Extended Abstract:

Direct Fit for SVI Implied Volatilities

Wolfgang Schadner*

May 17, 2023

Abstract

The stochastic volatility inspired (SVI) formula is one of the mainstream models for fitting the option implied volatility smile. It is traditionally calibrated using multi-parameter non-linear optimization routines. In contrast, in this study I reveal that the SVI equation represents a conic section, limited to the geometric shape of a hyperbola. This insight enables a full linearization of the SVI formula, making the fitting task significantly simpler. By using the linearized equation, we can fit the empirical data directly with a non-iterative and type-specific least-squares method, resulting in a computational efficient and numerical stable closed-form solution. To prove the effectiveness of this approach, I conduct calibration experiments on empirical data of various asset classes, demonstrating the robustness and accuracy of the method.

Keywords: implied volatility, volatility smile, calibration, SVI, hyperbola
JEL: C58, G12

1 Introduction

The *stochastic volatility inspired* (SVI) formula was introduced by Jim Gatheral in 2004 [4]. It quickly became one of the leading references for implied volatility modeling due to its inherent simplicity, asymptotic relation to stochastic volatility models (like Heston, cf. [3, 6]), and virtually arbitrage-free nature (cf. [2]). Not only does it enjoy an ongoing vivid academic discussion [7, 9, 13, 14], it also became widely recognized within the practitioners literature [1, 10, 11] and found its way as a prime example into finance textbooks [5, 8]. Gatheral's [4] original "raw" parameterization is defined as

$$\omega(x) = a + b\left[\rho(x-m) + \sqrt{(x-m)^2 + \sigma^2}\right] \tag{1}$$

*University of St. Gallen. E-Mail: wolfgang.schadner@unisg.ch. Web: www.w-schadner.com

with $\omega(x)$ as the total implied variance for the level of forward log-moneyness x and $\{a, b, \rho, m, \sigma\} \in \mathbb{R}$ as the five model parameters. The total implied variance per se follows the standard notation of $\omega(x) = \tau\sigma_{BS}(x)^2$, having $\sigma_{BS}(x)$ as the Black-Scholes implied volatility for the time-to-maturity τ. Obviously, the SVI formula (Eq.1) involves multiple parameters and exhibits a non-linear relationship. Hence, the traditional approach for fitting the SVI equation uses sophisticated non-linear optimization algorithms. These come with numerous drawbacks such as a strong reliance on initial parameter choice, convergence towards local minima, and computational inefficiency. As a result, it is generally preferred to simplify and linearize optimization problems whenever feasible. This is precisely what this work aims to achieve, proposing a numerically stable and efficient method for fitting the SVI formula.

A partial linearization was already introduced by [12], which became a calibration reference on its own as it reduced the problem from a five- down to a two-parameter optimization problem. Different to this quasi-explicit solution I derive a fully-explicit, non-iterative solution. The key to my approach lies in recognizing that the SVI formula represents a conic section, with only hyperbolic or flat geometric shapes being feasible. Leveraging this conic representation, I draw upon fitting methods from the pattern recognition literature, specifically designed for this type of task. To evaluate its efficacy, I conduct calibration experiments using both simulated and empirical data from the S&P 500 index. These experiments demonstrate a high level of goodness of fit, alongside excellent computational performance, highlighting the considerable potential of the proposed direct least-squares method for future applications.

The reading is set up as follows. Section 2 details the methodology employed, and the corresponding calibration experiments are presented in Section 3. The study is summarized in Section 4, along with a discussion on possible avenues for future research.

2 Method

2.1 Conic Representation

The SVI model is frequently called a hyperbola (e.g., [16]), however, within the existing literature I could not identify any formal proof for that statement. Hence, this section briefly elaborates on the geometry of SVI as it motivates the usage of conic fitting algorithms for SVI calibration.

