



A Probabilistic Approach for Denoising Option Prices

Djibril Gueye^{1*}, Kokulo Lawuobahsumo²

¹QUANTLABS: The Innovative Subsidiary of the QUANTEAM Group, a Consulting Firm Specializing in Banking and Insurance, in Financial Services and IT Professions, 75008, Paris, France, ²Department of Economics, Statistics and Finance, University of Calabria, Ponte Bucci, 87030 Rende, Italy. *Email: djibril.gueye@quantlabs

Received: 11 December 2022

Accepted: 22 February 2023

DOI: <https://doi.org/10.32479/ijefi.13993>

ABSTRACT

This paper aims to directly denoise option price while adhering to the no-arbitrage conditions. To achieve our goal, we propose the Gaussian Process (GP) method that entails training the GP on noisy data of option prices as a linear function of the pair of maturity and strike. Utilizing the GP approach not only allows for removing noises on the option price surface by verifying the no-arbitrage conditions but also is a probabilistic approach that allows quantifying the uncertainty on the quantity of interest by constructing confidence bands around the estimate. The GP further permits forecasting out-of-the-sample prices without needing to compute the risk-neutral density of the option price surface. To investigate the efficiency of GP in removing the noise from option prices, we tested it on a simulated dataset. The overall MSE between the computed Black-Scholes prices and the GP denoised is 0.10, and between the Black-Scholes prices and the noisy prices is 2.21 - a 95.33% noise removal. The curves of the graphs for the denoised prices are all convex and non-increasing in strikes, upholding the no-arbitrage conditions. To our best knowledge, the challenge of directly denoising option prices has led to little interest in this area, and our work is the first to undertake this task.

Keywords: Gaussian Process, Denoising, Wavelet, Arbitrage, Option Price

JEL Classifications: C1, G1

1. INTRODUCTION

The options market is noisy, with large bid/ask spreads, making determining a fair price for any given option extremely difficult. Noise in the prices arises from errors in the recording and reporting process and also from price discreteness, liquidity, and non-synchronicity.

The literature on denoising noisy raw data is populated with the application of the wavelet method. Due to its multi-scaling property, the efficiency of the wavelet method in dealing with noisy data series has resulted in its adoption as a cross-disciplinary tool (Averkamp and Houdr'e, 2003; Lada and Wilson, 2006; Asgharian, 2011; Capobianco, 2001; Sun and Meinl, 2012; Jammazi et al., 2015; Liu et al., 2019, etc.)

To obtain the filtered (denoised) price, implied volatilities (IV) for each option are computed and then denoised. The denoised

volatility and stock price are plugged into an option model to get cleaner option prices. Haven et al. (2012) used this approach to denoise option prices applying the wavelet method. The wavelet denoising procedure involves decomposing the noisy data series by computing the discrete wavelet transform at a given level to obtain the wavelet coefficient vectors and the scaling coefficient vector. Appropriate thresholding is performed so that wavelet coefficients with magnitudes less than the threshold are set to zero or shrunk towards zero by the amount of the threshold using the hard or soft thresholding rule, respectively. We reconstruct the entire series to obtain the denoised series by replacing the wavelet coefficient vectors with the thresholded wavelet coefficient vectors (Haven et al., 2012 for detailed procedure).

However, by using the wavelet method for denoising implied volatility, Haven et al. (2012) did not consider the corresponding

option price to avoid arbitrage opportunities. The denoised implied volatility is used as input to price those financial derivatives, but using the implied volatility should be made consistently considering the no-arbitrage conditions. In addition, the wavelet method is deterministic, and its use does not allow quantifying the uncertainty of the estimated implied volatility.

Given the drawbacks of the wavelet approach, the paper aims to contribute to the literature on denoising option prices by directly denoising the option price while honoring the no-arbitrage conditions. We propose the Gaussian Process (GP) for denoising option prices.

Our approach consists in training the GP on the noisy option prices and considering it as a linear function of the pair maturity strike. In order to take into account the no-arbitrage conditions of the market, we use the constrained GP developed in Cousin et al. (2016). Thus, the filtered option prices and the associated noises appeared respectively by the most probable response price surface and the most probable noise vector corresponding to the Maximum a Posteriori (MAP) of the GP and that of the noise presented in Chataigner et al. (2021). Using the GP approach not only contributes to the construction of the option price surface verifying the no-arbitrage conditions but also, being a probabilistic approach, allows quantifying the uncertainty on the quantity of interest by constructing confidence bands around that estimate. Furthermore, the GP allows for forecasting out-of-the-sample prices without computing the RND, unlike the wavelet method.

Let's recall that the GPs regression known also as kriging is basically used in geostatistics for estimating the distribution of mineral resources in the ground given the relatively small set of boreholes, (Matheron, 1963; Cressie, 1990; Krige and Magri, 1982) and has for some years begun to gain popularity in quantitative finance. These works concern, among others the paper of Sousa et al. (2012) for calibrating the Vasicek interest rate model under the risk-neutral measure, Asgharian et al. (2013) in the analysis of stock market linkages, Ludkovski (2015) for improving the Monte Carlo Least square method for the valuation of Bermuda, De Spiegeleer et al. (2018) comparing to Monte Carlo method for pricing options and approximating implied volatility, Dixon and Crepey (2018) in derivative portfolio modeling, Gonzalvez et al. (2019) for fitting the yield curve, the recent paper of Ludkovski and Saporito (2021) for approximating the option Greeks.

It is only thereafter the paper of Cousin et al. (2016) that constrained kriging has been applied in actuarial science where authors constructed the term structure of interest rate by imposing market no-arbitrage conditions. Their technique has been used in Cousin and Gueye (2021) for implied volatility surface and by Chataigner et al. (2021) who compared constrained GPs and Neural Networks for local volatility construction.

Compared to previous work on options such as that of Ludkovski and Saporito (2021), we set ourselves a double objective. On the one hand, we take into account the no-arbitrage conditions in the

approximation of the option prices. On the other hand, we seek to determine the associated noise, focusing on the MAP of the noise vector combined with the GP.

