Arbitrage-free Interpolation of Call Option Prices

Christian Bender[1] and Matthias Thiel[1]

August 22, 2019

Abstract

In this paper we introduce a new interpolation method for call option prices and implied volatilities with respect to the strike, which first generates, for fixed maturity, an implied volatility curve that is smooth and free of static arbitrage. Our interpolation method is based on a distortion of the call price function of an arbitrage-free financial 'reference' model of one's choice. It reproduces the call prices of the reference model, if the market data is compatible with the model. Given a set of call prices for different strikes and maturities, we can construct a call price surface by using this one-dimensional interpolation method on every input maturity and interpolating the generated curves in the maturity dimension. We obtain the algorithm of Kahalé [2004] as a special case, when applying the Black-Scholes model as reference model.

1 Introduction

European call options on liquidly traded assets are one of the fundamental financial products on the market. The price evaluation of an option depends on the choice of the financial model describing the behavior of the asset price. Since the work of Black and Scholes [1973] their model was continuously adapted to better describe the real world behavior of asset prices. Some examples for these adaptations are jump-diffusion or stochastic volatility models. In constrast to the Black-Scholes model these models are better suited to calibrate the smile effect. Another financial model which is able to deal with the volatility smile is the local volatility model introduced by Dupire [1994] and Derman and Kani [1994]. To calibrate the local volatility model it is necessary to know the price of call options for every non-negative strike and maturity. As there are only option prices observable for a finite set of strikes and maturities, the quoted prices need to be inter-and extrapolated. The major problem which arises by interpolating a finite set of call prices is the possibility that arbitrage may exist in the interpolated surface. There are plenty approaches to generate a call price surface in an arbitrage-free way, like the stochastic volatility inspired (SVI) method by Gatheral and Jacquier [2014] or different smoothing approaches using splines by e.g. Laurini [2011], Orosi [2015] or Fengler [2009].

[1]Saarland University, Department of Mathematics, Postfach 151150, D-66041 Saarbrücken, Germany, bender@math.uni-sb.de; thiel@math.uni-sb.de.

Another method for an arbitrage-free inter-and extrapolation of call prices is suggested by Kahalé [2004]. In his one-dimensional interpolation with respect to the strike he uses a linear transformation of the Black-Scholes formula for call prices and calibrates parameters to fit his call price function to the option prices quoted in the market. Moreover he chooses the slopes of the price function at input strikes in a way that guarantees the absence of static arbitrage. His use of the Black-Scholes formula for the local interpolation facilitates the proof of absence of static arbitrage in the interpolated implied volatility surface, but otherwise appears to be somewhat arbitrary. In this paper we significantly generalize Kahalé's method for a single maturity by interpolating the observed call prices with a linear transformation of the option price formula of a reference model of one's choice. Our interpolation method reproduces the call prices of the reference model, if the market data is compatible with the model, and otherwise linearly disturbs the prices of the reference model. In this way, one can incorporate his preferred model into the interpolation procedure. Application of this method on every quoted maturity and linear interpolation of the constructed curves in the total implied variance in the maturity dimension leads to an implied volatility surface that is, under certain conditions, free of static arbitrage. The generated implied volatility surface is sufficiently smooth to calibrate the local volatility model. The results of Kahalé are included in our method as a special case, where the Black-Scholes model with arbitrary volatility is chosen as reference model.

The paper is structured as follows: In Section 2 we recall conditions for a set of call prices to be free of static arbitrage by quoting the results of Carr and Madan [2005] and Roper [2010]. In Section 3 our construction of a C^1-interpolation function with respect to the strike for a single maturity will be explained. The structure and some properties of the generated call price function will be the topic of Section 4, in which the connection between our approach and the approach of Kahalé will be clarified. To obtain better smoothness properties of our call price function we will introduce in Section 5 a C^2-interpolation method. A short explanation on the interpolation in the maturity dimension and on the calibration of the Dupire model is given in Section 6. In Section 7 we introduce a transformation for call prices that allows us to use our method on call prices on assets that pay continuous dividend yields in a market with a constant and non-negative interest rate. Simulations on different data sets in order to to outline the practical applicability of our method will be made in Section 8. Section 9 concludes.

2 Conditions for the Absence of Static Arbitrage

The two key properties, which an interpolated call price surface should satisfy, is the absence of static arbitrage and a sufficient smoothness to calibrate the local volatility model. In this section we want to point out the conditions for a set of call prices to be free of static arbitrage. The term static arbitrage was introduced by Carr, Geman, Madan, and Yor [2003]. A set of call prices is said to be free of static arbitrage, if there exists a filtered probability space containing a non-negative martingal S, in which the quoted call prices can be expressed as expectation of the discounted payoff of the option at maturity. There are various works on conditions for a set of call prices or implied volatilitys to be free of static arbitrage, see e.g., Carr and Madan [2005], Roper [2010], or Davis and Hobson [2007]. There are different settings in which we need conditions for the absence of static arbitrage. The first one is to identify whether a finite set of call prices for different strikes at a single maturity is free of static arbitrage, because that is a necessary condition for our method to work. This condition is also used in the construction of our one-dimensional interpolation curve. The conditions we give in this section apply in markets without interest rates and on assets that do not pay dividends. We quote the conditions proven in Carr and Madan [2005].

Lemma 2.1. *Let $(k_i, c_i)_{0 \le i \le n+1}$ be a sequence satisfying*

$$0 = c_{n+1} = k_0 < k_1 < \cdots < k_n < k_{n+1} = \infty,$$

where c_i, $0 \le i \le n$ is the price of a call with strike k_i for a single maturity. These prices are free of static arbitrage, if and only if c_0 equals the current spot and, for every $1 \le i \le n$,

$$-1 \le \frac{c_i - c_{i-1}}{k_i - k_{i-1}} \le \frac{c_{i+1} - c_i}{k_{i+1} - k_i} \le 0.$$

In practice, we will most likely have a grid of call prices for different strikes and maturities. Our two dimensional interpolation method is only applicable, if for every fixed maturity the conditions of Lemma 2.1 hold and it will only result in a call price surface that is free of static arbitrage, if the quoted price grid is free of static arbitrage. A rigorous discussion on when such a grid of prices is free of static arbitrage can be found in Davis and Hobson [2007]. If the grid is free of static arbitrage, we will use our one-dimenisonal interpolation method on every single quoted maturity and afterwards interpolate the constructed slices in the maturity dimension. To know if our method generated a call price surface with the properties we desire, we need conditions for such a complete price surface to be free of static arbitrage. These conditions can be found in Roper [2010] and we quote them without proof.

Lemma 2.2. *A call price surface $C : [0, \infty) \times [0, \infty) \to \mathbb{R}^+$, where $C(T, 0)$ equals the current spot price $S_0 > 0$ is free of static arbitrage, if $C(T, k)$ satisfies the following conditions:*

- $C(T, \cdot)$ *is a convex function for all* $T \ge 0$.
- $C(\cdot, k)$ *is non-decreasing for all* $k \ge 0$.
- $\lim_{k \to \infty} C(T, k) = 0$, *for all* $T \ge 0$.
- $(S_0 - k)^+ \le C(T, k) \le S_0$ *for all* $k, T \ge 0$.
- $C(k, 0) = (S_0 - k)^+$ *for all* $k \ge 0$.

3 C^1-Interpolation

We first consider the case of a single fixed maturity T and fix a random variable S_T as model for the asset price at time T on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. Assuming zero interest rates and no dividend payments, S_T induces a call price curve via

$$C_T(k) = \mathbb{E}\left[(S_T - k)^+\right],$$

which, by definition, is free of static arbitrage. We assume that the cumulative distribution function F of S_T under the measure $\mathbb{Q}$ is continuous on $[0,\infty)$ and satisfies $\int_0^\infty |x|\ dF(x) < \infty$. The last equation can then be expressed as

$$C_T(k) = \int_k^\infty (x-k)\ dF(x).$$

The following calculations will be made for a fixed maturity T, so the time index will be omitted. Let $(k_i, c_i)_{0\leq i\leq n+1}$ be the sequence of call prices c_i on an underlying asset with strikes k_i for $i = 1, \ldots, n$ with a fixed maturity T and additional elements (k_0, c_0) and (k_{n+1}, c_{n+1}), where $c_0 := S_0 := \mathbb{E}[S_T]$ equals the current price of the underlying. The sequence is assumed to satisfy

$$0 = c_{n+1} = k_0 < \cdots < k_n < k_{n+1} = \infty. \tag{1}$$

Our objective is to construct a cumulative distribution function F on $[0,\infty)$ which fulfills

$$C(k_i) = \int_{k_i}^\infty (x-k_i)\ dF(x)\ = c_i, \tag{2}$$

for $i = 0, \ldots, n$ and $k_0 = 0$, $c_0 = S_0$, $k_{n+1} = \infty$, $c_{n+1} = 0$. To preserve the arbitrage condition in Lemma 2.1, one chooses c'_i for $i = 0, \ldots, n+1$, such that

$$c'_0 = -1\ ,\ c'_n < c'_{n+1} = 0 < c_n \tag{3}$$

and

$$\begin{aligned} -1 < \frac{c_i - c_{i-1}}{k_i - k_{i-1}} < c'_i < \frac{c_{i+1} - c_i}{k_{i+1} - k_i} \qquad &, i = 1, \ldots, n-1, \\ \frac{c_n - c_{n-1}}{k_n - k_{n-1}} < c'_n < 0. \end{aligned} \tag{4}$$

Assumption (3)-(4) is supposed to be in force throughout the paper.