Proposition 2.1. *The SVI model of Eq.1 has the geometry of a conic section.*

Proof. The square-root within SVI fades away when rearranging terms and taking squares,

$$[\omega(x) - a - b\rho(x - m)]^2 = b^2[(x - m)^2 + \sigma^2] \tag{2}$$

. Next, expanding the above equation, bringing all terms to one side, and introducing the conic coefficients $\mathbf{z}$

$$\mathbf{z} = \begin{pmatrix} z_{[1]} \\ z_{[2]} \\ z_{[3]} \\ z_{[4]} \\ z_{[5]} \\ z_{[6]} \end{pmatrix} := \begin{pmatrix} b^2(\rho^2 - 1) \\ 1 \\ -2b\rho \\ 2mb^2 - 2b\rho(b\rho m - a) \\ 2(b\rho m - a) \\ (b\rho m - a)^2 - b^2(m^2 + \sigma^2) \end{pmatrix} \tag{3}$$

results in the quadratic equation

$$z_{[1]}x^2 + z_{[2]}\omega(x)^2 + z_{[3]}x\omega(x) + z_{[4]}x + z_{[5]}\omega(x) + z_{[6]}1 = 0 \tag{4}$$

. With respect to definitions from existing literature (e.g., [15]), this is known as the algebraic representation of a conic section. □

The great advantage of the conic representation is that it fully linearizes the otherwise non-linear SVI equation. Therefore, it drastically simplifies the fitting task. Under the conic representation the total implied variance $\omega(x)$ is calculated from

$$\omega(x) = \frac{-(z_{[3]}x + z_{[5]}) + \sqrt{(z_{[3]}x + z_{[5]})^2 - 4z_{[2]}(z_{[1]}x^2 + z_{[4]}x + z_{[6]})}}{2z_{[2]}} \tag{5}$$

.[1]

Remark 2.1. *It is known (cf. [15]) that the algebraic conic equation produces four different geometric shapes. These are distinguished by their quadratic coefficients:*

$$z_{[3]}^2 - 4z_{[1]}z_{[2]} \begin{cases} > 0 \ldots \textit{ hyperbola} \\ = 0 \ldots \textit{ parabola (or linear)} \\ < 0 \ldots \textit{ ellipse} \end{cases} \tag{6}$$

Proposition 2.2. *Given the SVI parameters $\{a, b, \rho, m, \sigma\}$ are real, the SVI model produces only two different shapes: flat or hyperbolic.*

Proof. Follows from Lemma 2.1-2.3. □

Lemma 2.1. *The SVI model produces a hyperbolic shape if, and only if, $b \neq 0$.*

[1] Eq.5 is a quadratic equation which comes with two solutions for $w(x)$. However, by the convexity of implied variance only the upper solution is of relevance.

Proof. Take Remark 2.1 and substitute for Eq.3. What we observe is

$$b^2 = \tfrac{1}{4} z_{[3]}^2 - z_{[1]} z_{[2]} \tag{7}$$

Hence, with $b \in \mathbb{R}$ and $b \neq 0$ it follows that $0 < b^2 \implies z_{[3]}^2 - 4 z_{[1]} z_{[2]} > 0$. Therefore, from Remark 2.1 we know that SVI forms a hyperbola whenever $b \neq 0$. □

Lemma 2.2. *The SVI model has no parabolic solution.*

Proof. From Remark 2.1 and Eq.7 we know that $b = 0$ is necessary for SVI to produce a parabolic shape. However, by Eq.1 we recognize that $b = 0$ implies $\forall x : \; w(x) = a$. Hence, $b = 0$ produces the flat Black-Scholes implied volatility surface, but not a parabola. □

Lemma 2.3. *The SVI model has no elliptic solution.*

Proof. Since $b \in \mathbb{R}$ it follows that $b^2 \geq 0$. Therefore, given Eq.7, we know that $z_{[3]}^2 - 4 z_{[1]} z_{[2]} < 0$ does not exist within the SVI model. □

The conic equation of Eq.4 can be arbitrarily scaled, which allows for different representations of the SVI parameters in terms of its conic coefficients $\mathbf{z}$. In Proposition 2.3 we observe that this is particularly useful for introducing parameter constraints.

