One more no-arbitrage parametric fit of volatility smile

Andrey Itkin

New York University, School of Engineering,
6 Metro Tech Center, RH 517E, Brooklyn NY 11201,
aitkin@nyu.edu

July 31, 2014

Abstract

During last 15 years various parameterizations of the implied volatility (IV) surface were proposed in the literature to address few goals: a) given a set of market quotes for some options build a no-arbitrage local volatility (Dupire’s) surface to further exploit it for calibration of a local stochastic volatility model; b) obtain volatilities for pricing OTC options and other derivatives with strikes and maturities other than that offered by the option exchanges; c) produce a volatility forecast over future periods of time, which is helpful in Value-at-Risk models, computing forward IVs and exposures, etc.; d) assess an adequacy of an option pricing model based on the shape of the IV surface.

Among various existing static parameterizations SVI model of Gatheral is the most elaborated one which takes into account a correct asymptotic behavior at wings, and respects no-arbitrage conditions as well as provides a no-arbitrage interpolation and extrapolation. In this paper we propose another class of the static parameterizations that deliver same functionality but sometimes with a better quality of the fit.

Keywords: volatility surface, static parametrization, no-arbitrage interpolation and extrapolation

JEL classification: C6, C61, G17

1 Overview

An implied volatility surface is useful in many ways. Quoting [Borovkova & Parmana (2009)], first of all it can be used to obtain volatilities for pricing OTC options (and other derivatives) with strikes and maturities other than that offered by the option exchanges. Another important area is a risk management. Recall that the implied volatility (IV) corresponding to an option maturing in, e.g., three months is in fact a forecast of the average volatility over the next three month period. With this observation in mind, we can use the IV surface for producing a volatility forecast over future periods of time, which is helpful in Value-at-Risk
models, computing forward IVs and exposures, etc. Finally, the shape of the IV surface can be used to assess an adequacy of an option pricing model. In other words, if some model is able to fit a market generated surface with a non-constant volatility across strikes and maturities (which is an evidence for, e.g., jumps or a stochastic volatility in the price of the underlying security) one can consider this as a validation of the model.

There exist two major approaches to construct a no-arbitrage IV surface. The first one uses some stochastic model for the underlying spot or forward price which is calibrated to the market data. For instance, in the equity world one can use the popular Heston model, calibrate it to the market data and then use this model to find the IVs for the missing strikes and expirations where the market quotes are not available. By construction the IVs found in such a way are no-arbitrage. However, the main problem with this approach is that it is difficult to come up with a model which is reach enough to well fit the observed market data.

Another interesting rectification of this approach was proposed in Lipton & Sepp (2011) who calibrate a model with tiled local volatility to sparse market data by using the direct and inverse Laplace transforms. The main idea is to have a parametric form for the local volatility with as many parameters as there are market quotes. This allows finding an exact solution to the calibration problem at each forward time step, rather than to solve it in the least-squares sense. So the advantage of the method is that it is pretty fast. On the other hand, as the local volatility is built using a no-arbitrage model, the corresponding IV surface is no-arbitrage as well.

The other approach does not consider any model of the underlying, but instead uses some parametric fit of the implied volatility surface. Parametric models of the IV came to a regular consideration at the end of 1990s. Several parametric models for the IV surface were suggested by Dumas et al. (1998) and adapted and tested for FTSE options by Alentorn (2004). In the Dumas parametric model the IV surface is modeled as a quadratic function\(^2\) of the so-called normalized strike (rather than the strike price), which is defined as\(^3\) \(z = \log(K/F)/(\sigma \sqrt{T})\), \(K\) is the option strike, \(F\) is the forward price, \(T\) is the time to expiration, and \(\sigma\) is the normalization constant\(^4\). The normalized strike is a unit-less quantity. Zero moneyness corresponds to the At-The-Money (ATM) option. For a call option, positive moneyness corresponds to the In-The-Money option and negative moneyness corresponds to the Out-of-The-Money option. Usually the normalized strike is used under an assumption of ”sticky moneyness”\(^5\) (also known as ”sticky delta”) which allows elimination of refitting the volatility smile within some postulated period of time even when the underlying price changes. This is different from another popular assumption which is called a ”sticky strike” rule (see Derman & Kani (1994)).

\(^1\)He actually suggested a model for the local volatility, which, however, could be re-mapped to the implied volatility.

\(^2\)Later this approach was further extended by Tompkins (2001), Kotzé et al. (2013).

\(^3\)Some people also call it moneyness or log-moneyness, however we reserve this word for a standard definition of the forward moneyness as \(M = K/F\).

\(^4\)Usually, it is set either to 1, or to the ATM implied volatility.

\(^5\)i.e., the IV doesn’t change when \(z\) stays constant.
That being said it is important to underline that the IV parametrization is just a static fit to the current market snapshot of the option prices (IVs). In other words, such a parametrization doesn’t contain any dynamics of the IV. For instance, the above mentioned “sticky” assumptions about the future dynamics of the IVs are irrelevant to the parametrization itself. The argument about the normalized strike to be a convenient underlying variable if ones rely on a ”sticky log-moneyness” dynamics doesn’t mean that any other dynamics cannot be used having today’s IV surface as the initial condition represented as a function of the normalized strike. Another important clarification is that as the parametrization is static by nature, it is impossible to forecast the future IVs per se using this static fit. Rather, in this approach by forecasting the practitioners usually mean that when the IV of some option with time horizon (maturity) \( T \) is known, it provides some average value of volatility from today to \( T \). This, however, is not a property of the parametrization, but rather the property of the current option market to provide some ”on average” information about the future behavior of the stock market.

An extended work on modeling the IV surface using the static approach has been done by Gatheral (see, e.g., Gatheral (2004)). He used a different parametrization of the smile, known as stochastic-volatility-inspired (SVI) model, which is driven by a forward log-moneyness \( \chi = \log(K/F) \). Gatheral and co-workers also proposed some empirical dependencies of how parameters of the fit evolve with time, Gatheral (2006). Later in Gatheral & Jacquier (2011) it was shown that the SVI parametrization and the large-time asymptotic of the Heston implied volatility agree algebraically, which provided an additional theoretical justification for the above parametrization.

Some other static parameterizations were also proposed in the literature, e.g., see Fengler (2005), Romo (2010), Zhao & Hodges (2013) and references therein. In the next section we discuss the main highly desirable features that any such a parametrization should provide with. It could be observed that in contrast to the old approaches, recent models, e.g. the extended SVI model, and models in Kotzé et al. (2013), Zhao & Hodges (2013) do make account for these features, and thus could be useful in practice.

As far as a demand for the dynamic models of the IV is concerned, Cont & Fonseca (2002) considered the prices of the index options at a given date (they are usually represented via the corresponding IV surface) that clearly demonstrated skew/smile features and also a term structure, the behavior that several IV models have attempted to reproduce. They underlined that the IV surface also changes dynamically over time in a way that is not taken into account by the existing modeling approaches, giving rise to a Vega risk in option portfolios. Using time series of option prices on the S&P500 and FTSE indices, they studied the deformation of this surface and showed that it may be represented as a randomly fluctuating surface driven by a small number of orthogonal random factors. Then Cont and Fonseca identified and interpreted the shape of each of these factors, studied their dynamics and their correlation with the underlying index. A simple factor model compatible with the empirical observations was proposed. The authors illustrated how this approach simulates and improves the well-known ”sticky moneyness” rule used by option traders for updating the IVs. Their approach gave justification for using Vega when measuring the volatility risk, and provided decomposition of the volatility risk as a sum of contributions from empirically
identifiable factors.

