Universiteit Utrecht

Faculteit Bètawetenschappen

From Volatility Smiles to Probability Distributions: A Local Jacobian Approach

Bachelor Thesis

Joost van Koesveld

Natuur- en Sterrenkunde

Supervisors:

Dr. Farshid Jafarpour
Theoretical Physics

Prof. dr. R.A. (Rembert) Duine
Theoretical Physics

June 12, 2026

Abstract

Options are widely traded financial instruments whose prices can be expressed as expectation values over the probability distribution of the future price of an underlying asset. Recovering this probability distribution would provide insight into market expectations and is relevant for pricing financial instruments and managing risk. In practice, these option prices are often represented by implied volatilities, whose dependence on strike price forms the volatility smile. This thesis investigates to what extent a local parametric Jacobian mapping can recover the probability distribution of the underlying asset from this volatility smile.

We develop a parametric approach. By parameterizing both the distribution and the volatility smile, we construct a Jacobian matrix that captures how perturbations around a log-normal reference distribution affect the volatility smile in a local linear manner.

We find that the Jacobian mapping and its inverse are locally accurate, but break down for larger deviations from the log-normal reference distribution. Applying the inverse Jacobian to observed volatility smiles of the S&P 500 yields less accurate reconstructed distributions than the Breeden–Litzenberger benchmark. A Hermite parameterization of the smile adds structure, but does not diagonalize the Jacobian.

The Jacobian is therefore useful as a sensitivity tool, but its inverse is insufficient to recover implied probability distributions from observed volatility smiles. We attribute this limitation to a combination of locality, ill-conditioning, and noise in the volatility smile. Future work should focus on regularization; we propose a Kalman filtering approach.

Contents

1 Introduction

Option contracts are among the most widely traded financial instruments in the market. They derive their value from an underlying asset, whose future price is uncertain. Since the payoff of an option depends directly on the future asset price, the option price can be interpreted as an expectation value over the probability distribution of the future price of the underlying asset. Recovering this probability distribution from observed option prices is therefore a natural inverse problem. Such a distribution describes more than a single expected price. It is therefore useful for the accurate pricing of financial instruments, the interpretation of market expectations, and risk management.

The connection between option prices and the probability distribution of the underlying asset is known. Breeden and Litzenberger [1] showed that the probability distribution is proportional to the second derivative of the option price with respect to the strike price[1]:

$$p(K) \propto \frac{\partial^2 V}{\partial K^2}, \tag{1.1}$$

where V is the option price and K is the strike price. However, in practice, computing the second derivative is numerically unstable due to the discreteness and noise in market data. This motivates us to investigate an alternative route, while using the Breeden–Litzenberger approach as a benchmark.

The Black–Scholes model [2] provides a standard way to represent option prices through implied volatility: the volatility parameter which, when inserted into the Black–Scholes formula, reproduces the observed option price. Empirically, implied volatility varies with the strike price, forming the so-called volatility smile. This observation is inconsistent with the Black–Scholes model, where asset prices follow a geometric Brownian motion and implied volatility is constant. The central question of this thesis is to what extent a local parametric Jacobian mapping can recover the implied probability distribution of the underlying asset from an observed volatility smile.

To understand our approach, we first consider the forward direction of the problem: how we obtain a volatility smile from a given distribution of the underlying asset. It is summarized in the forward model:

$$\text{Probability Distribution} \longrightarrow \text{Option Prices} \longrightarrow \text{Volatility Smile}. \tag{1.2}$$

Given a distribution, the option prices can be calculated as expectations of the option payoff under this distribution. Given the option prices, we can convert them to implied volatility by numerically inverting the Black–Scholes model. The implied volatility as a function of the strike price is then called the volatility smile.

We aim to approximate the forward model by linearizing around a reference distribution. The reference distribution is the log-normal distribution, which corresponds to a flat volatility smile. Within this framework, we can study how perturbations around a log-normal distribution result in the shape of the volatility smile. Controlled perturbations allow us to construct a local linear mapping between the probability distribution and the volatility smile.
To be more precise, we use a Taylor-type expansion around the log-normal distribution to parameterize the distribution [3]. This allows us to write deviations from the Gaussian distribution in log-price space in a systematic way. By using the forward model in Eq. (1.2), we can calculate the volatility smile and observe how perturbations of the distribution manifest in the volatility smile. By fitting the smile with a power series expansion, we can capture the influence of the perturbations of the distribution in a Jacobian matrix. This Jacobian matrix thus provides a linear approximation of the forward model.
The inverse of this matrix provides a local mapping from the volatility smile parameters back to the distribution parameters. This enables us to investigate whether the inverse Jacobian can accurately recover distribution parameters from observed volatility smile parameters.

We evaluate the domain in which the Jacobian mapping and its inverse remain accurate. Both the forward mapping and its inverse are found to hold locally but break down for larger deviations from the reference distribution. Applying the inverse mapping to the observed volatility smiles of the S&P 500[2] yields poor results compared with the Breeden–Litzenberger/Tavin benchmark, indicating that our method does not provide an accurate standalone tool for retrieving probability distributions from market data.

[1]The strike price is the predetermined price of an option contract, at which the holder of the option has the right to buy or sell the underlying asset when the option is exercised.

[2]The S&P 500 is a stock market index representing the 500 largest publicly traded companies in the United States. It commonly serves as a proxy for the broad U.S. stock market.

We further attempt to change the power-series basis of the smile into a Hermite basis to diagonalize the Jacobian, but find no sufficient improvement. We use singular value decomposition to diagnose the ill-conditioning of the inverse problem.
Finally, we discuss an extended Kalman filtering model as a possible regularized approach to the inverse problem.

This thesis is structured as follows. Starting with the background in chapter 2, we broadly explain the concepts and tools needed to understand the problem this thesis is trying to solve. In chapter 3, the exact methodology of our approach is laid out in a reproducible manner. The performance of the Jacobian matrix and its inverse, including the application on market data are given in chapter 4. Chapter 5 discusses the results and diagnoses the main limitations of our approach. Finally, chapter 6 concludes this thesis with an outlook for future work. The three appendices contain supporting material: derivation of the Black–Scholes model, details on the Hermite basis of the smile, and the extended Kalman model.

2 Background

In this chapter, we introduce the financial and mathematical concepts that are essential in this thesis. We start by introducing the concepts of stocks and option contracts, and how they contain information about the underlying probability distribution of the stock. We then introduce implied volatility smiles as the representation of market option prices. After defining the forward model, we discuss existing approaches that attempt to find the underlying probability distribution of the stock. Finally, we introduce parameterizations for the probability distribution and the volatility smile, which provide the framework used in the methodology chapter.

2.1 Options and Implied Probability Distributions

This section explains how option prices contain information about the probability distribution for the future price of an asset.

Options as Financial Instruments

The financial market can be seen as a large system with many interacting components and enormous complexity. Generally, a financial market is defined as a marketplace where buyers and sellers trade financial instruments. Widely traded instruments include, for example, stocks, bonds, currencies, and options. For the purpose of this thesis, we focus only on stocks and options.

Simply put, stocks represent fractional ownership in a corporation, where the holders have claims to its assets and earnings. Stocks are associated with a price that is determined by supply and demand in the market. Supply and demand are influenced by many components with complex relationships. The price of a stock is therefore best modeled as a random variable, implying that the future value of a stock is uncertain. This inherent uncertainty motivates a probabilistic description of future asset prices. From a dynamical perspective, the price evolution is best described by stochastic processes.

The second type of financial instrument central to this thesis is the option contract. There exist many types of options. We will focus on European option contracts, which we will simply refer to as 'options'. Historically, options were introduced as insurance against unfavorable price movements of an owned asset [4]. Early forms of these option-like contracts were used to mitigate risks associated with uncertain harvests. Modern companies may use them to protect their investment portfolios. Today, however, options are also widely used for speculation, offering more flexible ways to bet on the price movements of assets. To understand why, we will proceed with a formal definition of options.

There are two classes of European option contracts: call options and put options. A call option gives the holder of the option the right, but not the obligation, to buy a specific asset at a specific time T at a predetermined price K.

A put option gives the holder of the option the right, but not the obligation, to sell a specific asset at a specific time T at a predetermined price K.

Formally, the variable T is called the maturity of the option and the variable K is called the strike price. From these definitions, it follows that the payoff of an option depends on the underlying asset price at maturity and the strike price of the options.

The payoff for a call option is defined as:

$$V_T = \max(S(T) - K, 0), \tag{2.1}$$

and similarly for a put option:

$$V_T = \max(K - S(T), 0), \tag{2.2}$$

where $S(T)$ is the value of the asset price at time T. Fig. 1 clarifies the concept of a payoff.

Options thus derive their value from the underlying asset; in this thesis, we consider stocks as the underlying asset. It is evident how one can speculate on stocks by directly trading them, as profits and losses are determined by the change in the asset price. Options allow for more flexible strategies as their value depends on the difference between the strike price and the stock price.

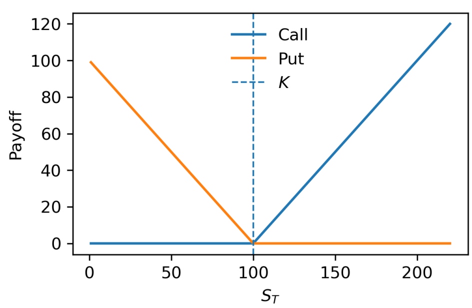

Figure 1: Payoff functions of European call and put options for a fixed strike $K = 100$. As the stock price S_T rises above the strike price, the call payoff rises linearly and the put payoff remains zero. Vice versa, if the stock price decreases below the strike price, the call payoff remains zero and the put payoff rises linearly.

From Option Prices to Probability Distributions

Just like the price of a stock, the prices of options are determined by supply and demand in the financial market. We can also determine the price of an option by deriving it from the stock price. Because the stock price $S(T)$ is a random variable, it is natural to describe it with a probability distribution. At an intuitive level, the price of an option can then be written as an expected value of the payoff. For a call option, this leads to an expression of the form:

$$\langle V_{\text{call}}(T)\rangle \sim \int_K^\infty (S-K)p(S,T)\,\mathrm{d}S, \tag{2.3}$$

and similarly for a put option:

$$\langle V_{\text{put}}(T)\rangle \sim \int_0^K (K-S)p(S,T)\,\mathrm{d}S. \tag{2.4}$$

The precise formula for an option requires an additional correction, which will be introduced in the next section. Eqs. (2.3) and (2.4) illustrate the key idea that prices depend on future outcomes of the stock price. As a result, these prices must contain information about the underlying probability distribution of the stock. This is clarified in Fig. 2.

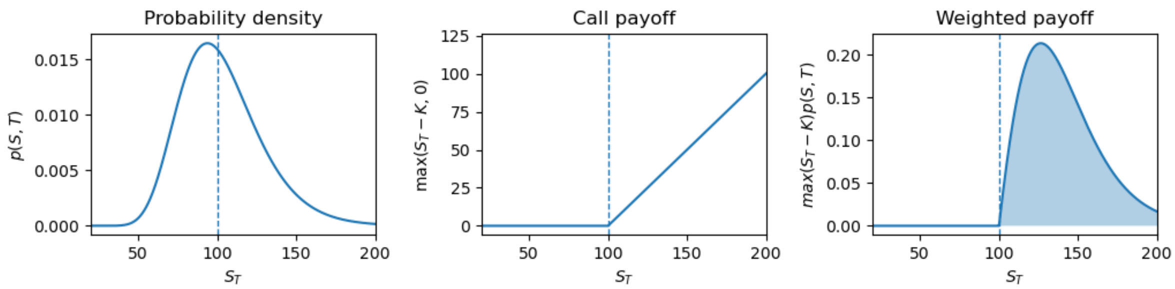

Figure 2: Illustration of a call option price as an expectation value with $K = 100$. The probability density gives the likelihood of future stock prices, the payoff gives the value of the option for each outcome, and their product determines the contribution to the expected payoff. In practice, we do not know the exact shape of the probability density. The weighted payoff curve is therefore also unknown. However, the blue area under the weighted payoff curve is approximately given by the market option price.

In practice, we cannot simply invert (2.3) and (2.4) to retrieve the distribution. We must consider a set of options written for the same underlying stock. One stock can have options with many different maturities, each with a range of strike prices. Each maturity corresponds to a probability distribution of the future stock price. Given all the options with the same maturity but different strikes, we can retrieve the probability distribution analytically with the Breeden–Litzenberger formula [1]:

$$p(S,T) \sim \left.\frac{\partial^2 V(K,T)}{\partial K^2}\right|_{K=S}, \tag{2.5}$$

where $V(K,T)$ is the market option price and S is the stock price. However, this formula assumes the strike price K is a continuous variable, which it is not. In practice, we only have a discrete set of strikes for each maturity. Applying second derivatives directly to raw option prices is therefore numerically unstable. This motivates either smoothing the option data before applying the Breeden–Litzenberger relation, or exploring alternative approaches to the inverse problem.

The relations introduced so far clarify how option prices contain information about the probability distribution of the future stock price. However, the expectation values in Eqs. (2.3) and (2.4) do not form a precise pricing framework yet. In the next section, we will introduce the risk-free framework required for consistent option pricing.

