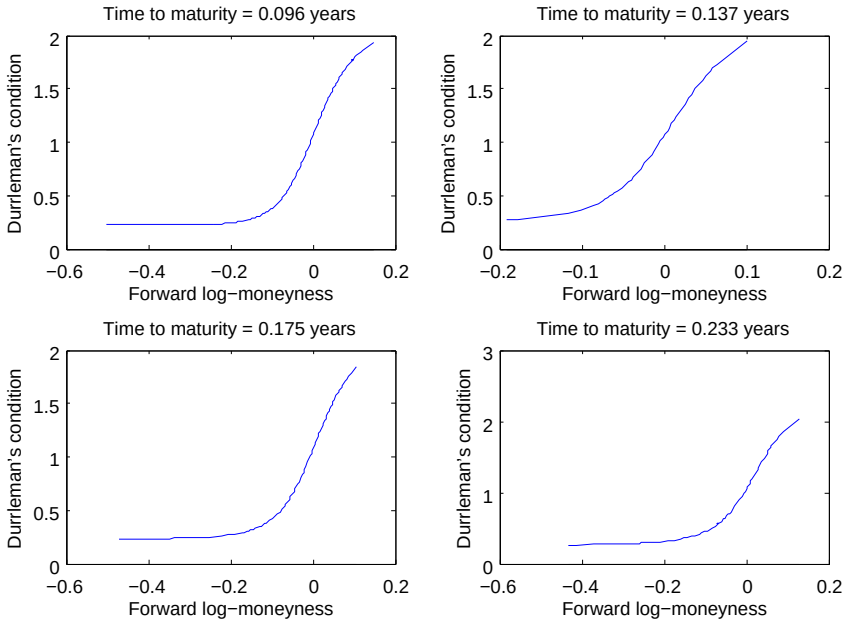

FIGURE 5.13: Durrleman's condition plotted corresponding to the new SVI implied volatility for the four times to maturity.

Bibliography

[1] Gatheral, J. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. *Presentation at Global Derivatives & Risk Management, Madrid*, 2004.

[2] Gatheral, J.,Jacquier, A. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.

[3] Wilmott, P. Paul Wilmott on Quantitative Finance. *John Wiley & Sons*, 2000.

[4] Gatheral, J. The Volatility Surface: A Practitioner's Guide. *Wiley Finance*, 2006.

[5] Madan, D.B., Yor, M. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 2002.

[6] Carr, P., Geman, H., Madan, D.B., Yor, M. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13(3):345–382, July 2003.

[7] Kellerer, H.G. Markov-Komposition und eine Anwendung auf Martingale. *Mathematische Annalen*, 198(3):99–122, 1972.

[8] Durrett, R. Probability, Theory and Examples. *Cambridge University Press*, 2010.

[9] Hirsch, F., Roynette, B. A new proof of Kellerer's theorem. *ESAIM: Probability and Statistics*, 16:48–60, 2012.

[10] Hirsch, F., Roynette, B., Yor, M. Kellerer's theorem revisited. *Unpublished*, 2012.

[11] Föllmer, H., Schied, A. Stochastic Finance: An Introduction in Discrete Time, 2nd ed. *Walter de Gruyter GmbH & Co*, 2004.

[12] Rooper, M. Arbitrage Free Implied Volatility Surfaces. *Preprint; avaliable at http://www.maths.usyd.edu.au/u/pubs/publist/preprints/2010/roper-9.pdf.*

[13] Hirsch, F., Profeta, C., Roynette, B., Yor, M. Peacocks and Associated Martingales, with Explicit Constructions. *Springer Milan*, 2011.

[14] Lowther, G. Limits of one-dimensional diffusions. 37:78–106, 2009.

[15] Durrleman, V. A note on initial volatility surface. *Unpublished*, February 2003.

[16] Fengler, R. Arbitrage-Free Smoothing of the Implied Volatility Surface. *Quantitative Finance*, 9(4):417–428, 2009.

[17] Broadie, M., Cvitanic, J., Soner, H.M. Optimal Replication of Contingent Claims under Portfolio Constraints. *The Review of Financial Studies*, 11(1):59–79, 1998.