It is worth mentioning that the popular assumptions of "sticky strike" or "sticky moneyness" are just an empirical rule-of-thumb. For instance, Ciliberti et al. (2008) analyzed these assumptions by considering in detail the skew of some stock option smiles, which is induced by the so-called leverage effect on the underlying, i.e., the correlation between past returns and future square returns. This naturally explains the anomalous dependence of the skew as a function of the option maturity. The market cap dependence of the leverage effect is analyzed using a one-factor model. The authors show how this leverage correlation gives rise to a non-trivial smile dynamics, which turns out to be intermediate between the sticky strike and the sticky delta rules. Finally, they compare their result with stock options data, and find that the option markets overestimate the leverage effect by a large factor, in particular, for the long-dated options. This subject requires some further investigation.

Another interesting idea was proposed in Carr & Wu (2010). This paper considers the future dynamics of the Black-Scholes implied volatility surface, and derives no-arbitrage constraints on the current shape of the volatility surface. Under the specified proportional volatility dynamics, the shape of the surface can be cast as solutions to a simple quadratic equation. Furthermore, corresponding to the option implied volatility for each contract, the paper defines a new, option-specific expected volatility measure that can be estimated from the historical sample price path of the underlying security. The measure is defined as the volatility input that generates zero expected delta-hedged gains from holding this option and can thus differ across different option strikes and expiries. Applying the new theoretical framework to the S&P500 index options market, the authors extract volatility risk and volatility risk premium from the two volatility surfaces, and find that the extracted volatility risk premium significantly predicts future stock returns. Thus, knowledge of the future dynamics also eliminates the necessity in any artificial assumptions like "stickiness", etc. See, also a recent paper of Sepp (2014).

So far, most of the IV researchers have been focused on Equity and FX derivatives. However, Borovkova & Parmana (2009) applied this idea to the option price data from oil markets. They combined the simplicity of the Gatheral parametric method with the flexibility of a non-parametric approach. The authors claim that the method can successfully deal with a limited amount of the option price data. Performance of the method was investigated by applying it to prices of the exchange-traded crude oil and gasoline options, and the results were compared with those obtained by a purely parametric approach. Furthermore, investigation of the relationship between volatilities implied from the European and Asian options showed that the Asian options in oil markets are significantly more expensive than the theoretical arguments imply.

It is interesting to note that traditional parametric models represent the smile as some polynomial function of $z$. One of the reasons for doing this is that according to Cont & Fonseca (2002), the IV patterns across moneyness vary less in time than when expressed as a function of the strike. Also, there is an additional computational benefit by regressing at moneyness rather than at the strike prices, since the function is of a simpler form, and, therefore, the estimation algorithm converges faster. A typical study is that of Alentorn
where using data in the FTSE 100 index, the following models were tested:

\[ \sigma(z, T) = \beta_0 + \beta_1 z + \beta_2 z^2 + \epsilon \]

\[ \sigma(z, T) = \beta_0 + \beta_1 z + \beta_2 z^2 + \beta_3 T + \beta_4 T z \epsilon, \]

where \( \beta_i, \ i \in [0, 4] \) are the regression parameters. Borovkova & Parmana (2009) use a similar regression. They also noticed that the parabolic shape of the implied volatility function for a fixed maturity is the average shape of the actual volatility functions. Note that increasing the power of the polynomial volatility function (from two to three or higher) does not really offer a solution here, since this volatility function will still be the same for all maturities.

Comparison of these models with the market data showed that they are able to capture a form of the volatility smile in the ATM region while often fail at wings. Another problem is fitting the smile close to expiration. Here \( T \to 0 \) implies \( z \to \infty \), and the volatility at wings tends to infinity which is not supported by the market data. Therefore, the regression coefficients \( \beta_1, \beta_2 \) must tend to zero, and the fitting function degenerates in this limit. This poses a real problem for the optimization routine (it never converges to such a limit).

Gatheral in his SVI parametrization uses another form

\[ w(\chi; a, b, \sigma, \rho, m) = a + b \left\{ \rho(\chi - m) + \sqrt{(\chi - m)^2 + \sigma^2} \right\}. \]

Here \( w \) is the total implied variance, \( a \) gives the overall level of variance, \( b \) gives the angle between the left and right asymptotes, \( \sigma \) determines how smooth the vertex is, \( \rho \) determines the orientation of the graph, and changing \( m \) translates the graph.

This form of the implied volatility surface is motivated by an asymptotic no-arbitrage argument, pioneered by Hodges (1996), Gatheral (1999) and later Lipton (2001) who mentions that the resulting IV(\( \chi \)) bounds are \( O(|\chi|^{1/2}) \) for large |\( \chi \)|. This was then further extended by the familiar results of Lee (2004) that the total implied variance should be linear in \( \chi \) at wings \( \chi \to \pm \infty \) with the slope \( 0 < \phi(\infty) < 2 \). As applied to the SVI Gatheral (2004) derives necessary and sufficient conditions for the IV surface to be arbitrage free and shows how this parametrization fits the IV surfaces generated by various currently popular models, including the stochastic volatility and jump models. Also some examples are provided where the SVI well fits the actual IV surfaces - even the notoriously hard-to-fit very short expirations.

However, as practitioners observed in their day-to-day trading, the SVI model sometimes fails to fit well the market data. The author's own experience also justifies a failure to fit this model to the data sets, obtained from some data providers. Also, according to Biscamp (2008) this model was thoroughly tested by practitioners in recent years and did not prove to work well for all products (like the index options, dispersion, equity options etc.). Therefore, some trading firms run their own proprietary models that exploit an idea of building a piecewise polynomial smile in the \( z \) space. This approach also has some problems, namely:

- Determining a boundary point between two pieces of the smile, where in addition the smile is \( C_2 \) continuous. Usually it requires to solve some non-linear equation, which is expensive. The necessity of solving the nonlinear equation slows down the volatility
smile fit, and especially computation of derivatives of the smile with respect to the model parameters which usually are computed by the bump-and-grind method.

- This approach still does not resolve the problem of fitting maturities close to expiration.
- This functional form does not fit the market data well.
- The asymptotic behavior of the smile at wings in $z$ does not agree with the result of Lee (2004) that the variance should be asymptotically linear in $z$.

All the above suggests that a new model suitable to better fit the static market volatility data could be helpful.

With allowance for the above our main goal in this paper is to propose another parametric fit which amounts to resolving the discussed issues with the existing approaches. We also show how to construct a no-arbitrage IV surface by using no-arbitrage interpolation and/or extrapolation if necessary.

We emphasize that according to Carr (2014) any such a formula must provide the following three properties:

1. It analytically describes implied volatilities instead of option prices.
2. It exactly fits any set of arbitrage-free mid-market implied volatilities.
3. It does not produce arbitrage.

While the first one is obvious, it is usually hard to guarantee the last two properties. In the approach of this paper first, we don’t guarantee an exact fit to the given mid-market quotes since we use a least-square optimization. However, we do guarantee, that the regressed implied volatility is in between of the given bid and ask, and is close (in some norm) to the mid price. Second, by construction we guarantee no-arbitrage in time. We also guarantee no-arbitrage on a given grid of strikes. However, we don’t suggest a no-arbitrage interpolation/extrapolation in the strike space. Therefore, definition of the reasonable grid of strikes is left up to the user of this approach. For instance, if one wants to use the local stochastic volatility model, and by doing that needs a local volatility (LV) surface calibrated to the market data, the appropriate LV grid could, e.g., coincide with the finite-difference grid in the spot space. Under this approach we are not interesting in the values of the implied volatilities in between the grid nodes, and therefore, the proposed method could be applied. We also guarantee a correct asymptotic behavior of the smile at both large positive and negative normalized strikes.

