

EXTREME-STRIKE ASYMPTOTICS FOR GENERAL GAUSSIAN STOCHASTIC VOLATILITY MODELS

ARCHIL GULISASHVILI, FREDERI VIENS, AND XIN ZHANG

ABSTRACT. We consider a stochastic volatility asset price model in which the volatility is the absolute value of a continuous Gaussian process with arbitrary prescribed mean and covariance. By exhibiting a Karhunen-Loève expansion for the integrated variance, and using sharp estimates of the density of a general second-chaos variable, we derive asymptotics for the asset price density for large or small values of the variable, and study the wing behavior of the implied volatility in these models. Our main result provides explicit expressions for the first five terms in the expansion of the implied volatility. The expressions for the leading three terms are simple, and based on three basic spectral-type statistics of the Gaussian process: the top eigenvalue of its covariance operator, the multiplicity of this eigenvalue, and the L^2 norm of the projection of the mean function on the top eigenspace. The fourth term requires knowledge of all eigen-elements. We present detailed numerics based on realistic liquidity assumptions in which classical and long-memory volatility models are calibrated based on our expansion.

JEL Classification: G13, C63, C02.

AMS 2010 Classification: 60G15, 91G20, 40E05.

Keywords: stochastic volatility, implied volatility, large strike, Karhunen-Loève expansion, chi-squared variates.

1. INTRODUCTION

In this article, we characterize wing behavior of the implied volatility for uncorrelated Gaussian stochastic volatility models. This introduction contains a careful description of the problem's background and of our motivations. Before going into details, we summarize some of the article's specificities; all terminology in the next two paragraphs is referenced, defined, and/or illustrated in the remainder of this introduction.

We hold calibration of volatility smiles as a principal motivator. Cognizant of the fact that non-centered Gaussian volatility models can be designed in a flexible and parsimonious fashion, we adopt that class of models, imposing no further conditions on the marginal distribution of the volatility process itself, beyond pathwise continuity. The spectral structure of the integrated variance allows us to work at that level of generality. We find that the first five terms in the extreme-strike implied volatility asymptotics – which is typically amply sufficient in applications – can be determined explicitly thanks to three parameters characterizing the top of the spectral decomposition of the integrated variance, with the exception of a factor appearing in the coefficient of the 4th term in these asymptotics, which depends on higher-order eigen-elements. In order to prove such a

Date: Jan 28, 2017.

precise statement while relying on a moderate amount of technicalities, we make use of the simplifying assumption that the stochastic volatility is independent of the asset price's driving noise.

When considering the trade-off between this restriction and calibration considerations, we observe that our model flexibility combined with known explicit spectral expansions and numerical tools may allow practitioners to compute the said spectral parameters in a straightforward fashion based on smile features, while also allowing them to select their favorite Gaussian volatility model class. Specific examples of Gaussian volatility processes are non-centered Brownian motion, Brownian bridge, and Ornstein-Uhlenbeck processes. This last sub-class can be particularly appealing since it contains stationary volatilities, and includes the well-known Stein-Stein model. We also mention how any Gaussian model specification, including long-memory ones, can be handled, thanks to the numerical ability to determine its spectral elements. We understand that the assumption of the stochastic volatility model being uncorrelated implies the symmetry of the implied volatility on either side of the money, which in some applications, is not a desirable feature. Moreover, while in many option markets, liquidity considerations limit the ability to calibrate using the large-strike wing (see the calibration study on SPX options in [21, Section 5.4]), the ability to work with a correlated volatility model is nonetheless important as soon as one uses the result of the calibration to, say, price illiquid options such as out-of-the-money calls. Hence a fully functional general Gaussian model would require a method for estimating the volatility's correlation with the asset using liquid options data. Such a study is beyond the scope of our article, since the case of general correlated Gaussian stochastic volatility models presents additional mathematical challenges which may require completely new methods and techniques. We will investigate them separately from this article. An important step toward a better understanding of the asymptotic behavior of the implied volatility in some correlated stochastic volatility models is found in the articles [14, 15].

Another problem which is mathematically interesting and important in practice is the asymptotics for implied volatility in small or large time to maturity. The techniques developed in the present paper are used in the subsequent paper [26] to study the small-time asymptotics of densities, option pricing functions, and the implied volatility in Gaussian self-similar stochastic volatility models.

1.1. Background and heuristics. Studies in quantitative finance based on the Black-Scholes-Merton framework have shown awareness of the inadequacy of the constant volatility assumption, particularly after the crash of 1987, when practitioners began considering that extreme events were more likely than what a log-normal model will predict. Propositions to exploit this weakness in log-normal modeling systematically and quantitatively have grown ubiquitous to the point that implied volatility (IV), or the volatility level that market call option prices would imply if the Black-Scholes model were underlying, is now a *bona fide* and vigorous topic of investigation, both at the theoretical and practical level. The initial evidence against constant volatility simply came from observing that IV as a function of strike prices for liquid call options exhibited non-constance, typically illustrated as a convex curve, often with a minimum near the money as for index options, hence the term 'volatility smile'.

Asset price models where the volatility is a stochastic process are known as stochastic volatility models; the term ‘uncorrelated’ is added to refer to the submodel class in which the volatility process is independent of the noise driving the asset price. In a sense, the existence of the smile for any uncorrelated stochastic volatility model was first proved mathematically by Renault and Touzi in [32]. They established that the IV as a function of the strike price decreases on the interval where the call is in the money, increases on the interval where the call is out of the money, and attains its minimum where the call is at the money. Note that Renault and Touzi did not prove that the IV is locally convex near the money, but their work still established stochastic volatility models as a main model class for studying IV; these models continued steadily to provide inspiration for IV studies.