Risk-Neutral Distribution

Eqs. (2.3) and (2.4) calculate the expectation value at time T when the payoff occurs. However, the value of the future payoff is not the same as its value at the present time. This is because the value of money itself evolves. A given amount of money has a deterministic growth characterized by the risk-free rate. Formally, a risk-free investment must grow at the risk-free rate:

$$I(T) = I(0)\mathrm{e}^{rT}, \tag{2.6}$$

where $I(T)$ is the value of the investment at $t = T$, $I(0)$ is the present value of the investment, and r is the risk-free rate, given in units yr^{-1}.
Applying this principle of the risk-free rate to option prices, the value of a call option at any time t before the option expires is given by:

$$V_{\mathrm{call}}(t) = \mathrm{e}^{-r(T-t)} \langle V_{\mathrm{call}}(T) \rangle, \tag{2.7}$$

and similarly for a put option:

$$V_{\mathrm{put}}(t) = \mathrm{e}^{-r(T-t)} \langle V_{\mathrm{put}}(T) \rangle. \tag{2.8}$$

These expressions relate the present option value to the expected future payoff through the discount factor. More precisely, the expectation value is taken under the risk-neutral[3] probability distribution implied by market option prices. These expressions provide the correct analytical prices for call and put options. The question remains how we find the underlying probability distribution of the stock. In the next section, we introduce the Black–Scholes model that provides an alternative way of representing option prices.

2.2 Black–Scholes, Implied Volatility, and the Forward Model

Geometric Brownian motion (GBM) is a core assumption of the Black–Scholes model [2]. We therefore first show how the GBM assumption leads to a log-normal distribution of future asset prices. We then introduce the Black–Scholes pricing framework. Next, we show how using the Black–Scholes model, option prices can be converted into implied volatility. Finally, the forward model summarizes how we convert a probability distribution of the future stock price into a volatility smile.

Geometric Brownian Motion and the Log-Normal Distribution

In GBM, the dynamics of the stock price is given by the stochastic differential equation:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W, \tag{2.9}$$

where S denotes the asset price, μ is the constant drift term, σ is the constant volatility, and W is a Wiener process. The Wiener process has independent, normally distributed increments satisfying:

$$\langle \mathrm{d}W \rangle = 0, \quad \langle (\mathrm{d}W)^2 \rangle = \mathrm{d}t. \tag{2.10}$$

The parameters μ and σ are assumed to be time-independent. Applying Itô's lemma[4] to the function $f(S) = \ln S$ gives:

$$\mathrm{d} \ln S = \frac{1}{S} \mathrm{d}S - \frac{1}{2S^2} (\mathrm{d}S)^2. \tag{2.11}$$

Substituting Eq. (2.9) and Eq. (2.10) yields:

$$\mathrm{d} \ln S = \left(\mu - \frac{1}{2}\sigma^2 \right) \mathrm{d}t + \sigma \mathrm{d}W. \tag{2.12}$$

Integrating gives the solution:

$$S(t) = S(0) \exp\left[\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t) \right], \tag{2.13}$$

where $S(0)$ is the asset price at $t = 0$. Taking the logarithm of Eq. (2.13) yields

$$\ln S(t) = \ln S(0) + \left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t). \tag{2.14}$$

[3] The risk-neutral distribution generally differs from the (physical) real-world distribution. For example, people are willing to pay extra for options to protect against market crashes. For this reason, the risk-neutral distribution can inflate the probabilities for extreme stock prices (fatter tails). The real-world distribution is significantly harder to determine, as it requires additional assumptions about investors' preferences. Throughout this thesis, we will refer to the risk-neutral distribution simply as the probability distribution unless otherwise stated.

[4] Itô's lemma is a stochastic calculus rule. It is analogous to the ordinary chain rule, but it contains an additional second-order term because Wiener increments satisfy $\langle (\mathrm{d}W)^2 \rangle = \mathrm{d}t$.

Since $W(t) \sim \mathcal{N}(0, t)$, it follows that:

$$\ln S(t) \sim \mathcal{N}\left(\ln S(0) + \left(\mu - \frac{1}{2}\sigma^2\right) t,\ \sigma^2 t\right). \tag{2.15}$$

Since $\ln S(t)$ is normally distributed, the distribution of $S(t)$ follows by a change of variables:

$$p(S, t) = \frac{1}{S\,\sigma\sqrt{2\pi t}} \exp\left[-\frac{\left(\ln\left(\frac{S}{S(0)}\right) - \left(\mu - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t}\right]. \tag{2.16}$$

The future stock price thus follows a log-normal distribution under the assumption of geometric Brownian motion.

Black–Scholes Model

The Black–Scholes model [2] provides a closed-form solution for European option prices. The model consists of a partial differential equation and its solution. The full derivation of the PDE and its solution is given in Appendix A. For the purpose of this thesis, knowing that the underlying stock follows geometric Brownian motion, as given in Eq. (2.9), is sufficient. The full solution to the Black–Scholes PDE depends on the risk-free rate r, the current asset price S, the time to maturity $\tau = T - t$, the strike price K, and the volatility σ. Time is measured in years, such that r has units yr^{-1} and σ has units $\mathrm{yr}^{-1/2}$. For a call option, the solution is:

$$V_{\text{call}}(S, t) = S\,\Phi(d_1) - K\mathrm{e}^{-r\tau}\Phi(d_2), \tag{2.17}$$

and similarly for a put option:

$$V_{\text{put}}(S, t) = K\mathrm{e}^{-r\tau}\Phi(-d_2) - S\,\Phi(-d_1), \tag{2.18}$$

where

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)(\tau)}{\sigma\sqrt{\tau}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{\tau}. \tag{2.19}$$

$\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The solution to the Black–Scholes PDE can thus be written as:

$$V(S, t) = V_{\text{BS}}(S, K, \tau, r, \sigma). \tag{2.20}$$

In practice, the parameters S, K, and τ are known exactly, and r can be estimated to great accuracy. $V(S, t)$ is given by the observed option prices in the financial market. The volatility parameter σ is hard to measure, but since we have all the parameters except the volatility, we can numerically invert $V_{\text{BS}}(S, K, \tau, r, \sigma)$ to obtain this parameter. According to the assumption of geometric Brownian motion, it should be constant. However, empirically, it is not constant and varies with respect to both the strike price and the maturity. This motivates the introduction of 'implied volatility' as an alternative representation of option prices.

Implied Volatility Smile

Formally, the implied volatility is the volatility that reproduces the observed market option price within the Black–Scholes model [4]. For a fixed maturity T, we can view the implied volatility as a function of the strike price K. Commonly, this is referred to as the implied volatility smile, or simply the 'smile'. In this representation, option prices V are expressed in implied volatility σ_{imp}. The transformation of option prices into implied volatility provides a more structured representation of the data. This is illustrated in Fig. 3.

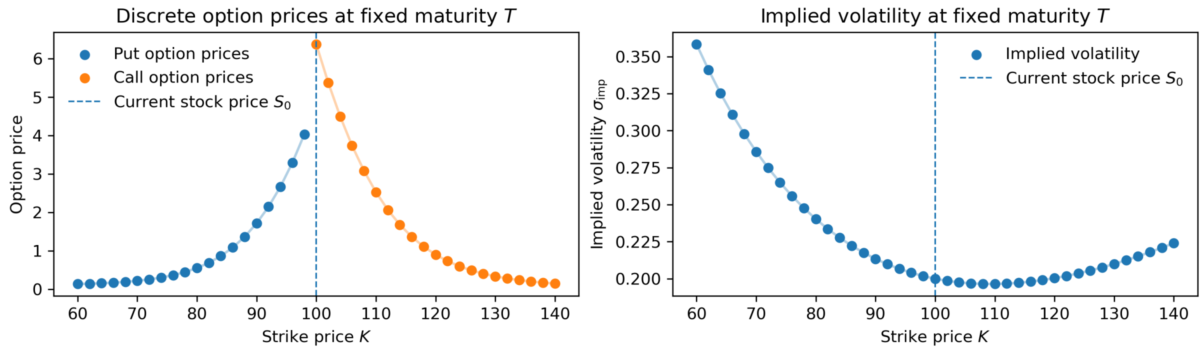

Figure 3: Comparison between option prices and implied volatility for the same synthetic option data at fixed maturity $T = 0.5$ years. The left panel shows put prices for strikes below the current stock price and call prices for strikes above it. This resembles a realistic data structure; calls with strikes below the stock price are traded less frequently and therefore have unreliable prices; vice versa, puts with strikes above the stock price are not reliable. The right panel shows the implied volatility obtained by inverting the Black–Scholes formula.

Whereas the option price might exhibit a more non-linear and irregular dependence on the strike price, the implied volatility typically varies more smoothly. This smoother structure is preferable for interpolation of the data points, which is essential when working with discrete data points. Additionally, expressing option prices in terms of implied volatility allows for direct comparison between put and call options, and also between different strike prices and maturities. It is important to note that this transformation does not solve the problem of finding the underlying probability distribution of the stock. It merely reorganizes our data points in a manner that makes subsequent analysis easier. The widespread use of implied volatility in the financial industry further motivates this choice [5].

Forward Model

So far, we have established how option prices can be calculated with a given probability distribution of the stock in Eqs. (2.7) and (2.8). Given these option prices, we can invert them to obtain the implied volatility smile. This procedure can be shown schematically as:

$$\text{Probability Distribution} \longrightarrow \text{Option Prices} \longrightarrow \text{Volatility Smile}, \tag{2.21}$$

and we will henceforth refer to this as the 'forward model'. The forward model provides us with a continuous implied volatility smile, given that we know the underlying distribution of the stock. In practice, we can only observe discrete sets of option prices and therefore we only have a discrete volatility smile. The inverse problem is to obtain the probability distribution given a discrete volatility smile. The log-normal distribution generated by geometric Brownian motion provides a natural reference case: it corresponds to the flat smile of the Black–Scholes model. The forward model forms the basis of our approach to approximate the inverse problem. Before introducing our approach, we will briefly review existing methods in the literature that attempt to solve the inverse problem.

2.3 Existing Approaches to the Inverse Problem

There exist multiple approaches to finding the probability distribution for a future stock price. A substantial body of literature is devoted to the analysis and development of these methods. It is worthwhile to briefly cover papers that present distinct approaches.

Buchen and Kelly proposed selecting a distribution with the highest entropy [6]. They view option prices as constraints on the set of possible distributions for the stock price. Given option prices for a finite number of strikes and maturities, there is still an infinite family of distributions that satisfy these constraints. The maximum entropy principle selects for the distribution that contains the least additional information while obeying the option constraints [7]. This allows for the construction of a Lagrangian:

$$L(p) = -\int_0^\infty p(x)\log p(x)\,\mathrm{d}x + (1+\lambda_0)\int_0^\infty p(x)\,\mathrm{d}x + \sum_{i=1}^{m}\lambda_i\int_0^\infty p(x)c_i(x)\,\mathrm{d}x, \tag{2.22}$$

where $p(x)$ is the probability distribution of the stock price, the first term is the entropy functional, the second term enforces normalization of the distribution, and the third term represents the option price constraints with corresponding Lagrange multipliers λ_i.
Setting the first variation equal to zero yields the maximum entropy distribution:

$$p(x) = \exp\left(\lambda_0 + \sum_{i=1}^{m} \lambda_i c_i(x)\right). \tag{2.23}$$

The Lagrange multipliers must then be solved numerically. The maximum entropy method produces the least biased distribution according to available option price information. It also works with a limited amount of option data. Solving the Lagrange multipliers numerically is a disadvantage as it requires additional numerical optimization techniques.

A second paper that we already mentioned is from Breeden and Litzenberger [1]. They derived the analytical solution for the probability distribution of a stock:

$$p(K) = \mathrm{e}^{rT} \frac{\partial^2 V(K,T)}{\partial K^2}. \tag{2.24}$$

We already mentioned that in practice, this closed-form solution is unstable due to the discrete nature of the option prices. One could attempt to smooth the option prices directly, after which Eq. (2.24) can be applied to obtain the probability distribution. In this thesis, the Breeden–Litzenberger/Tavin reconstruction of the probability distribution from option prices is used as a benchmark.
However, many papers first use the volatility smile [8, 9]. One fits the smile, after which it can be converted into option prices by using the Black–Scholes equations (2.17) and (2.18). Tavin [10] gives the density directly as a function of the volatility smile:

$$p(K) = \mathrm{e}^{rT} \frac{\partial^2 V_{\mathrm{BS}}(K, \sigma(K,T))}{\partial K^2}. \tag{2.25}$$

It is in general more stable than applying second derivatives directly to raw option prices. The main drawback of this method is that the resulting probability distribution relies heavily on the chosen interpolation technique.

While the existing methods in the literature are effective, we attempt a different approach to the inverse problem. Our method attempts to connect the probability distribution and volatility smile in a linearized way.

2.4 Parametric Framework

The forward model maps an entire probability distribution to an implied volatility smile. To study this mapping numerically, we must represent both objects with a finite set of parameters. In this section, we introduce these parametrizations.

Relating Distributions and Smiles

The log-normal distribution serves as a base distribution that we can perturb to study the influence on the smile. The idea is to perform a local approximation of the forward model in Eq. (2.21) around the log-normal distribution. This allows us to perform controlled perturbations. The effect of these perturbations on the smile can be calculated with the forward model itself. In this framework, we can quantify how different deviations from the log-normal distribution affect the volatility smile. The next step is to construct a parametric form of the distribution in which perturbations can be introduced.

Parametric Representation of the Distribution

A parametric form of the distribution can be constructed by expanding around a Gaussian distribution in terms of its cumulants. This leads to a series of Hermite polynomials called a Gram-Charlier series [3]:

$$f(x) = \psi(x)\left[1 + \frac{\kappa_3}{6\sigma^3} H_3(z) + \frac{\kappa_4}{24\sigma^4} H_4(z) + \dots\right], \tag{2.26}$$

where $\psi(x)$ is the base Gaussian, $H_n(z)$ are Hermite polynomials, $z = \frac{x-\mu}{\sigma}$ is the standardized variable, and κ_n are cumulants. This expansion provides a systematic way to describe deviations from a Gaussian.