Our approach is not so far from that of Chataigner et al. (2021), except that we exploit the noise MAP to contribute to the literature of noise removal on option prices which, to our best knowledge, remains a challenge until now and that there are few articles that are interested in it, which is the novelty of our technique.

Throughout the paper, we restrict our attention to European calls on a stock (or index) S in an economy with a constant interest rate r .

The remainder of this paper is organized as follows. Section 2 focuses on the methodology of denoising option prices using GPs. We briefly recall GPs modeling with linear inequality constraints and emphasize the solution to the denoising problem. Section 3 is devoted to numerical experiments.

2. MATERIALS AND METHODS

We consider at time θ the market call price $C(T, K)$ on the underlying S as a 2D surface in the pair (T, K) where T represents its maturity and K its strike price. Given n noisy observations $\hat{y} = [\hat{y}_1, \dots, \hat{y}_n]^T$ of the function C at input points $X = [x_1, \dots, x_n]^T$, where $x_i = (T_i, K_i)$ represents the observed maturities and strikes, we aim at removing the noise $\varepsilon = [\varepsilon_1, \dots, \varepsilon_n]^T$ associated to \hat{y} to obtain the denoised prices y that respect the no-arbitrage conditions of the market. In others terms, we aim at estimating C at the input points y based on the model $\hat{y} = C(X) + \varepsilon$ such that C verifies the no-arbitrage conditions. These conditions are interpreted by a finite number of linear inequalities constraints such as the convexity of C in the strikes and the increase of C in the maturities. Furthermore, in addition to these points, C should also be non-increasing in strikes. That is the call price surface C should fulfill the following conditions,

$$\partial_T C(T, K) \geq 0, \partial_{K^2}^2 C(T, K) \geq 0, \partial_K C(T, K) \leq 0 \quad (1)$$

For this purpose, we consider the call price function C to be a zero-mean bivariate GP with a 2-dimensional isotropic covariance kernel function k and assume the noise term ε to be a zero-mean Gaussian vector, independent from $C(X)$, and with a covariance matrix given as $\sigma^2 I_n$, where I_n is the identity matrix of dimension n . The covariance kernel function κ is defined for any $x = (T, K)$ and $x' = (T', K')$ as $\kappa(x, x') = \sigma^2 R_T(T - T', \theta_T) R_K(K - K', \theta_K)$

Where represents the length scale hyper-parameter of κ and s corresponds its the variance hyper-parameter, R_T and R_K are kernel correlation functions. For instance, in our numerical experiments we use some *mate'rn* correlation functions which can be written, for example in the maturity direction, by

$$R_T(T - T') = \left(1 + \frac{\sqrt{5}|T - T'|}{\theta_T} + \frac{5(T - T')^2}{3\theta_T^2} \right) \exp\left(-\frac{\sqrt{5}|T - T'|}{\theta_T} \right)$$

To handle the condition (1) in the estimation, we approximate the Gaussian prior C by a finite-dimensional GP C^N presented in the following subsection (e.g., Chataigner et al., 2021; Cousin and Gueye, 2021 for more details).

2.1. Approximation of the Bivariate GP C

Let Ω the rectangular domain in time and space of the GP C , we consider a discretized version of Ω as a $N=(N_T+1)\times(N_K+1)$ regular grid $\Omega^N = \{(u_p, v_j) \mid u_i = ih_T, v_j = jh_K, i = 0, \dots, N_T, j = 0, \dots, N_K\}$ where $h_T = \frac{1}{N_T}$ and $h_K = \frac{1}{N_K}$. For each knot (u_p, v_j) , we introduce the hat basis function $\phi(T, K)$ over Ω , as the following tensor product

$$\phi_{i,j}(T, K) = \max\left(1 - \frac{|T - u_i|}{h_T}, 0\right) \max\left(1 - \frac{|K - v_j|}{h_K}, 0\right)$$

So that $\text{supp}(\phi_{i,j}) = [K_{i-1}, K_{i+1}] \times [T_{j-1}, T_{j+1}]$

The finite-dimensional approximation C^N of C is the piecewise-bilinear quadrilateral interpolation of C at the knots $(u_p, v_j)_{i,j}$ so that

$$C^N(T, K) = \sum_{i=0}^{N_T} \sum_{j=0}^{N_K} \phi_{i,j}(T, K) \beta_{i,j}, \quad \forall (T, K) \in \Omega \quad (2)$$

Where $\beta = [\beta_{0,0}, \dots, \beta_{ij}, \dots, \beta_{N_T, N_K}]^T$, with $\beta = (u_p, v_j)$ is a zero-mean Gaussian vector with $N \times N$ covariance matrix Γ^N such that $\Gamma_{i_1 i_2}^N = \kappa([u_{i_1}, v_{j_1}], [u_{i_2}, v_{j_2}])$, for any two grid index pairs (i_1, j_1) and (i_2, j_2) corresponding to global indices I_1 and I_2 respectively. If we denote $\phi(T, K) = [\phi_{0,0}(T, K), \dots, \phi_{i,j}(T, K), \dots, \phi_{N_T, N_K}(T, K)]$, one has the matrix form of the equality (2) given as $C^N(x) = \Phi(x) \beta$, where $C^N(x) = [C^N(x_1), \dots, C^N(x_n)]^T$ and $\Phi(x)$ denotes the $n \times N$ matrix of basis function in which, each row i corresponds to the vector $\phi(T_p, K_p)$.

The advantage of using the approximation (2) is the fact that since C converges to C^N when N goes to infinity, no-arbitrage conditions (by means condition [1]) could be verified on C^N . However, using C^N leads to a finite number of inequality constraint checks unlike using the original GP which is infinite-dimensional. By consequence, we may restate the condition (1) as follows

(i) $\partial_T(T, K) \geq 0$ if and only if $\beta_{i,j+1} \geq \beta_{i,j} \quad \forall (i,j)$,

(ii) $\partial_{K^2} C(T, K) \geq 0$ if and only if $\beta_{i,j+2} - \beta_{i,j+1} \geq \beta_{i,j+1} - \beta_{i,j} \quad \forall (i,j)$,

(iii) $\partial_T(T, K) \leq 0$ if and only if $\beta_{i,j} \geq \beta_{i,j+1} \quad \forall (i,j)$
 Of course, in order to better adjust the GP to real market prices, it is necessary to also take into account certain market evidence such as the following equalities,

(iv) $C(0, K) = \max(S_0 - K, 0)$, for all K ,

(v) $C(T, 0) = S_0$, for all T .