A current emphasis, which has become fertile mathematical ground, is on IV asymptotics, such as large/small-strike, large-maturity, or small-time-to-maturity behaviors. These are helpful to understand and select models based on smile shapes. Several techniques are used to derive IV asymptotics. For instance, by exploiting a method of moments and the representation of power payoffs as mixtures of a continuum of calls with varying strikes, in a rather model-free context, R. Lee proved in [30] that, for models with positive moment explosions, the squared IV’s large strike behavior is of order the log-moneyness $\log\left(\frac{K}{s_0 e^{rT}}\right)$ times a constant which depends explicitly on supremum of the order of finite moments. A similar result holds for models with negative moment explosions, where the squared IV behaves like $K \mapsto \log\left(\frac{s_0 e^{rT}}{K}\right)$ for small values of K . More general formulas describing the asymptotic behavior of the IV in the ‘wings’ ($K \rightarrow 0$ or $+\infty$) were obtained in [4, 5, 6, 23, 24, 27, 18] (see also the book [22]).

From the standpoint of modeling, one of the advantages of Lee’s original result is the dependence of IV asymptotics merely on some simple statistics, namely as we mentioned, in the notation in [30], the maximal order \tilde{p} of finite moments for the underlying S_T , i.e.

$$\tilde{p}(T) := \sup \left\{ p \in \mathbb{R} : \mathbb{E} \left[(S_T)^{p+1} \right] < \infty \right\}.$$

This allows the author to draw appropriately strong conclusions about model calibration. A special class of models in which \tilde{p} is positive and finite is that of Gaussian volatility models, which we introduce next.

1.2. Gaussian Stochastic volatility models. Let W be a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let X be a continuous Gaussian process on the same space that is independent of W . We have $X(t) = m(t) + \tilde{X}(t)$, where m is a continuous deterministic function on $[0, T]$ (the mean function) and \tilde{X} is a continuous centered Gaussian process on $[0, T]$ independent of W , with covariance Q . Suppose $\{\mathcal{F}_t\}$ is a filtration such that W is a Brownian motion with respect to $\{\mathcal{F}_t\}$, and the process X is adapted to $\{\mathcal{F}_t\}$.

In the present paper, we study the following asset price model:

$$dS_t = rS_t dt + |X_t| S_t dW_t : t \in [0, T] \tag{1}$$

on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$, where the filtration $\{\mathcal{F}_t\}$ is such as above. It is also assumed that the short rate r is constant. The initial condition for the

asset price process will be denoted by s_0 . Note that the initial condition X_0 for the process X may be a nonconstant random variable.

We will next provide a typical example of a filtration $\{\mathcal{F}_t\}$ satisfying the conditions mentioned above. Let \mathcal{N} be the σ -algebra generated by the events of probability zero, and let $\{\mathcal{F}_t^W\}$ and $\{\mathcal{F}_t^X\}$ be the augmentations by the family \mathcal{N} of the filtrations generated by the processes W and X , respectively. Consider the filtration $\{\mathcal{F}_t\}$ such that for every $t \geq 0$, $\mathcal{F}_t = \sigma(\mathcal{F}_t^W, \mathcal{F}_t^X)$. Then the process W is a Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$, and the process X is adapted to $\{\mathcal{F}_t\}$. Note that if $X_0 = \text{const}$ a.s., then \mathcal{F}_0 is a sub- σ -algebra of \mathcal{N} , while if X_0 is a random variable, then $\mathcal{F}_0 = \sigma(X_0; \mathcal{N})$.

Note that it is not supposed in (1) that the process X is a solution to a stochastic differential equation as is often assumed in classical stochastic volatility models. A well-known special example of a Gaussian stochastic volatility model is the Stein-Stein model introduced in [36], in which the volatility process X is the mean-reverting Ornstein-Uhlenbeck process satisfying

$$dX_t = \alpha (m - X_t) dt + \beta dZ_t \quad (2)$$

where m is the level of mean reversion, α is the mean-reversion rate, and β is level of uncertainty on the volatility; here Z is another Brownian motion, which may be correlated with W . In the present paper, we adopt an analytic technique, encountered for instance in the analysis of the uncorrelated Stein-Stein model by this paper's first author and E.M. Stein in [25] (see also [22]).

Returning to the question of the value of \tilde{p} , for a Gaussian volatility model, it can sometimes be determined by simple calculations, which we illustrate here with an elementary example. Assume S is a geometric Brownian motion with random volatility, i.e. a model as in (1) where (abusing notation) $|X_t|$ is taken the non-time-dependent $\sigma|X|$ where σ is a constant and X is an independent unit-variance normal variate (not dependent on t). Thus, at time T , with zero discount rate, $S_T = s_0 \exp(\sigma|X|W_T - \sigma^2 X^2 T/2)$. To simplify this example to the maximum, also assume that X is centered; using the independence of X and W , we get that we may replace $|X|$ by X in this example, since this does not change the law of S_T (i.e. in the uncorrelated case, X 's non-positivity does not violate standard practice for volatility modeling). Then, using maturity $T = 1$, for any $p > 0$, the p th moment, via a simple change of variable, equals

$$\mathbb{E}[(S_1)^p] = \frac{s_0^p}{2\pi\sqrt{1+p\sigma^2}} \iint_{\mathbb{R}^2} dy dw \exp\left(-\frac{1}{2}\left(y^2 + w^2 - 2\frac{p\sigma}{\sqrt{1+p\sigma^2}}wy\right)\right)$$

which by an elementary computation is finite, and equal to $s_0^p / \sqrt{1+p\sigma^2 - p^2\sigma^2}$, if and only if

$$p < \tilde{p} + 1 = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\sigma^2}}.$$

In the cases where the random volatility model X above is non-centered and is correlated with W , a similar calculation can be performed, at the essentially trivial expenses of invoking affine changes of variables, and the linear regression of one normal variate against another.