Gram–Charlier type expansions have been used in option-pricing formulas because they allow deviations from a Gaussian to be expressed through interpretable parameters such as skewness, kurtosis, and higher-order cumulants [11, 12]. In practice, we must truncate the series. This can lead to negative distribution values, a known drawback of the series [11, 13]. To resolve this problem, we introduce semi-nonparametric probability distributions (SNP) [14].

$$f(x) = \frac{\psi(x)}{\mathbf{n}^T \mathbf{n}} \left(\sum_{i=0}^{N} n_i H_i(z) \right)^2 , \tag{2.27}$$

where n_i are arbitrary shape parameters. The key idea of this expansion is that it retains the structure of the Gram-Charlier expansion, but it enforces positivity by squaring the Hermite series. This results in a guaranteed non-negative distribution. However, normalization is no longer satisfied automatically and we must enforce it with a normalization factor.
The use of SNP densities has its trade-offs. The main advantage is that it guarantees a valid probability density. However, the coefficients n_i are arbitrary and no longer represent cumulants. Instead, cumulants must be computed from the resulting distribution:

$$\langle x^n \rangle = \int x^n f(x) \, \mathrm{d}x. \tag{2.28}$$

The set of allowed distributions is changed by the squaring of the Hermite series, i.e., the coefficients no longer affect the distribution in the same way as in the Gram–Charlier expansion. While increasing the amount of Hermite terms can improve flexibility, it also increases the dimensionality of our parameter space.
To complete the parametric framework, we must also introduce a parametric representation of the volatility smile, allowing us to quantify how perturbations in the distribution manifest in the smile.

Parametric Representation of the Smile

Observed market smiles consist of discrete points, but can typically be modeled with smooth functions. This motivates the use of a power series expansion:

$$\sigma(k) = a_0 + a_1 k + a_2 k^2 + \cdots + a_m k^m, \tag{2.29}$$

where the log-moneyness k is defined as:

$$k = \log \left(\frac{K}{F_0} \right), \tag{2.30}$$

which is a dimensionless quantity. Here, K is the strike price and F_0 is the forward price, defined as $F_0 = \mathrm{e}^{rT} S_0$, where S_0 is the current stock price. Both K and F_0 have the same currency units. The log-moneyness is denoted by the letter k, while cumulants are denoted by the Greek letter κ_n.

Together, the parameterizations of the distribution and the volatility smile allow us to study how local perturbations around the log-normal distribution affect the smile. This framework is used in the methodology chapter.

3 Methodology

This section explains how the Jacobian matrix is constructed to obtain a local mapping between the probability distribution and the volatility smile. We construct this mapping by perturbing each distribution parameter around a Gaussian base in log-price space and measuring the resulting changes in the fitted volatility smile parameters. This allows us to quantify how deviations from the log-normal base affect the smile.

3.1 Numerical Framework

In this section we use the parameterizations introduced in Sec. 2.4 as a starting point.

Overview of the Procedure

We start with a set of distribution parameters. For the Gaussian reference density, all perturbation parameters are zero. We perturb each parameter individually with a fixed value ϵ and the other parameters are kept at zero. The resulting smile coefficients are computed through the following chain:

$$\begin{aligned}\text{Distribution parameters} &\to \text{Probability distribution} \to \text{Option prices} \\ &\to \text{Implied volatility smile} \to \text{Smile parameters.}\end{aligned} \tag{3.1}$$

Perturbing each distribution parameter individually and repeating this computational chain allows us to construct a Jacobian matrix.

Parameterization of the Probability Distribution

We describe deviations from the Gaussian reference density by a finite set of distribution parameters

$$\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_m)^T. \tag{3.2}$$

Here, $\boldsymbol{\theta} = \mathbf{0}$ corresponds to the Gaussian reference distribution. In this thesis, $\boldsymbol{\theta}$ denotes either the free parameters of the SNP series in Eq. (2.27), or the corresponding Hermite parameters in the Gram–Charlier series in Eq. (2.26). Each component of $\boldsymbol{\theta}$ defines one perturbation direction in the probability distribution.

Parameterization of the Volatility Smile

As introduced in Eq. (2.29), the volatility smile is parameterized by a power series in log-moneyness. The coefficients are obtained by performing a least-squares fit of the power series expansion to the computed volatility smile over a fixed range of log-moneyness. The resulting coefficients

$$\mathbf{a} = (a_0, a_1, \ldots, a_p)^T \tag{3.3}$$

form the smile-parameter vector used in the Jacobian construction.

3.2 Forward Model and Jacobian Construction

Option Prices from the Density

We start from a given probability distribution for the future price of a stock at a future time T. Given this distribution, we can calculate call and put option prices using Eqs. (2.7) and (2.8). The integral is computed numerically. The option prices are calculated for a discrete set of strike values K_i corresponding to an evenly spaced grid in log-moneyness k_i. For strikes below the forward price, $K_i < F_0$ we calculate only put prices. For strikes above or equal to the forward price, $K_i \geq F_0$ we calculate only call prices. The resulting set is given by:

$$V_i = \begin{cases} P(K_i, T), & K_i < F_0, \\ C(K_i, T), & K_i \geq F_0, \end{cases} \qquad K_i = F_0 e^{k_i}, \qquad i = 1, \ldots, N. \tag{3.4}$$

Implied Volatility Inversion

Using the set of option prices given in Eq. (3.4), we can convert each price into an implied volatility by numerically inverting the Black–Scholes solution for the call and put options as given in Eqs. (2.17) and (2.18). We then obtain a set of implied volatilities for a fixed maturity T on the evenly spaced grid in log-moneyness k_i:

$$\sigma_{\mathrm{imp},i} = \begin{cases} \sigma^P_{\mathrm{BS}}(V_i, K_i, T), & K_i < F_0, \\ \sigma^C_{\mathrm{BS}}(V_i, K_i, T), & K_i \geq F_0, \end{cases} \qquad i = 1, \ldots, N. \tag{3.5}$$

Finite-Difference Construction of the Jacobian

We systematically perturb a single distribution parameter θ_j with $+\epsilon$ and $-\epsilon$, with $\epsilon > 0$. For both perturbations we calculate the full volatility smile and its fitted smile parameters. Using central finite differences, a single entry of the Jacobian is then given by:

$$J_{ij} = \left.\frac{\partial a_i}{\partial \theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(0)}} \approx \frac{a_i(\boldsymbol{\theta}^{(0)} + \epsilon \mathbf{e}_j) - a_i(\boldsymbol{\theta}^{(0)} - \epsilon \mathbf{e}_j)}{2\epsilon}. \tag{3.6}$$

Here, $\boldsymbol{\theta}^{(0)}$ denotes the parameter vector corresponding to the Gaussian reference distribution. For a single distribution parameter, this gives one column of the Jacobian matrix. After systematically perturbing each distribution parameter individually, we can construct a full Jacobian matrix:

$$\delta \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} \approx \left.\begin{pmatrix} \frac{\partial a_0}{\partial \theta_0} & \frac{\partial a_0}{\partial \theta_1} & \cdots & \frac{\partial a_0}{\partial \theta_m} \\ \frac{\partial a_1}{\partial \theta_0} & \frac{\partial a_1}{\partial \theta_1} & \cdots & \frac{\partial a_1}{\partial \theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial a_p}{\partial \theta_0} & \frac{\partial a_p}{\partial \theta_1} & \cdots & \frac{\partial a_p}{\partial \theta_m} \end{pmatrix}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(0)}} \delta \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_m \end{pmatrix}. \tag{3.7}$$

Each element of this matrix quantifies how a perturbation of a specific distribution parameter affects a specific smile parameter in a linear approximation.

3.3 Inverse Mapping and Stability Analysis

Inverse Jacobian Mapping.

If $p = m$, the Jacobian in Eq. (3.7) is square. Its inverse can then be numerically computed:

$$\delta \boldsymbol{\theta} \approx J^{-1} \delta \mathbf{a}. \tag{3.8}$$

This inverse matrix can be used to convert smile parameter perturbations into a set of distribution parameter perturbations.

Singular Value Decomposition.

We introduce singular value decomposition (SVD), which finds orthogonal bases for the smile and the distribution in which the Jacobian mapping is diagonal. This is useful in analyzing the stability of the inverse mapping. We start from the Jacobian forward mapping

$$\delta \mathbf{a} = J \delta \boldsymbol{\theta}, \tag{3.9}$$

where $\delta \boldsymbol{\theta}$ denotes a perturbation in the distribution parameters and $\delta \mathbf{a}$ denotes the corresponding perturbation in the smile parameters. SVD decomposes J as:

$$J = U \Sigma V^T. \tag{3.10}$$

Here, V is an orthogonal matrix whose columns form a basis in the distribution-parameter space, U is an orthogonal matrix whose columns form a basis in the smile-parameter space, and Σ is a diagonal matrix containing the singular values.

This decomposition is computed using standard numerical linear-algebra routines. Distribution modes are then mapped to smile modes as follows:

$$J \mathbf{v}_i = \sigma_i \mathbf{u}_i. \tag{3.11}$$

For a square, invertible Jacobian, the inverse is then given by:

$$\delta\boldsymbol{\theta} = J^{-1}\delta\mathbf{a} = V\Sigma^{-1}U^T\delta\mathbf{a}, \tag{3.12}$$

where Σ^{-1} is obtained by replacing each singular value σ_i by $1/\sigma_i$. The singular values can therefore be used to diagnose error amplification in the inverse mapping.

3.4 Application to Market Data

Construction of Market Smiles

Using the `yfinance` module in Python, we extract option chain data at the present time. The option data consist of a discrete set of maturities T. For each maturity T, we have a discrete set of strike prices K. We filter the option prices such that below the forward price F_0 we only use put prices and at and above the forward price we only use call prices:

$$V_i^{\text{mkt}} = \begin{cases} P_i^{\text{mid}}, & K_i < F_0, \\ C_i^{\text{mid}}, & K_i \geq F_0, \end{cases} \qquad i = 1, \dots, N_T. \tag{3.13}$$

For each option price, the price is determined by taking the midpoint of the best bid and the best ask[5]:

$$P_i^{\text{mid}} = \frac{P_i^{\text{bid}} + P_i^{\text{ask}}}{2}, \qquad C_i^{\text{mid}} = \frac{C_i^{\text{bid}} + C_i^{\text{ask}}}{2}. \tag{3.14}$$

The stock price is also calculated from the midpoint of the best bid and best ask:

$$S_0^{\text{mid}} = \frac{S_0^{\text{bid}} + S_0^{\text{ask}}}{2}. \tag{3.15}$$

Because the Black–Scholes equation we use does not account for dividends, we subtract all future dividend payments before maturity from the current stock price:

$$S_{\text{adj}} = S_0 - \sum_{d:\, t_d \leq T} D_d e^{-rt_d}, \tag{3.16}$$

where D_d is the value of the dividend payment, t_d is the time until the dividend payment, and e^{-rt_d} is the discount factor.

This approximation is reasonable because expected dividend payments before maturity reduce the forward value of the stock. We now have $S(0) = S_{\text{adj}}$, the strike values $\{K_i\}_{i=1}^{N_T}$, the maturity T, the market option prices $\{V_i^{\text{mkt}}\}_{i=1}^{N_T}$, and an approximate risk-free rate $r = 0.03$. We can numerically invert the Black–Scholes solutions in Eqs. (2.17) and (2.18) to obtain a discrete volatility smile for each maturity T:

$$\{(k_i, \sigma_{\text{mkt},i})\}_{i=1}^{N_T}. \tag{3.17}$$

Application of the Inverse Mapping

The discrete market smiles given in Eq. (3.17) can be fitted by using the power series in Eq. (2.29). We then obtain a parameter set for the market smile:

$$\mathbf{a}^{\text{mkt}}(T) = \begin{pmatrix} a_0^{\text{mkt}}(T) \\ a_1^{\text{mkt}}(T) \\ \vdots \\ a_p^{\text{mkt}}(T) \end{pmatrix}. \tag{3.18}$$

We can approximate the corresponding distribution-parameter perturbation by using the inverse Jacobian matrix:

$$\delta\boldsymbol{\theta}^{\text{mkt}}(T) \approx J^{-1}\left(\mathbf{a}^{\text{mkt}}(T) - \mathbf{a}^{(0)}\right). \tag{3.19}$$

[5]The best bid is the highest price offered by the buyers of the option. Similarly, the best ask is the lowest price requested by the sellers of the option. The midpoint of the best bid and best ask is used as an estimate of the option price.

Here, $\mathbf{a}^{(0)}$ denotes the smile-parameter vector of the log-normal reference case. Since the log-normal reference corresponds to zero distribution-parameter perturbations, we set

$$\boldsymbol{\theta}^{\mathrm{mkt}}(T) = \delta\boldsymbol{\theta}^{\mathrm{mkt}}(T). \tag{3.20}$$

The resulting parameter vector is inserted into the chosen distribution series (Eq. (2.26) or Eq. (2.27)) to reconstruct the market-implied probability distribution. This distribution is then compared against the Breeden–Litzenberger/Tavin benchmark from Eq. (2.25) in the results section.

4 Results

In this chapter we evaluate the accuracy of both the Jacobian mapping and its inverse. The Jacobian is tested on perturbed Gaussian distributions in log-price space, and predicted volatility smiles are compared to those obtained from the forward model. This allows us to determine a domain in which the linear approximation is valid. We find that both the mapping and its inverse are accurate only for small perturbations. In addition, the inverse mapping is applied to volatility smiles of the S&P 500 index and compared to the Breeden-Litzenberger/Tavin benchmark [1, 10], as introduced in Sec. 2.3.

4.1 Structure of the Jacobian Matrix

We start by computing the Jacobian matrix in which the distribution has been parameterized with the Gram-Charlier series given in Eq. (2.26). This allows for an interpretable Jacobian heatmap. We then compute the Jacobian matrix in which the distribution has been parameterized with the SNP series given in Eq. (2.27). This series will later be used for the inversion as it cannot produce negative probability distributions.