The above mentioned means that our model of implied volatility is a discrete model defined at a given given set of “states” (strikes), similar to, say, a discrete Markov chain model. And we are not aware of any continuous limit of this model at the moment.

---

6 This grid could be non-uniform, and it consists of the nodes with different strikes and/or maturities
7 Compare this with the SVI model of J. Gatheral where a nice result is available that the model structurally coincides with the hight $T$ asymptotic of the Heston model.
The rest of the paper is organized as follows. In section 2 we give a general outline of the model. Next section provides an asymptotic analysis of the model behavior at extreme strikes and expirations as well as ATM and some critical strikes. Based on this analysis we also are able to give a financial interpretation of some model parameters. In section 4 we explain how to construct a no-arbitrage IV surface and describe in detail our approach to no-arbitrage interpolation and extrapolation. In section 5 some numerical experiments are presented as well as stability of the fitted parameters with time is analyzed. The last section discusses some remaining issues.

2 Model

The construction of a new parametric model to fit the IV surface implied by some market is done based on the following assumptions.

1. As the independent variables of the parametric regression we choose the normalized strike $z$ and time to maturity $T$.

2. An asymptotic behavior of the variance smile at wings should be linear in $z^8$.

3. As there exist multiple justifications that the smile is not symmetric in the $z$ space, it is highly desirable to fit the call and put wings independently.

4. The parametric function must be continuous in $z$.

5. It should be well-behaved close to expiration.

6. We fit the term structure of the IV term-by-term, i.e., first the variance curve at the first maturity $T_1$, than the variance curve at the second maturity $T_2 > T_1$, etc. Therefore, we do not consider the dependence of the regression parameters on time. However, we do discuss how to build the whole no-arbitrage IV surface.

7. As a possible extension of this approach one can rely on the definition of $z$ where the calendar clock $T$ is replaced with a business clock $T_v$. Here we just mention this opportunity which apparently improves the fitting capability of the model, especially close to expiration, but don’t discuss it in detail.

8. The number of parameters must be minimal.

9. The parametric function must be fast to evaluate.

Note, that this model could be naturally extended to simulate a different behavior of the smile at call and put wings. Such a situation could be helpful when modeling commodities where one wing could demonstrate a linear behavior while the other one - sublinear, see e.g., Zhao & Hodges (2013). Here, however, this discussion is omitted to be presented elsewhere.

For instance, a recent extension of the SVI model proposed in Zhao & Hodges (2013) utilizes the Kummer hypergeometric functions which makes the computation expensive.
10. The whole IV surface should respect the no-arbitrage conditions.

Despite, at the first glance, the number of requirements is too big to easily satisfy all of them, below we provide a construction which seems to address every point on this list.

Given $T$, our new parametrization of one term at the IV surface reads

$$w(z) = w_c + SC\frac{y}{1 + y^2} + \sum_{i=1}^n a_i Y^i(y)$$

$$y = z - C, \quad Y(y) \equiv \begin{cases} \frac{1}{\alpha} \mathcal{G}[-\alpha y], & y \leq 0 \\ \frac{1}{\beta} \mathcal{G}[-\beta y], & y > 0 \end{cases}$$

where $w(z)$ is the total implied variance, $w(z) = I^2(z)T$, $I(z)$ is the implied volatility, $n$ determines the maximum degree of the polynomial on $Y(y)$, and $\mathcal{G}(x)$ belongs to the class of the so-called sigmoid functions, see von Seggern (2007). The sigmoid functions tend to some constant at both ends when the argument $x$ tends to $\pm\infty$, and vanish at $x = 0$. Many natural processes, including those of complex system learning curves, exhibit a progression from small beginnings that accelerates and approaches a climax over time. Besides the logistic function, sigmoid functions include the ordinary arctangent, the hyperbolic tangent, the Gudermannian function, and the error function, but also the generalized logistic function and algebraic functions like $x/\sqrt{1 + x^2}$.

From the performance point of view, we want $w(z)$ to be computed with the minimal possible number of computer operations. From this prospective using $\mathcal{G}(x) = \text{erf}(x)$ is apparently a good choice, since an existence of the approximate formula

$$\text{erf}(x) \approx 1 - (1 + a_1 x + a_2 x^2 + \ldots + a_6 x^6)^{-16},$$

with the maximum error $3 \cdot 10^{-7}$, where $a_1 = 0.0705230784, a_2 = 0.0422820123, a_3 = 0.0092705272, a_4 = 0.0001520143, a_5 = 0.0002765672, a_6 = 0.0000430638$. This approximation is valid for $x \geq 0$. To use it for the negative $x$, exploit the fact that $\text{erf}(x)$ is an odd function, so $\text{erf}(x) = -\text{erf}(-x)$, see Abramowitz & Stegun (1964).

It is worth mentioning that using polynomial functions in arctan($z$) was a popular choice among practitioners a while ago, however we don’t put this restriction. Also for clarity we fix $n = 2$ and provide a special notation for $S \equiv a_1, K \equiv a_2$. The reason for this notation will become clear right below.

Under these assumptions, $w(z)$ in Eq. (3) has 7 parameters:

- $C$ - shift. This is an edge point between the left and right branches of the smile. For equity options $C \approx 0$, i.e. this is close to the ATM point. Then, the left branch is a put wing while the right branch is a call wing. For index options the minimum of the smile is usually shifted into positive $z$. The parameter $C$ just reflects the value of this shift. Note, that the smile is $C_2$ at $z = C$. Indeed, the direct differentiation of $Y(z)$ in the Eq. (3) shows that the first derivative is continuous and reads $Y'(z)|_{z=C} = -1$, while the second derivative vanishes.
• $w_C$ - this is the variance at $z = C$.
• $S_C$ - this parameter determines skew of the smile at $z = C$.
• $\alpha$ - this is a put wing parameter which determines how steep the put wing should be.
• $\beta$ - this is a call wing parameter which determines how steep the call wing should be.
• $S$ - this parameter determines skew of the smile outside of the region $0 \leq z \leq C$.
• $K$ - this parameter determines kurtosis of the smile outside of the region $0 \leq z \leq C$.

In the limit $\alpha \to 0$ or $\beta \to 0$ we obtain $Y(z) \to C - z$.

Further on, for getting better results we need a minor refinement of the model. As the fitted variance $w(z)$ is expected to be at least $C^2$ continuous in $z$, it would be better to eliminate such a non-continuous function as $|z-C|$. This could be relatively easy done if we find a continuous approximation of the function $|z-C|$. Among various possible functions we chose that

$$|y| \approx y \tanh(py),$$

where $p$ is some constant parameter. Choosing $p$ big enough, say 1000, gives us highly accurate approximation of $|y|$ which is infinitely continuous.

3 Asymptotic analysis and meaning of the parameters

Below we provide an asymptotic analysis of the model to reveal the financial meaning of all the model parameters.

3.1 Behavior at $z = C$

To better understand why one needs a linear correction term, consider the asymptotic behavior of the smile. As $z \to C$ the function $w(z)$ behaves like

$$w(z) \approx w_C + S_C y - \left( S_C + pS \sqrt{T} Y_0(0) \right) y^3 + O(y^4).$$

(5)

Thus, this is a polynomial function of $z-C$ which is similar to what Dumas et al. (1998) model does. More rigorously, it is linear in $y$ at small $y$ if $p < 1/y$, and quadratic if we choose $p \approx 1/y$ at small $y$. Also as the parameter $\alpha$ determines the steepness of the smile in the put wing, it is reasonable to have an independent parameter to better shape the linear part of the smile near $z = C$. That is why in the Eq. (3) we introduced an extra term which is proportional to $zS_C$ at $z \approx C$, and vanishes at $z \to \infty$.