The above example illustrates heuristically that, by Lee's moment formula, the computation of \tilde{p} might be the quickest path to obtain the leading term in the large-strike

expansion of the IV, for more complex Gaussian volatility models, namely ones where the volatility X is time-dependent. However, computing \tilde{p} is not necessarily an easy task, and appears, perhaps surprisingly, to have been performed rarely. For the Stein-Stein model, the value of \tilde{p} can be computed using the sharp asymptotic formulas for the asset price density near zero and infinity, established in [25] for the uncorrelated Stein-Stein model, and in [15] for the correlated one. These two papers also provide asymptotic formulas with error estimates for the IV at extreme strikes in the Stein-Stein model. Beyond the Stein-Stein model, little was known about the extreme strike asymptotics of general Gaussian stochastic volatility models. In the present paper, we extend the above-mentioned results from [25] and [15] to such models.

1.3. Motivation and summary of main result. Adopting the perspective that an asymptotic expansion for the IV can be helpful for model selection and calibration, our objective is to provide an expansion for the IV in a Gaussian volatility model relying on a minimal number of parameters, which can then be chosen to adjust to observed smiles. The restriction of non-correlated volatility means that the asset price distribution is a mixture of geometric Brownian motions with time-dependent volatilities, whose mixing density at time T is that of the square root of a variable in the second-chaos of a Wiener process. That second-chaos variable is none other than the integrated variance $\Gamma_T := \int_0^T X_s^2 ds$. By relying on a general Hilbert-space structure theorem which applies to the second Wiener chaos, we prove that, for a wide class of non-centered Gaussian stochastic volatility processes with a possible degeneracy in the eigenstructure of the covariance Q of X viewed as a linear operator on $L^2([0, T])$ (i.e. when the top eigenvalue λ_1 is allowed to have a multiplicity n_1 larger than 1), the large-strike IV asymptotics can be expressed with three terms and an error estimate. These terms depend explicitly on T and on the following three parameters: λ_1 , n_1 , and the ratio $\delta = \|P_{E_1} m\|^2 / \lambda_1$, where $\|P_{E_1} m\|$ is the norm in $L^2([0, T])$ of the orthogonal projection of the mean function m on the first eigenspace of Q . We also push the expansion to five terms, and notice that the fifth term also only depends on λ_1 , n_1 , and δ , while the fourth term depends on all other eigenvalues and the action of m on all other eigenfunctions. Specifically, with $I(K)$ the IV as a function of strike K , letting $k := \log(K/s_0) - rT$ be the discounted log-moneyness, as $k \rightarrow +\infty$, we prove

$$\begin{aligned} I(K) = & M_1(T, \lambda_1)\sqrt{k} + M_2(T, \lambda_1, \delta) + M_3(T, \lambda_1, n_1)\frac{\log k}{\sqrt{k}} \\ & + M_4(T, \lambda_1, n_1, V)\frac{1}{\sqrt{k}} + M_5(T, \lambda_1, n_1, \delta)\frac{\log k}{k} + O\left(\frac{1}{\sqrt{k}}\right), \end{aligned} \quad (3)$$

where the constants M_1, M_2, M_3, M_4 , and M_5 depend explicitly on T and λ_1 , M_2 also depends explicitly on δ , while M_3 also depends explicitly on n_1 , M_5 depends explicitly also on both n_1 and δ , and M_4 has an additional rather complex dependence on all the eigen-elements through a factor V ; this is all stated in Theorem 13 and formula (47). A similar asymptotic formula is obtained in the case where $k \rightarrow -\infty$, using symmetry properties of uncorrelated stochastic volatility models (see (55)). The specific case of the Stein-Stein model is expanded upon in some detail.

1.4. Practical implications. The first-order constant M_1 is always strictly positive. The second-order term (the constant M_2) vanishes if and only if m is orthogonal to the first eigenspace of Q , which occurs for instance when $m \equiv 0$. The third-order and fifth-order terms vanish if and only if the top eigenvalue has multiplicity $n_1 = 1$, which is typical (the case $n_1 > 1$ can be considered degenerate, and does not occur in common examples). The behavior of M_1 and M_2 as functions of T is determined partly by how the top eigenvalue λ_1 depends on T , which can be non-trivial. In the present paper, we assume T is fixed.

For fixed maturity T , assuming that Q has lead multiplicity $n_1 = 1$ for instance, a practitioner will have the possibility of determining a value λ_1 and a value δ to match the specific root-log-moneyness behavior of small- or large-strike IV; moreover in that case, choosing a constant mean function m , one obtains $\delta = m^2 \lambda_1^{-1} \left| \int_0^T e_1(t) dt \right|^2$ where e_1 is the top eigenfunction of Q . Market prices may not be sufficiently liquid at extreme strikes to distinguish between more than two parameters; this is typical of calibration techniques for implied volatility curves for fixed maturity, such as the ‘stochastic volatility inspired’ (SVI) parametrization disseminated by J. Gatheral: see [19, 20] (see also [21] and the references therein). Our result shows that Gaussian volatility models with non-zero mean are sufficient for this flexibility, and provide equivalent asymptotics irrespective of the precise mean function and covariance eigenstructure, since modulo the disappearance of the third-order term in the unit top multiplicity case $n_1 = 1$, only λ_1 and δ are relevant. The fourth-order term in our expansion can provide additional precision in calibration. Its use is illustrated in Section 7.