This choice would be judicious since these two relations will behave as being equality constraints of the GP type interpolation model.

By denoting M set of inequality constraints ((i), (ii), (iii)), i.e., M is the set of 2D continuous functions that are non-decreasing in T , convex in K , and non-increasing in K , our problem consists in estimating β and the noise ε associated to C_N conditionally to,

$$\begin{cases} \tilde{y} = \Phi \cdot \beta + \varepsilon \\ \beta \in M \end{cases}$$

Given the inequality constraints, the best estimator of the denoised price is the MAP of C^N and the one of the noise is the MAP of ε .

2.2. The Solution of the Denoising Problem

As we mentioned above, the best approximators of the filtered price and the noise are given by the joint MAP $(m_{C^N}, \hat{\varepsilon})$ of the truncated Gaussian C^N and the Gaussian noise vector ε . Note that $m_{C^N} = \Phi \cdot \hat{\beta}$ where $\hat{\beta}$ is the MAP of β . Hence the solution to the denoising problem is to find the joint MAP $(\hat{\beta}, \hat{\varepsilon})$ of the truncated Gaussian coefficient β and the Gaussian noise vector which is a solution of the problem

$$\max_{\beta, \varepsilon} \mathbb{P}(\beta \in [\beta, \beta + d\beta], \varepsilon \in [\varepsilon, \varepsilon + d\varepsilon] \mid \Phi \cdot \beta + \varepsilon = \tilde{y}, \beta \in M)$$

Since (β, ε) and is Gaussian centered with block-diagonal covariance matrix with blocks Γ^N and $\zeta^2 I_n$ the MAP $(\hat{\beta}, \hat{\varepsilon})$ satisfies

$$(\hat{\beta}, \hat{\varepsilon}) = \min_{\beta, \varepsilon} \mathbb{P}_{\beta, \varepsilon}(\beta \in [\beta, \beta + d\beta], \varepsilon \in [\varepsilon, \varepsilon + d\varepsilon] \mid \Phi \cdot \beta + \varepsilon = \tilde{y}, \beta \in M) \quad (3)$$

With this, we reduced the denoising problem to a quadratic problem.

3. NUMERICAL EXPERIMENTS

In this section, we propose to illustrate the behavior and the efficiency of GP to remove the noise of option prices on a simulated dataset. These data come from the simulation of the Black–Scholes model for 512 call options with 32 different strikes ranging from 30 to 100 and, 16 maturities ranging from 0.02 to 3 years listed on $t = 0$ with $S = 70$. The risk-free interest rate is set at 0.05 and the volatility at 0.2. To measure the effectiveness of the GP, we propose to disturb the call option prices obtained with a given noise. Of course, the efficiency and behavior of the model chosen to remove noise will depend on the level of these noises. Considering this last aspect, it is natural to train the GP on different noisy data characterized by their noise level. To do this, for a call price vector $C = [c_1, \dots, c_n]$ of size n obtained from the Black–Scholes model, we choose a spread of $\delta\%C$ from the initial prices and we define the noise vector covariance matrix (assumed to be a

zero-mean Gaussian vector) as $\zeta^2 I_n$ with $\zeta = \sqrt{\frac{1}{n} \sum_i (\delta\%c_i)^2}$.

This choice of deviation for the prices is completely natural although that of the noise level δ is arbitrary. As we will see later, the interest in this noise level will allow us, during the calibration phase of the GP on real data, to characterize with precision the nature of the noise linked to any data set and to be able to locate the robustness of the GP in removing such noise.

Once the noise level has been chosen, we adjust the GP on each set of data comprising $n = 512$ noisy call prices obtained after disturbance using a mat'ern 5/2 covariance structure. The estimation of the GP parameters is performed using the Maximum Likelihood (MLE) method implemented in the R package Dicekriging developed by Roustant et al. (2012). The results of this estimation are summarized in Table 1 where we have chosen different levels of noise $\delta \in \{2,3,5,8,10,15,20\}$. To better adjust the GP on the call price surface, we added the prices with maturities 0 and strike prices 0 on the equality constraints of the GP since these are always accessible on the market regardless of the types of data. Regarding the basic functions of the finite-dimensional approximation, we choose 30 subdivisions in the direction of the maturities and 20 in the direction of the strikes, which will make a number of 20×30 nodes of subdivision. To study the performance evaluation of the behavior of the model, we propose to present the quantitative results on the mean square error (MSE) metric.

The analysis of this metric for the different retained noise levels δ was first carried out on the entire surface and then constructed on certain maturities. In Table 2, we represent the MSE (named MSE Denoised Price) between the true call option prices and the prices estimated by the GP, the MSE (named MSE Noised Price) between the true prices and the noised prices as well as the percentage reduction of MSE Noised Prices compared to MSE Denoised Prices in the Reduction column. Table 3 illustrates the same quantities as Table 2 with a much more detailed study by searching the MSEs by maturity. A wide range of low MSEs is usually noted, which leads to strong noise reduction.