Proposition 2.3. *A necessary condition for the positivity of $w(x)$ is that $\rho \in [-1, 1]$ (cf. [13]), which holds whenever $-z_{[1]} z_{[2]} \geq 0$.*

Proof. This can be seen when normalizing $\mathbf{z}$ by $z_{[3]}$. Let the normalized coefficients be denoted by $\tilde{\mathbf{z}}$, from which follows that $-4 \tilde{z}_{[1]} \tilde{z}_{[2]} = \frac{1-\rho^2}{\rho^2}$. Therefore, $-z_{[1]} z_{[2]} \geq 0$ implies $\rho \in [-1, 1]$ and vice versa. □

The constraint of $\rho \in [-1, 1]$ describes that the branches towards which $w(x)|_{x \to \pm\infty}$ converges will have a non-negative slope. Hence, is obligatory for $w(x)$ being positive. With an eye on Remark 2.1 we recognize that $-z_{[1]} z_{[2]} \geq 0$ strictly produces a subset of possible hyperbolas. For fitting the SVI equation it is thus useful to limit ρ to ± 1, and work with the more stringent constraint $-z_{[1]} z_{[2]} \geq 0$ instead of Remark 2.1.

The conic representation can be also used for direct calibrations of the SSVI and vanishing SVI submodels, briefly discussed below.

2.1.1 Vanishing SVI

The vanishing SVI is defined as the implied variance surface becoming flat within the left or the right wing (cf. [13]), which is equivalent to set $\rho = \pm 1$. From this follows that $z_{[1]} = 0$ while the rest of $\mathbf{z}$ remains equal to Eq.3. With respect to Remark 2.1 we recognize that this always produces the hyperbolic shape (or flat if

$z_{[3]} = 0$). Therefore, the vanishing SVI can be directly fit by an unconstrained ordinary least-squares with $-w(x)^2$ as the dependent variable and the remaining terms (excluding x^2) as the explanatory variables. The sign of ρ is thereby automatically detected and the fitted variance will converge towards the parameter a at $x \to \pm\infty$. Note that a represents the minimum of $w(x)$ only in the vanishing SVI model, but not for the general cases where $|\rho| < 0$. Following, to avoid the vanishing $w(x)$ to converge towards a negative value it is possible to pre-specify the parameter a while still directly fitting it without any constraints. In that case, the dependent variable is $\{-(w(x) - a)^2\}$ and the three explanatory variables are $\{(w(x) - a)x,\ w(x),\ \mathbf{1}\}$, related to the coefficients $\mathbf{z}_v$. The link between $\mathbf{z}_v$ from the fit and the original SVI parameters is given by

$$\begin{pmatrix} b & \rho & m & \sigma \end{pmatrix} = \begin{pmatrix} \frac{1}{2}|z_{v[1]}| & \frac{z_{v[1]}}{|z_{v[1]}|} & \frac{z_{v[3]}}{z_{v[1]}} & 2\sqrt{-\frac{a z_{v[3]} + z_{v[2]}}{|z_{v[1]}|}} \end{pmatrix} \tag{8}$$

2.1.2 SSVI

The key advantage of the SSVI model [7] is that it has traceable conditions under which it produces an arbitrage-free volatility surface. For a slice of implied variances, the model is defined by

$$\omega(x) = \frac{\theta}{2}\left(1 + \rho\phi x + \sqrt{(\phi x + \rho)^2 + 1 - \rho^2}\right) \tag{9}$$

. This equation is linearized by the conic representation using the coefficients

$$\begin{pmatrix} z_{[1]} \\ z_{[2]} \\ z_{[3]} \end{pmatrix} = \begin{pmatrix} \frac{1}{4}\theta^2\phi^2(\rho^2 - 1) \\ 1 \\ -\theta\rho\phi \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} z_{[4]} \\ z_{[5]} \\ z_{[6]} \end{pmatrix} = \begin{pmatrix} 0 \\ -\theta \\ 0 \end{pmatrix} \tag{10}$$

and substituting back into Eq. 4.