Modelers wishing to stick to well-known classes of processes for X may then adjust the value of λ_1 by exploiting any available invariance properties for the desired class. For example, if X is standard Brownian motion, or the Brownian bridge, on $[0, T]$, we have $\lambda_1 = 4T^2/\pi$ or $\lambda_1 = T^2/\pi$ respectively, and these values scale quadratically with respect to a multiplicative scaling constant for X , beyond which an arbitrary mean value m may be chosen. If X is the mean-zero stationary OU process, we have $\lambda_1 = \beta^2 / (\omega_T + \alpha^2)$ where ω_T is the smallest positive solution of $2\alpha\omega \cos(\omega T) + (\alpha^2 - \omega) \sin(\omega T) = 0$, in which case, for a fixed arbitrarily selected rate of mean reversion α , a scaling of λ_1 is then equivalent to selecting the variance of X , while a constant mean value m can then be selected independently. [10, Chapter 1] can be consulted for the eigenstructure of the covariance of Brownian motion and the Brownian bridge, which are classical results, and for a proof of the eigenstructure of the OU covariance (see also [12]). The top eigenfunctions in all three of these cases are known explicit trigonometric functions (see [10, Chapter 1]), and need to be referenced when selecting m . For the OU bridge, the eigenstructure of Q (equivalently known as the Karhunen-Loève expansion of Q) was found in [11], while in [13], such an expansion was characterized for special Gaussian processes generated by independent pairs of exponential random variables. On the other hand, fractional Brownian motion and OU processes driven by fractional Brownian motion (also known as fOU processes) do not fall in the class of Gaussian processes for which the Karhunen-Loève expansion is known explicitly.

However, efficient numerical techniques allowing to compute the eigenfunctions and eigenvalues in these cases were developed by S. Corlay (see Chapter 2 in [10]). Corlay uses the trapezoidal Nyström method and the three-step Richardson-Romberg method to

approximate the five highest Karhunen-Loève eigenvalues of various Gaussian processes; in principle, eigenvalues and eigenfunctions of arbitrarily high order can be obtained using his method. He starts with such estimates for Brownian motion, Brownian bridge, and Ornstein-Uhlenbeck process, for which explicit expressions for the eigenvalues are known. The resulting approximations are very close to the values obtained from the explicit formulas for the eigenvalues, which shows that the method used by Corlay is rather powerful. Corlay also estimates the five highest Karhunen-Loève eigenvalues of fractional Brownian motion on $[0, 1]$ with the Hurst exponent $H = 0.7$. Of special interest to the context of the present paper is the largest Karhunen-Loève eigenvalue λ_1 of fractional Brownian motion, for which Corlay obtains the approximation $\lambda_1 \approx 0.374532521757236$.

While we do not need this value, and instead use Corlay's method to compute λ_1 for several fOU processes, we are confident that the values we obtain for the various λ_1 's we use have similar levels of accuracy to what is illustrated in [10]. Corlay's method is thus one of the main ingredients in the numerical part of our paper (see the discussion after (80) in Section 7). Fractional OU processes were proposed early on for option pricing, and recently analyzed in [9, 8]; these processes are versions of the volatility process in the Stein-Stein model. Therefore, the resulting stochastic volatility models may be called fractional Stein-Stein models. Section 7 illustrates how, in the case of the classical and fractional Stein-Stein models (OU and fOU processes), the explicit, semi-explicit, or numerically accessible Karhunen-Loève expansion of X can be used in conjunction with the asymptotics (3) for calibrating parameters. We find that market liquidity considerations limit the theoretical range of applicability of calibration strategies, but that significant practical results are nonetheless available.

The remainder of this article is structured as follows. Section 2 sets up a convenient second-chaos representation for the model's integrated volatility. In Section 3, we generalize some results from [7, 28, 38], concerning the asymptotic behavior of densities of infinite linear combinations of chi-squared random variables, and derive precise asymptotics for the density of the mixing distribution. Section 4 converts these asymptotics into sharp asymptotic formulas for the density of the asset price S_T , thanks to the analytic tools developed in [25, 22]. In Section 5, we characterize the wing behavior of the implied volatility in Gaussian stochastic volatility models. We find sharp asymptotic formulas for the implied volatility with five explicit terms and an error estimate. The special case of the uncorrelated Stein-Stein model is studied in more detail in Section 6. Finally, our practical study of calibration strategies, with numerics, is in Section 7.

2. GENERAL SETUP AND SECOND-CHAOS EXPANSION OF THE INTEGRATED VARIANCE

Let X be an almost-surely continuous Gaussian process on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ with mean and covariance functions denoted by $m(t) = \mathbb{E}[X_t]$ and

$$Q(t, s) = \text{cov}(X_t, X_s) = \mathbb{E}[(X_t - m(t))(X_s - m(s))],$$

respectively, and suppose the restrictions imposed in (1) are satisfied.