From Eq.(5) it is clear that $w_C$ is the total variance at $z = C$, and $S_C$ is the skew at $z = C$, while the kurtosis at $z = C$ vanishes. Also, it is seen that varying $p$ one can change the value of higher moments, which, however, for our analysis is not that important.
Thus, we can interpret the coefficients $w_C, S_C$ and $C$ as some form of adjustment for the critical point not being at $z = 0$.

Note that since the derivatives bear no dependence on $\beta$ (or $\alpha$), the model is indefinitely continuous around $z = C$.

### 3.2 Behavior ATM

Let’s consider the behavior of our model when $K = F \iff z = 0$. To simplify the analysis, we assume $C > 0$, $pC \gg 1$. The value of $p$ could be always chosen such that $pC \gg 1$ unless $C = 0$. Then $\tanh(pC) \approx 1$ if $C > 0$, and $\tanh(pC) \approx -1$ if $C < 0$. For easy of notation, denote $A \equiv Y(-C)$, $A' \equiv Y_y(-C)$, $A'' \equiv Y_{yy}(-C)$.

From the Eq.(3) the ATM variance is given by

$$w_0 = w_C - \frac{C}{1+C^2}S_C + AC\sqrt{T}\tanh(pC)\frac{S_K + AK}{\alpha^2} + O(y + C).$$

Accordingly, the ATM skew is approximately given by

$$S_{ATM} = -\frac{(C^2 - 1)}{(C^2 + 1)^2}S_C + \tanh(pC)\sqrt{T}[-A(S + KA) + CA'(S + 2KA)] + O(y + C),$$

and the ATM kurtosis is

$$K_{ATM} = -2S_C\frac{C(C^2 - 3)}{(C^2 + 1)^2} + \sqrt{T}\tanh(pC)[-2A'(K(2A - CA') + S) + CA''(2AK + S)] + O(y + C).$$

Further on, we want to determine a connection between the inflection point $C$ and parameters of the smile ATM. In order to do that, first suppose $C$ is small, but our assumption $pC \gg 1$ is still preserved because of a big $p$. Also, from our experiments with this model we know that a typical value of $K$ is about 1.0, $S$ is of the order of 1.0, and $\alpha$ varies from 0 to 5. Therefore, from Eq.(7) we obtain

$$C^2 = \frac{S_C - S_{ATM}}{3S_C - 2p\sqrt{T}Y_y(0)}$$

Thus, our assumption that the value of $C$ is small is true if $S_C - S_{ATM} \ll 3S_C - 2p\sqrt{T}Y_y(0)$. As $C$ is small (in other words, close to the ATM) the difference $S_C - S_{ATM}$ also has to be small.

At very small $T$ the above solution transforms to

$$S_{ATM} = -\frac{(C^2 - 1)}{(C^2 + 1)^2}S_C$$

At small $C$ this equation has the root

$$C = \sqrt{\frac{S_C - S_{ATM}}{3S_C}}$$
If $C$ is positive, the ATM point belongs to the put wing, and $S_{ATM} < 0$. Therefore, $S_C$ has to be positive in order for the Eq. (9) to be consistent.

Note, that the Eq. (9) does not contain $\alpha$ or $\beta$, therefore it is valid regardless of whether $C$ is positive or negative. Also from the Eq. (7) it follows that the minimum of the smile at $C = 0$ does not coincide with the ATM point.

Here we provide an example of a real smile computed using the proposed model. We run this test on Oct. 7, 2010 and fit the implied volatility of "EGO" options which mature on Oct. 15, 2010. The results of fitting are given in Fig. 1.

![Figure 1: Fitting of the IV smile for EGO, $T = 10/15/2010$.](image1.png)

![Figure 2: Fitting of the IV smile for EGO, $T = 11/19/2010$.](image2.png)

Parameters of the fit found by calibration are given in Tab. 1:

| $w_C$   | $S_C$   | $C$   | $S$   | $K$   | $\alpha$ | $\beta$ |
|---------|---------|-------|-------|-------|-----------|----------|
| 0.1652  | -0.04302| 0     | -0.20 | 1.035 | -0.42623  | 0.60308  |

Table 1: Experiment 1, parameters of the fit.

Thus, this smile does not demonstrate any shift of the minimum from ATM.

In the second example we fitted the next term of the same product. The results are given in Fig. 2. Parameters of the fit found by calibration are given in Tab. 2.
Table 2: Experiment 2, parameters of the fit.

| $w_C$ | $S_C$ | $\mathcal{C}$ | $S$ | $K$ | $\alpha$ | $\beta$ |
|-------|-------|--------------|-----|-----|----------|--------|
| 0.16775 | 0 | 0.5769 | -0.003 | 0.11 | -0.0004 | 2.7457 |

It is seen that in the test $\mathcal{C}$ is not a small parameter, therefore simple approximations suggested in the above cannot be used in this case.

The third example is given for XLF, also for the front term Oct.15, 2010. The results are given in Fig. 3.

Figure 3: Fitting of the IV smile for EGO, $T = 10/15/2010$.

Parameters of the fit found by calibration are given in Tab. 3.

| $w_C$ | $S_C$ | $\mathcal{C}$ | $S$ | $K$ | $\alpha$ | $\beta$ |
|-------|-------|--------------|-----|-----|----------|--------|
| 0.0703 | -0.038 | 0.0032 | -0.2 | 0.99 | 0.5741 | -0.80175 |

Table 3: Experiment 3, parameters of the fit.

In this test the calibrated value of $\mathcal{C}$ is small, so one can use the proposed approximations which connect the ATM skew and kurtosis with the value of $\mathcal{C}$.
3.3 Behavior At Infinity

As $K \to \infty$ at fixed $S$, so does $z$. Assume, for example, that $\mathcal{G}(x) \equiv \arctan(x)$. Expanding variance in series around positive infinity, we have

$$w(z) \approx w_C + \frac{\sqrt{T}}{\beta^2} (\beta S - \pi K) + \pi \sqrt{T} \frac{1}{2} \frac{\pi K - S \beta}{2 \beta^2} y + O(1/y). \quad (11)$$

Hence, the variance is linear in log-moneyness $\chi$ at positive infinity, with the slope

$$\phi(\infty) = \frac{1}{2} \pi \beta^{-2} \left( \frac{1}{2} \pi K - S \beta \right). \quad (12)$$

This well agrees with the result of [Lee (2004)]. Thus, our interpretation of $\beta$ follows: this parameter controls the slope of the smile at the infinite strike.

At $K \to 0$, $z \to -\infty$. Expanding variance in series around negative infinity, we have

$$w(x) \approx w_C - \frac{\sqrt{T}}{\alpha^2} (\alpha S + \pi K) - \pi \sqrt{T} \frac{1}{2} \frac{\pi K + S \alpha}{2 \alpha^2} y + O(1/y). \quad (13)$$

Hence, the variance is also linear in log-moneyness $\chi$ at negative infinity, with the slope

$$\phi(-\infty) = -\frac{1}{2} \pi \alpha^{-2} \left( \frac{1}{2} \pi K + S \alpha \right). \quad (14)$$

This also agrees with the result by [Lee (2004)]. Accordingly, our interpretation of $\alpha$ is: this parameter controls the slope at strike close to zero.

Close to expiration $z$ tends to infinity. However, for our function in the Eq. (3) this is not a problem. Indeed, at $T \to 0$, $z \to \infty$, the product $z \sqrt{T} \to \log K/F$, therefore from Eq. (11)

$$w(z) \to w_C + \frac{\pi}{2 \beta^2} \left( \frac{1}{2} \pi K - \beta S \right) \log \frac{K}{F}, \quad (15)$$

As mentioned in [Medvedev (2008), Ledoit et al. (2002)] in diffusion models, the ATM implied volatility is known to converge to the spot volatility when $T$ goes to zero. In our case from Eq. (15) the ATM value $w(z) |_{z=0} = w_C$, which implies $I(z) = \sigma_C = \text{const}$, $w_C = \sigma_C^2 T$. This provides another interpretation of the parameter $\sigma_C$ as the IV of the underlying stock at $T = 0$.