For the time being we take the values c'_i as an input, but later on these values will be determined by the algorithm in order to enhance the smoothness of the interpolation function C. Integration by parts yields

$$C(k) = \int_k^\infty (1 - F(x))dx$$

and consequently the first derivative of the call price function $C'(k) = F(k) - 1$.
We thus define

$$F(k_i) := 1 + c'_i \text{ for } i = 0\ , \ldots,\ n. \tag{5}$$

With this choice of F at the given strikes, it only remains to determine the conditional distribution functions $F(\cdot\,|(k_i, k_{i+1}])$. Since

$$F(x) = \sum_{i=0}^{n} F(x\,|(k_i, k_{i+1}])(F(k_{i+1}) - F(k_i)), \tag{6}$$

it follows that for every $i = 0, \ldots, n$,

$$C(k_i) = \sum_{j=i}^{n-1} (c'_{j+1} - c'_j) y_j - c'_n y_n - k_i(1 - F(k_i)), \tag{7}$$

where $y_j = \int_{k_j}^{k_{j+1}} x\ dF(x\,|(k_j, k_{j+1}])$ and $y_n = \int_{k_n}^{\infty} x\ dF(x\,|(k_n, \infty))$. The following lemma shows how to determine the y_i, such that equation (2) holds. This lemma can be shown with elementary manipulations and its proof is omitted.

Lemma 3.1. *With the introduced notations, the equality*

$$c_i = \int_{k_i}^{\infty} (x - k_i)\ dF(x)$$

holds for all i, if and only if

$$y_i = \frac{c_i - c'_i k_i - c_{i+1} + c'_{i+1} k_{i+1}}{c'_{i+1} - c'_i} \qquad \textit{for} \ \ i = 0, \ldots, n-1$$

and

$$y_n = \frac{c_n - c'_n k_n}{-c'_n}.$$

Consequently our goal is to construct a cumulative distribution function F that, given sequences (k_i), (c_i), (c'_i), satisfies

$$\int_{k_i}^{k_{i+1}} x\ dF(x\,|(k_i, k_{i+1}]) = y_i \ \text{ for } i = 0, \ldots, n.$$

Our construction takes a distribution function G on $\mathbb{R}$ as an input, which we think of as a reference distribution of one's choice, for the log-price of the stock at time T. We assume that G has a continuous and strictly positive density g and has exponential moments of every order $m \in \mathbb{N}$, i.e.,

$$\int_{-\infty}^{\infty} e^{mx}\ dG(x) < \infty \text{ for every } m \in \mathbb{N}. \tag{8}$$

We divide the construction of the function F into three parts, the construction in between input strikes, before the first input strike and after the last input strike. So first we want to find distribution functions $F^{(i)}$ on the intervals $(k_i, k_{i+1}]$ for $i \in \{1, \ \dots \ , n-1\}$, which satisfy the equations

$$\int_{k_i}^{k_{i+1}} x \; dF^{(i)}(x) = y_i \quad \text{for } i = 1, \dots, n-1. \tag{9}$$

Such an $F^{(i)}$ will, in view of the previous considerations, serve as the conditional distribution function $F(\cdot\,|(k_i, k_{i+1}])$.

Construction 1: For a fixed $i \in \{1, \dots, n-1\}$ and $a \in (c'_{i+1} - 1, c'_i)$ define:

$$\begin{aligned}
d_j(a) &:= d_j(a; c'_i, c'_{i+1}) := G^{-1}(c'_{i+j} - a) \quad \text{for} \quad j = 0, 1, \\
\alpha(a) &:= \alpha_i(a; c'_i, c'_{i+1}) := \frac{d_1(a) - d_0(a)}{\log(k_{i+1}) - \log(k_i)}, \\
\beta(a) &:= \beta_i(a; c'_i, c'_{i+1}) := d_1(a) - \alpha(a)\log(k_{i+1}) = d_0(a) - \alpha(a)\log(k_i)
\end{aligned}$$

and

$$F_a^{(i)}(x) := F_a^{(i)}(x; c'_i, c'_{i+1}) \tag{10}$$

$$= \begin{cases} 0 & \text{for } 0 \le x < k_i \\ \dfrac{G(\alpha(a)\log(x) + \beta(a)) + a - c'_i}{c'_{i+1} - c'_i} & \text{for } k_i \le x \le k_{i+1} \\ 1 & \text{for } x > k_{i+1}. \end{cases} \tag{11}$$

Note that, for every $a \in (c'_{i+1} - 1, c'_i)$, $F_a^{(i)}$ is a continuous distribution function on $[0, \infty)$.

We now consider the function

$$I^{(i)} : (c'_{i+1} - 1, c'_i) \to \mathbb{R}, \; a \mapsto a + \frac{1}{k_{i+1} - k_i} \int_{k_i}^{k_{i+1}} G(\alpha(a)\log(x) + \beta(a))dx,$$

which we extend to the closed interval $[c'_{i+1} - 1, c'_i]$ by setting $I^{(i)}(c'_{i+1} - 1) := c'_{i+1}$ and $I^{(i)}(c'_i) := c'_i$. We sometimes write $I^{(i)}(a; c'_i, c'_{i+1})$ to emphasize the dependence of $I^{(i)}$ on the slopes.

Lemma 3.1 and integration by parts imply that $F_a^{(i)}$ satisfies equation (9), if and only if the real number $a \in (c'_{i+1} - 1, c'_i)$ solves

$$I^{(i)}(a) = (c_{i+1} - c_i)/(k_{i+1} - k_i). \tag{12}$$

The following Lemma shows the existence of such a solution.

Lemma 3.2. *The function $I^{(i)}$ is continuous on $[c'_{i+1} - 1, c'_i]$. In particular, there exists a solution $a_i^* \in (c'_{i+1} - 1, c'_i)$ of the equation $I^{(i)}(a_i^*) = (c_{i+1} - c_i)/(k_{i+1} - k_i)$.*

Proof. The function $I^{(i)}$ is obviously continuous on the open interval $(c'_{i+1}-1, c'_i)$ by dominated convergence and continuity of G, α, and β. Note that

$$\lim_{a\to c'_{i+1}-1} G(\alpha(a)\log(x)+\beta(a)) = \lim_{a\to c'_{i+1}-1} G(\alpha(a)(\log(x)-\log(k_i))+d_0(a)).$$

With $\alpha(a)\to\infty$ for $a\to c'_{i+1}-1$ and $\log(x)-\log(k_i)>0$ for $x\in(k_i,k_{i+1})$ it follows

$$\lim_{a\to c'_{i+1}-1} G(\alpha(a)\log(x)+\beta(a)) = 1.$$

Similar argumentation leads to

$$\lim_{a\to c'_i} G(\alpha(a)\log(x)+\beta(a)) = 0.$$

Applying the dominated convergence theorem again, we obtain

$$\lim_{a\to c'_{i+1}-1} I^{(i)}(a) = c'_{i+1} \text{ and } \lim_{a\to c'_i} I^{(i)}(a) = c'_i.$$

As a result the function $I^{(i)}$ is continuous on $[c'_{i+1}-1, c'_i]$. In view of the inequalities (4), the intermediate value theorem proves the existence of a solution $a^*_i\in(c'_{i+1}-1, c'_i)$ of the equation $I^{(i)}(a^*_i) = (c_{i+1}-c_i)/(k_{i+1}-k_i)$. □

In general, the solution a^*_i is not unique, but by setting additional conditions on the distribution function G in Construction 1 we obtain uniqueness. A sufficient condition for the uniqueness of the solution is the strict convexity of $1/g$, where g denotes the density of the function G in Construction 1.

Theorem 3.3. *Assume that the distribution function G satisfies*

$$1/g \text{ is strictly convex (uniqueness condition)}, \tag{13}$$

Then, for every $i\in\{1,\ldots,n-1\}$ there exists exactly one $a^*_i\in(c'_{i+1}-1, c'_i)$, satisfying $I^{(i)}(a^*_i) = (c_{i+1}-c_i)/(k_{i+1}-k_i)$.

Proof. By setting $\lambda = (\log(x)-\log(k_i))/(\log(k_{i+1})-\log(k_i))$ for fixed $x\in(k_i,k_{i+1})$, the following inequality holds (cp. Kahalé [2004], Lemma 3)

$$\begin{aligned}
&\frac{d}{da}G(\alpha(a)\log(x)+\beta(a)) = \frac{d}{da}(G(\lambda G^{-1}(c'_{i+1}-a)+(1-\lambda)G^{-1}(c'_i-a)))\\
&= - g(\lambda d_1(a)+(1-\lambda)d_0(a))\left(\frac{\lambda}{g(d_1(a))}+\frac{1-\lambda}{g(d_0(a))}\right)\\
&< - 1.
\end{aligned}$$

Hence, the mapping $a \mapsto a + G(\alpha(a)\log(x) + \beta(a))$ is strictly decreasing for $a \in (c'_{i+1} - 1, c'_i)$. Now, the monotonicity of the integral yields, that $I^{(i)}(a)$ is strictly decreasing on the interval $(c'_{i+1} - 1, c'_i)$. This clearly implies uniqueness. □

The numbers a_i^* for $i = 1, \dots, n-1$ can be calculated numerically.
Set $F^{(i)}(x) = F^{(i)}_{a_i^*}(x)$ for all $i = 1, \ \dots \ , n-1$.