Define the centered version of X : $\tilde{X}_t := X_t - m(t)$, $t \geq 0$, and fix a time horizon $T > 0$. It is not hard to see that $Q(s, s) > 0$ for all $s > 0$. Since the Gaussian process X is almost surely continuous, the mean function $t \mapsto m(t)$ is a continuous function on $[0, T]$,

and the covariance function $(t, s) \mapsto Q(t, s)$ is a continuous function of two variables on $[0, T]^2$. Indeed, the continuity of the process X implies its continuity in probability on Ω . Hence, the process X is continuous in the mean-square sense (see, e.g., [29], Lemma 1 on p. 5, or invoke the equivalence of L^p norms on Wiener chaos, see [31]). Mean-square continuity of X implies the continuity of the mean function on $[0, T]$. In addition, the autocorrelation function of the process X , that is, the function $R(t, s) = \mathbb{E}[X_t X_s]$, $(t, s) \in [0, T]^2$, is continuous (see, e.g., [3], Lemma 4.2). Finally, since $Q(t, s) = R(t, s) - m(t)m(s)$, the covariance function Q is continuous on $[0, T]^2$. We refer the interested reader to [2] for more information on the continuity problems for general Gaussian processes.

In our analysis, it will be convenient to refer to the Karhunen-Loève expansion of \tilde{X} . We will next provide certain details concerning the Karhunen-Loève expansion and introduce notation that will be used throughout the paper.

Consider the covariance operator defined by

$$\mathcal{K}(f)(t) = \int_0^T f(s)Q(t, s)ds, \quad f \in L^2([0, T]), \quad 0 \leq t \leq T.$$

The operator \mathcal{K} is a nonnegative compact self-adjoint operator on $L^2([0, T])$. The non-zero eigenvalues of the operator \mathcal{K} are of finite multiplicity, and we assume that they are rearranged so that

$$\lambda_1 = \lambda_2 = \dots = \lambda_{n_1} > \lambda_{n_1+1} = \lambda_{n_1+2} = \dots = \lambda_{n_1+n_2} > \dots$$

In particular, λ_1 is the top eigenvalue, and n_1 is its multiplicity. It is known that the series $\sum_{n=1}^{\infty} \lambda_n$ converges. The system of eigenfunctions $E = \{e_n\}_{n \geq 1}$, corresponding to the system $\{\lambda_n\}_{n \geq 1}$, is orthonormal, and each function e_n is continuous on $[0, T]$. The number $\lambda_0 = 0$ always belongs to the spectrum of the covariance operator, and it may happen so that λ_0 is an eigenvalue of \mathcal{K} . The spectral subspace associated with λ_0 may be infinite-dimensional, and we choose a basis \tilde{E} in this subspace. Then (E, \tilde{E}) is a complete orthonormal system in $L^2([0, T])$. Note that the eigenvalues and eigenfunctions of \mathcal{K} depend on T .

The classical Karhunen-Loève theorem (see, e.g., [37], Section 26.1) states that there exists an i.i.d. sequence of standard normal variates $\{Z_n : n = 1, 2, \dots\}$ such that

$$\tilde{X}_t = \sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n. \quad (4)$$

Remark 1. *The number of positive eigenvalues may be finite. We will assume throughout the paper that the set of positive eigenvalues is infinite; this is the case for all illustrative examples we use, such as the OU and fOU processes. It is easy to understand how the parameters used in the paper change if the number of positive eigenvalues is finite.*

Using (4), we obtain

$$\int_0^T \tilde{X}_t^2 dt = \int_0^T \left(\sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n \right)^2 dt = \sum_{n=1}^{\infty} \lambda_n Z_n^2. \quad (5)$$

It is worth pointing out that the previous expression for the integrated variance in a Gaussian model with centered volatility is in fact the most general form of a random

variable in the second Wiener chaos with half-bounded support, with mean adjusted to ensure almost-sure positivity of the integrated variance. This is established using a classical structure theorem on separable Hilbert spaces, as explained in [31, Section 2.7.4]. In other words (also see [31, Section 2.7.3] for additional details), any prescribed mean-adjusted integrated variance in the second chaos is of the form

$$V(T) := \iint_{[0,T]^2} G(s,t) dZ(s) dZ(t) + 2 \|G\|_{L^2([0,T]^2)}^2$$

for some standard Wiener process Z and some function $G \in L^2([0,T]^2)$. Moreover one can find a centered Gaussian process \tilde{X} such that $V(T) = \int_0^T \tilde{X}_t^2 dt$ and one can compute the coefficients λ_n in the Karhunen-Loève representation (5) as the eigenvalues of the covariance of \tilde{X} .

Let us set

$$s = \int_0^T m(t)^2 dt \quad \text{and} \quad \delta_n = \int_0^T m(t) e_n(t) dt, \quad n \geq 1. \quad (6)$$

Then, it follows from (5) and (6) that, for the non-centered process X ,

$$\begin{aligned} \int_0^T X_t^2 dt &= \sum_{n=1}^{\infty} \lambda_n Z_n^2 + 2 \sum_{n=1}^{\infty} \sqrt{\lambda_n} \delta_n Z_n + s \\ &= \sum_{n=1}^{\infty} \lambda_n \left[Z_n + \frac{\delta_n}{\sqrt{\lambda_n}} \right]^2 + \left(s - \sum_{n=1}^{\infty} \delta_n^2 \right). \end{aligned} \quad (7)$$

Remark 2. *It is easy to see, using (7) that if the function $t \mapsto m(t)$ belongs to the subspace of $L^2[0, T]$ generated by the orthonormal system E , then*

$$\int_0^T X_t^2 dt = \sum_{n=1}^{\infty} \lambda_n \left[Z_n + \frac{\delta_n}{\sqrt{\lambda_n}} \right]^2. \quad (8)$$

For instance, the equality in (8) holds if $\lambda = 0$ is not an eigenvalue of the operator \mathcal{K} . In the case where the process X is centered, we have

$$\int_0^T X_t^2 dt = \sum_{n=1}^{\infty} \lambda_n Z_n^2. \quad (9)$$

Note that the right-hand sides of (8) and (9) are infinite linear combinations of chi-square random variables.