Table 1: Hyperparameter estimates using the maximum likelihood estimator for different noise levels δ

δ (%)	$\hat{\theta}_K$	$\hat{\theta}$	$\hat{\sigma}$	$\hat{\zeta}$	$\hat{\zeta}$
3	43.1337	5.7651	15.6026	0.6942	0.7082
5	49.0177	5.9600	17.1289	1.1185	1.1803
7	50.0195	5.9600	17.1847	1.5596	1.6525
10	51.6419	5.9600	17.2343	2.1807	2.3607
15	53.7653	5.9600	17.4755	3.1706	3.5410
20	54.8400	5.9600	17.6194	4.0880	4.7213

Table 2: The global mean square error given a noise level δ between the noised prices and the true ones, and the mean square error between the Gaussian process estimated prices (the denoised prices) and the true prices

δ (%)	MSE noised price	MSE denoised price	Reduction (%)
3	0.43	0.04	90.54
5	1.13	0.07	93.71
7	2.21	0.10	95.33
10	4.36	0.19	95.59
15	9.31	0.59	93.61
20	15.73	2.24	85.79

He last column compute the reduction the MSE after denoising. MSE: Mean square error

For the convenience of the reader, we present the case where $\delta = 7\%$. In that case, the overall MSE between the true prices and the GP estimated ones is 0.10 while that between the true prices and the noisy prices is 2.21 (Table 2), which corresponds to a reduction of the latter by 95.33% after estimation by using the GP. Such an observation on the performance of the reduction can also be done in Table 4. These results also correspond to the Figure 1 representing the estimated prices according to the strikes for the maturities $\{0.02, 0.22, 0.62, 1.41, 2.01, 3.00\}$ and where one can observe a good correspondence between the GP and the real prices of the call by Black–Scholes. In addition, all these curves are convex and non-increasing in strikes, which means that the no-arbitrage conditions are being fulfilled by them.

We also tested the behavior of the GP on out-of-sample forecasts, in particular on the strike set $\{10.15, 20.25\}$ and the maturity $T = 0.5$. The results represented in the table 4 reveal a good behavior of the GP in the forecast with a low MSE equal to 0.03.

To show the effectiveness of our method, we follow the approach of Haven et al. (2012) to denoise the option prices. It is worth noting that the main objective of Haven et al. (2012) was to improve the

Table 3: Mean square error on the noised and denoised prices

Noise Level (δ)	Maturities	MSE noise price	MSE denoised price	Reduction (%)
$\delta=3\%$	0.02	0.72	0.00	99.51
	0.22	0.21	0.04	78.47
	0.62	0.49	0.03	93.23
	1.41	0.28	0.04	84.84
	2.01	0.29	0.02	91.87
	3.00	0.36	0.05	85.86
	0.02	2.01	0.00	99.76
$\delta=5\%$	0.22	0.60	0.09	84.46
	0.62	1.39	0.05	96.15
	1.41	0.81	0.09	88.52
	2.01	0.87	0.03	96.43
	3.00	1.10	0.09	91.74
	0.02	3.95	0.01	99.85
	0.22	1.21	0.15	87.40
$\delta=7\%$	0.62	2.74	0.05	98.26
	1.41	1.62	0.14	91.58
	2.01	1.76	0.02	98.68
	3.00	2.22	0.13	94.11
	0.02	8.08	0.01	99.91
	0.22	2.56	0.27	89.38
	0.62	4.74	0.17	96.47
$\delta=10\%$	1.41	3.23	0.31	90.38
	2.01	3.70	0.01	99.73
	3.00	4.64	0.27	94.24
	0.02	15.91	0.01	99.93
	0.22	6.17	0.80	87.03
	0.62	11.17	0.88	92.12
	1.41	7.61	0.84	89.01
$\delta=15\%$	2.01	7.25	0.22	97.00
	3.00	11.08	0.66	94.03
	0.02	28.19	0.03	99.91
	0.22	10.95	2.67	75.62
	0.62	16.22	4.23	73.95
	1.41	14.43	2.71	81.24
	2.01	12.75	1.75	86.29
$\delta=20\%$	3.00	20.25	2.36	88.33

MSE: Mean square error

Figure 1: Slices of GP with noise level $\delta = 7\%$. The red curve denotes the denoised prices, the green one represents the Black–Scholes prices and the black points denote the noisy prices.

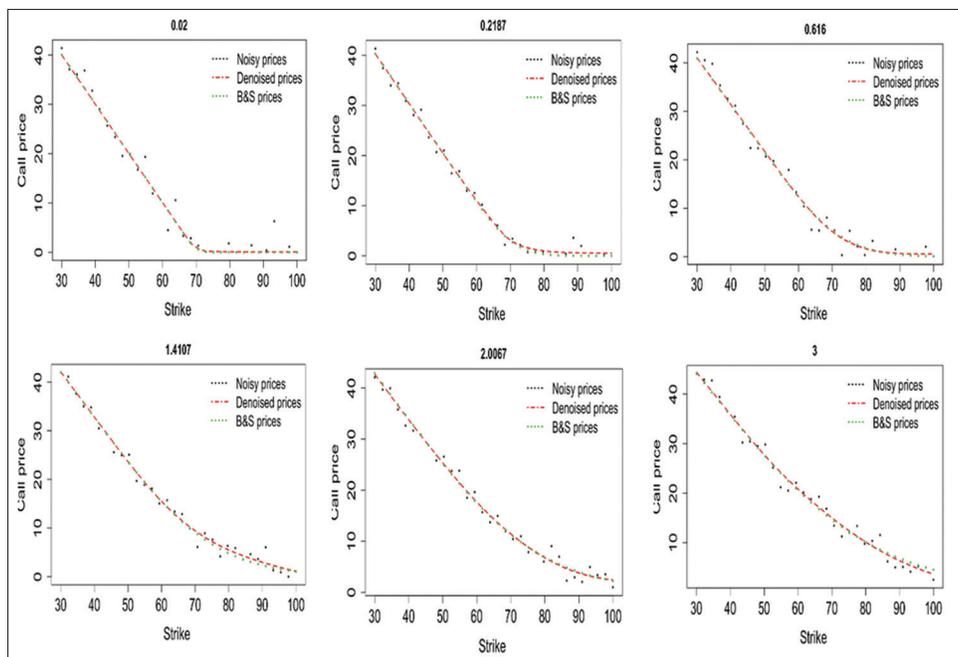


Table 4: Out of sample forecast with $\delta=7$ for the maturity 0.5 year

TTM forecast	Strike forecast	Denoised price forecast	Black–Scholes Price
0.50	10.00	60.33	60.25
0.50	15.00	55.50	55.37
0.50	20.00	50.68	50.49
0.50	25.00	45.85	45.62

TTM: Time to Maturity

performance of the smoothed implied volatility smile method in estimating the implied RND. They applied the wavelet method to denoise option prices to allow better estimation of the risk-neutral densities.