Gram-Charlier Expansion

We compute the Jacobian for a representative maturity $T = 1$ year, shown in Fig. 4. The resulting Jacobian is a 7×7 matrix, mapping 7 distribution parameters to 7 smile parameters.

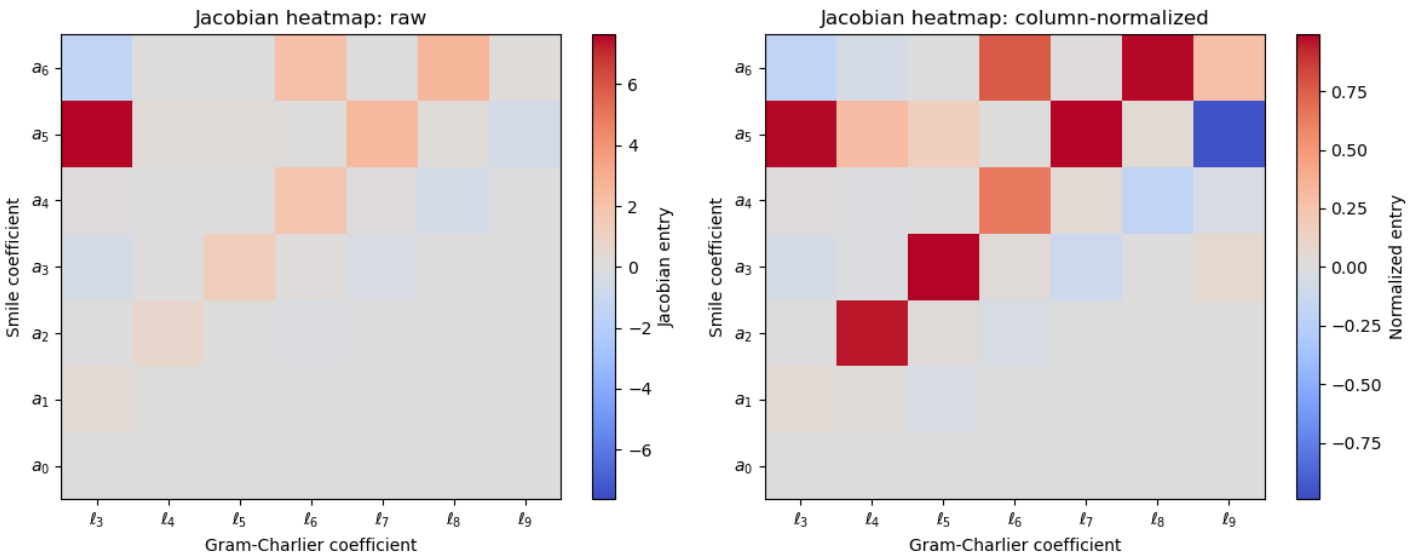

Figure 4: Heatmaps of the Gram–Charlier Jacobian matrix for $T = 1$ year. The left heatmap is the raw Jacobian, while the right heatmap has its columns normalized. The Jacobian relates the distribution parameters $\{l_3, l_4, l_5, l_6, l_7, l_8, l_9\}$ to the smile parameters $\{a_0, a_1, a_2, a_3, a_4, a_5, a_6\}$. The parameters l_j are standardized cumulants, $l_j = \kappa_j / v^{j/2}$, with $v = \sigma^2 T$. The Jacobian entries are computed using a finite perturbation of size $\epsilon = 10^{-4}$ applied to each distribution parameter. The normalized heatmap emphasizes the relative effect of each standardized cumulant on the smile parameters.

The heatmap shows moderate structure. Odd-order cumulant perturbations mainly affect odd-order smile coefficients, whereas even-order cumulants mainly affect even-order smile coefficients. This correspondence is not perfect and coupling into other smile coefficients is visible.

SNP Expansion

We also compute the Jacobian for the same representative maturity $T = 1$ year, now using the SNP parametrization of the distribution. The result is shown in Fig. 5.

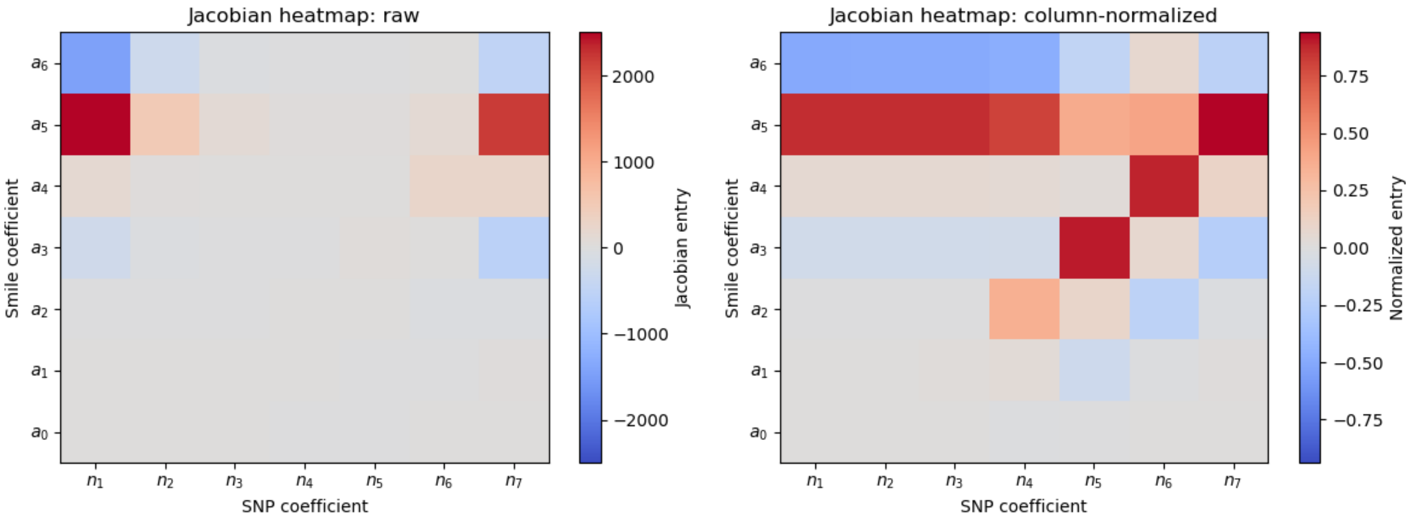

Figure 5: Heatmaps of the SNP Jacobian matrix for maturity $T = 1$. The left heatmap shows the raw Jacobian, while the right heatmap has its columns normalized. The Jacobian maps the shape parameters $\{n_1, n_2, n_3, n_4, n_5, n_6, n_7\}$ to the smile parameters $\{a_0, a_1, a_2, a_3, a_4, a_5, a_6\}$. The Jacobian entries are computed using a finite perturbation of size 10^{-4} applied to each shape parameter. The normalized heatmap emphasizes the relative effect of each SNP shape parameter on the smile parameters.

This Jacobian mapping is not diagonal. Perturbations in the shape parameters n_j affect multiple smile coefficients and do not separate cleanly into even and odd modes. The higher-order smile parameters are mostly affected by these perturbations. The SNP Jacobian is therefore less transparent than the Gram–Charlier Jacobian.

4.2 Accuracy of the Forward Jacobian Mapping

In this section, we quantify the accuracy of the Jacobian mapping for different perturbation sizes of the distribution parameters. The distribution has been parameterized with the SNP series. The root mean squared error (RMSE) measure is used to quantify the error:

$$\mathrm{RMSE}_\sigma = \left[\frac{1}{N} \sum_{i=1}^{N} \left(\sigma_{\text{target}}(k_i) - \sigma_{\text{approx}}(k_i) \right)^2 \right]^{1/2} . \tag{4.1}$$

Here, N is the number of log-moneyness grid points k_i, σ_{target} is the implied volatility calculated with the forward model in Eq. (3.1), and σ_{approx} is the implied volatility calculated with the Jacobian approximation of the forward model. RMSE has the same units as implied volatility: $yr^{-1/2}$. The results are shown in Fig. 6.

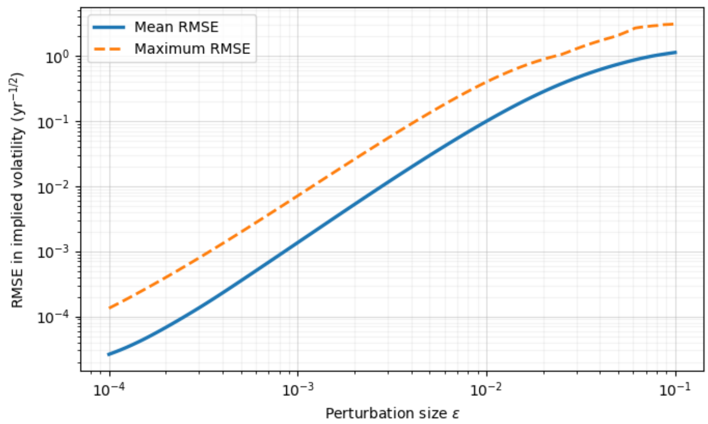

Figure 6: RMSE of the forward Jacobian plotted against the perturbation size ϵ. The distribution is parameterized with the SNP series, and all RMSE values are computed for $T = 1$ year on the fixed log-moneyness grid $k \in [-0.25, 0.25]$. The Jacobian is computed with a fixed perturbation size $\epsilon_J = 10^{-4}$. For each perturbation size ϵ, 100 random perturbation directions in the SNP parameter space are generated. The plot shows the mean and maximum RMSE between the full forward model smile and the smile computed with the Jacobian mapping. The increasing RMSE shows the local nature of the Jacobian approximation.

Fig. 7 shows three example perturbations along with the Jacobian reconstructed smile and the smile reconstructed with the exact forward model. This figure serves as an interpretation of the RMSE magnitude.

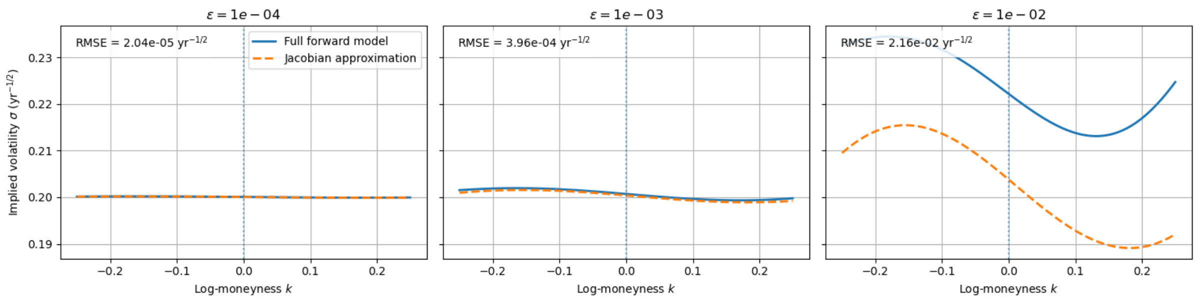

Figure 7: Three example perturbations are shown with implied volatility σ_{imp} as a function of the log-moneyness k. The distribution is parameterized with an SNP series and the Jacobian is computed with a fixed perturbation size $\epsilon_J = 10^{-4}$. In each panel, the distribution is perturbed in a general direction with perturbation size ϵ. The volatility smile for each perturbation is computed with the full forward model and the Jacobian model. The RMSE measures the difference between these two smiles. These graphs illustrate the magnitude of the RMSE.

Fig. 6 and Fig. 7 show the expected local validity of the Jacobian mapping. For small perturbations, the RMSE stays closer to zero, indicating that the Jacobian accurately reproduces the volatility smiles. For larger perturbation sizes, errors grow rapidly, showing that the Jacobian breaks down on a more global scale. Since the Jacobian depends on maturity, the size of this validity domain may differ for different maturities.

4.3 Accuracy of the Inverse Jacobian Mapping

In this section, we quantify the accuracy of the inverse Jacobian mapping. The distribution has been parameterized with the SNP series. The purpose is not to directly compare the errors of the forward and inverse Jacobian as perturbations in the distribution space and smile space are not equivalent.

To evaluate the accuracy of the inverse mapping, we start from the flat volatility smile corresponding to the log-normal distribution. We then perturb the smile parameters to obtain a target smile. The inverse Jacobian is applied to parameters of this target smile, giving an approximate set of distribution

parameters. From this reconstructed distribution, we calculate the volatility smile again using the full forward model. This serves as the approximate smile. The RMSE is then calculated between the target smile and the approximate smile.
The root mean squared error (RMSE) is used to quantify the error:

$$\mathrm{RMSE}_{\sigma}^{\mathrm{inv}} = \left[\frac{1}{N} \sum_{i=1}^{N} \left(\sigma_{\mathrm{target}}(k_i) - \sigma_{\mathrm{approx}}(k_i) \right)^2 \right]^{1/2} . \tag{4.2}$$

Here, N is the number of log-moneyness grid points, σ_{target} is the perturbed target smile, and σ_{approx} is the volatility smile obtained by applying the full forward model to the reconstructed distribution. The results are shown in Fig. 8.