Let us denote the chi-squared distribution with the number of degrees of freedom k and the parameter of noncentrality λ by $\chi^2(k, \lambda)$ (more information on such distributions can be found in [22] or in any probability textbook; the convention used here is that the mean of $\chi^2(k, \lambda)$ is $k + \lambda$). Set

$$\Lambda_T = \frac{1}{\lambda_1} \left(\int_0^T X_t^2 dt - s + \sum_{n=1}^{\infty} \delta_n^2 \right) \quad (10)$$

and denote

$$\xi_0 = \sum_{n=1}^{n_1} \delta_{n'}^2, \quad \xi_k = \sum_{n=n_1+\dots+n_{k+1}}^{n_1+\dots+n_{k+1}} \delta_n^2, \quad k \geq 1. \quad (11)$$

Denote also

$$\delta = \frac{\tilde{\xi}_0}{\lambda_1}. \quad (12)$$

Then, it is not hard to see, using (7), (10), (11), and (12), that

$$\Lambda_T = \chi^2(n_1, \delta) + \sum_{k=2}^{\infty} \frac{\lambda_{n_1+\dots+n_{k-1}+1}}{\lambda_1} \chi^2\left(n_k, \frac{1}{\lambda_{n_1+\dots+n_{k-1}+1}} \tilde{\xi}_{k-1}\right), \quad (13)$$

where the repeated chi-squared notation is used abusively to denote independent chi-square random variables. We will denote the distribution density of Λ_T by q_T .

3. ASYMPTOTICS OF THE MIXING DENSITY

The asymptotic behavior of the distribution density of an infinite linear combination of independent central chi-squared random variables was characterized by Zolotarev (see formula (5) in [38]). In [28], Hoeffding found more general and sharp formulas. The results obtained by Zolotarev and Hoeffding were generalized to the case of noncentral chi-squared variables by Beran (see [7]). Note that Beran considered infinite sums of chi-squared variables with all the noncentrality parameters strictly greater than zero. Since there is a gap between the results of Zolotarev, Hoeffding, and Beran, we decided to include a discussion of a similar result, where there are no restrictions on the noncentrality parameters. Keeping in mind the series in (13), we will study the asymptotic behavior of the density q of the following infinite sum:

$$\Lambda = \chi^2(n_1, \eta_1) + \sum_{k=2}^{\infty} \rho_k \chi^2(n_k, \eta_k), \quad (14)$$

where $n_k \geq 1$, $k \geq 1$, are integers, and $\eta_k \geq 0$ for all $k \geq 2$. If $\eta_k = 0$ for some k , then the corresponding chi-squared random variable is central. It is also assumed that $1 > \rho_2 > \rho_3 > \dots > 0$,

$$\sum_{k=2}^{\infty} n_k \rho_k < \infty, \quad \sum_{k=2}^{\infty} \eta_k \rho_k < \infty, \quad (15)$$

and the chi-squared random variables in (14) are independent. We will denote by q_Λ the distribution density of the random variable Λ .

The distribution density of a chi-squared random variable $\chi^2(n, \eta)$ will be denoted by $p_{\chi^2}(\cdot; n, \eta)$. It is known that if $\eta > 0$, then

$$p_{\chi^2}(x; n, \eta) = \frac{1}{2} \left(\frac{x}{\eta}\right)^{\frac{n}{4}-\frac{1}{2}} e^{-\frac{x+\eta}{2}} I_{\frac{n}{2}-1}(\sqrt{\eta x}), \quad x > 0, \quad (16)$$

where I_ν is the modified Bessel function of the first kind (see, e.g., [22], Theorem 1.31). For $\eta = 0$, we have

$$p_{\chi^2}(x; n, 0) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n-2}{2}} \exp\left\{-\frac{x}{2}\right\}, \quad x > 0 \quad (17)$$

(see, e.g., Lemma 1.27 in [22]). It is not hard to see that $\lim_{\eta \rightarrow 0} p_{\chi^2}(x; n, \eta) = p_{\chi^2}(x; n, 0)$. Let us also mention that

$$I_\nu(t) = \frac{e^t}{\sqrt{2\pi t}} \left(1 + O\left(t^{-1}\right)\right), \quad t \rightarrow \infty, \quad (18)$$

for all $\nu \geq 0$ (see, e.g., 9.6.7 in [1]).

It is known that for $t < \frac{1}{2}$, the moment generating function of a chi-squared random variable $\chi^2(n, \eta)$ with $\eta \geq 0$ is as follows:

$$t \mapsto \frac{1}{(1-2t)^{\frac{n}{2}}} \exp\left\{\frac{\eta t}{1-2t}\right\}. \quad (19)$$

In the formulation of the next result, we will use the following number:

$$A = \mathbb{E} \left[\exp \left\{ \frac{U}{2} \right\} \right],$$

where U is defined as Λ without the first term:

$$U = \sum_{k=2}^{\infty} \rho_k \chi^2(n_k, \eta_k). \quad (20)$$

Next, using (20) and (19), we obtain

$$A = \prod_{k \geq 2} (1 - \rho_k)^{-\frac{n_k}{2}} \exp \left\{ \frac{\eta_k \rho_k}{2(1 - \rho_k)} \right\}, \quad (21)$$

and it is not hard to see, by taking into account (15), that $A < \infty$.