Remark 3.4. Let $l_i := (c_{i+1} - c_i)/(k_{i+1} - k_i)$ for $i = 1, \dots, n-1$. We can also consider the functions $I^{(i)}(\cdot\,; l_i, c'_{i+1}) : [c'_{i+1} - 1, l_i] \to \mathbb{R}$ and
$I^{(i)}(\cdot\,; c'_i, l_{i+1}) : [l_{i+1} - 1, c'_i] \to \mathbb{R}$, replacing c'_i, respectively c'_{i+1}, by their limiting slopes in the definition of $I^{(i)}$ in the obvious way. The same argumentation as above show that under assumption (13), these two functions are continuous and strictly decreasing with values

$$I^{(i)}(l_i; l_i, c'_{i+1}) = l_i, \quad I^{(i)}(c'_{i+1} - 1; l_i, c'_{i+1}) = c'_{i+1},$$
$$I^{(i)}(c'_i; c'_i, l_{i+1}) = c'_i, \quad I^{(i)}(l_{i+1} - 1; c'_i, l_{i+1}) = l_{i+1}.$$

To construct the remaining functions $F^{(0)}$ and $F^{(n)}$ for the extrapolation, Construction 1 has to be modified. We proceed with the construction of a distribution function $F^{(0)}$ on the interval $[0, k_1]$ which satisfies the equation

$$\int_0^{k_1} x \; dF^{(0)}(x) = y_0. \tag{14}$$

Construction 2: Define for $a \in (0, \infty)$:

$$\begin{aligned} F_a^{(0)}(x) &:= F_a^{(0)}(x; c'_1) \\ &= \begin{cases} 0 & \text{for } x = 0 \\ \dfrac{G(a(\log(x) - \log(k_1)) + G^{-1}(c'_1 + 1))}{c'_1 + 1} & \text{for } 0 < x \le k_1 \\ 1 & \text{for } x > k_1, \end{cases} \end{aligned}$$

which again is a continuous distribution function on $[0, \infty)$ for every $a \in (0, \infty)$.

Similarly to Construction 1, we consider the function

$$I^{(0)} : (0, \infty) \to \mathbb{R}, \; a \mapsto \frac{1}{k_1} \int_0^{k_1} \big(G(a(\log(x) - \log(k_1)) + G^{-1}(c'_1 + 1)) - 1 \big) \; dx.$$

As in Construction 1, a solves the equation $I^{(0)}(a) = (c_1 - c_0)/(k_1 - k_0)$, if and only if the associated distribution function $F_a^{(0)}$ satisfies equation (14). The following theorem shows the existence and uniqueness of such a solution. In constrast to Construction 1, the uniqueness of such solution does not require any further conditions on the distribution function G.

Theorem 3.5. *By defining $I^{(0)}(0) = c_1'$ the function $I^{(0)}$ is continuous on the interval $[0,\infty)$ and $\lim_{a\to\infty} I^{(0)}(a) = 0$. Moreover, there exists a unique $a_0^* \in (0,\infty)$, which satisfies the equation $I^{(0)}(a_0^*) = (c_1 - c_0)/(k_1 - k_0)$.*

Proof. The map $a \mapsto G(a(\log(x) - \log(k_1)) + G^{-1}(c_1' + 1)) - 1$ is clearly strictly decreasing and continuous on $(0,\infty)$, and then so is $I^{(0)}$ by monotone convergence. Moreover, for every $x \in (0, k_1)$

$$\lim_{a\to 0} G(a(\log(x) - \log(k_1)) + G^{-1}(c_1' + 1)) = c_1' + 1$$

and

$$\lim_{a\to\infty} G(a(\log(x) - \log(k_1)) + G^{-1}(c_1' + 1)) = 0.$$

Application of the dominated convergence theorem leads to

$$\lim_{a\to 0} I^{(0)}(a) = c_1' \quad \text{and} \quad \lim_{a\to\infty} I^{(0)}(a) = -1.$$

As $-1 < (c_1 - c_0)/(k_1 - k_0) < c_1'$ by (4), the assertion follows immediately. □

The number a_0^* can be calculated numerically. Set $F^{(0)}(x) = F^{(0)}_{a_0^*}(x)$.

Finally, we want to to find a distribution function $F^{(n)}$ on the interval $[k_n, \infty)$ which satisfies the equation

$$\int_{k_n}^{\infty} x \; dF^{(n)}(x) = y_n. \tag{15}$$

Construction 3: Define for $a \in (0,\infty)$:

$$\begin{aligned} F_a^{(n)}(x) :&= F_a^{(n)}(x, c_n') \\ &= \begin{cases} 0 & \text{for } x < k_n \\ \dfrac{G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1)) - c_n' - 1}{-c_n'} & \text{for } x \geq k_n. \end{cases} \end{aligned}$$

The function $F_a^{(n)}$ is a continuous distribution function on $[0,\infty)$ for every $a \in (0,\infty)$.

Note that, for every $a \in (0,\infty)$, with $d_n := G^{-1}(c_n' + 1)$,

$$\int_{k_n}^{\infty} x \; dF_a^{(n)}(x) = -\frac{1}{c_n'} \int_{k_n}^{\infty} x \; dG(a(\log(x) - \log(k_n)) + d_n) = -\frac{k_n}{c_n'} \int_{d_n}^{\infty} e^{\frac{y-d_n}{a}} \; dG(y)$$

is finite by (8). Thus, we can apply integration by parts to the left-hand side, which shows

$$\int_{k_n}^{\infty} x \; dF_a^{(n)}(x) = k_n + \int_{k_n}^{\infty} (1 - F_a^{(n)}(x))dx.$$

Define

$$I^{(n)} : (0,\infty) \to \mathbb{R}, \; a \mapsto \int_{k_n}^{\infty} (1 - G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1)) \; dx.$$

Then, by Lemma 3.1, $a \in (0,\infty)$ solves the equation $I^{(n)}(a) = c_n$, if and only if the associated distribution function $F_a^{(n)}$ satisfies equation (15).

Theorem 3.6. *The function $I^{(n)}$ is continuous on $(0,\infty)$ with limits $\lim_{a\to\infty} I^{(n)}(a) = 0$ and $\lim_{a\to 0} I^{(n)}(a) = \infty$. Moreover there exists exactly one $a_n^* \in (0,\infty)$ satisfying $I^{(n)}(a_n^*) = c_n$.*

Proof. The map $a \mapsto 1 - G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1))$ is clearly strictly decreasing and continuous on $(0,\infty)$, and then so is $I^{(n)}$ by monotone convergence. Moreover, for every $x \in (k_n,\infty)$,

$$\lim_{a\to 0} 1 - G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1)) = -c_n'$$

and

$$\lim_{a\to\infty} 1 - G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1)) = 0$$

Application of the monotone convergence theorem leads to

$$\lim_{a\to 0} I^{(n)}(a) = \infty \quad \text{and} \quad \lim_{a\to\infty} I^{(n)}(a) = 0.$$

As $0 < c_n < \infty$ by (3), the assertion follows immediately. □

As in Construction 2, the uniqueness of the solution a_n^* is given without any further conditions on the distribution function G. Set $F^{(n)}(x) = F_{a_n^*}^{(n)}(x)$. The following Theorem, which summarizes the foregoing, is the first main result of this paper.

Theorem 3.7. *Choose a distribution function G with strictly positive continuous density function, which satisfies (8) and (13). Then, for all sequences $(k_i)_{0\le i\le n+1}$, $(c_i)_{0\le i\le n+1}$ and $(c_i')_{0\le i\le n+1}$, which satisfy the conditions (1),(3) and (4), there exist unique numbers $a_0^* \in (0,\infty)$, $a_i^* \in (c_{i+1}' - 1, c_i')$ for $i = 1,\dots,n-1$ and $a_n^* \in (0,\infty)$, such that*

$$\int_{k_i}^{k_{i+1}} x \; dF_{a_i^*}^{(i)}(x) = y_i \quad \text{for } i = 0,\dots,n.$$

(Here, the y_i are defined in Lemma 3.1.)

Moreover, the function $C : [0,\infty) \to \mathbb{R},\ k \mapsto \int_k^\infty (x-k)dF(x)$ with

$$F(x) := \sum_{i=0}^{n} F_{a_i^*}^{(i)}(x)(c_{i+1}' - c_i')$$

is continuously differentiable on $[0,\infty)$ and twice differentiable on $(0,\infty) \setminus \{k_1,\dots,k_n\}$ and satisfies

$$C(k_i) = c_i \text{ and } C'(k_i) = c_i' \text{ for every } i = 1,\dots n.$$

Proof. In view of the previous results, it only remains to argue the smoothness of C. As each of the distribution functions $F^{(i)}_{a^*_i}$ is continuous on $[0,\infty)$ and differentiable on $[0,\infty)\setminus\{k_i,k_{i+1}\}$, the distribution function F is continuous on $[0,\infty)$ and differentiable on $(0,\infty)\setminus\{k_1,\ldots,k_n\}$. The relation $C'(k)=F(k)-1$, then concludes. ☐

Remark 3.8.

- The convexity condition in Construction 1 guarantees the uniqueness of the constructed call price function for a chosen sequence $(c'_i)_{n\geq 0}$. But the interpolation method works even if the convexity condition does not hold.
- In principle, one can choose different distribution functions G on each interval $[0,k_1]$, $(k_i,k_{i+1}]$ for $i=1,\ldots,n-1$ and (k_n,∞). Then, the integrability condition (8) only needs to be satisfied for the distribution function in Construction 3.
- Note that our interpolated call price function C is by construction free of static arbitrage.

We close this section by stating an algorithm which describes the construction of a C^1-price function C which interpolates the quoted option prices for a fixed maturity.

Algorithm 3.9. Let $(k_i,c_i)_{1\leq i\leq n}$ be a sequence of strikes k_i and associated call prices c_i with additional elements $k_0=0;\ c_0=S_0;\ k_{n+1}=\infty$ and $c_{n+1}=0$.