[18] Lee, R. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.

[19] Benaim, S., Friz, P., Lee, R. On Black-Scholes Implied Volatility at Extreme Strikes. *Frontiers in Qualitative Finance: Volatility and Credit Risk Modeling*, 2008.

[20] Hagan, P.S.,Kumar, D.,Lesniewski, A.S.,Woodward, D.E. Managing Smile Risk. *Wilmott Magazine*, 1:84–108, 2002.

[21] Damghani, B.M., Kos, A. De-arbitraging With a Weak Smile: Application to Skew Risk. *Wilmott*, (64):40–49, 2013.

[22] Hodges, S., Zhao, B. Parametric modeling of implied smile functions: a generalized SVI model. *Review of Derivatives Research*, 16:53–77, april 2013.

[23] Gatheral, J., Jacquier, A. Convergence of Heston to SVI. *Quantitative Finance*, 11 (8):1129–1132, 2011.

[24] Castagno, A.,Mercurio, F. The vanna-volga method for implied volatilities. *Risk Magazine*, 20(1):106–111.

[25] Guo, G., Jacquier, A., Martini, C., Neufcourt, L. Generalised arbitrage-free SVI volatility surfaces. *Preprint at arXiv:1210.7111*, 2013.

[26] Zeliade Systems. Quasi-explicit calibration of Gatheral's SVI model. *Zeliade White Papers*, 2009.

[27] De Marco, S. On Probability Distributions of Diffusions and Financial Models with non-globally smooth coefficients. *Ph.D. thesis at Université Paris-Est*, November 2010.

[28] Rogers, L.C.G. Tehranchi, M.R. Can the implied volatility surface move by parallel shifts? *Finance and Stochastics*, 14(2):235, 2010.

[29] Leursen, M.A., Le Riche, R. Globalized Nelder-Mead method for engineering optimization. *Computers & Structures*, 82(23-26):2251–2260, 2004.

[30] Carr, P., Madan, D.B. Towards a Theory of Volatility Trading. *Volatility, risk publications*, Jarrow, R., 1998.

Appendix A

Nelder-Mead method

The Nelder-Mead method is a simplex method for finding a local minimum of a function of several variables. It employs a pattern search, which will be described below, that compares function values at the vertices of the simplex. The objective function is evaluated at all vertices. The worst vertex, that is the vertex that gives the highes objective function value, is rejected and replaced by a new vertex. A new simplex is formed and a new comparison is made. After itterating this algorithm, the simplices will zoom in on a local minium and will have decreased in size so that eventually, they approximate the local minimum. By rerunning the algorithm from various initial points, all local minima can be found and hence also the global minimum.

Probabilistic restarts

Local optimizers can make up a global search when repeatedly started from different initial points. In the SVI calibration, the Nelder Mead method is applied to a non-linear function that can have multiple minima. Therefore, a global search is necessary. In [29], the authors introduce a method to choose the initial points. Their procedure involve a random sampling of a set of points, and then a selection of the initial point from this set by measuring the distance, in some sense, to the previous initial points. This procedure saves us from the task of defining a grid of initial points, which would have been the case in a deterministic search. Thus we do not have to know how many times we have to restart the algorithm in advance. The procedure also uses information about past searches, which a pure random choice of initial points would not. Thus the precedure has a low chance of choosing a initail point close to some old initial point, and therefore finding the same local minimum twice.

If we have already sampled a set $\{x_i\}_{i=1}^N$ of initial points for the algorithms, let the probability of sampling a point $x \in \mathbb{R}^n$ to be the initial point be

$$p(x) = \frac{1}{N}\sum_{i=1}^{N} p_i(x), \tag{A.1}$$

where

$$p_i(x) = \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(x-x_i)^T\Sigma^{-1}(x-x_i)}. \tag{A.2}$$

Here, n is the dimension of $\mathbb{R}^n$, the space we run the optimize in, and Σ is the covariance matrix. It is suggested in [29] to estimate the covariances, σ_j^2, by

$$\sigma_j^2 = \alpha(x_j^{\max} - x_j^{\min})^2, \qquad j = 1, \ldots, n,$$

where α is a positive parameter that controls the width of the Gaussian bell in Equation (A.2) and $x_j^{\max}$ and $x_j^{\min}$ are the upper and lower bounds for the feasible region in the j:th direction. A value for α that works will with the SVI implementation is 0.1.