4 No arbitrage conditions

In contrast to the case where the IV surface is built based on some model, e.g., a stochastic volatility model, which guarantees no-arbitrage by construction, using regressions doesn’t provide such a nice feature per se. Therefore, in the latter cases a special care should be taken under calibration in order not to introduce arbitrage into the IV surface. See, for instance, [Andreasen & Hugel (2011), Lipton & Sepp (2011), Gatheral & Jacquier (2014)] and references therein.
The no-arbitrage conditions could be expressed in various forms. One of the approaches is to say that the local volatility function must be non-negative. The reason for that is that the local volatility function is directly related to the pdf (density) of the underlying, which in turn has to be non-negative. Then using Dupire’s formula for the local volatility as this is done in \cite{Gatheral2006}, i.e., representing it via the $w(z)$ function, one gets

$$
\sigma^2_{loc}(T, K) = \frac{\partial_T w}{\left(1 - \frac{\chi \partial_x w}{2w}\right)^2 - \frac{(\partial_x w)^2}{4} \left(\frac{1}{w} + \frac{1}{4}\right) + \frac{\partial_x^2 w}{2}},
$$

(16)

The nominator of this expression is the so-called calendar spread, and the denominator of it is equivalent to the so-called butterfly spread\footnote{The butterfly spread for the call option with price $C(T, K)$ is defined as $\frac{\partial^2 C(T, K)}{\partial K^2}$.} see \cite{Gatheral&Jacquier2014}. Both spreads must be non-negative for no-arbitrage.

However, as shown in \cite{Carr&Madan2005}, one more condition is required in addition to the above mentioned, which tells that so-called vertical call spread\footnote{The vertical spread for the call option with price $C(T, K)$ is defined as $\frac{\partial C(T, K)}{\partial K}$.} should be negative for the call options, or the vertical put spread should be positive for the put options. For the IV these conditions for the vertical spreads could be transformed to the following, see \cite{Carr2004}

$$
R(d_2) \sqrt{T} \leq K \frac{\partial I(K, T)}{\partial K} \leq R(-d_2) \frac{\sqrt{T}}{\sqrt{T}},
$$

(17)

where $R(d) \equiv \frac{1 - N(d)}{N'(d)}$ is Mill’s ratio, $N(d)$ is the normal cdf, and $d_2$ comes from the Black-Scholes formula. The convenience of such a representation lies in the fact that Mill’s ratio for the standard normal distribution reads

$$
R(x) = e^{x^2/2} \frac{\sqrt{\pi}}{2} \text{erfc} \left(\frac{x}{\sqrt{2}}\right).
$$

The latter can be efficiently computed by the particularly simple continued fraction representation at $x > 1$

$$
R(x) = \frac{1}{x + \frac{1}{x + \frac{2}{x + \ldots}}},
$$

(18)

or by using Taylor series expansion at $0 \leq x \leq 1$, see \cite{Gasull&Utzet2013}. The IV surface should also satisfy the asymptotic conditions discussed in the previous section, namely: the slope $\phi(\infty)$ of the call wing at $z \to \infty$ should be $0 \leq \phi(\infty) \leq 2$, and the slope of the put wing $\phi(\infty)$ at $z \to -\infty$ should be $0 \geq \phi(\infty) \geq -2$.

Being equipped with all these no-arbitrage conditions, the next step to consider is the construction of the IV surface in the domain $(T, K)$. This is not a problem if, say, we...
want to have the IV surface to be defined at some discrete grid in the \((T, K)\) space\(^{12}\) \(G : [T_i \times K_j], \ i \in [1, N], \ j \in [1, M]\) under two assumptions made: i) for every grid node \((i, j)\) there exists a market quote \(q(T_i, K_j)\) which is an option price (call or put, or both); ii) there is no need to ever know the IV at other possible values of \(T, K\) which don’t belong to \(G\). Certainly, in practice both assumptions are unrealistic. Therefore, some kind of interpolation/extrapolation which preserves no-arbitrage is necessary.

### 4.1 Finding parameters for one term

To obtain the values of the smile parameters, a non-linear least square optimization is used. Every market point is taken with some weight which is usually of the following form

\[
w(z) = \frac{1}{2} (w_c(z) + w_p(z)),
\]

\[
w_c(z) = (1 - |\Delta_c|) \min\left[0.1, \left(\frac{z}{\sigma_{atm}}\right)^{\nu}\right]
\]

\[
w_p(z) = (1 - |\Delta_p|) \min\left[0.1, \left(\frac{z}{\sigma_{atm}}\right)^{\nu}\right]
\]

Here \(\Delta_c\) and \(\Delta_p\) are the market call and put deltas of the option, \(\sigma_{atm}\) is the ATM market implied volatility, \(\nu\) is some parameter which is typically taken to be -2 or -3. Having these weights, the following optimization problem was solved to obtain parameters of the fit

\[
\min_{p_1...p_7} \sum_{i=1}^{N} W_i(z) \left[w_m(z_i) - w(z_i, p_1...p_7)\right]^2,
\]

where \(N\) is the total number of the raw option data, \(W_i(z)\) is the weight of the \(i\)th point, \(w_m\) is the market total implied variance of the data, \(\nu_i, i = 1, 7\) are the parameters of the model.

This minimization problem is solved under the whole bunch of no-arbitrage constraints discussed in the previous section. The no-arbitrage constraints are checked at every node on the grid \(G\), while the asymptotic slope is checked at two edge points on the grid for every time slice.

Further on, we calibrate all terms, provided as an input, that at least contain a single data point, by using bootstrap, i.e. term by term. We start with reordering all the market data in the ascending order, and then proceed with fitting the shortest term at \(T = T_1\). Then the next term at \(T = T_2\) is fitted, etc. To solve this optimization problem we use a genetic algorithm which guarantees to find a global minimum. A typical time necessary to get the values of the parameters for one term in C++ is about 0.5 secs with the maximum number of function evaluations set to \(10^4\).\(^{13}\) Note that as our algorithm belongs to the class of evolutionary optimization and is very suitable for parallelization.

\(^{12}\)That could be a grid where we want the local volatility function to be determined - a standard approach when one calibrates the LSV model to the market data.

\(^{13}\)Based on our experiments this value provides a very good fit, while it could be lowered to get a better performance.
We also need to underline that the above minimization problem is solved at points $z_i, i = 1...s$ where for the given term $T_j$ the market prices are available at $s$ strikes $\hat{K}_i, 1 = 1...s$ such that $z_i = (\log(\hat{K}_i/F(T_j)))/\sqrt{T_j}, i = 1...s$. However, the no-arbitrage constraints are checked at another set of points: that ones that belong to the $G$ grid. According to the definition of $G$ these points are $z_i = (\log(\hat{K}_i)/F(T_j))/\sqrt{T_j}, i = 1...M$.

**Smart initial guess.** When calibrating every term at the beginning we use a special algorithm to provide a good initial guess. The idea behind this algorithm is that as follows from Eq.(3) $w_{zz}(z) = 0$ at the point $z = C$. Therefore, one can look at the input IVs in the $z$ space, compute the second derivative and find where it vanishes. In case the second derivative is positive everywhere, as the value of $C$ one can take such $z$ where the IV is minimal among all values belonging to this term.