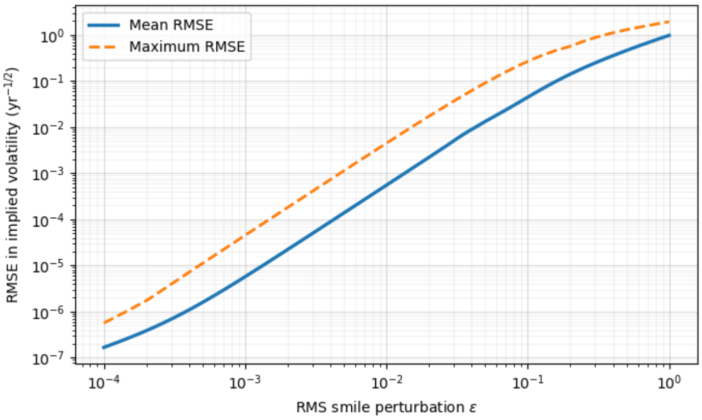

Figure 8: RMSE of the inverse Jacobian plotted against the perturbation size ϵ. The distribution is parameterized with the SNP series, and all RMSE values are computed for $T = 1$ year on the fixed log-moneyness grid $k \in [-0.25, 0.25]$. The Jacobian is computed with a fixed perturbation size of $\epsilon_J = 10^{-4}$. For each perturbation size ϵ, 100 random perturbation directions are generated in the parameter space of the smile. The RMSE is computed between the perturbed smile and the smile corresponding to the reconstructed distribution obtained through the inverse Jacobian. The increasing RMSE shows the local nature of the inverse Jacobian approximation.

Fig. 9 illustrates the magnitude of the RMSE for different perturbation sizes along with their corresponding distributions. It serves to give a rough interpretation of the size of the RMSE.

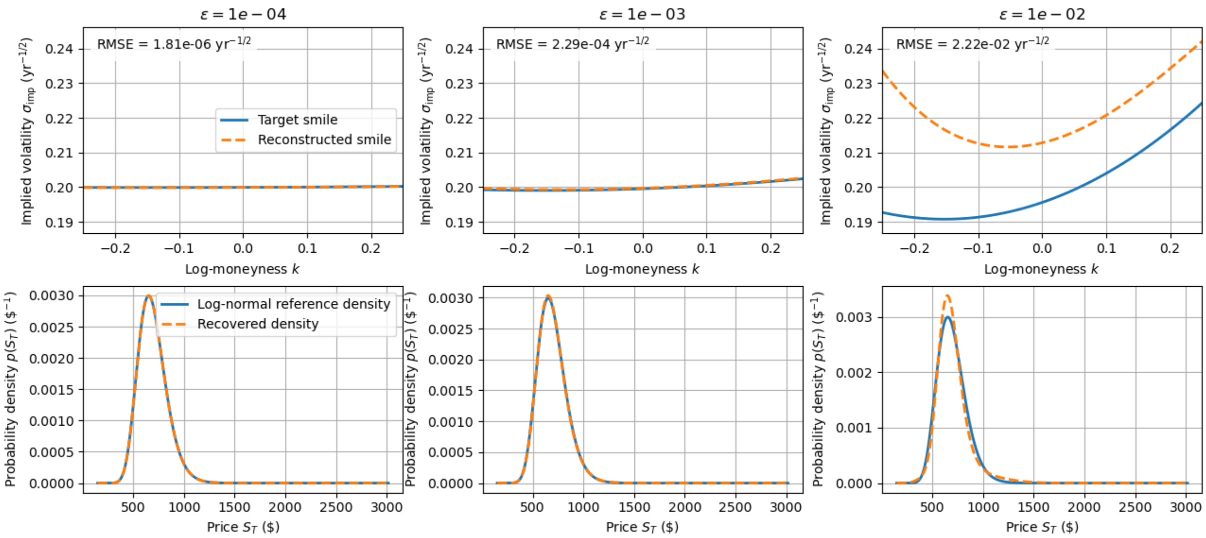

Figure 9: Three example perturbations in the volatility smile are shown in the top panels. The three bottom panels show the respective distributions obtained via the inverse Jacobian mapping. The distribution is parameterized with an SNP series and the Jacobian is computed with a fixed perturbation size $\epsilon_J = 10^{-4}$. In each panel, the target smile is the perturbed smile with perturbation size ϵ. The reconstructed distribution is recovered by applying the inverse Jacobian to the target smile parameters. The reconstructed smile is then obtained by computing the volatility smile from this reconstructed distribution using the exact forward model. The RMSE shown in each panel measures the discrepancy between the target smile and the reconstructed smile.

The performance of the inverse Jacobian is consistent with a linearized approximation. The inverse mapping performs well for small perturbations, but poorly for larger deviations from the flat reference smile, as shown in Fig. 8 and Fig. 9. Since the Jacobian is computed for the representative maturity $T = 1$ year, the size of this local validity domain may differ for other maturities.

4.4 Application to the S&P 500 Market Smiles

The previous sections showed that the forward and inverse Jacobian mappings are only accurate in a small perturbative domain. In this section, the inverse mapping is tested on real market data. We apply the inverse mapping to observed volatility smiles of the S&P 500 index. The market data are obtained from a single option-chain snapshot. The results in this section should therefore be interpreted as a snapshot analysis, since an analysis over longer time periods has not been performed. As a benchmark, we use the implied-volatility version of the Breeden–Litzenberger method described by Tavin [10], as introduced in Sec. 2.3. For each maturity T_j, the RMSE is computed between the fitted market smile and the reconstructed smile:

$$\mathrm{RMSE}_\sigma^{\mathrm{inv}}(T_j) = \left[\frac{1}{N_j} \sum_{i=1}^{N_j} \left(\sigma_{\mathrm{fit}}(k_i, T_j) - \sigma_{\mathrm{rec}}(k_i, T_j) \right)^2 \right]^{1/2} . \tag{4.3}$$

Here N_j is the number of log-moneyness grid points used in the fitted smile for maturity T_j. This is the same RMSE measure as in Sec. 4.3, but now applied to each maturity separately.
Three representative smiles are shown in Fig. 10.

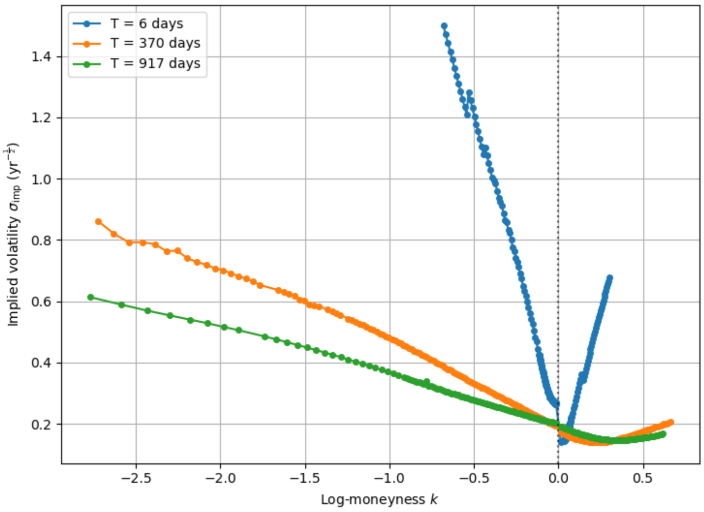

Figure 10: Observed implied volatility σ_{imp} of the S&P 500 index as a function of the log-moneyness k. Implied volatilities for three representative maturities of approximately 6, 370, and 917 days are plotted. Strike prices with too little data or unreliable implied volatility estimates are excluded from the graph.

Compared to the synthetic perturbations used in the previous sections, these smiles show stronger curvature. We therefore do not expect the inverse Jacobian to work as accurately as in the local synthetic tests. The reconstructed distributions for three representative smiles are shown in Fig. 11.

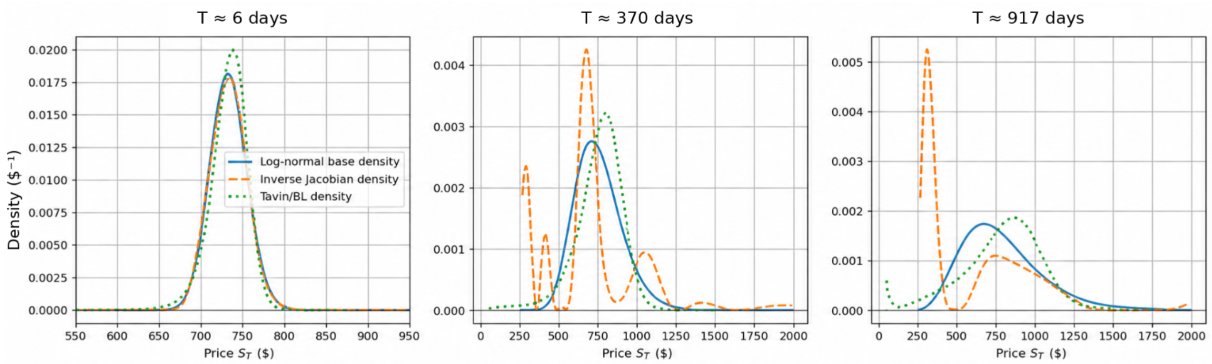

Figure 11: Reconstructed distributions as a function of the stock price S_T for three representative maturities. Each panel compares the log-normal reference distribution with the distribution reconstructed by the inverse Jacobian method and the Breeden–Litzenberger/Tavin benchmark. The benchmark is included as a reference for the inverse Jacobian reconstruction.

The distributions constructed by the inverse Jacobian differ from the log-normal reference distribution and the benchmark, especially for longer maturities. The reconstruction errors are shown in Fig. 12, where the fitted smiles of the S&P 500 are compared to the volatility smiles implied by the reconstructed distributions.

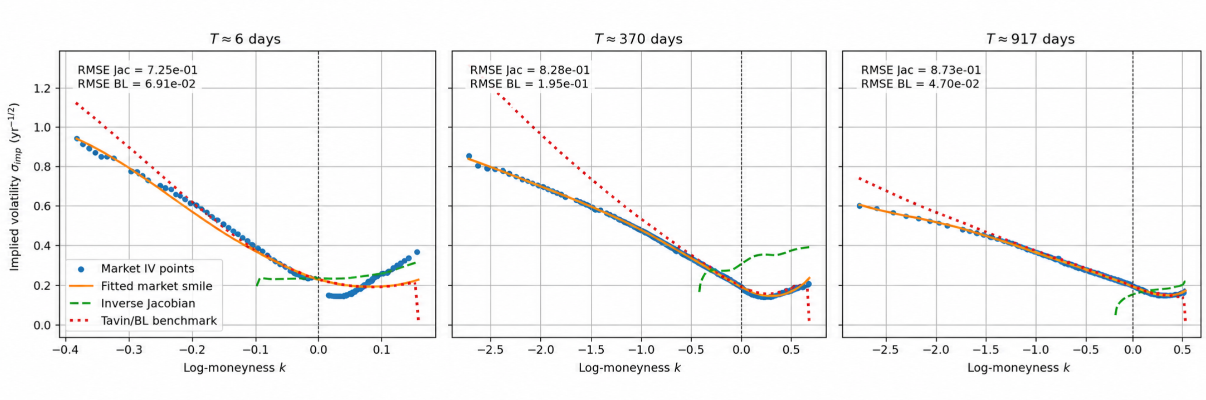

Figure 12: Fitted market smiles versus the smiles reconstructed from the distributions obtained via the inverse Jacobian mapping and the Breeden–Litzenberger/Tavin benchmark for three representative maturities. For each maturity, the RMSE is computed between the fitted market smile and each reconstructed smile. The benchmark shows greater accuracy than the inverse Jacobian reconstruction.

For these three representative maturities, the performance of the Jacobian inverse mapping is poor, while the benchmark reconstruction performs better. The RMSE for the benchmark and the Jacobian inverse mapping are plotted for each observed volatility smile from the S&P 500 in Fig. 13.

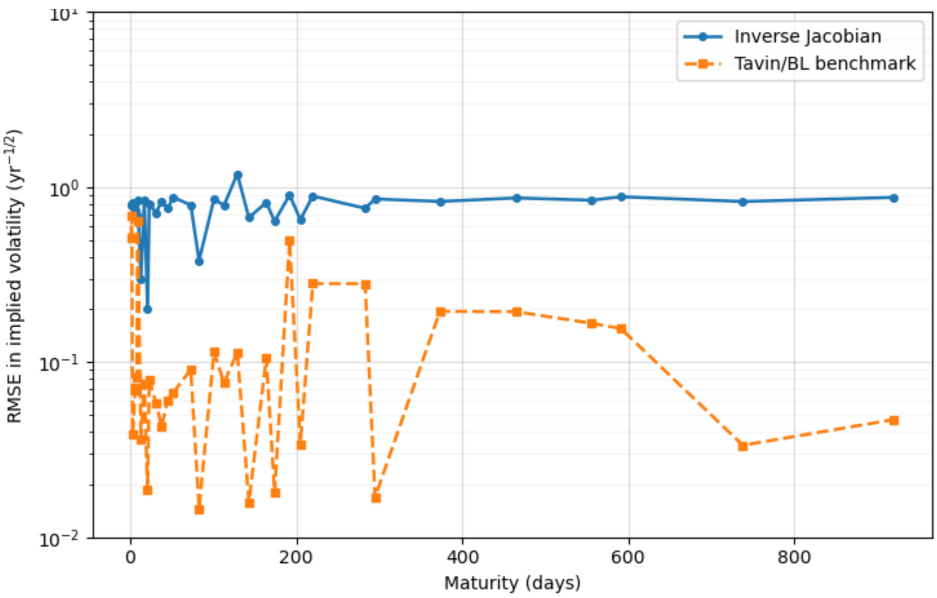

Figure 13: RMSE as a function of maturity. For each maturity, the RMSE is calculated between the fitted market smile and both the smiles reconstructed from the distributions obtained via the inverse Jacobian and the Breeden–Litzenberger/Tavin benchmark distribution. The RMSE should be interpreted as a maturity-wise diagnostic, since different maturities have different strike ranges. The benchmark shows greater accuracy than the inverse Jacobian method across all maturities.

The available strike range differs between maturities. Maturities with wider strike ranges include more wing behaviour, which is typically harder to reconstruct. Therefore, the RMSE should be interpreted as a maturity-wise diagnostic. We use the RMSE values given in the three representative smiles in Fig. 12 to interpret the magnitude of the RMSE values in Fig. 13. The RMSE is large across all maturities and consistently worse than the Breeden–Litzenberger/Tavin benchmark, suggesting that the inverse Jacobian fails to accurately reconstruct probability distributions of the underlying asset.