2.2 Direct Least-Squares

The linearized SVI equation Eq.4 might be directly fit via unconstrained OLS, which has the chance of already producing the desired solution. However, by Proposition 2.2 we now know that only the hyperbolic solutions are feasible, such that $z_{[3]}^2 - 4z_{[1]}z_{[2]} \geq 0$ is a necessary condition for SVI calibrations. Extending on that I work with the more stringent constraint of $-z_{[1]}z_{[2]} \geq 0$ (Proposition 2.3) to ensure that $\rho \in [-1, 1]$. The literature on effectively fitting a conic section is a field on its own. The approach I chose herein is based on [15]'s direct type-specific least-squares with slight modification to match the SVI-specific Proposition 2.3.

Using the input data of $\mathbf{x}$ and $\boldsymbol{\omega}$ (vector of x's and $\omega(x)$'s), I partition the design matrix D into the

constrained D_2 and unconstrained D_1 sub-matrices:

$$D = \Big(\underbrace{\mathbf{x}^2 \quad \boldsymbol{\omega}^2}_{D_2} \quad \underbrace{\mathbf{x}\boldsymbol{\omega} \quad \mathbf{x} \quad \boldsymbol{\omega} \quad \mathbf{1}}_{D_1}\Big) \tag{11}$$

. The same split is applied for the conic coefficients $\mathbf{z}_2 = \begin{pmatrix} z_{[1]} & z_{[2]} \end{pmatrix}'$ and $\mathbf{z}_1 = \begin{pmatrix} z_{[3]} & z_{[4]} & z_{[5]} & z_{[6]} \end{pmatrix}'$. The quadratic parameter constraint is implemented via the constraint matrix C,

$$\textit{constraint:} \qquad \mathbf{z}_2' C \mathbf{z}_2 > 0 \qquad \text{with} \qquad C = \begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} \tag{12}$$

which also ensures that the fitted coefficients produce a hyperbola.[2] The optimization problem now reads as

$$\min \mathbf{z}' D' D \mathbf{z} \quad \text{s.t.} \quad \mathbf{z}_2' C \mathbf{z}_2 \geq 0 \tag{13}$$

. Let us introduce the scatter matrices $S_{ij} = D_i' D_j$ to reduce notation. The Lagrangian then writes as

$$\mathcal{L} = \begin{pmatrix} \mathbf{z}_2 & \mathbf{z}_1 \end{pmatrix} \begin{pmatrix} S_{22} & S_{21} \\ S_{12} & S_{11} \end{pmatrix} \begin{pmatrix} \mathbf{z}_2 \\ \mathbf{z}_1 \end{pmatrix} - \lambda \mathbf{z}_2' C \mathbf{z}_2 \tag{14}$$

with the partial derivatives set to zero for the optima:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{z}_2} \overset{\text{set}}{=} 0 : \qquad S_{22}\mathbf{z}_2 + S_{21}\mathbf{z}_1 - \lambda C_2 z_2 = 0 \tag{15}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{z}_1} \overset{\text{set}}{=} 0 : \qquad S_{12}\mathbf{z}_2 + S_{11}\mathbf{z}_1 = 0 \quad \Longrightarrow \quad \mathbf{z}_1 = -S_{11}^{-1} S_{12} \mathbf{z}_2 \tag{16}$$