1. Set $c'_0=-1$ and $c'_{n+1}=0$. Choose c'_i for $i=1,\ldots,n$, such that (4) holds.
2. Choose a distribution function G on $\mathbb{R}$ with strictly positive continuous density, that satisfies the integrability condition (8) and (optionally) the uniqueness condition (13). Then construct the functions $F^{(i)}_a$ for $i=0,\ldots,n$.
3. Calculate a^*_i with $\int_{k_i}^{k_{i+1}} x\ dF^{(i)}_{a^*_i}(x)=y_i$ for $i=0,\ldots,n$ numerically (e.g., by bisection).
4. Set $F^{(i)}=F^{(i)}_{a^*_i}$ and compute the distribution function
   $F(x)=\sum\limits_{i=0}^{n}(c'_{i+1}-c'_i)F^{(i)}(x)$.
5. Compute the interpolated price function C.

4 Some Properties of the Interpolation Function

In this section, we discuss how the call price interpolation function C constructed in Theorem 3.7 relates to the call price function of the reference model. If the market prices are not compatible with the reference model (which is the typical case), the interpolation function can be interpreted as a piecewise distortion of the call price function of the reference model. In order to make this more precise, recall that G denotes the reference distribution of the log-price at time T. We denote by U a random variable with distribution function G and consider

$$Call(k;\sigma,\mu) = E[(e^{\mu+\sigma U} - k)^+], \quad \mu \in \mathbb{R}, \quad \sigma > 0$$

which is the call price function of a linear transformation of the log-price of the reference model.

Theorem 4.1. *Under the assumptions and with the notation of Theorem 3.7, the interpolation function C can be represented as*

$$C(k) = \begin{cases} Call(k;\sigma_0,\mu_0) + b_0, & k \in [0,k_1] \\ Call(k;\sigma_i,\mu_i) + A_i k + b_i, & k \in (k_i,k_{i+1}],\ i = 1,\dots,n-1 \\ Call(k;\sigma_n,\mu_n), & k \in (k_n,\infty) \end{cases}$$

for constants

$$\begin{aligned} \sigma_0 &= \frac{1}{a_0^*}, \quad \mu_0 = \log(k_1) - G^{-1}(c_1' + 1)/a_0^*, \quad b_0 = c_1 - Call(k_1;\sigma_0,\mu_0), \\ \sigma_i &= \frac{1}{\alpha_i(a_i^*, c_i', c_{i+1}')}, \quad \mu_i = \frac{-\beta_i(a_i^*, c_i', c_{i+1}')}{\alpha_i(a_i^*, c_i', c_{i+1}')}, \\ A_i &= (1 + a_i^*), \quad b_i = c_i - Call(k_{i+1};\sigma_i,\mu_i) - A_i k_{i+1}, \\ \sigma_n &= \frac{1}{a_n^*}, \quad \mu_n = \log(k_n) - G^{-1}(c_n' + 1)/a_n^* \end{aligned}$$

for $i = 1,\dots,n-1$.

Remark 4.2. If we choose U to be Gaussian, then C interpolates the observed call option prices on each of the intervals $[k_i, k_{i+1})$ by a linear perturbation of Black-Scholes prices. By the uniqueness result in Theorem 2 of Kahalé [2004], we, thus, obtain Kahalé's interpolation function as a special case of our construction, when U is Gaussian.

Proof of Theorem 4.1. We provide the proof for the case $i \in \{1,\dots,n-1\}$ and note that the other two cases can be treated similarly. Taking Theorem 3.7 and Lemma 3.1 into account, we get for, $k \in (k_i, k_{i+1}]$,

$$\begin{aligned} C(k) &= (c_{i+1}' - c_i') \int_k^{k_{i+1}} x dF_{a_i^*}^{(i)}(x) \\ &\quad + \sum_{j=i+1}^{n} (c_{j+1}' - c_j') \int_{k_j}^{k_{j+1}} x dF_{a_j^*}^{(j)}(x) - k(1 - F(k)) \\ &= (c_{i+1}' - c_i') \left(\int_k^{k_{i+1}} x dF_{a_i^*}^{(i)}(x) + k F_{a_i^*}^{(i)}(k) \right) \\ &\quad + k c_i' - k_{i+1} c_{i+1}' + c_{i+1}. \end{aligned}$$

Note that, for $x \in [k_i, k_{i+1}]$,

$$(c'_{i+1} - c'_i) F^{(i)}_{a^*_i}(x) = G(\alpha(a^*_i) \log(x) + \beta(a^*_i)) + a^*_i - c'_i$$

and that the distribution function of $e^{\mu_i + \sigma_i U}$ is given by
$x \mapsto G(\alpha(a^*_i) \log(x) + \beta(a^*_i))$. Hence, recalling that $F^{(i)}_{a^*_i}(k_{i+1}) = 1$,

$$\begin{aligned} C(k) = E[e^{\mu_i+\sigma_i U} \mathbf{1}_{\{e^{\mu_i+\sigma_i U} \in [k, k_{i+1}]\}}] \\ + kQ(\{e^{\mu_i+\sigma_i U} \leq k\}) + a^*_i k - k_{i+1} c'_{i+1} + c_{i+1} \\ = Call(k; \sigma_i, \mu_i) + A_i k - E[e^{\mu_i+\sigma_i U} \mathbf{1}_{\{e^{\mu_i+\sigma_i U} \in [k_{i+1}, \infty)\}}] \\ - k_{i+1} c'_{i+1} + c_{i+1} \\ = Call(k; \sigma_i, \mu_i) + A_i k + c_{i+1} - A_i k_{i+1} - Call(k_{i+1}; \sigma_i, \mu_i). \end{aligned}$$

□

The second result states that, under suitable assumptions, if the chosen reference model is compatible with the observed call option prices, our interpolation procedure reproduces the call price function of this model.

Theorem 4.3. *Choose a reference model with distribution function G for the log-price that satisfies the assumptions of Theorem 3.7 and let $\tilde{F} := G \circ \log$. Suppose that the observed call option prices are compatible with the call price function of the reference model, i.e. $c_i = \tilde{C}(k_i)$ for every $i = 0 \ldots, n$, where*
$\tilde{C}(k) := \int_k^\infty (x-k) d\tilde{F}(x)$ and that $c'_i = \tilde{F}(k_i) - 1$. Then, the call price interpolation function C in Theorem 3.7 satisfies $C(k) = \tilde{C}(k)$ for every $k \geq 0$.

Proof. Thanks to Lemma 3.1, we obtain, for $i = 1, \ldots, n-1$,

$$y_i = \frac{1}{c'_{i+1} - c'_i} \int_{k_i}^{k_{i+1}} x d\tilde{F}(x).$$

Hence, the equation $\int_{k_i}^{k_{i+1}} x \, dF^{(i)}_{a^*_i}(x) = y_i$ from Construction 1, can be equivalently reformulated as

$$\int_{k_i}^{k_{i+1}} x \, dG(\alpha(a^*_i) \log(x) + \beta(a^*_i)) = \int_{k_i}^{k_{i+1}} x \, d\tilde{F}(x).$$

Note that the functions α and β in Construction 1 satisfy $\alpha(-1) = 1$ and $\beta(-1) = 0$, because

$$G^{-1}(c'_i + 1) = \log(\tilde{F}^{-1}(c'_i + 1)) = \log(k_i).$$

Hence, $\int_{k_i}^{k_{i+1}} x\ dF_{-1}^{(i)}(x) = y_i$, which implies $a_i^* = -1$ and

$$F_{a_i^*}^{(i)}(x)(c'_{i+1} - c'_i) = \mathbf{1}_{(k_i,k_{i+1}]}(x)(\tilde{F}(x) - \tilde{F}(k_i)) + \mathbf{1}_{(k_{i+1},\infty)}(x)(\tilde{F}(k_{i+1}) - \tilde{F}(k_i)).$$

Analogously, it is easy to verify that $\int_0^{k_1} x\ dF_1^{(1)}(x) = y_0$ and $\int_{k_n}^{\infty} x\ dF_1^{(n)}(x) = y_n$, and for $a_0^* = a_n^* = 1$

$$\begin{aligned} F_{a_0^*}^{(0)}(x)(c'_1 - c'_0) &= \mathbf{1}_{[0,k_1]}(x)\tilde{F}(x) + \mathbf{1}_{(k_1,\infty)}(x)\tilde{F}(k_1), \\ F_{a_n^*}^{(n)}(x)(c'_{n+1} - c'_n) &= \mathbf{1}_{(k_n,\infty]}(x)(\tilde{F}(x) - \tilde{F}(k_n)). \end{aligned}$$

Hence, the distribution function F constructed in Theorem 3.7 coincides with the reference distribution $\tilde{F}$ for the stock price. □

Surely there would be no need to interpolate call prices, if we could simply choose a model which is compatible with the observed prices. But intuitively this property shown in Theorem 4.3 is a feature that an interpolation method that depends on the choice of a reference model should satisfy. Aditionally, Theorem 4.3 provides the unique parameters ($a_1^* = a_n^* = 1,\ a_i^* = -1$ for $i = 2, \ldots, n-1$) for which the observed prices are compatible with the chosen model. Therefore, we obtain an indicator of how close the reference model prices are to the observed ones.