4.5 Attempted Diagonalization in Hermite Modes

The Jacobian matrix connects two parameter spaces to each other. A natural question is to look for a basis in which the mapping becomes diagonal. A diagonal mapping would allow for a cleaner inversion of the mapping. Furthermore, a diagonalization permits a more direct interpretation of smile features and how they relate to distribution features. Our main attempt consisted of finding a different basis in which to express the volatility smiles. Naturally, we attempted the Hermite series as this is also used in the parameter space of the distribution. This basis did not provide an approximate diagonal mapping, indicating that the forward map mixes Hermite components. For detailed results and further elaboration, we refer the reader to Appendix B.

4.6 Singular Value Analysis

We have seen that the Jacobian matrix is not trivial to diagonalize analytically. To numerically find the bases in which the linearized mapping is diagonal, we can utilize singular value decomposition (SVD), explained in detail in Sec. 3.3. We apply the decomposition $J = U\Sigma V^T$ to the SNP forward Jacobian given in Fig. 5. The columns of V are distribution-space modes, the columns of U are smile-space modes, and the singular values in the diagonal matrix Σ measure the strength of each mode. Fig. 14 shows the modes of the distribution and smile after applying the SVD.

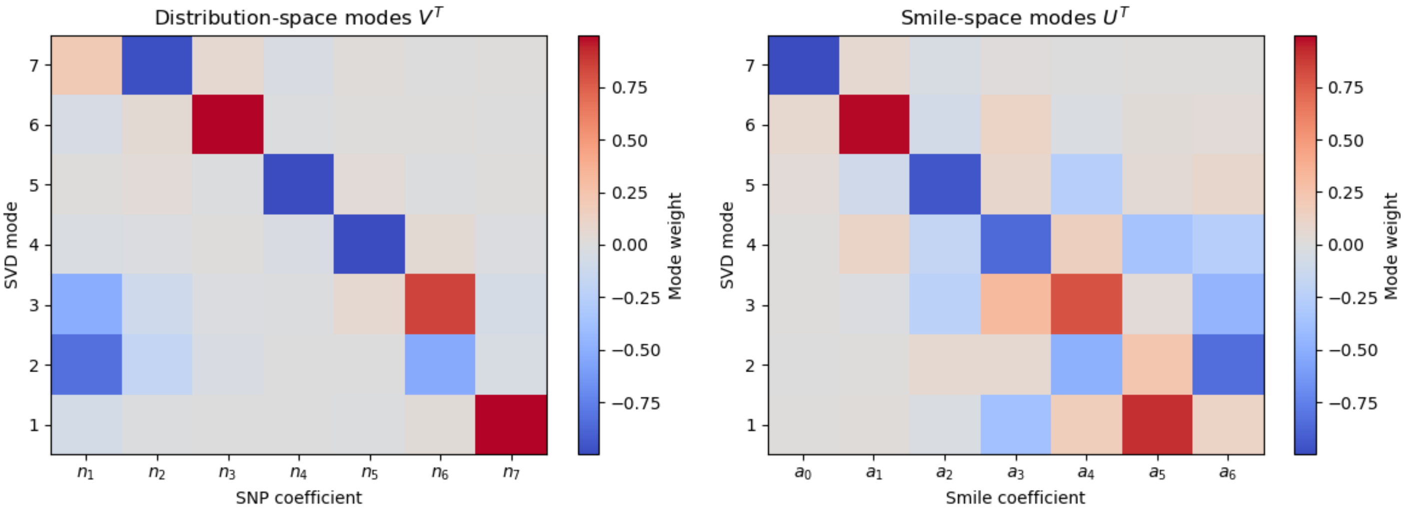

Figure 14: Heatmaps of the distribution-space modes V^T and the smile-space modes U^T after applying SVD to the SNP forward Jacobian in Fig. 5. The y-axis represents the mode numbers, ordered from largest to smallest singular value. The x-axis shows the composition of each mode in terms of the distribution parameters and the smile parameters, respectively. These modes form orthogonal bases in their respective spaces.

The diagonal entries (singular values) of Σ are shown in Fig. 15.

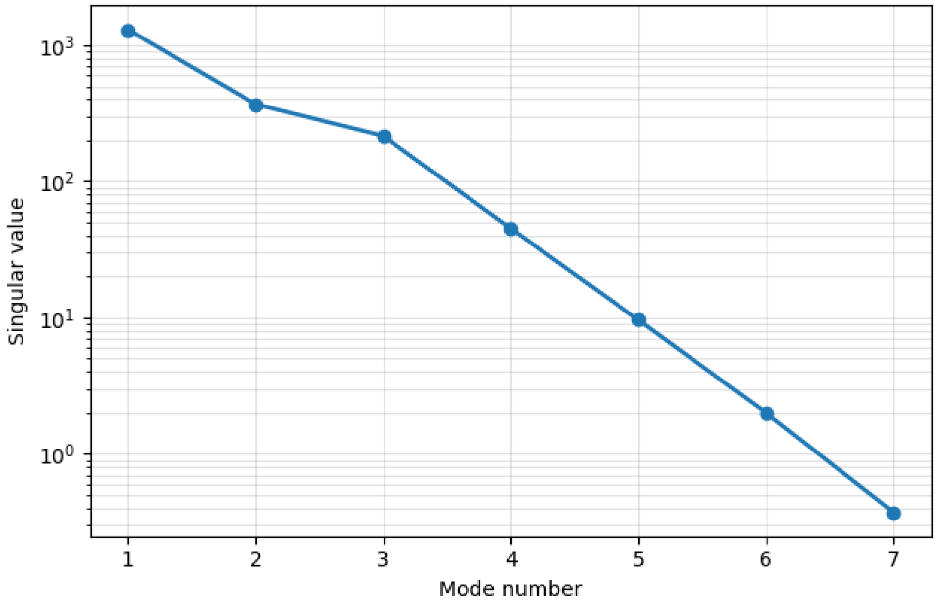

Figure 15: Singular values as a function of mode number. Singular values are obtained after applying SVD to the SNP forward Jacobian in Fig. 5. The singular values decay strongly, indicating that some distribution modes have a much weaker effect on the smile.

The singular values decay over several orders of magnitude, indicating that certain distribution directions have only a weak effect on the smile. This provides a quantitative indication that the Jacobian is ill-conditioned, which is further discussed in Sec. 5.2.

5 Discussion

In this chapter we discuss the results of the Jacobian mapping. We analyze the inverse of this mapping and examine why it fails to produce accurate distributions when tested on the S&P 500 option data as shown in Sec. 4.4. We interpret this failure in terms of three main mechanisms: locality, ill-conditioning, and noise in the market data. Finally, we propose a Kalman filtering extension as a possible way to regularize the inverse problem.

5.1 Locality of the Jacobian Approximation

The Jacobian mapping acts as a local linear approximation of the forward model around the log-normal reference distribution. The inverse mapping is therefore centered around a flat volatility smile. Because it is a local linear approximation, we cannot expect it to behave well for strongly curved smiles. The synthetic inverse test in Fig. 8 shows that the reconstruction error increases as the perturbation size grows. Fig. 10 shows three smiles from the S&P 500. They have visible curvature and are not small deviations from a flat smile. The poor reconstruction of the S&P 500 smiles is therefore consistent with the locality of the Jacobian approximation: the inverse Jacobian is expected to be accurate only close to the flat reference smile. This limitation is also reflected in Fig. 13, where the inverse Jacobian gives a higher RMSE than the Breeden–Litzenberger/Tavin benchmark for almost all maturities.

5.2 Ill-Conditioning and Regularization

A second mechanism that can explain why the inverse Jacobian fails is the ill-conditioning of the Jacobian. It means that small perturbations in certain smile parameters have a large effect on the distribution. To make this claim more rigorous, we look at the singular value analysis in Sec. 4.6. The singular values decay strongly for the singular modes. The singular modes are given by:

$$J\mathbf{v}_i = s_i\mathbf{u}_i,$$

where $\mathbf{v}_i$ is a direction in distribution-parameter space, $\mathbf{u}_i$ is the corresponding direction in smile-parameter space, and s_i is the singular value. The inverse operation is given by:

$$\delta\boldsymbol{\theta} = J^{-1}\delta\mathbf{a} = \sum_i \frac{\mathbf{u}_i^T\delta\mathbf{a}}{s_i}\mathbf{v}_i.$$

Here, $\boldsymbol{\theta}$ is the vector of distribution parameters and $\boldsymbol{a}$ is the vector of smile parameters. For small singular values, the $1/s_i$ term can blow up. Therefore, if noise is introduced in these specific modes, it can greatly change the outcome of the distribution. This explains the sensitivity of the inverse problem to numerical errors, fitting errors, and market noise. Since SVD is only a change of basis, we conclude that this instability is not a property of the decomposition, but rather of the Jacobian itself. SVD merely demonstrates the ill-conditioning of the Jacobian. This source of error can be reduced by regularization, i.e., suppressing modes in the smile-parameter space with small singular values. This can therefore reduce the amplification of noise.

5.3 Noise in Market Smile Data

A third source of error in the inverse method is the noise in the observed volatility smiles themselves. Market noise can have many origins; we list the main sources:

- *Trading activity and liquidity.* The more an option is traded, the more stable its price is expected to be. Options that are barely traded can be easily moved in price by individual trades. This is especially relevant for strike prices far away from the forward price, where trading activity is generally low. These parts of the smile are therefore more noisy.

- *Bid-ask spreads.* The best bid is the highest price someone is willing to pay for the option; the best ask is the lowest price someone is willing to sell the option for. We have used the midpoint of the best bid and best ask to calculate the option price. A larger spread between the best bid and best ask means our midpoint estimate of the option price becomes less certain. This uncertainty in the price estimation propagates to uncertainty in the smile.

- *Fitting and interpolation.* Market data are only available at discrete strike prices and maturities. The volatility smile is thus also discrete. We must fit this discrete smile to obtain a set of smile parameters. Different fitting methods can lead to different smile coefficients, especially for higher-order terms. The choice of fitting method thus introduces an additional degree of uncertainty in the volatility smile.

Noise in the market smiles affects the performance of the inverse mapping. We have divided it into two mechanisms:

- *Amplification by ill-conditioning.* As discussed in Sec. 5.2, small errors in the smile coefficients can be amplified by the ill-conditioned inverse matrix. If $\boldsymbol{\eta}$ denotes the error in the fitted smile coefficients, we can write
$$\mathbf{a}_{\text{market}} = \mathbf{a}_{\text{true}} + \boldsymbol{\eta}.$$
Applying the inverse Jacobian gives
$$\boldsymbol{\theta}_{\text{rec}} = J^{-1}\mathbf{a}_{\text{market}} = J^{-1}\mathbf{a}_{\text{true}} + J^{-1}\boldsymbol{\eta}.$$
The final term shows how errors in the smile coefficients are mapped into errors in the reconstructed distribution parameters. If the inverse is ill-conditioned, this error term can become large even when $\boldsymbol{\eta}$ is small.

- *Dependence on the fitted smile representation.* Because different fitting methods can create different smile parameters for the same discrete and noisy smile, the reconstructed distribution depends on the chosen smile representation.

We can interpret the observed market smiles as noisy measurements of the true underlying smiles. We can therefore propose a Kalman filtering extension, which acts as a filter for the noise in our observations.

5.4 Kalman Filtering as an Extended Model

Since the inverse mapping is sensitive to noise in the market smile, we propose a Kalman filtering approach as an extension of the direct inversion method. The Kalman filter can be interpreted as a linear-Gaussian Bayesian updating method [15, 16]. In practice, this filtering is performed over calendar time. However, we propose a model in which we use the maturities as our time steps. We treat the distribution parameters as hidden variables and connect them to the smile via the Jacobian. The observed smiles are then treated as noisy measurements of the hidden distribution parameters.
Let the available maturities be ordered as

$$T_1 < T_2 < \cdots < T_N.$$

We start with a prior belief for the distribution parameters at the longest maturity T_N. Since long maturities tend to be closer to the log-normal reference distribution, we take the initial prior to be

$$\boldsymbol{\theta}_{T_N}^{\text{prior}} = \mathbf{0}.$$

This prior is then updated using the observed smile:

$$\boldsymbol{\theta}_{T_N}^{\text{updated}} = \boldsymbol{\theta}_{T_N}^{\text{prior}} + G_{T_N}\left(\mathbf{a}_{T_N}^{\text{obs}} - J_{T_N}\boldsymbol{\theta}_{T_N}^{\text{prior}}\right).$$

Here, G_{T_N} is the Kalman gain. It determines how strongly the observed smile changes the prior belief. By assuming a linear, piecewise evolution of the distribution parameters from one maturity to the next shorter maturity, we can calculate a prior $\boldsymbol{\theta}_{T_{N-1}}^{\text{prior}}$. This prior can again be updated with the smile parameters at T_{N-1}. Repeating this procedure iteratively from T_N to T_1 gives us a filtered distribution-parameter estimate for each smile:

$$\left\{\boldsymbol{\theta}_{T_1}^{\text{updated}}, \boldsymbol{\theta}_{T_2}^{\text{updated}}, \ldots, \boldsymbol{\theta}_{T_N}^{\text{updated}}\right\}. \tag{5.1}$$

A more detailed description of the filtering model, including the Kalman gain, the evolution between neighbouring distribution parameters, and further implementation choices, is given in App. C.