. When defining $M := S_{22} - S_{21} S_{11}^{-1} S_{12}$, then from $\frac{\partial \mathcal{L}}{\partial \mathbf{z}_2} = 0$ follows that

$$C^{-1} M \mathbf{z}_2 = \lambda \mathbf{z}_2 \tag{17}$$

represents an eigenvalue problem. As the inverse matrix $C^{-1} = \left(\begin{smallmatrix} 0 & -2 \\ -2 & 0 \end{smallmatrix}\right)$ is known, solving Eq.17 is straightforward. The eigendecomposition of $C^{-1}M$ returns two optimal eigenvectors for $\mathbf{z}_2$. Of these two eigenvectors, the to one related to the smallest positive eigenvalue represent the optimal solution for $\mathbf{z}_2$. Next, $\mathbf{z}_1$ is calculated from Eq.16 and the SVI parameters are recovered from Eq.3.

[2] For the more general, hyperbola-only constraint (Remark 2.1) the matrix shall be partitioned into $D_2 = \begin{pmatrix} \mathbf{x}^2 & \boldsymbol{\omega}^2 & \mathbf{x}\boldsymbol{\omega} \end{pmatrix}$ and $D_1 = \begin{pmatrix} \mathbf{x} & \boldsymbol{\omega} & \mathbf{1} \end{pmatrix}$. The corresponding constraint matrix is $C = \left(\begin{smallmatrix} 0 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$, see [15].

3 Calibration Experiments

The empirical applicability of the direct least-squares fitting method is tested upon various different samples. The asset classes represented are stocks (Apple, Amazon, Microsoft), cryptocurrency (Bitcoin), foreign exchange (EUR/USD rate), metals (gold), commodities (WTI oil), bonds (10-year Treasury Note) and agriculture (Wheat). The target maturities between the samples vary dependent on availability, but roughly represent one-, three-, and six months. The data sources are OptionMetrics for stocks and CME's DataMine (accessed via QuickStrike) for all other asset classes. The data's observation dates are chosen as the most recent available ones, which are 12/31/2021 for OptionMetrics and 4/7/2023 for CME. For the OptionMetrics data I fit Put and Call implied volatility smirks separately. In contrast, the CME data are joint samples of OTM Puts and Calls.

For the calibration experiments the direct least-squares method is applied upon slices of implied variance (i.e., the squared Black-Scholes implied volatility). The SVI curves are fit under the constrain of Proposition 2.3, and the raw implied volatility slices together with the fitted curves are visualized in Figs.1-2. Tab. 1 reports the worst goodness-of-fit R^2 and the worst mean absolute error MAE per underlying as measures for the fitting quality.