Given $C$ other parameters could be obtained relatively straightforward. Indeed, from Eq.(3)

$$w_C = w(z)|_{z=C}, \quad S_C = w_z(z)|_{z=C}. \quad (21)$$

Now denote

$$\kappa(y) = \frac{1}{y \tanh(py)\sqrt{T}} \left( w(y+C) - w(C) - SC \frac{y}{1+y^2} \right), \quad y \equiv z - C,$$

so based on Eq.(21) $\kappa(y)$ is a known function of $y$. Accordingly Eq.(3) could be re-written in the form

$$SY(y) + KY^2(y) = \kappa(y). \quad (22)$$

To find the initial guess for the remaining parameters $\alpha, \beta, S, K$ we need 4 additional market IVs. At lest two of them should lie on the different sides of the IV curve with regard to the point $y = 0$. As an example, consider three points $y_1 > y_2 > y_3 > 0$. Then, $Y(y)$ in Eq.(22) is defined via parameter $\beta$, see Eq.(3). Using Eq.(22) with $y_1$ and $y_2$ we find

$$S = \frac{\kappa(y_1)Y^2(y_2) - \kappa(y_2)Y^2(y_1)}{Y(y_1)Y^2(y_2) - Y(y_2)Y^2(y_1)}, \quad K = \frac{\kappa(y_2)Y(y_1) - \kappa(y_1)Y(y_2)}{Y(y_1)Y^2(y_2) - Y(y_2)Y^2(y_1)}. \quad (23)$$

Now using the point $y_3$ we numerically solve the equation Eq.(22) with regard to $\beta$.

Since parameters $S, K$ are already found, the last point $y_4 < 0$ could be used together with Eq.(22) to numerically solve for $\alpha$. This finalizes computation of the initial guess.

In case the input data points are located as $y_1 > y_2 > 0 > y_3 > y_4$ the easiest way is to add an extra point to the negative $y$ by using interpolation, and then remove one point, e.g., $y_2$ from the positive $y$, thus getting back to the previous case.

\textsuperscript{14}This construction also works well when the IV surface has a skew, not a smile, which is typical for index options.
4.2 No-arbitrage interpolation on the grid

Various approaches were discussed in the literature with regard to this problem. For instance, a no-arbitrage interpolation was considered in Andreasen & Huge (2011), Fengler (2005), Gatheral & Jacquier (2014) (see also references therein). Here, however, we suggest another approach, which is similar in spirit to that in Gatheral & Jacquier (2014).

If one works in the $z$ space given $T$ the usual approach would be to choose some number $\gamma$ such that all $z = [z_1...z_M]$ for this term are in the range $-\gamma < z/\sigma^* < \gamma$. Here $\sigma^*$ is some normalization constant which doesn’t depend on $T$. By financial meaning, $\sigma^*$ could be chosen as the ATM IV which corresponds to the shortest maturity, This is the most liquid strike of the instrument, and usually it is pretty well-known from the market data. In other words, for the IV surface just a range of $\gamma$ standard deviations in both up and down directions from the ATM is taken into account. Outside of this domain the remaining strikes are treated to be illiquid, and, therefore, they are taken out of consideration. In practical applications $\gamma = 5$ could be chosen, but this assumption could be easily relaxed.

However, if one needs to construct the IV surface and use it is the input to Dupire’s formula in order to build the corresponding local volatility surface, e.g., for fast calibration of the local stochastic volatility model to the market data, this approach should be changed. In this case for a given instrument we define a fixed domain in $K$ space: $[K_1...K_M]$. Accordingly, for the $z$ variable we have a map $z_i = \log(K_i/F(T_j))/\sqrt{T_j}$ which depends on the current expiration $T_j$, $j = 1...N$. In other words, we work on the $G$ grid which was described in the above.

Having a set of the IV input data for expirations $T_1 < T_2 < ... < T_m$ (which don’t coincide with the temporal nodes of $G$, but could be a subset of that) we calibrate our model term by term with allowance for the entire set of the no-arbitrage constraints on the given grid $G$ in the $z$ space. However, by construction this set of $z$ points exactly matches a set of strikes provided to build our grid at. At the end we obtain all values $w(z, t)$ where $z \in [-\gamma \sigma^* = z_1, ..., z_N = \gamma \sigma^*], T \in [T_1, ..., T_m]$. The corresponding undiscounted call and put prices can be obtained by using the Black-Scholes formula afterwards.

Next we need to interpolate thus obtained IVs to the $G$ grid. It is more convenient to proceed in the pricing space despite this is a bit more computationally intensive as we need to convert the IVs force to the prices at the beginning of this step, and back to the IVs at the end of this step. For the fitted terms after calibration is done we already know all parameters of the fit, so we are able to compute the call option value at any point at the $G$ grid. And the no-arbitrage conditions were already respected in these points as well. Also we know the corresponding map $K_i \rightarrow F(T_j)e^{z_i\sqrt{T_j}}, i = 1...M, j = 1...N$.

Accordingly, in $K$ space the no-arbitrage conditions for the call option: non-negativity...
of the calendar and butterfly spreads and non-positivity of the vertical spread read

\[ \frac{\partial C(K, T)}{\partial T} \geq 0, \]
\[ \frac{\partial C(K, T)}{\partial K} \leq 0, \]
\[ \frac{\partial^2 C(K, T)}{\partial K^2} \geq 0 \]

(24)

Now choose a monotonic time interpolation of \( C(K, T) \) at \( K=\text{const} \) of the form

\[ C(K, T) = \alpha(T)C(K, T_1) + [1 - \alpha(T)]C(K, T_2) \]

(25)

where \( T_1 < T < T_2 \) and

\[ \alpha(T) = \frac{a(T_2) - a(T)}{a(T_2) - a(T_1)}, \]

(26)

where \( a(T) \) is some monotonic function. Obviously, \( \alpha(T) \in [0, 1] \), and \( \alpha(T) \) doesn’t depend on \( K \). And this is a valid interpolation formula in a sense that the values of \( C(K, T) \) at \( T = T_1 \) and \( T = T_2 \) coincide with \( C(K, T_1) \) and \( C(K, T_2) \). Also thus defined \( C(K, T) \) provides

\[ \frac{\partial C(K, T)}{\partial T} = \alpha(T)\frac{\partial C(K, T_1)}{\partial T} + [1 - \alpha(T)]\frac{\partial C(K, T_2)}{\partial T} \]

\[ + \frac{\partial \alpha(T)}{\partial T} [C(K, T_1) - C(K, T_2)] \geq 0 \]

(27)

if \( \partial_T \alpha(T) < 0 \), i.e., \( \partial_T a(T) < 0 \). That is because we constructed \( C(K, T_1) \), \( C(K, T_2) \) such that they obey the no-arbitrage condition \( \partial_T C(K, T)|_{T=T_i} > 0 \), \( i = 1, 2 \).

It is easy to see that Eq.(25) also solves the second and third lines in Eq.(24) provided that these conditions were met at \( T = T_1 \) and \( T = T_2 \). The latter follows from our construction at the previous step of the algorithm. Also it can be shown that this expression still preserves the extreme slopes of the interpolated terms (that are at \( z \to \infty \) and at \( z \to -\infty \)) to follow the asymptotic conditions provided by Lee (2004). To see that, note that the latter could be represented in the form \( C(K, T, I) < C(K, T, \sqrt{2|\chi|/T}) \). Therefore, our interpolation provides

\[ C(K, T, I) = \alpha(T)C(K, T_1, I_1(K, T_1)) + (1 - \alpha(T))C(K, T_2, I(K, T_2)) \]

\[ < \alpha(T)C(K, T_1, \sqrt{2|\chi_1|/T_1}) + (1 - \alpha(T))C(K, T_2, \sqrt{2|\chi_2|/T_2}) \]

\[ < C(K, T, \sqrt{2|\chi|/T}). \]

(28)

The last equality holds because we interpolate at \( K=\text{const}, S=\text{const} \), so \( C(K, T, \sqrt{2|\chi|/T}) \) is a function of \( T \) only, and this is a concave function of \( T \).