6 Conclusion and Outlook

Conclusion

The goal of this thesis was to investigate to what extent a local parametric Jacobian mapping can recover the implied probability distribution of the underlying asset from observed volatility smiles. The results show that the forward Jacobian matrix works as a local approximation of the forward model. The inverse of the Jacobian mapping is also valid as a local approximation, but breaks down for larger perturbations. The inverse Jacobian significantly underperforms compared to the Breeden–Litzenberger benchmark and therefore gives less reliable reconstructions of the probability distribution.
We therefore conclude that the inverse Jacobian cannot recover accurate probability distributions from observed volatility smiles. We attribute this underperformance to a combination of locality, ill-conditioning, and noise in the observed smile. The Jacobian matrix can act as a sensitivity tool to study perturbations around the log-normal reference distribution.

Outlook

The results of our work suggest several possible directions for future work. First, the accuracy of the inverse Jacobian on market smiles could be improved by finding a different reference distribution to linearize around. A probability distribution corresponding to a significantly curved volatility smile would be a good starting point. Second, the instability of the inverse mapping could be improved by using regularization. The singular value analysis showed that small singular values can cause major amplification of small errors in the smile. Whether to suppress or fully remove the weak modes remains a modelling choice and is up to the researcher applying the method. Finally, the Kalman filtering model proposed in Sec. 5.4 could be implemented on market data. This may provide a more stable way to obtain probability distributions from observed market smiles.

Acknowledgements and Declarations

I would like to thank Dr. Farshid Jafarpour for his guidance throughout the development of this thesis.

Readers interested in the code used for this thesis are welcome to contact the author at `joostvankoesveld@gmail.com`. The observed market data were obtained on 12 June 2026.

The author acknowledges the use of generative AI (ChatGPT-4o) for grammar inspection and stylistic improvements of the text. All AI-supported amendments have been thoroughly reviewed by the author.

A Black-Scholes Model

A.1 Assumptions and Derivation of the PDE

The Black–Scholes model provides a framework for pricing European option contracts [2, 4]. The model makes the following assumptions about the financial market:

- Arbitrage opportunities are assumed not to exist. In practice, this means that a dynamically traded portfolio whose evolution is completely deterministic must grow at the risk-free rate r.
- Trading is assumed to be frictionless, meaning transaction costs associated with buying and selling assets are neglected.
- Trading is assumed to occur continuously in time.
- The underlying stock follows geometric Brownian motion with constant drift μ and constant volatility σ.
- Only European options are considered, implying that the option can only be exercised at maturity.
- Dividend payments are neglected.

The price of the stock follows a stochastic process; therefore, the option price also follows a stochastic process. The model proceeds by constructing a portfolio containing both the option and the underlying stock, such that the stochastic component of the portfolio can be eliminated:

$$\Pi = V - \Delta S, \tag{A.1}$$

where Π is the value of the total portfolio, V is the price of the option, S is the price of the underlying stock, and Δ is the hedge ratio. The hedge ratio is chosen such that the stochastic component in the portfolio vanishes. The elimination of the stochastic component implies that the portfolio must evolve deterministically over an infinitesimal time interval. The change in the portfolio value is given by:

$$\mathrm{d}\Pi = \mathrm{d}V - \Delta \mathrm{d}S. \tag{A.2}$$

The stock price follows geometric Brownian motion:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W, \tag{A.3}$$

where μ is the constant drift rate, σ is the constant volatility, and the Wiener process W has independent, normally distributed increments satisfying:

$$\langle \mathrm{d}W \rangle = 0, \quad \langle (\mathrm{d}W)^2 \rangle = \mathrm{d}t \tag{A.4}$$

The option price is a function of the stock price and time, $V = V(S,t)$. Applying Itô's lemma to $V(S,t)$ gives:

$$\mathrm{d}V = \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial S}\mathrm{d}S + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\mathrm{d}S)^2 + \cdots, \tag{A.5}$$

where the omitted higher-order terms vanish in the limit $\mathrm{d}t \to 0$, except for $(\mathrm{d}S)^2$ due to the stochastic nature of the Wiener process. $(\mathrm{d}S)^2$ can be written as:

$$\begin{aligned}(\mathrm{d}S)^2 &= (\mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W)^2 \\ &= \mu^2 S^2 (\mathrm{d}t)^2 + 2\mu\sigma S^2\,\mathrm{d}t\,\mathrm{d}W + \sigma^2 S^2 (\mathrm{d}W)^2 \\ &\approx \sigma^2 S^2\,\mathrm{d}t,\end{aligned} \tag{A.6}$$

where terms of order $(\mathrm{d}t)^2$ and $\mathrm{d}t\,\mathrm{d}W$ vanish in the limit $\mathrm{d}t \to 0$, and we used $(\mathrm{d}W)^2 = \mathrm{d}t$.
Eq. (A.5) can now be written as:

$$\mathrm{d}V \approx (\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S^2)\mathrm{d}t + \sigma S\frac{\partial V}{\partial S}\mathrm{d}W \tag{A.7}$$

By substituting Eq. (A.3) and Eq. (A.7) into Eq. (A.2), we obtain:

$$\left(\frac{\partial V}{\partial \mathrm{t}} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - \Delta\mu S\right)\mathrm{d}t + \left(\sigma S\frac{\partial V}{\partial S} - \Delta\sigma S\right)\mathrm{d}W = \mathrm{d}\Pi \tag{A.8}$$

We now choose the hedge ratio:

$$\Delta = \frac{\partial V}{\partial S},$$

such that the stochastic term proportional to $\mathrm{d}W$ disappears:

$$\mathrm{d}\Pi = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - \mu S\frac{\partial V}{\partial S}\right)\mathrm{d}t. \tag{A.9}$$

The drift term μ in Eq. (A.3) represents the expected growth rate of the stock and depends on investor risk preferences. Remarkably, the same hedge ratio that eliminates the stochastic term also cancels the drift contribution proportional to μ. The portfolio evolution therefore becomes:

$$\mathrm{d}\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t. \tag{A.10}$$

Since the stochastic component has been eliminated, the portfolio no longer carries exposure to the stochastic stock price fluctuations. The portfolio will thus evolve deterministically over an infinitesimal time interval.
By the no-arbitrage assumption, this deterministic portfolio must grow at the risk-free rate:

$$\mathrm{d}\Pi = \Pi r\mathrm{d}t = (V - \Delta S)r\,\mathrm{d}t \tag{A.11}$$

Eq. (A.10) then becomes:

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t = (V - \Delta S)r\,\mathrm{d}t \tag{A.12}$$

Rearranging gives us the Black–Scholes partial differential equation [2]:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{A.13}$$

A notable property of the Black–Scholes equation is that the drift parameter μ no longer appears in the final result. Before constructing the hedged portfolio, one would generally expect the option price to depend on the expected return of the stock, since the stock itself follows the stochastic process in Eq. (A.3). In practice, however, the expected return of a stock is difficult to determine accurately and depends on investor expectations and risk preferences.
By choosing the hedge ratio appropriately, the stochastic fluctuations proportional to $\mathrm{d}W$ are eliminated from the portfolio evolution. Remarkably, this same hedge ratio also removes the drift contribution proportional to μ. The resulting Black–Scholes equation therefore no longer depends on the expected return of the stock.
This observation leads to the concept of risk-neutral valuation (see, e.g., [4]). Since the option pricing equation no longer depends on the real-world drift μ, option prices can instead be calculated by discounting the expected future payoff at the risk-free rate:

$$V(t) = e^{-r(T-t)}\mathbb{E}^Q[V(T)], \tag{A.14}$$

where $\mathbb{E}^Q$ denotes the expectation value under the risk-neutral distribution.
It is important to note that the risk-neutral distribution does not necessarily correspond to the true real-world probability distribution of future stock prices. Instead, it represents the probability distribution implied by market option prices under the no-arbitrage framework. For the purpose of this thesis, we focus on this implied distribution, since it is the distribution directly connected to the observed volatility smile.

A.2 Solution

We start from the Black–Scholes PDE [2, 4]:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{A.15}$$

where V is the option price, S is the stock price, σ is the volatility, and r is the risk-free rate. It is a second-order, parabolic PDE. It can be classified as a backwards PDE, as the terminal condition is known and we work towards the initial condition.

For a call option, the terminal condition at $t = T$ is given by:

$$V_{\text{call}}(S,T) = \max(S-K,0), \tag{A.16}$$

and similarly for a put option:

$$V_{\text{put}}(S,T) = \max(K-S,0). \tag{A.17}$$

For a call option, the boundary conditions are:

$$V_{\text{call}}(0,t) = 0, \qquad V_{\text{call}}(S,t) \sim S - Ke^{-r(T-t)} \quad \text{as } S \to \infty, \tag{A.18}$$

and similarly for a put option:

$$V_{\text{put}}(0,t) = Ke^{-r(T-t)}, \qquad V_{\text{put}}(S,t) \to 0 \quad \text{as } S \to \infty. \tag{A.19}$$

To solve the PDE, the general idea is to transform to the heat equation first, as its solution is known. We first define the log-price: $x = \ln \frac{S}{K}$, which turns multiplicative motion into additive motion. After substitution, we obtain a PDE with constant coefficients:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial V}{\partial x} - rV = 0. \tag{A.20}$$

To reverse time, we define the time-to-maturity variable: $\tau = T - t$. After substitution, we obtain:

$$-\frac{\partial V}{\partial \tau} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right)\frac{\partial V}{\partial x} - rV = 0. \tag{A.21}$$

To remove the advection term: $\left(r - \frac{1}{2}\sigma^2\right)\frac{\partial V}{\partial x}$ and the decay term: rV, we rescale the option value V as follows:

$$V(x,\tau) = Ke^{\alpha x + \beta \tau} u(x,\tau), \tag{A.22}$$

where

$$\alpha = -\frac{1}{2}\left(\frac{2r}{\sigma^2} - 1\right), \qquad \beta = -\frac{\sigma^2}{8}\left(\frac{2r}{\sigma^2} + 1\right)^2. \tag{A.23}$$

This substitution leaves us with the heat equation:

$$\frac{\partial u}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2}, \tag{A.24}$$

with initial condition:

$$u(x,0) = \frac{1}{K}e^{-\alpha x} V(Ke^x, T). \tag{A.25}$$

Remarkably, we have been able to transform the Black–Scholes PDE to a heat equation by a set of transformations. This heat equation can be solved by using the Green's function method. The standard solution is given by:

$$u(x,\tau) = \frac{1}{\sqrt{2\pi\sigma^2\tau}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{2\sigma^2\tau}\right] u(\xi,0)\, d\xi. \tag{A.26}$$

Transforming it back to the original variables, we find the solution to the Black–Scholes partial differential equation for a call option:

$$V_{\text{call}}(S,t) = S\,\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2), \tag{A.27}$$

Similarly for a put option:

$$V_{\text{put}}(S,t) = Ke^{-r(T-t)}\Phi(-d_2) - S\,\Phi(-d_1), \tag{A.28}$$

where

$$d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}. \tag{A.29}$$

Here, $\Phi(x)$ is the cumulative distribution function of the standard normal distribution:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2}\, dz. \tag{A.30}$$

The transformation to the heat equation shows us that the option price follows a diffusive process in logarithmic price space. This diffusive structure explains why the Black–Scholes model predicts a flat volatility smile, as it corresponds to a pure Gaussian distribution in the log-price space [17].

B Hermite Basis of the Smile

The smile was originally parametrized with a power series in log-moneyness $k = \log(K/F_0)$. We have made several attempts to diagonalize or at least simplify the Jacobian mapping. In this appendix, we show the result of changing the basis from a power series to Hermite polynomials. Since the probability distribution is expanded into Hermite polynomials, the hypothesis was that a smile in a Hermite basis might simplify the Jacobian mapping. For the change of basis, we assume the smile is described with a power series up to sixth order:

$$\sigma(k) = a_0 + a_1 k + a_2 k^2 + \cdots + a_6 k^6. \tag{B.1}$$

The Hermite polynomials up to sixth order are:

$$\mathrm{He}_0(k) = 1, \tag{B.2}$$
$$\mathrm{He}_1(k) = k, \tag{B.3}$$
$$\mathrm{He}_2(k) = k^2 - 1, \tag{B.4}$$
$$\mathrm{He}_3(k) = k^3 - 3k, \tag{B.5}$$
$$\mathrm{He}_4(k) = k^4 - 6k^2 + 3, \tag{B.6}$$
$$\mathrm{He}_5(k) = k^5 - 10k^3 + 15k, \tag{B.7}$$
$$\mathrm{He}_6(k) = k^6 - 15k^4 + 45k^2 - 15. \tag{B.8}$$

We can write the monomials in the power series in terms of Hermite polynomials:

$$1 = \mathrm{He}_0, \tag{B.9}$$
$$k = \mathrm{He}_1, \tag{B.10}$$
$$k^2 = \mathrm{He}_2 + \mathrm{He}_0, \tag{B.11}$$
$$k^3 = \mathrm{He}_3 + 3\mathrm{He}_1, \tag{B.12}$$
$$k^4 = \mathrm{He}_4 + 6\mathrm{He}_2 + 3\mathrm{He}_0, \tag{B.13}$$
$$k^5 = \mathrm{He}_5 + 10\mathrm{He}_3 + 15\mathrm{He}_1, \tag{B.14}$$
$$k^6 = \mathrm{He}_6 + 15\mathrm{He}_4 + 45\mathrm{He}_2 + 15\mathrm{He}_0. \tag{B.15}$$

The power series can then be rewritten in Hermite polynomials:

$$\sigma(k) = b_0\mathrm{He}_0(k) + b_1\mathrm{He}_1(k) + \cdots + b_6\mathrm{He}_6(k), \tag{B.16}$$

where the coefficients b_i are given by a linear combination of the originally fitted coefficients a_i:

$$b_0 = a_0 + a_2 + 3a_4 + 15a_6, \tag{B.17}$$
$$b_1 = a_1 + 3a_3 + 15a_5, \tag{B.18}$$
$$b_2 = a_2 + 6a_4 + 45a_6, \tag{B.19}$$
$$b_3 = a_3 + 10a_5, \tag{B.20}$$
$$b_4 = a_4 + 15a_6, \tag{B.21}$$
$$b_5 = a_5, \tag{B.22}$$
$$b_6 = a_6. \tag{B.23}$$

In matrix form, the change of basis is written as:

$$\mathbf{b} = M\mathbf{a}, \tag{B.24}$$

where $\mathbf{a}$ is the vector of power-series coefficients, $\mathbf{b}$ is the vector of Hermite coefficients, and M is the corresponding change-of-basis matrix. If the original Jacobian maps perturbations in the distribution parameters $\boldsymbol{\ell}$ to perturbations in the power-series smile coefficients,

$$\delta\mathbf{a} = J\delta\boldsymbol{\ell}, \tag{B.25}$$

then the Jacobian in the Hermite smile basis is

$$\delta\mathbf{b} = MJ\delta\boldsymbol{\ell}. \tag{B.26}$$

Thus,

$$J_{\mathrm{H}} = MJ. \tag{B.27}$$

In the computed Jacobian used for this test, the distribution has not been parameterized with the SNP series in Eq. (2.27), but with the original Gram-Charlier series in Eq. (2.26). In the Gram–Charlier series, the Hermite polynomials are decoupled, allowing for perturbations in individual Hermite polynomials. The result of this change of basis applied to the Jacobian is shown in Fig. 16.