The option-implied RND, derived from option prices, provides relevant data for pricing additional complicated derivatives written on the same underlying asset. Additionally, the implied RND in market prices exhibits substantial abnormalities, such as negative probability, due to microstructure noise from several sources. Shimko (1993) was the first to propose converting option prices into IV to interpolate and smooth the curve, then transform the smoothed IVs back into price space and proceed with the extraction of RND from the resulting set of option prices.

Shimko (1993) cubic spline or a low-order polynomial procedure does not assume that the Black–Scholes model holds for these option prices but uses the Black–Scholes equation as a computational tool for transforming the data into an appropriate space for smoothing. Following the method of Bliss and Panigirtzoglou (2002), Haven et al. (2012) applied the wavelet denoising technique on IV first and obtained option prices based on the denoised volatilities using the Black–Scholes model. One issue with implied volatility is that it varies significantly depending on "moneyness" when computed for options written on the same

underlying. The volatility smile or the implied volatility for options on equities and stock indexes becomes sufficiently asymmetrical over time, with higher implied volatility for low-exercise price options. Implied volatility is also affected by the valuation model used to calculate it. The presence of volatility smile in Black–Scholes implied volatility indicates that option market prices are not completely consistent with that model.

Using the wavelet method to denoise implied volatility does not guarantee the positiveness of implied volatility, meaning that we can get negative implied volatility, which will result in a negative denoised price. We also observe that when the noise is huge, the wavelet method uses the mean of the noisy series as the denoise series. Furthermore, the wavelet method does not work with data, not in the power of 2. Figure 2 shows the denoised prices for different maturities using the wavelet approach. In panel (a), the wavelet method denoised the noisy series for a short maturity of 0.02 years. The MSE between the computed Black–Scholes price and the denoised price is 0.0021. For panel (b-f), we see that the wavelet-denoised price violates the no–arbitrage constraints. We see a curve that is both convex and concave. The concavity of the curves is a violation of the structural form of the option price. Excluding the maturities of 0.02 and 1.4107 years, the denoising process resulted in negative IV. These negative values result in a negative price. For a maturity of 3 years, the MSE of the denoised price is larger than that of the noisy price. The MSE of the denoised price is 36.7285, while the noisy price is 1.1754.

4. EMPIRICAL RESULTS

The panel (a) of Figure 3 displays European data for MSFT call options traded on November 24th, 2022 with spot price $S_0 = 248.03$. This market data contains several available strike prices and maturities ranging from 105 to 400 and from 0.156 to 2.150 respectively. There are violations of the no–arbitrage conditions

Figure 2: Slices of Wavelet with noise level $d = 7\%$. The red curve denotes the denoised prices, the green one represents the Black–Scholes prices and the black points denote the noisy prices

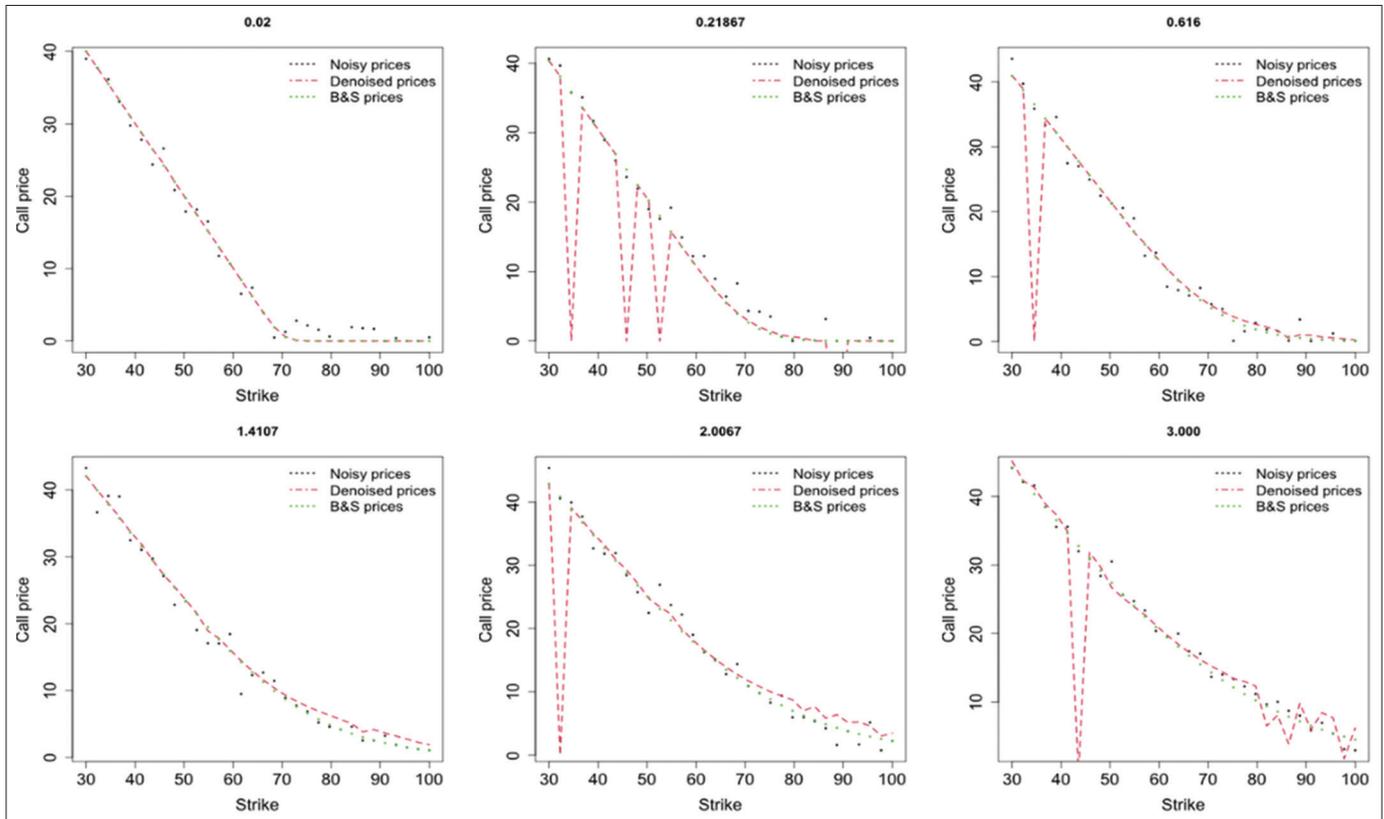
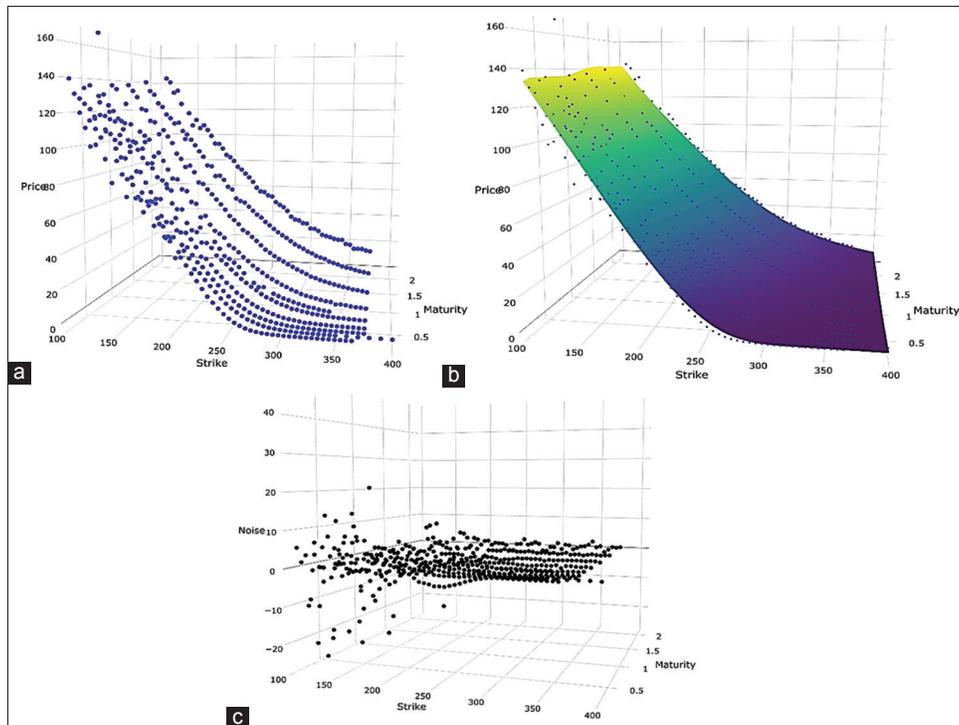