By casting points randomly from a uniform distribution on the feasible set, we get candidates to the initial point for the algorithm. We can then order these candidates by evaluating (A.1) at all the casted points. The point which generates the smallest value of Equation (A.1), the point which has the least probability of being casted given earlier initail points, is selected as initial point.

The algorithm, step by step

Let s_0 be the initial simplex with vertices $x_1, \ldots, x_n$. Let f be the objective function we want to minimize.

i) Reorder the vertices with respect to their objective function values, that is if

$$f(x_1) \leq \cdots \leq f(x_n)$$

then x_1 is the best point and x_n is the worst point.

ii) Calculate the center of mass of $s_0 \backslash x_n$, the initial simplex without the worst point,

$$x_c = \frac{1}{n-1}\sum_{i=1}^{n-1} x_i$$

iii) Calculate the reflection, x_r of the worst point, x_n with respect to x_m,

$$x_r = x_m + \rho(x_m - x_n).$$

The choice of ρ does not necessarly have to be 1, but has to be positive. If x_r is better that the second worst point x_n, but not better than the best point x_1, that is

$$f(x_1) \leq f(x_r) \leq f(x_n - 1),$$

then let the new simplex s_1 be s_0 but with the worst point x_n replaced by x_r. In this case, restart the algorithm from step i). Otherwise, move on to step iv).

iv) If x_r is better than the best point x_1, that is $f(x_r) < f(x_1)$, then calculate the expansion x_e of x_n with respect to x_m,

$$x_e = x_m + \eta(x_m - x_n), \quad \eta > \rho.$$

If x_e is better than x_r, that is $f(x_e) < f(x_r)$, then let the new simplex be s_0 but with the worst point x_n replaced by x_e and restart the algorithm from step i). If that is not the case, then the new simplex is s_0 with the worst point x_n replaced by x_r.

v) Now $f(x_r) \geq f(x_{n-1})$, so calculate the contraction x_c of x_n through x_m,

$$x_c = x_m - c(x_m - x_n), \quad 0 < c < \rho.$$

If x_c is better than the worst point, that is $f(x_c) < f(x_n)$, then let the new simplex be s_0 but with the worst point x_n replaced by x_c and restart the algorithm from step i). If that is not the case, go to step vi).

vi) Replace all vertices except the best one, $\{x_i\}_{i=2}^n$, with their contractions towards the best one, $\{x_i'\}_{i=2}^n$,

$$x_i' = x_1 + \sigma(x_i - x_1).$$

Now let the new simplex be the convex hull of $\{x_1\} \cup \{x_i'\}_{i=2}^n$ and return to step i).

Figure A.1 illustrates one iteration of the algorithm.