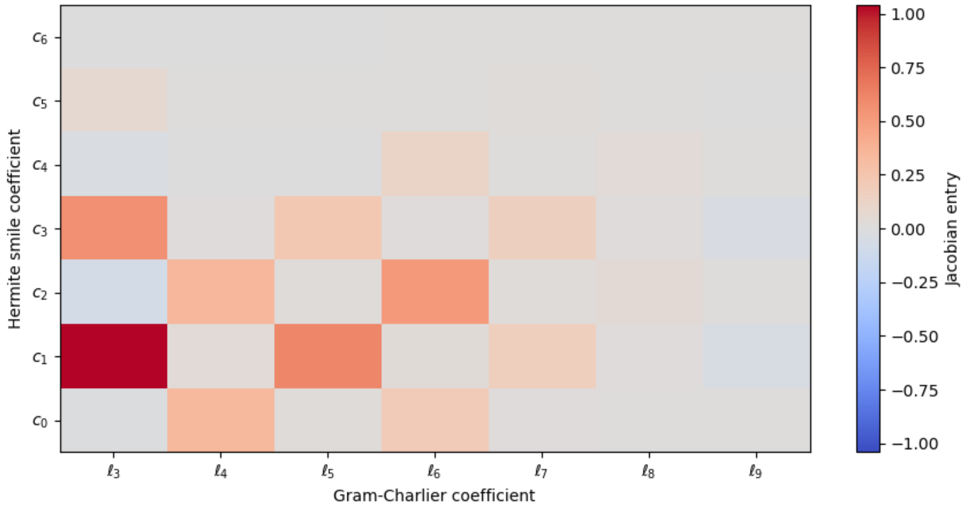

Figure 16: Heatmap of the Gram–Charlier Jacobian for $T = 1$ year. The vertical axis shows the smile coefficients and the horizontal axis shows the distribution coefficients. The distribution coefficients are denoted by ℓ_n, which are standardized Gram–Charlier coefficients and satisfy $\ell_n = \frac{\kappa_n}{v^{n/2}}$, $v = \sigma^2 T$, where κ_n are cumulants. The smile parameterization has been changed from a power series to a Hermite series. The Jacobian entries are computed using a finite perturbation of size 10^{-4} applied to each distribution parameter ℓ_n.

We observe a parity structure, where even-order distribution parameters predominantly influence even-ordered Hermite polynomials in the smile, while the odd-ordered distribution parameters predominantly influence odd-order Hermite polynomials. However, a one-to-one mapping where the m^{th} order distribution parameter affects only the m^{th} order Hermite polynomial in the smile is not observed. We conclude that the Hermite basis does make the Gram–Charlier Jacobian more structured, but it does not make the Jacobian diagonal.

C Kalman Filtering Extension

In this appendix, we lay out the Kalman filtering model in detail. This filtering model is motivated by the hypothesis that it regularizes the unstable inverse Jacobian. This model is not supported by numerical results, but rather serves as a possible direction for future work.

State-Space Model

We treat the observed smile parameters at a maturity T_j as noisy measurements of the hidden distribution parameters:

$$\mathbf{a}_{T_j}^{\text{obs}} = J_{T_j}\boldsymbol{\theta}_{T_j} + \boldsymbol{\varepsilon}_{T_j}. \tag{C.1}$$

Here, $\mathbf{a}_{T_j}^{\text{obs}}$ is the observed smile parameter vector, $\boldsymbol{\theta}_{T_j}$ is the hidden distribution-parameter vector, J_{T_j} is the Jacobian at maturity T_j, and $\boldsymbol{\varepsilon}_{T_j}$ represents market noise. A Kalman filter requires a prior belief. If we start the filtering at the longest maturity T_N, our prior belief will be:

$$\boldsymbol{\theta}_{T_N}^{\text{prior}} = \mathbf{0}. \tag{C.2}$$

This means our initial best guess for the distribution at the longest maturity is a Gaussian in log-price space, which is a reasonable starting point because longer maturities tend to have flatter volatility smiles and are thus closer to the log-normal reference distribution. We can then update this prior belief by using the observed smile parameters in Eq. (C.1):

$$\boldsymbol{\theta}_{T_N}^{\text{updated}} = \boldsymbol{\theta}_{T_N}^{\text{prior}} + G_{T_N}\left(\mathbf{a}_{T_N}^{\text{obs}} - J_{T_N}\boldsymbol{\theta}_{T_N}^{\text{prior}}\right), \tag{C.3}$$

where G_{T_N} is the Kalman gain. The Kalman gain determines how strongly the observed smile changes the prior estimate of the distribution parameters. Following the standard Kalman updating [18], where we take the observation matrix to be the Jacobian J_{T_j}, the Kalman gain is written as:

$$G_{T_j} = P_{T_j}^{\text{prior}} J_{T_j}^T \left(J_{T_j} P_{T_j}^{\text{prior}} J_{T_j}^T + R_{T_j}\right)^{-1}, \tag{C.4}$$

where $P_{T_j}^{\text{prior}}$ is the prior covariance matrix of the distribution parameters. It describes the uncertainty in the prior belief $\boldsymbol{\theta}_{T_j}^{\text{prior}}$. The matrix R_{T_j} is the noise covariance matrix of the smile coefficients. It describes the uncertainty in the observed smile parameters due to market noise. The relative size of $P_{T_j}^{\text{prior}}$ and R_{T_j} determines how strongly the prior belief is changed.

After updating $\boldsymbol{\theta}_{T_N}^{\text{prior}} = \mathbf{0}$, we need a prior for $\boldsymbol{\theta}_{T_{N-1}}^{\text{prior}}$. We obtain it by assuming a linear evolution of the distribution parameters between neighbouring maturities:

$$\boldsymbol{\theta}_{T_{N-1}}^{\text{prior}} = F_{T_N \to T_{N-1}} \boldsymbol{\theta}_{T_N}^{\text{updated}}, \tag{C.5}$$

where $F_{T_N \to T_{N-1}}$ is the evolution matrix. It can be a simple choice like the identity matrix or it can be more sophisticated. The same idea applies to the uncertainty in the prior belief $P_{T_j}^{\text{prior}}$. This covariance matrix is assumed to propagate in an analogous way:

$$P_{T_{N-1}}^{\text{prior}} = F_{T_N \to T_{N-1}} P_{T_N}^{\text{updated}} F_{T_N \to T_{N-1}}^T, \tag{C.6}$$

which is a simplified evolution model for the covariance matrix. The covariance matrix is updated following standard Kalman rules [19]:

$$P_{T_j}^{\text{updated}} = \left(I - G_{T_j} J_{T_j}\right) P_{T_j}^{\text{prior}}. \tag{C.7}$$

The observation-noise covariance R_{T_j} describes uncertainty in fitted smile coefficients at T_j. It is not propagated by $F_{T_N \to T_{N-1}}$ and it must be chosen for each maturity independently.
Repeating the procedure in Eqs. (C.5) and (C.6) iteratively from T_N to T_1 gives us a filtered distribution-parameter estimate for each smile:

$$\left\{\boldsymbol{\theta}_{T_1}^{\text{updated}}, \boldsymbol{\theta}_{T_2}^{\text{updated}}, \ldots, \boldsymbol{\theta}_{T_N}^{\text{updated}}\right\}. \tag{C.8}$$

Minimal Implementation

To turn this filtering process into an algorithm, we must make implementation choices for $F_{T_j \to T_{j-1}}$, $P^{\text{prior}}_{T_j}$, and R_{T_j}. A simple choice is to assume that the prior distribution parameters for T_{j-1} are closely related to the updated distribution parameters at T_j. We therefore set

$$F_{T_j \to T_{j-1}} = I. \tag{C.9}$$

For the prior covariance matrix, we can make a simple choice as well:

$$P^{\text{prior}}_{T_j} = p_j I, \tag{C.10}$$

where the scalar p_j controls how uncertain we are about the prior distribution parameters.
Similarly, for the noise covariance matrix:

$$R_{T_j} = r_j I, \tag{C.11}$$

where the scalar r_j controls how noisy the observed smile is.
For a first implementation, we can assume p_j and r_j to be constant.

Interpretation as Regularized Inverse Jacobian

To argue that this filtering model is a regularized inverse Jacobian, we compare it to the direct inverse. The noisy observation equation is:

$$\mathbf{a}^{\text{obs}}_{T_j} = J_{T_j} \boldsymbol{\theta}_{T_j} + \boldsymbol{\varepsilon}_{T_j}. \tag{C.12}$$

The direct inversion instead estimates:

$$\boldsymbol{\theta}_{T_j} \approx J^{-1}_{T_j} \mathbf{a}^{\text{obs}}_{T_j}. \tag{C.13}$$

This can potentially amplify the noise $\boldsymbol{\varepsilon}_{T_j}$ greatly. The Kalman filter instead combines the observation in Eq. (C.12) with a prior prediction for $\boldsymbol{\theta}_{T_j}$. It chooses parameters that reproduce the observed smile while not moving too far away from the prior prediction. It therefore acts like a regularization of the inverse Jacobian.
This regularization does not remove the locality limitation of the method. The matrix J_{T_j} is still a local linear approximation around the log-normal reference distribution. Therefore, the filtering approach may reduce the sensitivity to noisy smile coefficients, but it does not solve the breakdown of the linear Jacobian approximation for distributions too far from the reference distribution.

References

[1] Douglas T. Breeden and Robert H. Litzenberger. "Prices of State-Contingent Claims Implicit in Option Prices". In: *The Journal of Business* 51.4 (1978), pp. 621–651.

[2] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. DOI: 10.1086/260062.

[3] Carl V. L. Charlier. "Über die Darstellung willkürlicher Funktionen". In: *Arkiv för Matematik, Astronomi och Fysik* 2.20 (1905), pp. 1–34.

[4] John C. Hull. *Options, Futures, and Other Derivatives*. Pearson, 2018.

[5] Roger W. Lee. "Implied Volatility: Statics, Dynamics, and Probabilistic Interpretation". In: *Recent Advances in Applied Probability*. Springer, 2005, pp. 241–268.

[6] Peter W. Buchen and Michael Kelly. "The Maximum Entropy Distribution of an Asset Inferred from Option Prices". In: *Journal of Financial and Quantitative Analysis* 31.1 (1996), pp. 143–159. DOI: 10.2307/2331384.

[7] Edwin T. Jaynes. "Information Theory and Statistical Mechanics". In: *Physical Review* 106.4 (1957), pp. 620–630. DOI: 10.1103/PhysRev.106.620.

[8] Allan M. Malz. *A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions*. Staff Report 677. Federal Reserve Bank of New York, 2014.

[9] Bhupinder Bahra. "Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application". In: *Bank of England Working Paper* 66 (1997).

[10] Bertrand Tavin. "Implied Distribution as a Function of the Volatility Smile". In: *Working Paper* (2011).

[11] Pakorn Aschakulporn and Jin E. Zhang. "Option-pricing formulas with skewness and kurtosis". In: *Review of Derivatives Research* (2025). DOI: 10.1007/s11147-025-09224-5.

[12] R. Hu. "Option Pricing under Gram-Charlier Density". In: *Working Paper* (2024).

[13] Eric Jondeau and Michael Rockinger. "Gram–Charlier Densities". In: *Journal of Economic Dynamics and Control* 25.10 (2001), pp. 1457–1483. DOI: 10.1016/S0165-1889(99)00082-2.

[14] Ángel León, Javier Mencía, and Enrique Sentana. *Parametric Properties of Semi-Nonparametric Distributions, with Applications to Option Valuation*. Vol. 27. Journal of Business & Economic Statistics. Article published in Journal of Business & Economic Statistics, 27(2), 176–192. New York: Taylor & Francis, 2009.

[15] Mascia Bedendo and Stewart D. Hodges. "The Dynamics of the Volatility Skew: A Kalman Filter Approach". In: *Journal of Banking & Finance* 33.6 (2009), pp. 1156–1165. DOI: 10.1016/j.jbankfin.2008.12.014.

[16] Si Chen, Zhen Zhou, and Shenghong Li. "An Efficient Estimate and Forecast of the Implied Volatility Surface: A Nonlinear Kalman Filter Approach". In: *Economic Modelling* 58 (2016), pp. 655–664. DOI: 10.1016/j.econmod.2016.06.003.

[17] Jim Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley, 2006.

[18] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering* 82.1 (1960), pp. 35–45. DOI: 10.1115/1.3662552.

[19] Andrew C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press, 1989.