Figure 3: The observed market prices in panel (a), the Gaussian Process MAP estimate in panel (b), and the associated noises of the MAP in panel (c)



for certain maturities, in particular on the low maturities of deep-in-the-money options. These violations would certainly be due to an error related to pricing. Given this remark, we aim at removing

that error without violating the no–arbitrage constraints and at constructing the entire call price surface which also respects these conditions.

Figure 4: Slices of Gaussian Process (GP) MAP (red curve), observed prices (black points) and 1000 paths of the GP's posterior (grey envelop) for maturities $T = 0.2328$ and $T = 0.4053$.

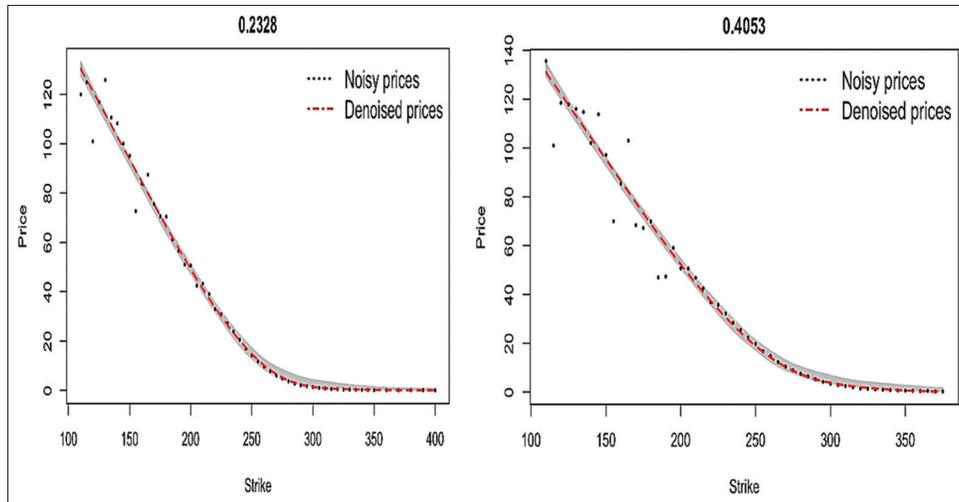
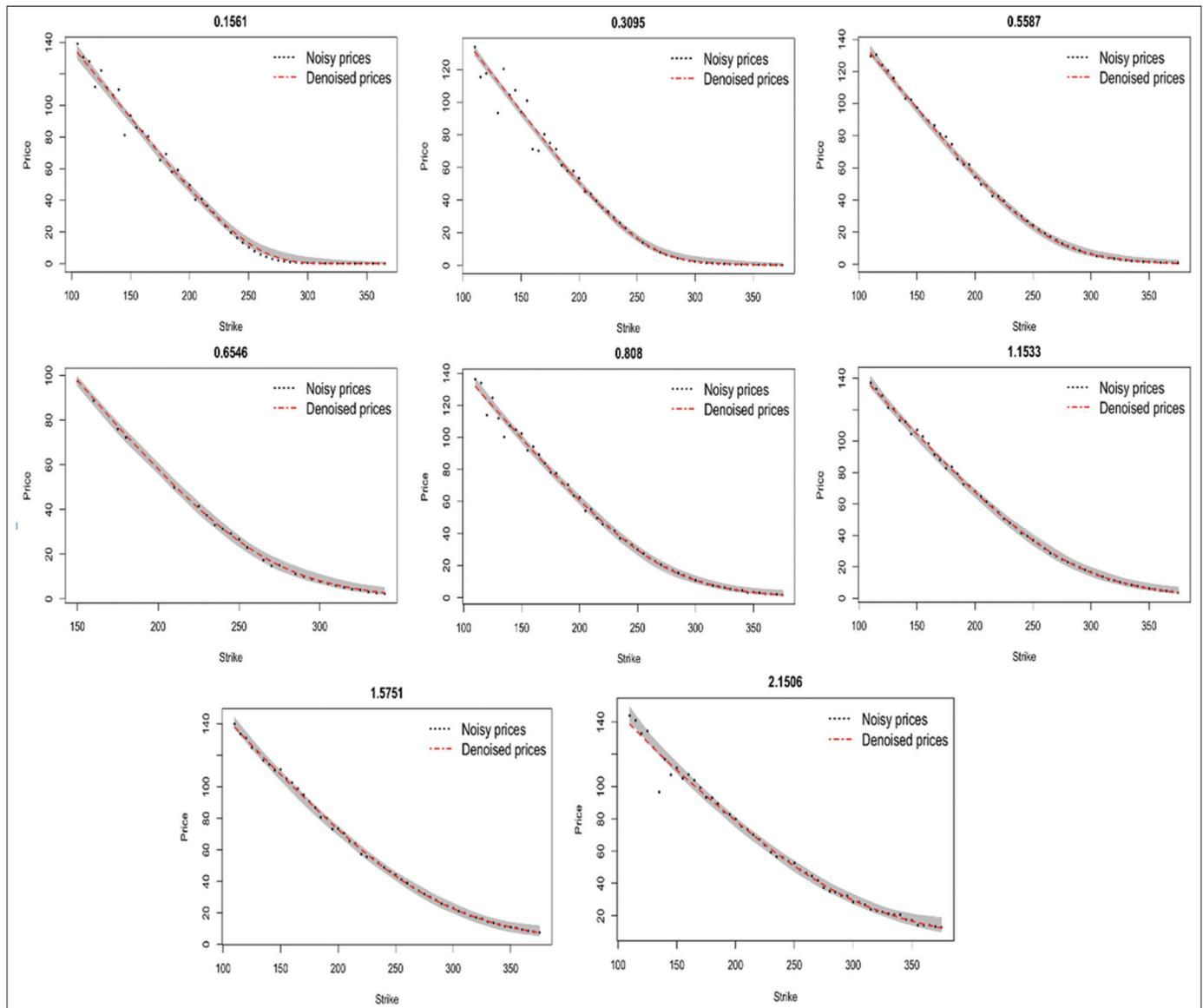


Figure 5: Slices of Gaussian Process (GP) MAP (red curve), observed prices (black points) and 1000 paths of the GP's posterior (grey envelop)



In this regard, we train the constrained GP on the grid of basis functions obtained by discretizing the grid of observation data by means of 50 nodes in the direction of strikes and 30 nodes in the direction of maturities. The GP is based on a mat'ern 5/2 covariance structure whose hyperparameters $\hat{\theta}_T = 3.9890$, $\hat{\theta}_K = 206.6507$, $\hat{\theta} = 55.4256$, as well as the noise variance $\hat{\sigma}^2 = 18.7634$, are estimated using the maximum likelihood estimator (MLE) developed in the Dicekriging R package. To solve the MAP quadratic program (3), we have used the function quadprog of the pracma R package of Borchers and Borchers (2022) which is developed for solving quadratic programming problems with linear and box constraints. The corresponding MAP surface as well as the associated noise vector are respectively given by panels (b) and (c) of Figure 3. Notice that the constructed call price surface is non-decreasing in maturities, and convex and non-increasing in strikes, hence respects the condition (1).