As far as extrapolation is concerned, in addition to the no-arbitrage conditions we need to prove that the extreme slopes of the extrapolated terms (that are at \( z \to \infty \) and at \( z \to -\infty \)) still preserve the asymptotic conditions provided by Lee (2004).
We show that an extrapolation formula

\[ T^k C(K, T) = \alpha(T) T_1^k C(K, T_1) + [1 - \alpha(T)] T_2^k C(K, T_2) \]  

(29)

with \( k \in \mathbb{R}, k \leq -0.5 \) is suitable for this purpose.

Indeed, similar to Eq.(25)

\[ T^k C(K, T, I) = \alpha(T) T_1^k C(K, T_1, I_1(K, T_1)) + (1 - \alpha(T)) T_2^k C(K, T_2, I(K, T_2)) \]  

(30)

is a function of \( T \) only, and \( k < 0 \) is chosen such that \( f(T) \) is a convex function. The latter condition depends on the value of the interest rate \( r \), and \( \chi \). Usually, \( k = -1 \) is sufficient even for the ATM strikes\(^{15}\) Also thus defined \( C(K, T) \) provides

\[
T^k \frac{\partial C(K, T)}{\partial T} = \alpha(T) T_{1}^k \frac{\partial C(K, T_1)}{\partial T} + [1 - \alpha(T)] T_{2}^k \frac{\partial C(K, T_2)}{\partial T} + \frac{\partial \alpha(T)}{\partial T} [T_{1}^k C(K, T_1) - T_{2}^k C(K, T_2)] - k T^{k-1} C(K, T) \geq 0
\]

(32)

if \( \partial_T \alpha(T) < 0 \). That is because we constructed \( C(K, T_1), C(K, T_2) \) such that they obey the no-arbitrage condition \( \partial_T C(K, T)|_{T=T_i} > 0, i = 1, 2, \) and \( T_{1}^k C(K, T_1) - T_{2}^k C(K, T_2) < 0 \) since \( T_1 < T_2 \) and \( k = -1 \). Thus, the calendar spread is non-negative for the call option. The other two no-arbitrage conditions obviously follow.

## 5 Numerical experiments

### 5.1 Stability of the fitted parameters

In a typical experiment an SPX volatility smile was fitted using the proposed model. The raw data are collected for \( T = 0.6247 \) (228 days to expiration), \( F = 76.58, T_v = 0.6215 \). We find that our model fits the data pretty well, with parameters of the fit, obtained by running the above described minimization algorithm, given in Tab. 4.

\[
\begin{array}{cccccc}
  w_C & S & K & \alpha & \beta & C \\
  0.0435 & -0.763 & 74.5 & 3.12468 & 1.5000 & 0.2739 & -0.05921
\end{array}
\]

Table 4: Values of the parameters obtained in the test

\(^{15}\)As it could be easily seen from analysis of the Black-Scholes formula for the call option prices this is the most sensitive region. Therefore, the choice of, e.g., \( k = -0.5 \) could make \( f(T) \) to be concave close to ATM.
It is interesting to see, however, how the fitting parameters behave as a function of time. In other words, what is the sensitivity of the fit to changes in time. To investigate this we use the same data on the SPX closing IVs given for 133 sequential days and plot time-series of the model parameter values. These results are given in Fig. 4-10.

Figure 4: Sensitivity of $w_C$ to the time change.

Figure 5: Sensitivity of $S$ to the time change.

Figure 6: Sensitivity of $K$ to the time change.

Figure 7: Sensitivity of $C$ to the time change.
As one can see, the most time sensible parameters are $\alpha$ and $\beta$. So they need to be refitted more often, probably few times a day. At the time scale of a few days other parameters change just within 10-20%; therefore, they could be refitted less often.
We want to emphasize that by construction our model provides just the static fit of the current market snapshot of the options IVs, and does not consider any dynamics of the IV surface. Therefore, the dependence of the model parameters on time serves just to the illustrative purposes and helps in organizing a rapid calibration procedure. This does not mean that looking at the time dependence of the model parameters one can make a predictive conclusion on how the future IV behaves with time.

5.2 Constructing a local volatility surface

In this example we take data from http://www.optionseducation.org on XLF traded at NYSEArca on March 25, 2014. The spot price of the index is $S = 22.64$, the interest rate $r = 0.0148$. The option IVs are given in Tab. 5. We take all OTM quotes and some ITM quotes which are very close to the ATM. At the overlapped strikes for calls and puts we take an average of $I_{\text{call}}$ and $I_{\text{put}}$ with weights proportional to $1 - |\Delta|_c$ and $1 - |\Delta|_p$ correspondingly. We use the proposed parametric fit to construct the IV surface at all given expirations and strikes in a range $K \in [17, 28]$ with the step 0.5. In doing so we calibrate the first and the last term using the above described algorithm. The other terms are found on the grid

| T       | K, Put |
|---------|--------|
|         | 18     | 19 | 20 | 21 | 22 | 23 |
| 4/19/2014 | - 32.90 | 26.79 | 20.14 | 15.19 | 12.93 |
| 5/17/2014 | 33.27 | 26.88 | 23.08 | 18.94 | 16.12 | 13.86 |
| 6/21/2014 | 27.84 | 23.90 | 21.07 | 18.88 | 16.95 | 15.82 |
| 7/19/2014 | 26.09 | 22.81 | 20.29 | 18.13 | 16.30 | 14.93 |
| 9/20/2014 | 24.20 | 22.23 | 20.32 | 18.76 | 17.40 | 16.41 |
| 12/20/2014 | 23.75 | 22.09 | 20.67 | 19.44 | 18.36 | 17.60 |

| T       | K, Call |
|---------|---------|
|         | 21     | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 4/19/2014 | - 15.79 | 13.38 | 15.39 | - | - | - | - |
| 5/17/2014 | 16.71 | 14.48 | - 13.75 | - | - | - | - |
| 6/21/2014 | 16.31 | 14.78 | - 13.92 | 14.28 | 16.58 | - | - |
| 7/19/2014 | 16.82 | 15.24 | - 14.36 | 14.19 | 15.20 | - | - |
| 9/20/2014 | 17.02 | 15.84 | - 14.99 | 14.56 | 14.47 | 14.97 | 16.31 |
| 12/20/2014 | 17.63 | 16.61 | - 15.86 | 15.47 | 15.12 | 15.18 | 15.03 |

Table 5: XLF option IVs: C - call options, P - put options.

\[16\] By doing so we do take into account effects reported in Ahoniemi (2009) that the IVs calculated from identical call and put options have often been empirically found to differ, although they should be equal in theory. However, our weights are a pure empirical rule of thumb, and more detailed investigation of this is required.
by applying the no-arbitrage interpolation with $a(T) = C(K, T, I(K, T))$ and $K = 17$ (in this case $I(K, T)$ is assumed to be provided in the given set of data).

The results of this fitting are given term-by-term in Fig. 11. Accordingly, thus constructed the IV surface is represented in Fig. 12, and the local volatility surface and the implied density obtained from the IV surface by applying the Dupire’s and Breeden-Litzenberger formulas are given in Fig. 13, 14.

FIGURE 11: Term-by-term fitting of the IV surface constructed using the whole set of data in Tab. 5.

FIGURE 12: The IV surface obtained using the first and last terms in Tab. 5 and interpolation.

FIGURE 13: The local volatility surface produced from the IV surface.

FIGURE 14: The density implied from the IV surface in Fig. 12.