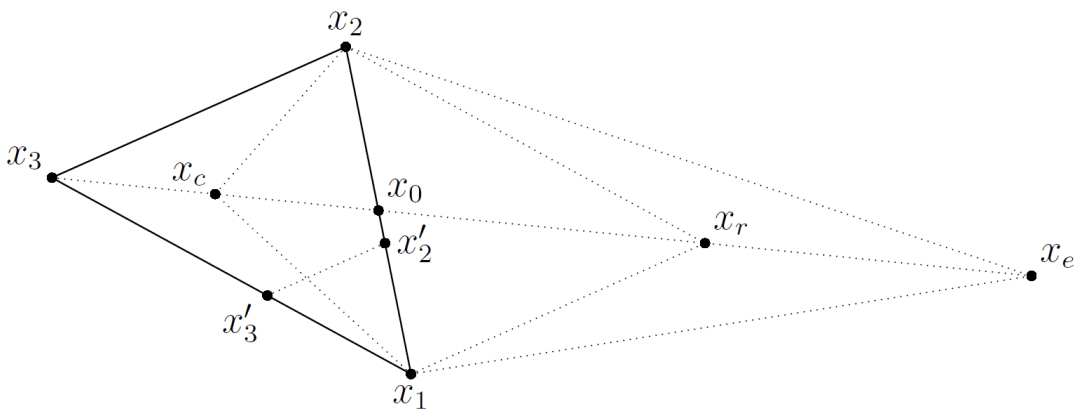

FIGURE A.1: One iteration of the Nelder-Mead algorithm with $\rho = 1, \eta = 2, c = 1/2$ and $\sigma = 1/3$.

Appendix B

Spanning a payoff with bonds and options

The following spanning formula was derived in [30] by Carr and Madan.

Using Dirac's delta function, we can for any function $f(s)$ write for some $k \geq 0$.

$$\begin{aligned} f(s) &= \int_0^\infty f(K)\delta(s-K)\mathrm{d}K \\ &= \int_0^k f(K)\delta(s-K)\mathrm{d}K + \int_k^\infty f(K)\delta(s-K)\mathrm{d}K, \end{aligned}$$

Recall that $\int_a^b \delta(s-K)\mathrm{d} = H(s<K)|_a^b$, where H is the Heaviside function. Integrating the previous equation by parts gives

$$\begin{aligned} f(s) &= f(K)H(s<K)|_0^k - \int_0^k f'(K)H(s<K)\mathrm{d}K \\ &\quad + f(K)H(s\geq K)|_k^\infty + \int_k^\infty f'(K)H(s\geq K)\mathrm{d}K. \end{aligned}$$

Integrating the integrals by parts one more time yields,

$$\begin{aligned} f(s) &= f(K)H(s<K)|_0^k + f'(K)(K-s)^+\big|_0^k + \int_0^k f''(K)(K-s)^+\mathrm{d}K \\ &\quad + f(K)H(s\geq K)|_k^\infty + f'(K)(s-K)^+\big|_k^\infty + \int_k^\infty f''(K)(s-K)^+\mathrm{d}K. \end{aligned}$$

Making the economic interpretation of $f(s)$ as a payoff on a contract written on an underlying s, and K as the strike, we can impose the boundary condition $f(0) = 0$. Any

other payoff at 0 would create arbitrage. Then the last equation becomes

$$\begin{aligned} f(s) &= f(k) + f'(k)\left((s-k)^+ - (k-s)^+\right) \\ &\quad + \int_0^k f''(K)(K-s)^+ \mathrm{d}K + \int_k^\infty f''(K)(s-K)^+ \mathrm{d}K. \end{aligned}$$

Appendix C

Call price derivative

Using the notation from Chapter 2.4, this appendix presents the calculations behind all used derivatives of the Black-Scholes call price function.

Since we assume that implied volatility is non-constant over strikes, we have to be careful when calculating the derivative of the call price with respect to strike. All dependencies and subscripts will be dropped in the calculations that follow to avoid messyness, so lets de a review of the dependencies of all involved functions,

$$\begin{aligned}
C^{\mathrm{B}} : (\tau, K, \tau\sigma^2_{\mathrm{imp}}; F, r, t) &\mapsto F_{[t,t+\tau]}\mathcal{N}(d_1) - K\mathcal{N}(d_2), \\
d_1 : (\tau, K, \tau\sigma^2_{\mathrm{imp}}; F, r, t) &\mapsto -\frac{x}{\theta_{\mathrm{imp}}} + \frac{\theta_{\mathrm{imp}}}{2}, \\
d_2 : (\tau, K, \tau\sigma^2_{\mathrm{imp}}; F, r, t) &\mapsto -\frac{x}{\theta_{\mathrm{imp}}} + \frac{\theta_{\mathrm{imp}}}{2}, \\
x : (F, K; t, \tau) &\mapsto \log(K/F_{[t,t+\tau]}), \\
F : (t, \tau; S, r) &\mapsto S_t e^{r\tau}, \\
\sigma_{\mathrm{imp}} : (\tau, K) &\mapsto \sigma_{\mathrm{imp}}(\tau, K), \\
\theta_{\mathrm{imp}} : (\tau, K) &\mapsto \sqrt{\tau}\sigma_{\mathrm{imp}}(\tau, K), \\
w_{\mathrm{imp}} : (\tau, K) &\mapsto \theta^2_{\mathrm{imp}}(\tau, K), \\
\mathcal{N} : d &\mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-t^2/2} \mathrm{d}t, \\
n : d &\mapsto \frac{1}{\sqrt{2\pi}} e^{-d^2/2}.
\end{aligned}$$