Of course, a major advantage of using GP for denoising option prices is its ability to interpolate prices in areas where there is little observation, as we saw in panel (a) of Figure 3 which constitutes a lot of missing data between the points of the input grid. This feature of GP gives it a special status that other denoising methods, such as the wavelet, do not have, but it is reinforced by its ability to quantify uncertainty by constructing confidence intervals around the estimated quantity. By way of illustration, we have represented the slices of the surface estimated call prices at the maturities $T = 0.2328$ and $T = 0.4053$ in Figure 4 (other slices are presented in Figure 5). The red line denotes the GP MAP estimate, the grey shaded envelopes represent the posterior uncertainty bands under 1000 samples per observation and the black points denote the observed noisy prices. As we mentioned above, most of the noises are concentrated in the deep-in-the-money options. However, the bands deviate a little more for the deep out-of-the-money options. By looking at Figure 5, we observe, for certain maturities (e.g., $T = 1.533$, $T = 1.5751$, $T = 2.1506$), that the envelope captures most of the observation data which implies low noise in these areas. For others where there are very few observations (e.g., $T = 0.6546$), we see that the envelope deviates more. Regarding the noise level of the surface in the observed prices, given the most likely call price surface (i.e., the MAP of the GP) m_{C^N} and the estimated variance of the noise $\hat{\zeta}$, we have computed the noise level $\hat{\delta}$ considered by the model which is a solution of $\hat{\zeta} = \sqrt{\frac{1}{2} \sum_i^n (\hat{\delta} \% m_{C^N})^2}$. The results show a noise level of around 7% treated by the model.

5. CONCLUSION

The aim of this paper was to contribute to the literature on denoising option prices by directly denoising the option price while honoring the no-arbitrage conditions. We proposed the GP framework for this purpose. Using the GP approach not only contributes to the construction of the entire option price surface verifying the no-arbitrage conditions but also allows quantifying the uncertainty on the quantity of interest. Furthermore, the GP permits forecasting out of the sample prices without computing the risk-neutral density.