One can see that the local volatility is positive everywhere on the grid which is provided
by i) using the no-arbitrage constraints when calibrating each term, and ii) using the no-arbitrage interpolation instead of calibration for some terms.

Also the above results clearly show that calibration provides a very good fit to the market data for the first and the last term. However, for the other terms the no-arbitrage conditions could be very restrictive. Therefore, the genetic algorithm requires many evaluations of the objective function, and could be slow. In contrast, the no-arbitrage interpolation is very fast but doesn’t give such a good fit to the market data.

6 Discussion

When fitting the IV surface we rely on raw quotes for liquid options provided by the market. Unfortunately, markets are different. For instance, in the oil market only a few strikes are traded as European listed options, usually the ATM, one ITM and one OTM option. Other strikes are traded OTC and, moreover, as Asian options. Also, the smile behavior at wings could be different for index and equity options. The variance smile could still be linear in $z$ at wings but with a very different skew. And it seems there is no a clear theoretical reason why it could not be. Therefore, if somebody has just an intuition on how the smile wings should behave, he/she could better rely on this intuition rather than on some unreliable illiquid data, and treat the latter as outliers.

Fortunately, the proposed model is able to address such an intuition by doing the following trick. Suppose that we want to have the new model for the index options smile at the call wing being as close as possible to the existing smile produced by some proven (reference) model. Then we can move the value of the call wing parameter $\beta$ into a different region by imposing a special constraint. By doing that, we make the fit a bit worse, but thus found slope (after the minimization is done) turns to be closer to the corresponding reference model skew. And we still preserve the continuity of the model.

Another issue with the model is as follows. Suppose, for a given term the number of strikes for which the market quotes are available, is less than the number of the model parameters, i.e. 7. In this case the parametrization is over determined. The no-arbitrage constraints and the asymptotic behavior of the smile help to resolve this, however could not be sufficient. For instance, if only the ATM quote is liquid and available. In this situation we have either to reduce the number of parameters, or to use some tricks.

To give an example of such a trick consider the case when only a single quote $I$ corresponding to the strike $K$ is available given the time to expiration $T$. Under this situation it doesn’t make sense to calibrate our parametrization to this single point. Instead, we treat the entire term as fully unknown, and find the IV values by using the no-arbitrage interpolation in time as it was described in the above. However, to fit exactly thus found IVs to the given quote we exploit the remaining flexibility in the definition of the function $a(T)$. We remind that $a(T)$ was not yet defined explicitly, and only indirectly via the condition $\partial_T a(T) > 0$. Then taking $a(T) = C(K, T, I)$ provides this inequality on one hand. On the other hand, as it could be easily checked, thus defined $a(T)$ matches exactly the given quote $C(K, T, I)$. 

24
When 3 or 4 quotes are available for the given term, the kurtosis $\mathcal{K}$ is a natural candidate to be removed from the parametrization. In other words, we fix $\mathcal{K} = 0$ and so reduce the total number of parameters to 6. Also, $\mathcal{S}_C$ could be the next preferable choice to omit.
Acknowledgments

I thank Peter Carr, Alex Lipton, Dmitry Kreslavsky, Roza Galeeva, Antonie Kotzé, Lewis Biscamp and Ping Sun for fruitful discussions.

References

Abramowitz, M., & Stegun, I. 1964. *Handbook of Mathematical Functions*. Dover Publications, Inc.

Ahoniemi, K. 2009. *Modeling and forecasting implied volatility*. Ph.D. thesis, Helsinki School of Economics.

Alentorn, A. 2004. *Modelling the implied volatility surface: An empirical study for FTSE options*. Tech. rept. Centre of Computational Finance and Economics Agents, University of Essex.

Andreasen, J., & Huge, B. 2011. Volatility Interpolation. *Risk Magazine*, March, 76–79.

Biscamp, L. 2008. *private communication*.

Borovkova, S., & Parmana, F. 2009. Implied volatility in oil markets. *Computational Statistics and Data Analysis*, 53(6), 2022–2039.

Carr, P. 2004 (November). *Implied Vol Constraints*. available at http://www.javaquant.net/papers/impvolconstrs3.pdf

Carr, P. 2014. *Private communication*.

Carr, P., & Madan, D. 2005. A note on sufficient conditions for no arbitrage. *Finance Research Letters*, 2, 125–130.

Carr, P., & Wu, L. 2010 (Oct.). *A New Framework for Analyzing Volatility Risk and Premium Across Option Strikes and Expiries*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1701685

Ciliberti, S., Bouchaud, J.P., & Potters, M. 2008. Smile dynamics - a theory of the implied leverage effect. 0809.3375v1 [q-fin.PR]: arXiv.

Cont, R., & Fonseca, J.D. 2002. Dynamics of implied volatility surfaces. *Quantitative Finance*, 45–60.

Derman, E., & Kani, I. 1994. *The Volatility Smile and Its Implied Tree*. Tech. rept. GS Quantitative Strategies Research Notes.

Dumas, B., Fleming, J., & Whaley, R.E. 1998. Implied volatility functions: Empirical tests. *The Journal of Finance*, LIII(6), 2059–2106.
Fengler, M.R. 2005. *Arbitrage-Free Smoothing of the Implied Volatility Surface*. Tech. rept. SFB 649 Discussion Paper. Trading & Derivatives.

Gasull, A., & Utzet, F. 2013 (July). *Approximating Mills ratio*. available at [http://arxiv.org/pdf/1307.3433v1.pdf](http://arxiv.org/pdf/1307.3433v1.pdf).

Gatheral, J. 1999. *The volatility skew: Arbitrage constraints and asymptotic behaviour*. Tech. rept. Merrill Lynch.

Gatheral, J. 2004. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. Global Derivatives And Risk Management.

Gatheral, J. 2006. *The volatility surface*. Wiley finance.

Gatheral, J., & Jacquier, A. 2011. Convergence of Heston to SVI. *Quantitative Financet, 11*(8), 1129–1132.

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

Hodges, H.M. 1996. Arbitrage Bounds of the Implied Volatility Strike and Term Structures of European-Style Options. *Journal of Derivatives, 3*(4), 23–32.

Kotzé, A., Labuschagne, C.C.A., Nair, M.L., & Padayachi, N. 2013. Arbitrage-free implied volatility surfaces for options on single stock futures. *North American Journal of Economics and Finance, 26*, 380–399.

Ledoit, O., Santa-Clara, P., & Yan, S. 2002. *Relative pricing of options with stochastic volatility*. Tech. rept. Eller College of Business and Public Administration, The University of Arizona.

Lee, R. 2004. The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance, 14*(3), 469–480.

Lipton, A. 2001. *Mathematical Methods For Foreign Exchange: A Financial Engineer’s Approach*. World Scientific.

Lipton, A., & Sepp, A. 2011. Filling the gaps. *Risk Magazine*, October, 86–91.

Medvedev, A. 2008. *Implied Volatility at Expiration*. Research Paper Series 08 - 04. Swiss Finance Institute.

Romo, J.M. 2010 (April). *Dynamics of the Implied Volatility Surface: Theory and Empirical Evidence*. Available at [http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2052215](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2052215)
Sepp, A. 2014. *Log-Normal Stochastic Volatility Model: New Insight and Closed-form Solution for Vanilla Options*. Tech. rept. BAML.

Tompkins, R. 2001. Implied volatility surfaces: Uncovering regularities for options on financial futures. *The European Journal of Finance*, 7(3), 198–230.

von Seggern, D. 2007. *CRC Standard Curves and Surfaces with Mathematics*. 2nd edn. CRC Press.

Zhao, B., & Hodges, S. D. 2013. Parametric modeling of implied smile functions: a generalized SVI model. *Review of Derivative Research*, 16(53–77).