5 C^2-Interpolation

One problem of the algorithm 3.9 is the arbitraryness in choosing the sequence $(c'_i)_{1\le i\le n}$. The only constraint that the sequence must obey is the arbitrage condition

$$\frac{c_i - c_{i-1}}{k_i - k_{i-1}} < c'_i < \frac{c_{i+1} - c_i}{k_{i+1} - k_i} \ \text{ for } i = 1, \ldots, n-1 \text{ and } \frac{c_n - c_{n-1}}{k_n - k_{n-1}} < c'_n < 0.$$

This still leaves a lot of freedom in the choice of the sequence. In this section we explain how the sequence $(c'_i)_{1\le i\le n}$ can be chosen to enhance the smoothness of the interpolation. Recall that function C constructed with Theorem 3.7 is continuously differentiable on $\mathbb{R}_+$ and twice continuously differentiable on $\mathbb{R}_+ \setminus \{k_1, \ldots, k_n\}$. Ideally, we would wish to choose the sequence $(c'_i)_{1\le i\le n}$, such that

$$\lim_{k\nearrow k_i} C''(k) := C''(k_i-) = C''(k_i+) =: \lim_{k\searrow k_i} C''(k) \text{ for all } i = 1, \ldots, n.$$

The following two lemmas explain the behavior of $C''(k_i-)$ and $C''(k_i+)$ at an input strike k_i if c'_i converges to its upper and lower bounds given by the arbitrage condition.

Lemma 5.1.
Let $(k_i)_{0\le i\le n+1}$, $(c_i)_{0\le i\le n+1}$ and $(c'_i)_{0\le i\le n+1}$ be sequences, which satisfy the conditions (1), (3)

and (4). Also let $C(k)$ be the call price function constructed in Theorem 3.7 on these sequences. Assume that the density functions g of G satisfies the limit conditions

$$\lim_{x\to-\infty} x\,g(x) = \lim_{x\to\infty} g(x) = 0. \tag{16}$$

Moreover define

$$l_i = \begin{cases} \frac{c_{i+1}-c_i}{k_{i+1}-k_i} & , \textit{ for } i = 0,\dots,n-1 \\ 0 & , \textit{ for } i = n. \end{cases}$$

Then, for every $i = 1,\dots,n$, the following limit properties holds:

$$\lim_{k\searrow k_i} C''(k) := C''(k_i+) = 0 \textit{ and } \lim_{k\nearrow k_i} C''(k) := C''(k_i-) > 0$$

if $c'_i \to l_i$.

Proof. First we take a look at the behavior of $C''(k_i+)$ for $i = 1,\dots,n-1$. Fix some $i \in \{1,\dots,n-1\}$. The first derivative of the function C, constructed in Theorem 3.7 on the interval (k_i, k_{i+1}) can be expressed as

$$C'(k) = (c'_{i+1} - c'_i)F^{(i)}_{a^*_i}(k) + c'_i = G(\alpha(a^*_i)\log(k) + \beta(a^*_i)) + a^*_i,$$

because of the identities $C'(k) = F(k) - 1$ and $F(k) = \sum_{i=0}^n F^{(i)}_{a^*_i}(k)(c'_{i+1} - c'_i)$. Consequently, we obtain

$$\begin{aligned} C''(k) &= \frac{G^{-1}(c'_{i+1} - a^*_i) - G^{-1}(c'_i - a^*_i)}{(\log(k_{i+1}) - \log(k_i))k} g(\alpha(a^*_i)\log(k) + \beta(a^*_i)) \\ C''(k_i+) &= \frac{G^{-1}(c'_{i+1} - a^*_i) - G^{-1}(c'_i - a^*_i)}{(\log(k_{i+1}) - \log(k_i))k_i} g(G^{-1}(c'_i - a^*_i)). \end{aligned} \tag{17}$$

Recall that a^*_i depends on c'_i and c'_{i+1}, and we thus need to study its behavior as $c'_i \to l_i$. To this end, we write $I^{(i)}(\cdot\,; c'_i, c'_{i+1})$ instead of $I^{(i)}$ to emphazise the dependence of $I^{(i)}$ on the chosen slopes c'_i and c'_{i+1}. Now let $(c'_{i,m})_{m\in\mathbb{N}}$ be a sequence with $c'_{i,m} \to l_i$ for $m \to \infty$. By Theorem 3.3 there is, for every $m \in \mathbb{N}$, a unique $a^*_{i,m} \in (c'_{i+1} - 1, c'_{i,m})$ such that

$$I^{(i)}(a^*_{i,m}, c'_{i,m}, c'_{i+1}) = \frac{c_{i+1} - c_i}{k_{i+1} - ki} = l_i.$$

The sequence $a^*_{i,m}$ takes values in $[c'_{i+1} - 1, l_i]$ and, hence, has a convergent subsequence with limit $a^*_i \in [c'_{i+1} - 1, l_i]$. By continuity of G, α and β, if
$a^*_{i,m} \in (c'_{i+1} - 1, l_i)$ (resp., by the same argument as in Lemma 3.2, if
$a^*_{i,m} \in \{c'_{i+1} - 1, l_i\}$), it follows, that

$$I^{(i)}(a^*_{i,m}; c'_{i,m}, c'_{i+1}) \to I^{(i)}(a^*_i; l_i, c'_{i+1}) \text{ for } m \to \infty,$$

and hence $I^{(i)}(a_i^*; l_i, c_{i+1}') = l_i$. Remark 3.4 now implies that $a_i^* = l_i$. Hence, every subsequence of $(a_{i,m}^*)_{m\in\mathbb{N}}$ converges to l_i and, then, so does the sequence itself. Consequently, by taking (16) and (17) into account, we obtain

$$\lim_{c_i' \to l_i} C''(k_i+) = 0.$$

We next show that this limiting identity also holds for $i = n$. The first and second derivative of the function C on the interval (k_n, ∞) can be expressed analogously to the previous case as

$$\begin{aligned} C'(k) &= (-c_n')F_{a_n^*}^{(n)}(k) + c_n' = G(a_n^*(\log(k) - \log(k_n)) + G^{-1}(c_n' + 1)) + 1, \\ C''(k) &= \frac{a_n^*}{k}\, g(a_n^*(\log(k) - \log(k_n)) + G^{-1}(c_n' + 1)), \\ C''(k_n+) &= \frac{a_n^*}{k_n}\, g(G^{-1}(c_n' + 1)). \end{aligned}$$

Here we used Construction 3 (instead of Construction 1). Note that
$g(G^{-1}(c_n' + 1)) \to 0$ by (16), as c_n' tends to zero. Thus, we only need to show that a_n^* does not converge to infinity, as c_n' approaches zero. Recall that, by Theorem 3.6, $a_n^* \in (0, \infty)$ is the unique solution to

$$I^{(n)}(a; c_n') = \int_{k_n}^{\infty} (1 - G(a(\log(x) - \log(k_n)) + G^{-1}(c_n' + 1))\, dx = c_n. \tag{18}$$

Hence a_n^* is monotonically decreasing in c_n' and, thus, cannot converge to infinity, as c_n' goes to 0 from the left.

Now we examine $C''(k_i-)$ for $i = 1, \ldots, n-1$. Fix $i \in \{2, \ldots, n\}$ and let again $(c_{i,m}')_{m\in\mathbb{N}}$ be a sequence with $c_{i,m}' \to l_i$ for $m \to \infty$. Applying Theorem 3.3 again, there is, for every $m \in \mathbb{N}$, a unique $a_{i-1,m}^* \in (c_{i,m}' - 1, c_{i-1}')$ such that

$$I^{(i-1)}(a_{i-1,m}^*; c_{i-1}', c_{i,m}') = \frac{c_i - c_{i-1}}{k_i - ki - 1} = l_{i-1}.$$

The sequence $a_{i-1,m}^*$ takes values in $[l_i - 1, c_{i-1}']$ and, hence, has a convergent subsequence with limit $a_{i-1}^* \in [l_i - 1, c_{i-1}']$. Argueing as in the first part of the proof, it follows, that

$$I^{(i-1)}(a_{i-1,m}^*, c_{i-1}', c_{i,m}') \to I^{(i)}(a_{i-1}^*, c_{i-1}', l_i), \quad (m \to \infty),$$

and hence $I^{(i-1)}(a_{i-1}^*, c_{i-1}', l_i) = l_{i-1}$. Remark 3.4 now implies that the limit a_{i-1}^* does not depend on the choice of the subsequence and belongs to the open interval $(l_i - 1, c_{i-1}')$. In view of (17), we may conclude that

$$\lim_{c_i' \to l_i} C''(k_i-) = \frac{G^{-1}(l_i - a_{i-1}^*) - G^{-1}(c_{i-1}' - a_{i-1}^*)}{(\log(k_i) - \log(k_{i-1}))k_i} g(G^{-1}(l_i - a_{i-1}^*)) > 0.$$

It remains to show that $\lim_{c'_1 \to l_1} C''(k_1-) > 0$. Applying Construction 2, we obtain for $k \in (0, k_1)$:

$$
\begin{aligned}
C''(k) &= \frac{a_1^*}{k}\, g(a_1^*(\log(k) - \log(k_1)) + G^{-1}(c'_1 + 1)), \\
C''(k_1-) &= \frac{a_1^*}{k_1}\, g(G^{-1}(c'_1 + 1)).
\end{aligned} \tag{19}
$$

As $l_1 + 1 > 0$ by (4), we observe that $\lim_{c'_1 \to l_1} g(G^{-1}(c'_1 + 1)) > 0$. A similar monotonicity argument as above shows that a_1^* is increasing in c'_1. Hence, a_1^* tends to a non-zero limit, as c_1^* approaches l_1 from the left. Now (19) concludes. □

Lemma 5.2.
Let $(k_i)_{0 \le i \le n+1}$, $(c_i)_{0 \le i \le n+1}$ and $(c'_i)_{0 \le i \le n+1}$ be sequences, which satisfy the conditions (1), (3) and (4). Also let $C(k)$ be the call price function constructed in Theorem 3.7 on these sequences. Assume that the density functions g of G satisfies the limit conditions

$$\lim_{x \to \infty} x\, g(x) = 0. \tag{20}$$

Moreover define

$$l_i = \begin{cases} \frac{c_{i+1} - c_i}{k_{i+1} - k_i} & , \text{ for } i = 0, \dots, n-1 \\ 0 & , \text{ for } i = n. \end{cases}$$

Then, for every $i = 1, \dots, n$, the following limit properties holds:

$$\lim_{k \nearrow k_i} C''(k) := C''(k_i-) = 0 \text{ and } \lim_{k \searrow k_i} C''(k) := C''(k_i+) > 0$$

if $c'_i \to l_{i-1}$.