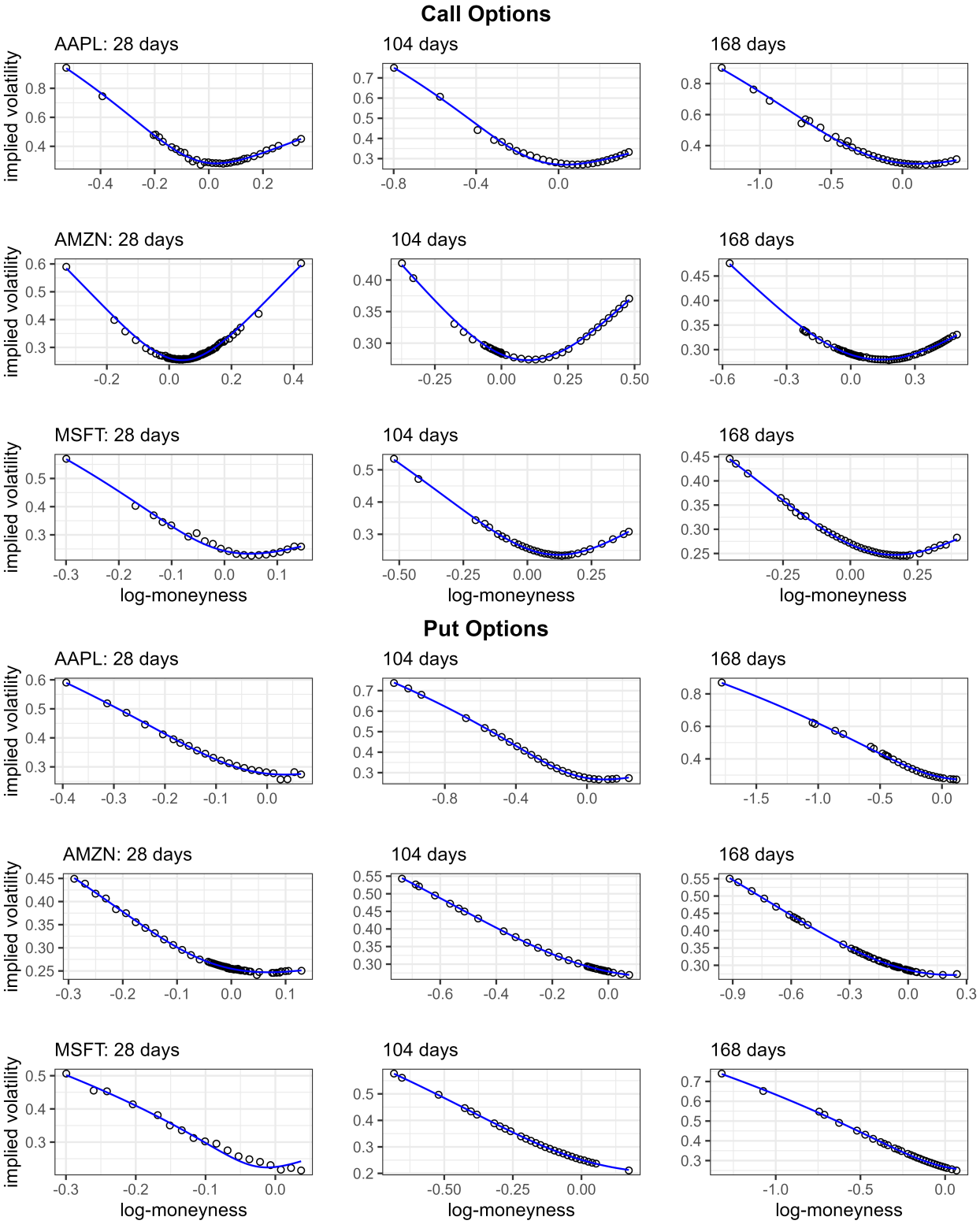

Figure 1: Additional empirical calibration experiments for the stocks of Apple, Amazon, and Microsoft. Here, the direct least-squares method is fitted under the constraint of $-z_{[1]}z_{[2]} > 0$ to ensure that $\rho \in [-1, 1]$.