To investigate the efficiency of GP in removing the noise from option prices, we tested it on a simulated dataset. The overall MSE between the computed Black–Scholes prices and the GP denoised is 0.10, and between the Black–Scholes prices and the noisy prices is 2.21 – a 95.33% noise removal. The curves of the graphs for the denoised prices are all convex and non-increasing in strikes, upholding the no-arbitrage conditions.

We test the effectiveness of our method following the approach of Haven et al. (2012) to denoise the option prices. They applied the wavelet method to denoise option prices to allow better estimation of the risk-neutral densities. This method involves using the Black–Scholes model to extract the implied volatility from the option, then denoise it, and then use the denoise price to recalculate the noise-free option price.

Our analysis showed that using the wavelet method to denoise implied volatility does not guarantee the positiveness of implied volatility, meaning that we can get negative implied volatility, which will result in a negative denoised price. We also observe that when the noise is huge, the wavelet method uses the mean of the noisy series as the denoise series. Furthermore, the wavelet method does not work with data, not in the power of 2. The wavelet denoised prices failed to honor the no-arbitrage condition.

For the empirical application of our method, we used Microsoft (MSFT) call options traded on November 24, 2022, with a spot price of $S_0 = 248.03$. This market data contains several available strike prices and maturities ranging from 105 to 400 and from 0.156 to 2.150, respectively. We observed violations of the no-arbitrage conditions for low maturities of deep-in-the-money options. These violations would be due to an error related to pricing. We used the GP to correct these pricing errors without violating the no-arbitrage constraints and constructed the entire call price surface, which also respects these conditions. To achieve this, train the constrained GP on the grid of basis functions obtained by discretizing the grid of observation data by means of 50 nodes in the direction of strikes and 30 nodes in the direction of maturities.

REFERENCES

- Asgharian, H. (2011), A conditional asset-pricing model with the optimal orthogonal portfolio. *Journal of Banking and Finance*, 35(5), 1027-1040.
- Asgharian, H., Hess, W., Liu, L. (2013), A spatial analysis of international stock market linkages. *Journal of Banking and Finance*, 37(12), 4738-4754.
- Averkamp, R., Houdre, C. (2003), Wavelet thresholding for non-necessarily Gaussian noise: Idealism. *The Annals of Statistics*, 31(1), 110-151.
- Bliss, R.R., Panigirtzoglou, N. (2002), Testing the stability of implied probability density functions. *Journal of Banking and Finance*, 26(2-3), 381-422.
- Borchers, H.W., Borchers, M.H.W. (2022), BWorld Robot Control Software. Vol. 4. Available from: <https://www.cran.r-project.org/web/packages/pracma/pracma.pdf>
- Capobianco, E. (2001), Wavelet transforms for the statistical analysis of returns generating stochastic processes. *International Journal of Theoretical and Applied Finance*, 4(3), 511-534.

- Chataigner, M., Cousin, A., Crépey, S., Dixon, M., Gueye, D. (2021), Beyond surrogate modeling: Learning the local volatility via shape constraints. *SIAM Journal on Financial Mathematics*, 12(3), SC58-SC69.
- Cousin, A., Gueye, D. (2021), Kriging for Implied Volatility Surface. Work in Progress.
- Cousin, A., Maatouk, H., Rulliere, D. (2016), Kriging of financial term-structures. *European Journal of Operational Research*, 255(2), 631-648.
- Cressie, N. (1990), The origins of kriging. *Mathematical Geology*, 22(3), 239-252.
- De Spiegeleer, J., Madan, D.B., Reyners, S., Schoutens, W. (2018), Machine learning for quantitative finance: Fast derivative pricing, hedging and fitting. *Quantitative Finance*, 18(10), 1635-1643.
- Dixon, M.F., Crepey, S. (2018), Multivariate Gaussian Process Regression for Portfolio Risk Modeling: Application to CVA. Department of Applied Mathematics. United States: Illinois Institute of Technology.
- Haven, E., Liu, X., Shen, L. (2012), De-noising option prices with the wavelet method. *European Journal of Operational Research*, 222(1), 104-112.
- Jammazi, R., Lahiani, A., Nguyen, D.K. (2015), A wavelet-based nonlinear ARDL model for assessing the exchange rate pass-through to crude oil prices. *Journal of International Financial Markets, Institutions and Money*, 34(C), 173-187.
- Krige, D.G., Magri, E.J. (1982), Geostatistical case studies of the advantages of lognormal-de Wijsian Kriging with mean for a base metal mine and a gold mine. *Journal of the International Association for Mathematical Geology*, 14(6), 547-555.
- Lada, E.K., Wilson, J.R. (2006), A wavelet-based spectral procedure for steady-state simulation analysis. *European Journal of Operational Research*, 174(3), 1769-1801.
- Liu, X., Cao, Y., Ma, C., Shen, L. (2019), Wavelet-based option pricing: An empirical study. *European Journal of Operational Research*, 272(3), 1132-1142.
- Ludkovski, M. (2018). Kriging metamodells and experimental design for bermudan option pricing. *Journal of Computational Finance* 22 (1), 37-77
- Ludkovski, M., Saporito, Y. (2021), Kri hedge: Gaussian process surrogates for delta hedging. *Applied Mathematical Finance*, 28(4), 330-360.
- Matheron, G. (1963), Principles of geostatistics. *Economic Geology*, 58(8), 1246-1266.
- Roustant, O., Ginsbourger, D., Deville, Y. (2012), Dicekriging, diceoptim: Two R packages for the analysis of computer experiments by kriging-based metamodelling and optimization. *Journal of Statistical Software*, 51, 1-55.
- Shimko, D. (1993), Bounds of probability. *Risk*, 6, 33-37.
- Sousa, J.B., Esquivel, M.L., Gaspar, R.M. (2012), Machine learning Vasicek model calibration with Gaussian processes. *Communications in Statistics-Simulation and Computation*, 41(6), 776-786.
- Sun, E.W., Meinl, T. (2012), A new wavelet-based denoising algorithm for high frequency financial data mining. *European Journal of Operational Research*, 217(3), 589-599.