All first and second order partial derivatives of C^{B} with respect to moneyness and implied volatility will be needed in the differentiation, so first these will be calculated

explicitly. A useful trick in the differentiation is the following,

$$\begin{aligned} n\left(d_1\right) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x/\theta_{\text{imp}})^2 - x + (\theta_{\text{imp}}/2)^2}{2}} \\ &= \frac{e^x}{\sqrt{2\pi}} e^{-\frac{(x/\theta_{\text{imp}})^2 + x + (\theta_{\text{imp}}/2)^2}{2}} \\ &= e^x n\left(d_2\right). \end{aligned}$$

Lets calculate the partial derivatives.

$$\begin{aligned} \partial_x C^{\text{B}} &= F\Big(n(d_1)\partial_x d_1 - e^x \mathcal{N}(d_2) - e^x n(d_2)\partial_x d_2\Big) \\ &= F\left(n(d_1)\left(-\frac{1}{\theta_{\text{imp}}}\right) - e^x \mathcal{N}(d_2) - e^x n(d_2)\left(-\frac{1}{\theta_{\text{imp}}}\right)\right) \\ &= -F e^x \mathcal{N}(d_2). \end{aligned}$$

$$\begin{aligned} \partial_{\sigma_{\text{imp}}} C^{\text{B}} &= F\Big(n(d_1)\partial_{\theta_{\text{imp}}}(d_1)\partial_{\sigma_{\text{imp}}}(\theta_{\text{imp}}) - e^x n(d_2)\partial_{\theta_{\text{imp}}}(d_2)\partial_{\sigma_{\text{imp}}}(\theta_{\text{imp}})\Big) \\ &= F n(d_1)\sqrt{\tau}\left(\frac{x}{\theta_{\text{imp}}^2} + \frac{1}{2} - \frac{x}{\theta_{\text{imp}}^2} + \frac{1}{2}\right) \\ &= F n(d_1)\sqrt{\tau}. \end{aligned}$$

$$\begin{aligned} \partial_{xx} C^{\text{B}} &= -F e^x \mathcal{N}(d_2) - F e^x n(d_2)\partial_{xx}(d_2) \\ &= F\left(-e^x \mathcal{N}(d_2) + n(d_1)\left(\frac{1}{\theta_{\text{imp}}}\right)\right). \end{aligned}$$

$$\begin{aligned} \partial_{x\sigma_{\text{imp}}} C^{\text{B}} &= -F e^x n(d_2)\partial_{\sigma_{\text{imp}}}(d_2) \\ &= F n(d_1)\sqrt{\tau}\left(-\frac{x}{\theta_{\text{imp}}^2} + \frac{1}{2}\right). \end{aligned}$$

$$\begin{aligned} \partial_{\sigma_{\text{imp}}\sigma_{\text{imp}}} C^{\text{B}} &= F\sqrt{\tau}\partial_{d_1}\big(n(d_1)\big)\partial_{\sigma_{\text{imp}}}(d_1) \\ &= F\sqrt{\tau} n(d_1)(-d_1)\sqrt{\tau}\left(\frac{x}{\theta_{\text{imp}}^2} + \frac{1}{2}\right) \\ &= F n(d_1)\tau\left(\frac{x}{\theta_{\text{imp}}} - \frac{\theta_{\text{imp}}}{2}\right)\left(\frac{x}{\theta_{\text{imp}}^2} + \frac{1}{2}\right) \\ &= F n(d_1)\tau\left(\frac{x^2}{\theta_{\text{imp}}^3} - \frac{\theta_{\text{imp}}}{4}\right). \end{aligned}$$