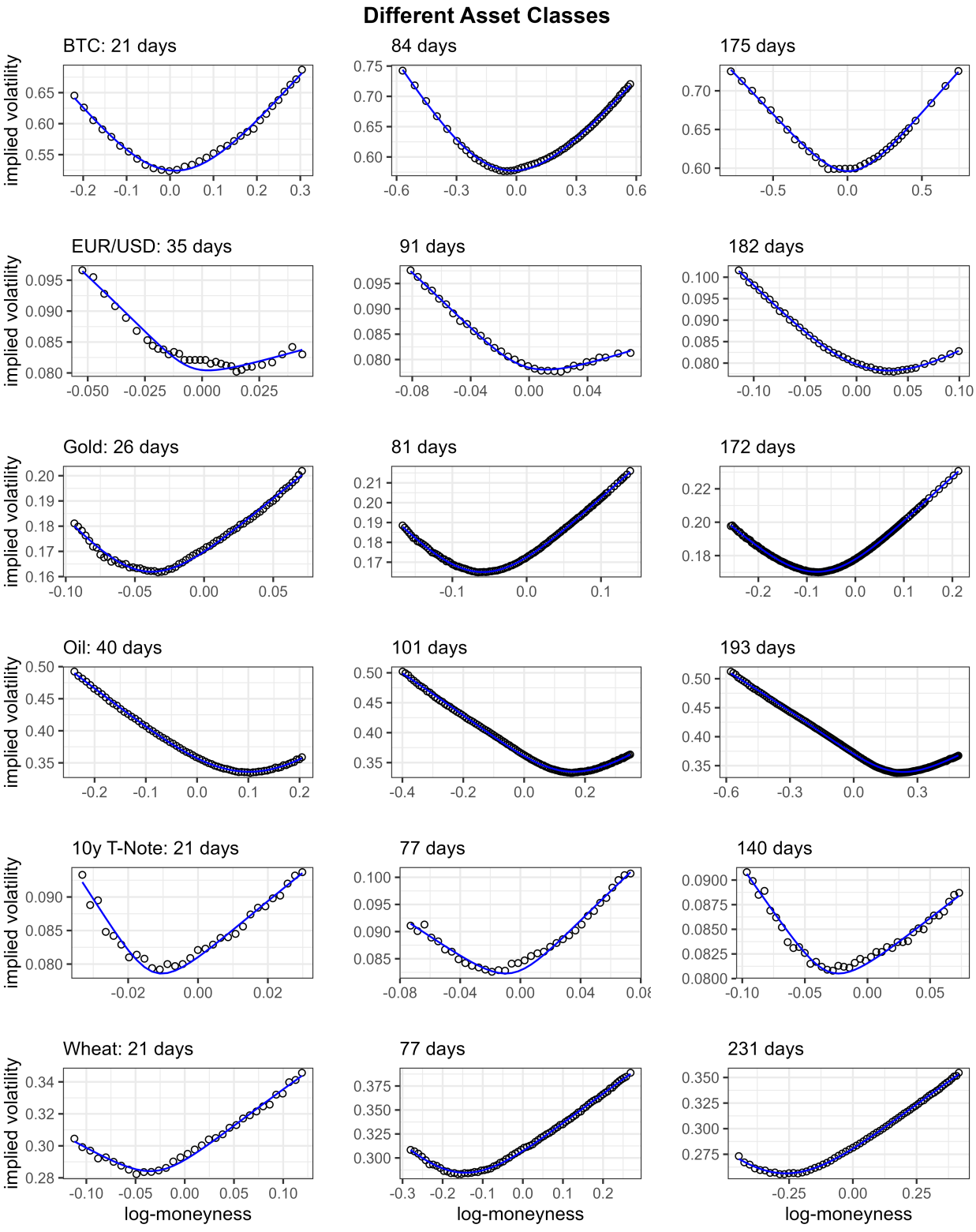

Figure 2: Empirical calibration experiments for various asset classes. Here, the direct least-squares method is fitted under the constraint of $-z_{[1]}z_{[2]} > 0$ to ensure that $\rho \in [-1, 1]$.

Table 1: Fitting quality of the directly estimated SVI curve. Worst goodness-of-fit R^2 and worst mean absolute error MAE in implied volatility per underlying.

	AAPL	AMZN	MSFT	Bitcoin	EUR/USD	Gold	Oil	T-Note	Wheat
R^2	0.993	0.994	0.983	0.993	0.960	0.995	0.999	0.970	0.991
MAE	7.98e-3	3.07e-3	9.84e-3	3.23e-3	7.98e-4	6.68e-4	1.29e-3	7.04e-4	1.40e-3

4 Summary and Outlook

In this work I have identified the SVI formula as a conic section, resulting in a fully linearized equation that simplifies the calibration process to empirical data. This approach offers a significant advantage over the traditional non-linear, multi-parameter optimization problem. Using the conic representation, I have presented a direct and non-iterative approach for fitting the SVI formula via hyperbolic-specific ordinary least-squares. By performing calibration experiments using both simulated and empirical data, I have demonstrated the high accuracy and reliability of this proposed method, which holds significant potential for future implementations.