The next assertion is based on the results of Zolotarev, Hoeffding, and Beran.

Theorem 3. *Suppose the conditions formulated after formula (14) hold. If $\eta_1 > 0$, then*

$$\left| \frac{q_\Lambda(x)}{p_{\chi^2}(x; n_1, \eta_1)} - A \right| = O\left(x^{-\frac{1}{2}}\right) \quad (22)$$

as $x \rightarrow \infty$, while if $\eta_1 = 0$, then

$$\left| \frac{q_\Lambda(x)}{p_{\chi^2}(x; n_1, 0)} - A \right| = O\left(x^{-1}\right) \quad (23)$$

as $x \rightarrow \infty$. In the formulas above, the constant A is given by (21).

Remark 4. *Theorem 3 is a minor generalization of similar propositions obtained in [28] and [7]. The difference between those propositions and our Theorem 3 is that [28] assumes that all the chi-squared variables in (14) are central, in Theorem 2 in [7] they are all assumed noncentral, while in our Theorem 3, we may have any combination of central and non-central chi-squared variables.*

Theorem 2 in [7] provides an asymptotic formula for the complementary distribution function (tail) of an infinite linear combination of independent noncentral chi-square random variables. A sharper formula for the distribution density of such a linear combination can be extracted from the proof of Theorem 2 in [7] (see the very end of that proof).

Sketch of the proof of Theorem 3. We follow the proof of Theorem 2 in [7]. Let us denote by p_U the distribution density of the random variable U in (20). Then

$$q_\Lambda(x) = \int_0^x p_{\chi^2}(x-y; n_1, \eta_1) p_U(y) dy, \quad x > 0. \quad (24)$$

Let us fix $0 < \alpha < 1$. We have

$$\frac{q_\Lambda(x)}{p_{\chi^2}(x; n_1, \eta_1)} - A = V_1 + V_2 + V_3 + V_4,$$

where

$$\begin{aligned} V_1 &= \int_0^{\alpha x} \left[\left(1 - \frac{y}{x}\right)^{\frac{n_1}{2}-1} - 1 \right] W(x, y) \exp\left\{\frac{y}{2}\right\} p_U(y) dy, \\ V_2 &= \int_{\alpha x}^x \left(1 - \frac{y}{x}\right)^{\frac{n_1}{2}-1} W(x, y) \exp\left\{\frac{y}{2}\right\} p_U(y) dy, \\ V_3 &= \int_0^{\alpha x} [W(x, y) - 1] \exp\left\{\frac{y}{2}\right\} p_U(y) dy, \\ V_4 &= - \int_{\alpha x}^{\infty} \exp\left\{\frac{y}{2}\right\} p_U(y) dy. \end{aligned}$$

In the formulas above, the function W is defined by

$$W(x, y) = \left(1 - \frac{y}{x}\right)^{-\frac{n_1}{4} + \frac{1}{2}} \frac{I_{\frac{n_1}{2}-1}(\sqrt{\eta(x-y)})}{I_{\frac{n_1}{2}-1}(\sqrt{\eta x})}$$

if $\eta_1 > 0$, while if $\eta_1 = 0$, then $W(x, y) = 1$. Note that $\eta_1 = 0$ implies $V_3 = 0$. Then, using calculations similar to those in the proof of Theorem 2 in [7], we find that when $\eta_1 > 0$, V_3 is the leading term and is of order $x^{-1/2}$, while when $\eta_1 = 0$, this term vanishes, and the next highest-order term is of order x^{-1} . This explains the different error estimates in the formulas in Theorem 3. We include two auxiliary statements below (Lemmas 5 and 6). They are needed to perform the above-mentioned calculations. This finishes the sketch of the proof of Theorem 3. \square

Lemma 5. *Under the assumptions in Theorem 3, the following holds:*

$$\mathbb{E} \left[U \exp\left\{\frac{U}{2}\right\} \right] < \infty.$$

Proof. This follows in a straightforward way (details omitted), using (20), differentiating the function in (19), and taking into account the resulting formula and (21), implying that:

$$\mathbb{E} \left[U \exp\left\{\frac{U}{2}\right\} \right] = A \sum_{k=2}^{\infty} \rho_k \left[\frac{n_k}{1 - \rho_k} + \frac{\eta_k}{(1 - \rho_k)^2} \right]$$

so that that Lemma 5 clearly follows from (15) and the finiteness $A < \infty$. \square

Lemma 6. *Under the restrictions in Theorem 3, there exists a number $\varepsilon > 0$, depending on the constants in (20), and such that*

$$p_U(y) = O\left(\exp\left\{-\left(\frac{1}{2} + \varepsilon\right)y\right\}\right)$$

as $y \rightarrow \infty$.