Proof. Completely analogous to the proof of Lemma 5.1. □

The following theorem shows that it is possible to construct a C^1-price function C that is twice differentiable at an input strike k_i for a fixed $i \in \{1, \dots, n\}$.

Theorem 5.3.
Fix $j \in \{1, \dots, n\}$. For all sequences $(k_i)_{0 \le i \le n+1}$, $(c_i)_{0 \le i \le n+1}$ and $(c'_i)_{0 \le i \le n}$ satisfying (1), (3) and (4), according to Theorem 3.7 there exists a continuously differentiable function $C(k)$ for $k \ge 0$, so that $C(k_i) = c_i$ and $C'(k_i) = c'_i$ for $i = 0, \dots, n$. If the chosen distribution functions satisfy

$$\lim_{x \to \infty} x\, g(x) = \lim_{x \to -\infty} x\, g(x) = 0, \tag{21}$$

then there exists a

$$\gamma \in \left(\frac{c_j - c_{j-1}}{k_j - k_{j-1}}, \frac{c_{j+1} - c_j}{k_{j+1} - k_j} \right),$$

so that the constructed price function C becomes twice differentiable at k_j, if we exchange c'_j with γ.

Proof. Fix $j \in \{1, \dots, n\}$. Set $l_i = (c_{i+1} - c_i)/(k_{i+1} - k_i)$ for $0 \leq i \leq n-1$ and $l_n = 0$. According to Theorem 3.7, there exists a continuously differentiable, convex fiunction C on the positive real axis that satisfies $C(k_i) = c_i$ for $i = 1, \dots, n$, $C'(k_i) = c'_i$ for $1 \leq i \leq n$, $i \neq j$ and $c'_j = \gamma$ for any $\gamma \in (l_j, l_{j+1})$. Taking Lemma 5.1 and 5.2 into account, the following inequalities hold

$$C''(k_j+) - C''(k_j-) > 0 \text{ if } \gamma \to l_{j-1} \text{ and } C''(k_j+) - C''(k_j-) < 0 \text{ if } \gamma \to l_j.$$

Because of the continuity of $C''(k_j+)$ and $C''(k_j-)$ with respect to γ, the intermediate value theorem ensures the existence of a $\gamma \in (l_{j-1}, l_j)$, so that
$C''(k_j+) = C''(k_j-)$. □

Remark 5.4.
Theorem 5.3 shows that there exists a slope c'_i, which generates (using Theorem 3.7) a price function that is twice continuously differentiable in k_i, but not how to find it. That slope can be calculated numerically by finding the root of a non-linear function.

Our goal is to construct a function C that is twice continuously differentiable at all input strikes $(k_i)_{0 \leq i \leq n}$, therefore the previous results of Theorem 5.3 are not sufficient. Taking a look at the input strikes k_1 and k_2, where c'_1 had been chosen, such that $C''(k_1+) = C''(k_1-)$. By choosing c'_2 to make C twice continuously differentiable at k_2, we adapt the function C on the interval $(k_1, k_2]$ and consequently the previous equation $C''(k_1+) = C''(k_1-)$ loses its validity. We assume, that the iterative application of Theorem 5.3 on the sequence $(k_i, c_i)_{1 \leq i \leq n}$ leads to a function $C \in C^2$. Numerical experiments confirm that conjecture. The following algorithm and the associated conjecture extend the formulations of Kahalé [2004] beyond the Gaussian case.

Algorithm 5.5.
Given an error parameter ε and a sequence $(k_i, c_i)_{1 \leq i \leq n}$ of strikes k_i and associated call prices c_i with additional elements $k_0 = 0,\ c_0 = S_0,\ k_{n+1} = \infty$ and $c_{n+1} = 0$, which preserves the condition

$$-1 < \frac{c_i - c_{i-1}}{k_i - k_{i-1}} < \frac{c_{i+1} - c_i}{k_{i+1} - k_i} < 0 \text{ for } , i = 1, \dots, n.$$

1. Set $c'_0 = -1$, $c'_{n+1} = 0$ and $c'_i = (l_i + l_{i-1})/2$ for $i = 1, \dots, n$, where $l_i = (c_{i+1} - c_i)/(k_{i+1} - k_i)$ for $i = 0, \dots, n-1$ and $l_n = 0$.

2. *Iteration*: Apply Theorem 5.3, with c'_i and c'_{i+1} defined in step 1, on every input strike k_i, for $i = 1, \dots, n$. This generates a sequence $\tilde{c}_i{}'$, which we use for the construction of the call price function $C(k)$ of Theorem 3.7.

3. *Validation*: Check if $|C''(k_i+) - C''(k_i-)| < \varepsilon$ for all $i = 1, \dots, n$. If the inequality holds for all i, terminate the algorithm. If not set $c'_i := \tilde{c}'_i$ for $i = 1, \dots, n$ and repeat step 2 and 3.

Conjecture 5.6.
For all sequences $(k_i)_{1\leq i\leq n}$, $(c_i)_{1\leq i\leq n}$, which satisfy the conditions (1), (3) and (4), the sequence $(c'_i)_{1\leq i\leq n}$ generated by Algorithm 5.5 converges to a sequence $(c'_i)^*_{1\leq i\leq n}$. The function C, constructed with Theorem 3.7 on this sequence $(c'_i)^*_{1\leq i\leq n}$ is twice continuously differentiable.

6 Two-dimensional Interpolation and Calibration of the Dupire Model

In this section we will shortly explain how our two dimensional interpolation works and how to use the generated implied volatility surface to calibrate the Dupire model. These ideas are mostly adopted from Kahalé [2004] and for more detailed discussion we advise the reader to take a look at his paper.
Given a set of maturities $T_1,\dots,T_m$, strikes $k_1,\dots,k_{n(i)}$ for every maturity T_i and the corresponding call prices $c_{i,j}$ for $i=1,\dots,n$ and $j=1,\dots,n(i)$. For every input maturity T_i we apply our C^2-interpolation method of section 5, which leads to to a set of call price functions $C(T_i,k)$ for $i=1,\dots,m$ and $k\in[0,\infty)$, that are free of static arbitrage. We calculate the corresponding implied volatility functions $\sigma_{\text{imp}}(T_i,k)$ and interpolate these functions in the maturity direction linear in implied total variance $\sigma^2_{\text{imp}}(T,k)T$. The linear interpolation in implied total variance ensures that the interpolated surface is increasing in the maturity dimension, if the input curves $\sigma_{\text{imp}}(T_i,k)$ are increasing with respect to maturity for every strike. Also it generates a smooth surface exept at input maturities. But there are also some problems which arise by interpolating the input curves linear in implied total variance. First, it is not clear if the functions $\sigma_{\text{imp}}(T_i,k)$ are increasing with respect to maturity for every strike, which is a condition for the surface to be free of static arbitrage. Also it is not verified that a function $\sigma_{\text{imp}}(T,k)$ for $T\in(T_i,T_{i+1})$, generated by linear interpolation in implied total variance of the functions $\sigma_{\text{imp}}(T_i,k)$ and $\sigma_{\text{imp}}(T_{i+1},k)$ is free of static arbitrage. So some adjustments have to be made to guarantee that the implied volatility surface is free of static arbitrage and these are detailed in Kahalé [2004].

For the calibration of the Dupire model it is necessary to calculate the local volatility $\sigma_{\text{loc}}(T,k)$. Usually one would calculate the derivatives of the call price surface with respect to the strike and maturity and use the standard Dupire formula

$$\sigma_{loc}(T,k)^2 = 2\frac{\frac{\partial C}{\partial T}(T,k)}{k^2\frac{\partial C}{\partial k^2}(T,k)} \tag{22}$$

to calculate the local volatility. We instead apply the Dupire formula with respect to implied volatility σ

$$\sigma^2_{loc} = \frac{\sigma^2 + 2\sigma T\left(\frac{\partial\sigma}{\partial T}\right)}{\left(1+\frac{k\log\left(\frac{k}{S_0}\right)}{\sigma}\frac{\partial\sigma}{\partial k}\right)^2 + k\sigma T\left(\frac{\partial\sigma}{\partial k} - \frac{1}{4}k\sigma T(\frac{\partial\sigma}{\partial k})^2 + k\frac{\partial^2\sigma}{\partial k^2}\right)}, \tag{23}$$

see Andersen and Brotherton-Ratcliffe [1998], because our two-dimensional interpolation is based on implied volatilities and not on call prices and this version of the Dupire formula leads to less numerical instabilities.