Using the partial derivatives from above together with the chain rule for differentiation of multivariable functions, we can calculate the derivatives that we are interested in.

$$
\begin{aligned}
\partial_{KK}C^{\mathrm{B}} &= \partial_K\Big(\partial_x\left(C^{\mathrm{B}}\right)\partial_K(x)+\partial_{\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\partial_K\left(\sigma_{\mathrm{imp}}\right)\Big) \qquad \text{(C.1)}\\
&= \partial_{xx}\left(C^{\mathrm{B}}\right)\left(\partial_K(x)\right)^2+\partial_{x\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\partial_K(x)\partial_K\left(\sigma_{\mathrm{imp}}\right)+\partial_x\left(C^{\mathrm{B}}\right)\partial_{KK}(x)\\
&\quad+\partial_{\sigma_{\mathrm{imp}}\partial x}\left(C^{\mathrm{B}}\right)\partial_K\left(\sigma_{\mathrm{imp}}\right)\partial_K\left(x\right)+\partial_{\sigma_{\mathrm{imp}}\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\left(\partial_K\left(\sigma_{\mathrm{imp}}\right)\right)^2\\
&\quad+\partial_{\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\partial_{KK}\left(\sigma_{\mathrm{imp}}\right)\\
&= \frac{1}{K^2}\partial_{xx}\left(C^{\mathrm{B}}\right)+\frac{1}{K^2}\partial_{x\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\partial_x\left(\sigma_{\mathrm{imp}}\right)-\frac{1}{K^2}\partial_x\left(C^{\mathrm{B}}\right)\\
&\quad+\frac{1}{K^2}\partial_{\sigma_{\mathrm{imp}}x}\left(C^{\mathrm{B}}\right)\partial_x\left(\sigma_{\mathrm{imp}}\right)+\frac{1}{K^2}\partial_{\sigma_{\mathrm{imp}}\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\left(\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2\\
&\quad+\frac{1}{K^2}\partial_{\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\left(-\partial_x\left(\sigma_{\mathrm{imp}}\right)+\partial_{xx}\left(\sigma_{\mathrm{imp}}\right)\right)\\
&= \frac{1}{K^2}\Big(\partial_{xx}\left(C^{\mathrm{B}}\right)+2\partial_x\left(\sigma_{\mathrm{imp}}\right)\partial_{x\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)+\left(\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2\partial_{\sigma_{\mathrm{imp}}\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\\
&\quad-\partial_x\left(C^{\mathrm{B}}\right)+\Big(-\partial_x\left(\sigma_{\mathrm{imp}}\right)+\partial_{xx}\left(\sigma_{\mathrm{imp}}\right)\Big)\partial_{\sigma_{\mathrm{imp}}}\left(C^{\mathrm{B}}\right)\Big)\\
&= \frac{Fn(d_1)}{K^2}\left(\frac{1}{\theta_{\mathrm{imp}}}+\sqrt{\tau}\left(1-\frac{2x}{\theta_{\mathrm{imp}}^2}\right)\partial_x\left(\sigma_{\mathrm{imp}}\right)\right.\\
&\quad\left.+\tau\left(\frac{x}{\theta_{\mathrm{imp}}^3}-\frac{\theta_{\mathrm{imp}}}{4}\right)\left(\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2+\sqrt{\tau}\Big(-\partial_x\left(\sigma_{\mathrm{imp}}\right)+\partial_{xx}\left(\sigma_{\mathrm{imp}}\right)\Big)\right)\\
&= \frac{Fn(d_1)}{K^2\theta_{\mathrm{imp}}}\left(\left(1-\sqrt{\tau}\frac{x}{\theta_{\mathrm{imp}}}\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2-\tau\frac{\theta_{\mathrm{imp}}^2}{4}\left(\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2+\sqrt{\tau}\theta_{\mathrm{imp}}\partial_{xx}\left(\sigma_{\mathrm{imp}}\right)\right).
\end{aligned}
$$