While these initial results are promising, further investigations are warranted to explore the full potential of the linearized SVI equation. Specifically, I plan to analyze and technically implement the Durrleman condition to ensure that the fitted equation is free of arbitrage. Additionally, I intend to test the empirical fit on a broader range of option data and extend the implied variance smile for the maturity domain to enable an overall surface fit.

References

[1] Damghani, B. M. and Kos, A. [2013], 'De-arbitraging with a weak smile: Application to skew risk', *Wilmott* **2013**, 40–49.

[2] Floc'h, F. L. and Oosterlee, C. W. [2019], 'Model-free stochastic collocation for an arbitrage-free implied volatility: Part i', *Decisions in Economics and Finance* **42**, 679–714.

[3] Friz, P., Gerhold, S., Gulisashvili, A. and Sturm, S. [2011], 'On refined volatility smile expansion in the heston model', *Quantitative Finance* **11**, 1151–1164.

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

[5] Gatheral, J. [2011], *The volatility surface: a practitioner's guide*, John Wiley & Sons.

[6] Gatheral, J. and Jacquier, A. [2011], 'Convergence of heston to svi', *Quantitative Finance* **11**, 1129–1132.

[7] Gatheral, J. and Jacquier, A. [2014], 'Arbitrage-free svi volatility surfaces', *Quantitative Finance* **14**, 59–71.

[8] Gulisashvili, A. [2012], *Analytically Tractable Stochastic Stock Price Models*, Springer.

[9] Guo, G., Jacquier, A., Martini, C. and Neufcourt, L. [2016], 'Generalized arbitrage-free svi volatility surfaces', *SIAM Journal on Financial Mathematics* **7**, 619–641.

[10] Itkin, A. [2014], 'One more no-arbitrage parametric fit of the volatility smile', *North Am. J. Econ. Financ* .

[11] Itkin, A. [2020], *Fitting Local Volatility: Analytic and Numerical Approaches in Black-Scholes and Local Variance Gamma Models*, World Scientific.

[12] Marco, S. D. and Martini, C. [2009], 'Quasi-explicit calibration of gatheral's svi model', *Zeliade White Paper* pp. 1–15.

[13] Martini, C. and Mingone, A. [2022], 'No arbitrage svi', *SIAM Journal on Financial Mathematics* **13**, 227–261.

[14] Nagy, L. and Ormos, M. [2019], 'Volatility surface calibration to illiquid options', *The Journal of Derivatives* **26**, 87–96.

[15] O'Leary, P. and Zsombor-Murray, P. [2004], 'Direct and specific least-square fitting of hyperbolæ and ellipses', *Journal of Electronic Imaging* **13**, 492–503.

[16] Wystup, U. [2020], 'The minimum of the smile fit is not even in the market!'.

Appendix

A Additional Proofs

Lemma A.1. *The SVI model is a vertical hyperbola, rotated through its linear term.*

Proof. Consider $b \neq 0$ and split Eq.1 into its linear and quadratic parts

$$\omega(x) = \underbrace{a + b\rho(x-m)}_{\omega_l} + \underbrace{b\sqrt{(x-m)^2+\sigma^2}}_{\omega_h} \tag{18}$$

. ω_h can be rearranged into the form of a unit hyperbola,

$$\frac{\omega_h^2}{b^2\sigma^2} - \frac{(x-m)^2}{\sigma^2} = 1 \tag{19}$$

. Since $\{b, \sigma\} \in \mathbb{R}$ implies $b^2, \sigma^2 > 0$, it follows that ω_h is a hyperbola of vertical orientation. □

From the split into the linear and hyperbolic part we observe that ω_l causes the rotation and vertical shift of the otherwise vertically oriented ω_h. Furthermore, as both ω_l and ω_h are centered around m, it follows that m represents the x-coordinate center of the conic shape.