7 Interest Rates and Dividend Yields

To this point we only concentrated on call prices on an asset that does not pay dividends in a market where the interest rate is zero. Clearly this setting is not very realistic, therefore we will present a method for including call prices on assets that pay a non-negative dividend yield in a market with a non-negative and constant interest rate. Let S_t be an underlying with constant dividend yield $q > 0$ and let $r > 0$ be the constant interest rate. We assume that under the risk-neutral measure, S_t follows the stochastic differential equation

$$dS_t = (r - q)S_t dt + \sigma(t, S_t)S_t dW_t. \tag{24}$$

We introduce a new price process S_t^* satisfying

$$S_t^* = e^{-(r-q)t} S_t.$$

Using Ito's Lemma we obtain

$$\begin{aligned} dS_t^* &= e^{-(r-q)t} dS_t - (r - q)S_t e^{-(r-q)t} dt \\ &= ((r - q)S_t dt + \sigma(t, S_t)S_t dW_t)e^{-(r-q)t} - (r - q)S_t e^{-(r-q)t} dt \\ &= \sigma(t, e^{(r-q)t} S_t^*)S_t^* dW_t. \end{aligned}$$

Therefore S_t^* can be seen as an dividend free asset in an interest free market. Next we develop a relation between the call prices on S_t and S_t^*. By using the risk-free call price formula we obtain

$$C_S(T, k) = e^{-rT} E[(S_T - k)^+] \text{ and } C_{S^*}(T, k) = E[(S_T^* - k)^+],$$

which yields

$$C_S(T, k) = e^{-rT} E[(S_T - k)^+] = e^{-rT} E[(e^{(r-q)T} S_T^* - k)^+] = e^{-qT} C_{S^*}(T, k^*),$$

where $k^* = e^{-(r-q)T} k$. Using this relation we obtain the formula

$$C_{S^*}(T, e^{-(r-q)T} k) = e^{qT} C_S(T, k), \tag{25}$$

which allows us to transform the call prices on a dividend paying asset S_t in a market with interest rate $r > 0$ to call prices on the non-dividend paying asset S_t^* in a market with interest rate $r = 0$.

Calibrating the Dupire model on the prices $C_S(T,k)$ can be done by using Dupire's formula on the call prices $C_{S^*}(T,k)$ (respectively the implied volatilities) to calculate the local volatility $\sigma^*_{loc}(T,k)$ and obtaining the desired local volatilities $\sigma_{loc}(T,k)$ corresponding to the prices $C_S(T,k)$ through the relation

$$\sigma_{loc}(T,k) = \sigma^*_{loc}(T, e^{-(r-q)T}k).$$

8 Simulations

In this section we use our interpolation method on two different data sets. We mainly want two show three features of our interpolation method, namely that different reference models (distributions) lead two different interpolation functions, applicability to real data sets and that, if the observed prices are compatible with the reference model used, the method generates the call price function of this model. First we generate prices using the Merton model with jump intensity $\lambda = 1$, expected jump heigth $a = -0.1$, jump heigth variance $b = 0.2$, volatility $\sigma = 0.3$, maturity $T = 0.2$, $S_0 = 500$ and 6 strikes reaching from 300 to 800. We interpolate by taking two different distribution functions as reference, the distribution function of the Merton model, with the same parameters used for the simulation of the prices and the Laplace distribution function with mean 0 and variance 1.

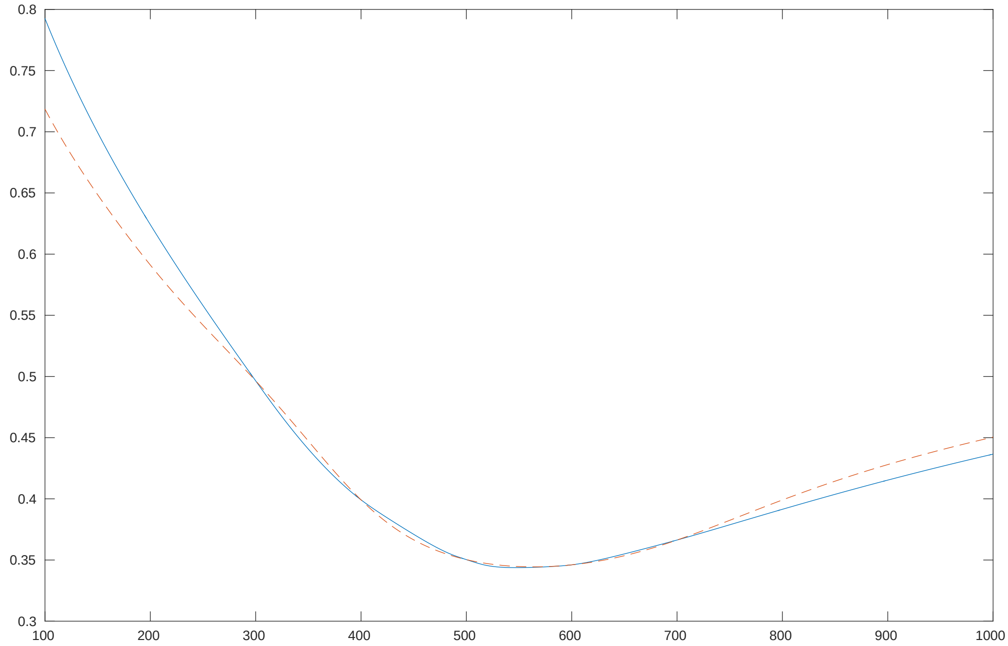
Remark 8.1. The Laplace distribution function with mean 0 and variance 1 is given by $L(x) = 0.5 + 0.5\ \mathrm{sgn}(x)(1 - e^{-|x|})$. Applying this distribution as reference has several advantages. The functions required in Constructions 1, 2 and 3 are easy to calculate and there is a closed form for the inverse of the distribution function, which significantly speeds up the numerical calculations compared to the usage of the Black-Scholes or Merton model as reference. One problem of the Laplace distribution is that the integrability condition (8) clearly does not hold, which could lead to problems in the extrapolation to infinity. It can still be possible that the method works with the Laplace distribution, namely when the root, that we have to find lies in the area where the function calculated in Construction 3 exists. So for practice we suggest using the Laplace distribution as reference for all intervals and if the method fails in the extrapolation, then one should switch the reference distribution function for the extrapolation deep out of the money to a distribution which fulfills the integrability condition, e.g. the distribution function of the Black-Scholes model.

The implied volatility functions corresponding to the interpolated call price functions with the two different reference distributions and their absolute difference is shown in Figure 1. Within the input strikes there is little difference between the two interpolated curves, but the difference increases deep in, or deep out of the money. So by changing the reference model one can influence the behavior of the interpolated call prices or respectively implied volatilities, especially at extreme strikes. As described in Theorem 4.3, using our C^1-interpolation method on input prices and slopes, that are compatible with the reference model, should give the exact price function of the model and the calculated sequence $(a_i)_{0 \leq i \leq n}$ should be given by $a_0 = a_n = 1$ and $a_i = -1$ for $i = 1, \ldots, n-1$. Here we used our C^2-interpolation method (in this method the slopes at input strikes are calculated in the course of the algorithm). The absolute difference between the calculated parameters $(a_i)_{0 \leq i \leq n}$ and the desired values 1 respectively -1 is at most $3 \cdot 10^{-4}$ with the compatible Merton model as reference, which shows that the interpolation function and the price function of the reference model are practically identical. With the Laplace distribution as reference the differences to the values 1 and -1 range from 0,0022 to 13,1. As described in Remark 8.1, using the laplace distribution as reference leads to a way faster numerical calculation. In this example the calculation of the slopes was 20 times faster using the laplace distribution (in comparison to the Merton model) as reference.

For the second simulation we use the dataset from Andersen and Brotherton-Ratcliffe [1998], that is shown in Table 1. Again we interpolate by taking two different distribution functions as reference. The first distribution function we use is the Merton model with maturity and spot according to the quoted prices and $\sigma = 0.3$, $\lambda = 1$, $a = -0.1$, $b = 0$ and compare it to the interpolation with the Laplace distribution function with mean 0 and variance 1. For both interpolations we first have to transform the quoted prices as described in section 7. We compare the local volatility curves corresponding to our interpolated call price functions, using both distribution functions, in Figure 2, where we plot the relative error between the two local volatility curves for four different maturities. Figure 3 shows the corresponding interpolated implied and local volatility surfaces, where we used the Laplace distribution function as reference. The relative error between input prices and the prices in the local volatility model computed by Crank Nicholson scheme is shown in Figure 4.

Maturity \Strike	501.5	531	560.5	590	619.5	649	678.5	708	767	826
0.175	0.190	0.168	0.133	0.113	0.102	0.097	0.120	0.142	0.169	0.200
0.425	0.177	0.155	0.138	0.125	0.109	0.103	0.100	0.114	0.130	0.150
0.695	0.172	0.157	0.144	0.133	0.118	0.104	0.100	0.101	0.108	0.124
0.940	0.171	0.159	0.149	0.137	0.127	0.113	0.106	0.103	0.100	0.110
1	0.171	0.159	0.150	0.138	0.128	0.115	0.107	0.103	0.099	0.108
1.5	0.169	0.160	0.151	0.142	0.133	0.124	0.119	0.113	0.107	0.102
2	0.169	0.161	0.153	0.145	0.137	0.130	0.126	0.119	0.115	0.111
3	0.168	0.161	0.155	0.149	0.143	0.137	0.133	0.128	0.124	0.123
4	0.168	0.162	0.157	0.152	0.148	0.143	0.139	0.135	0.130	0.128
5	0.168	0.164	0.159	0.154	0.151	0.148	0.144	0.140	0.136	0.132

Table 1: Implied volatilities on the S&P 500 index with interest rate 0.06, continuous dividend yield 0.026 and spot price 590$.