It is useful to translate this expression into derivatives with respect to total implied variance.

$$
\begin{aligned}
\partial_{KK}\left(C^{\mathrm{B}}\right) &= \frac{Fn(d_1)}{K^2\theta_{\mathrm{imp}}}\left(\left(1-\sqrt{\tau}\frac{x}{\theta_{\mathrm{imp}}}\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2-\tau\frac{\theta_{\mathrm{imp}}^2}{4}\left(\partial_x\left(\sigma_{\mathrm{imp}}\right)\right)^2\right. \qquad \text{(C.2)}\\
&\quad\left.+\sqrt{\tau}\theta_{\mathrm{imp}}\partial_{xx}\left(\sigma_{\mathrm{imp}}\right)\right)\\
&= \frac{Fn(d_1)}{K^2\theta_{\mathrm{imp}}}\left(\left(1-\frac{x}{\theta_{\mathrm{imp}}}\partial_x\left(\theta_{\mathrm{imp}}\right)\right)^2-\frac{\theta_{\mathrm{imp}}^2}{4}\left(\partial_x\left(\theta_{\mathrm{imp}}\right)\right)^2+\theta_{\mathrm{imp}}\partial_{xx}\left(\theta_{\mathrm{imp}}\right)\right)\\
&= \frac{Fn(d_1)}{K^2\sqrt{w_{\mathrm{imp}}}}\left(\left(1-\frac{x}{\sqrt{w_{\mathrm{imp}}}}\frac{1}{2\sqrt{w_{\mathrm{imp}}}}\partial_x\left(w_{\mathrm{imp}}\right)\right)^2-\frac{w_{\mathrm{imp}}}{4}\frac{1}{4w_{\mathrm{imp}}}\left(\partial_x\left(w_{\mathrm{imp}}\right)\right)^2\right.\\
&\quad\left.+\sqrt{w_{\mathrm{imp}}}\left(\frac{1}{2\sqrt{w_{\mathrm{imp}}}}\partial_{xx}\left(w_{\mathrm{imp}}\right)-\frac{1}{4w_{\mathrm{imp}}^{3/2}}\left(\partial_x\left(w_{\mathrm{imp}}\right)\right)^2\right)\right)\\
&= \frac{Fn(d_1)}{K^2\sqrt{w_{\mathrm{imp}}}}\left(\left(1-\frac{x}{2w_{\mathrm{imp}}}\partial_x\left(w_{\mathrm{imp}}\right)\right)^2-\frac{1}{4}\left(\frac{1}{w_{\mathrm{imp}}}+\frac{1}{4}\right)\left(\partial_x\left(w_{\mathrm{imp}}\right)\right)^2\right.\\
&\quad\left.+\frac{1}{2}\partial_{xx}\left(w_{\mathrm{imp}}\right)\right).
\end{aligned}
$$

Furthermore, in the proof of Theorem 5, the derivative of

$$f(w_{\text{imp}}) = \frac{F}{K}\mathcal{N}(d_1) - \mathcal{N}(d_2)$$

with respect to total implied variance is needed.

$$\begin{aligned} \partial_{w_{\text{imp}}} f(w_{\text{imp}}) &= \frac{F}{K} n(d_1)\partial_{w_{\text{imp}}}(d_1) - n(d_2)\partial_{w_{\text{imp}}}(d_2) \\ &= \frac{F}{K} e^x n(d_2)\partial_{w_{\text{imp}}}\left(-\frac{x}{\sqrt{w_{\text{imp}}}} + \frac{\sqrt{w_{\text{imp}}}}{2}\right) \\ &\quad - n(d_2)\partial_{w_{\text{imp}}}\left(-\frac{x}{\sqrt{w_{\text{imp}}}} - \frac{\sqrt{w_{\text{imp}}}}{2}\right) \\ &= n(d_2)\frac{1}{2\sqrt{w_{\text{imp}}}} \end{aligned} \tag{C.3}$$

TRITA-MAT-E 2014:53
ISRN-KTH/MAT/E—14/53-SE

www.kth.se