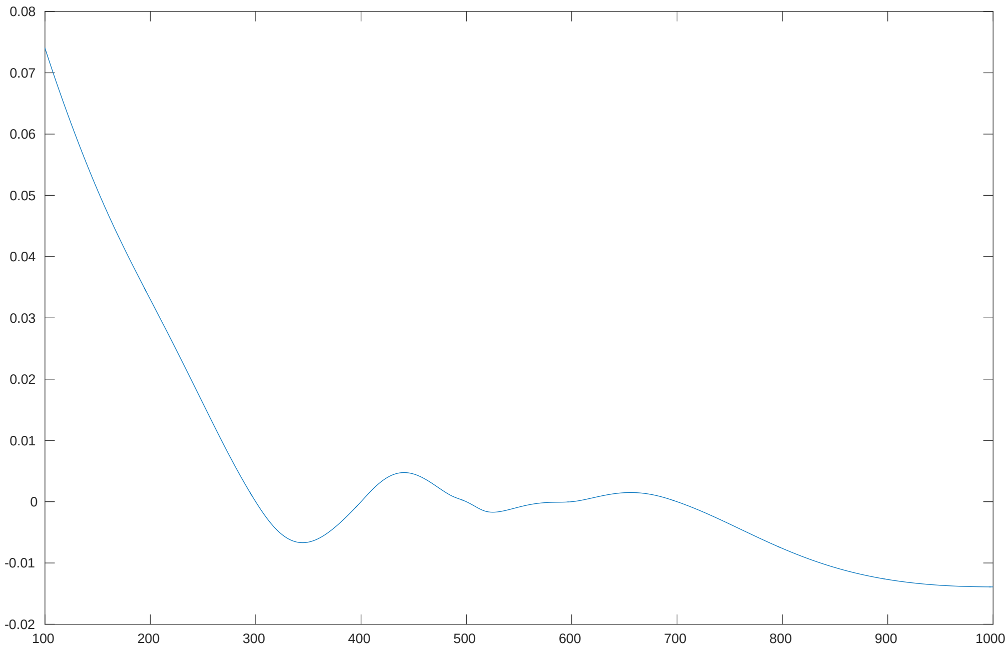


Figure 1: Implied volatilities of the interpolation function on the simulated Merton prices, where the Merton model (dashed) and the Laplace distribution (solid) were taken as reference model (distribution) for our interpolation and their absolute difference.

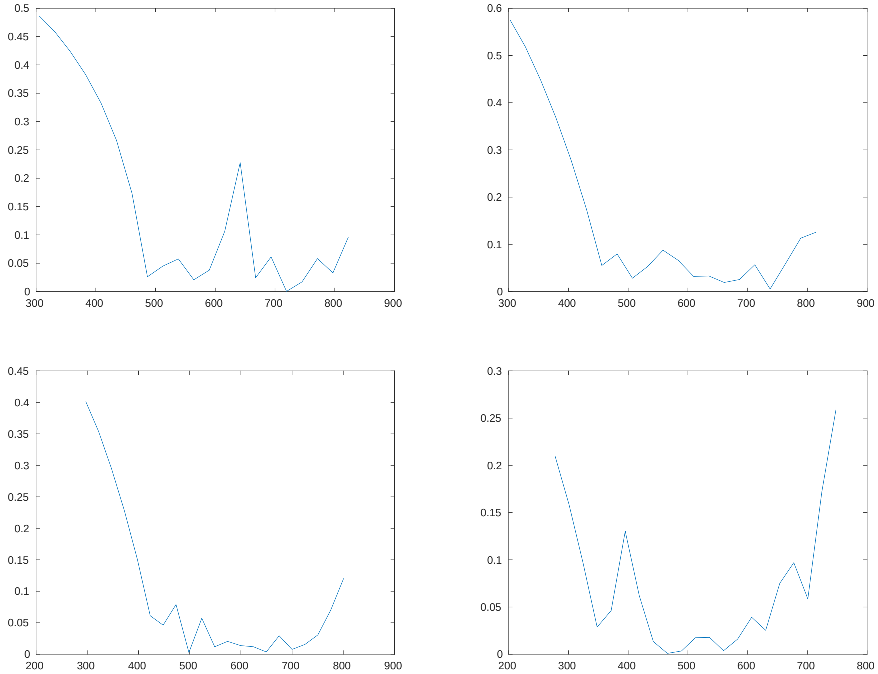


Figure 2: Relative error between local volatilitiy curves corresponding to the two different reference distribution functions at maturities $0.2, 0.5, 1, 3$ (from top left to bottom right), as a function of the strike price.

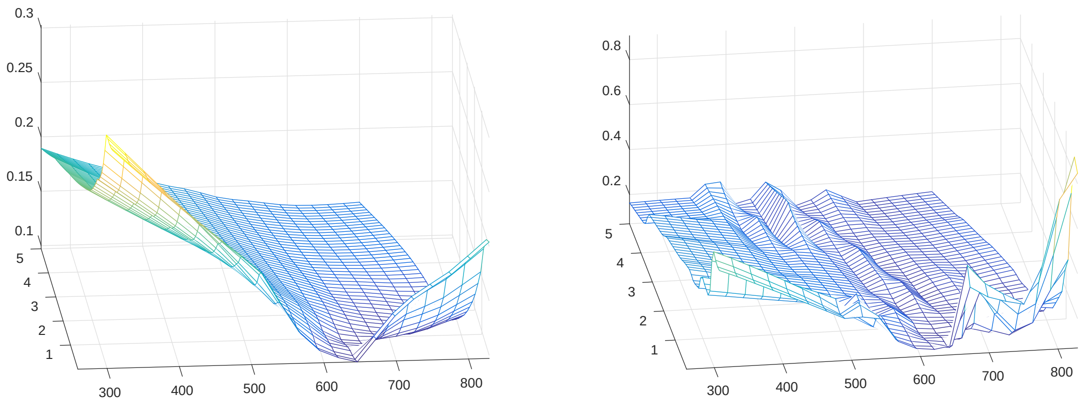


Figure 3: Interpolated implied (left) and local (right) volatility surface constructed on the S&P500 Data with the Laplace distribution as reference distribution.

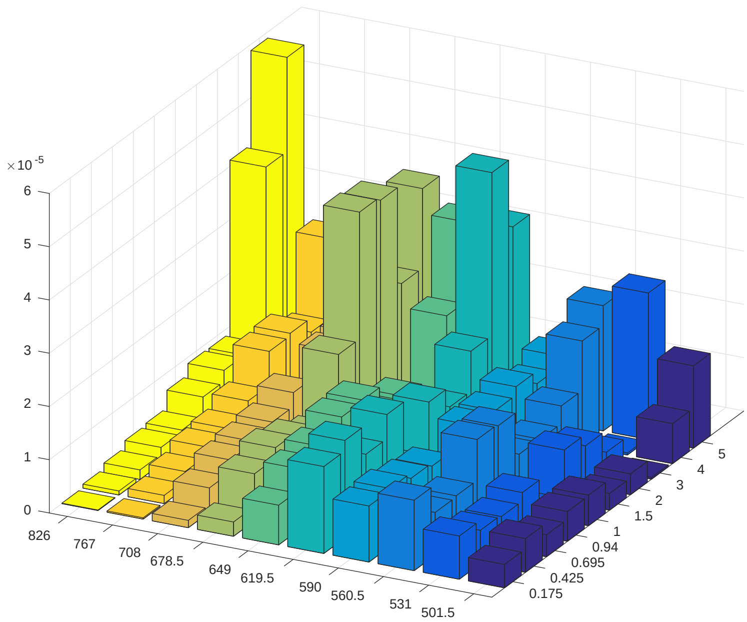


Figure 4: Error between input and local volatility prices using Crank Nicholson scheme in relation to the spot price.

9 Conclusion

In this paper we introduced a new one-dimensional interpolation method for call option prices, relying on a chosen refrence model (distribution), which by construction generates a smooth interpolation function that is free of static arbitrage. The calibration distorts the price function of the reference model only to the extent which is necessary to fit the observed call prices. This generalizes the results of Kahalé from the Black Scholes model to a much wider class of reference distributions. We have shown that the usage of different reference distributions leads to different interpolation functions, which enables practitioners to choose distributions according to their intentions. If the point of interest lies in the speed of the calibration one should use distributions whose inverse is easy to compute and for which the integrals in Construction 1, 2 and 3 can be handled, e.g. the Laplace distribution. If the user is more interested in a particular behavior of the implied volatilities, especially at extreme strikes, one can choose a reference model with the desired characteristics. As illustrated in Section 8, the smoothness properties of our one-dimensional interpolation method combined with the interpolation in implied total variance makes it possible to generate implied volatility surfaces that are free of static arbitrage and allow to calibrate a local volatility model, which almost perfectly fits the observed call prices.

References

Leif B. G. Andersen and Rupert Brotherton-Ratcliffe. The equity option volatility smile: an implicit finite-difference approach. *Journal of Computational Finance*, 1(2):5–37, 1998.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81:637–654, 1973.

Peter Carr and Dilip B. Madan. A note on sufficient conditions for no arbitrage. *Finance Research Letters*, 2:125–130, 2005.

Peter Carr, Hélyette Geman, Dilip B. Madan, and Marc Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13(3):345–382, 2003.

Mark H. A. Davis and David G. Hobson. The range of traded option prices. *Mathematical Finance*, 17(1):1–14, 2007.

Emanuel Derman and Iraj Kani. Riding on a smile. *Risk*, 7(2):32–39, 1994.

Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

Matthias Fengler. Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9(4):417–428, 2009.

Jim Gatheral and Antoine Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.

Nabil Kahalé. An arbitrage-free interpolation of volatilities. *Risk*, 17(5):102–106, 2004.

Márcio Poletti Laurini. Imposing no-arbitrage conditions in implied volatilities using constrained smoothing splines. *Applied Stochastic Models in Business and Industry*, 27(6):649–659, 2011.

Greg Orosi. Arbitrage-free call option surface construction using regression splines. *Applied Stochastic Models in Business and Industry*, 31(4):515–527, 2015.

Michael Roper. *Arbitrage Free Implied Volatility Surfaces.* PhD thesis, University of Sydney, NSW 2006, Australia, 2010.