Proof. We have $U = \rho_2 \tilde{U}$, where $\tilde{U} = \sum_{k=2}^{\infty} \tilde{\rho}_k \chi^2(n_k, \eta_k)$ with $\tilde{\rho}_2 = 1$ and $\tilde{\rho}_k = \frac{\rho_k}{\rho_2}$ for all $k \geq 3$. It follows that $p_U(x) = \frac{1}{\rho_2} p_{\tilde{U}}\left(\frac{1}{\rho_2} y\right)$. Since $\rho_2 < 1$, and the random variable \tilde{U} has the same structure as the random variable Λ in (14), it suffices to show that for every $\tau > 0$,

$$q_{\Lambda}(x) = O\left(\exp\left\{\left(-\frac{1}{2} + \tau\right)y\right\}\right) \quad (25)$$

as $x \rightarrow \infty$.

Let us first assume $n_1 \geq 2$. Then, using (24), (16), the fact that the function I_ν is increasing, and (18), we obtain (25). Next, let $n_1 = 1$. We have

$$\Lambda \leq \chi^2(n_1, \eta_1) + \chi^2(n_2, \eta_2) + \sum_{k=3}^{\infty} \rho_k \chi^2(n_k, \eta_k).$$

Next, we observe that $\chi^2(n_1, \eta_1) + \chi^2(n_2, \eta_2) = \chi^2(n_1 + n_2, \eta_1 + \eta_2)$ (the previous formula follows from (19)). This reduces the case where $n_1 = 1$ to the already considered case where $n_1 > 1$. It follows from the previous reasoning that (25) holds. This completes the proof of Lemma 6. \square

Theorem 3 will allow us to characterize the asymptotic behavior of the distribution density q_T of the random variable Λ_T defined by (13). Using Theorem 3, we see that if $\delta > 0$, then

$$\left| \frac{q_T(x)}{p_{\chi^2}(x; n_1, \delta)} - A \right| = O\left(x^{-\frac{1}{2}}\right) \quad (26)$$

as $x \rightarrow \infty$, while if $\delta = 0$, then

$$\left| \frac{q_T(x)}{p_{\chi^2}(x; n_1, 0)} - A \right| = O\left(x^{-1}\right) \quad (27)$$

as $x \rightarrow \infty$. In (26) and (27), the formula for A is

$$A = \prod_{j>n_1} \left(\frac{\lambda_1}{\lambda_1 - \lambda_j} \right)^{\frac{1}{2}} \exp\left\{ \frac{\delta_j^2}{2(\lambda_1 - \lambda_j)} \right\}. \quad (28)$$

It is clear that for $\delta > 0$, (26) gives

$$q_T(x) = A p_{\chi^2}(x; n_1, \delta) \left(1 + O\left(x^{-\frac{1}{2}}\right)\right) \quad (29)$$

as $x \rightarrow \infty$. Similarly, for $\delta = 0$, (27) implies that

$$q_T(x) = A p_{\chi^2}(x; n_1, 0) \left(1 + O\left(x^{-1}\right)\right) \quad (30)$$

as $x \rightarrow \infty$.

It is known that

$$I_\nu(t) = \frac{e^t}{\sqrt{2\pi t}} \left(1 + O\left(t^{-1}\right)\right) \quad t \rightarrow \infty,$$

(see 9.7.1 in [1]). Next, using the previous formula in (16), we obtain

$$p_{\chi^2}(x; n, \lambda) = \frac{1}{2\sqrt{2\pi}} \lambda^{-\frac{n-1}{4}} x^{\frac{n-3}{4}} e^{\sqrt{\lambda x}} e^{-\frac{x+\lambda}{2}} \left(1 + O\left(x^{-\frac{1}{2}}\right)\right) \quad (31)$$

as $x \rightarrow \infty$.

Recall that we denoted by q_T the distribution density of the random variable Λ_T defined by (10). Using (29) and (31), we see that for $\delta > 0$,

$$q_T(x) = \frac{A}{2\sqrt{2\pi}} \delta^{-\frac{n_1-1}{4}} x^{\frac{n_1-3}{4}} e^{\sqrt{\delta x}} e^{-\frac{x+\delta}{2}} \left(1 + O\left(x^{-\frac{1}{2}}\right)\right) \quad (32)$$

as $x \rightarrow \infty$. The constants A and δ in (32) are defined by (28) and (12), respectively.

We next turn our attention to the case where $\delta = 0$. In this case, it follows from (30), (28), and (17) that

$$\begin{aligned} q_T(x) &= \frac{1}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} \prod_{k>n_1} \left(\frac{\lambda_1}{\lambda_1 - \lambda_k}\right)^{\frac{1}{2}} \exp\left\{\frac{\delta_k^2}{2(\lambda_1 - \lambda_k)}\right\} x^{\frac{n_1-2}{2}} \exp\left\{-\frac{x}{2}\right\} \\ &\quad \times \left(1 + O\left(x^{-1}\right)\right) \end{aligned} \quad (33)$$

as $x \rightarrow \infty$.

Remark 7. In comparing (32) and (33), one notes that the latter cannot be obtained from the former by letting δ tend to 0: while the exponential terms would match, the power terms do not, and an additional discrepancy would occur when $n_1 > 1$ from the singular term $\delta^{-(n_1-1)/4}$.

Our next goal is to characterize the asymptotic behavior of the distribution density p_T of the integrated variance $\Gamma_T = \int_0^T X_t^2 dt$. The following statement holds.

Theorem 8. (i) If $\delta > 0$, then

$$p_T(x) = C x^{\frac{n_1-3}{4}} \exp\left\{\sqrt{\frac{\delta}{\lambda_1}} \sqrt{x}\right\} \exp\left\{-\frac{x}{2\lambda_1}\right\} \left(1 + O\left(x^{-\frac{1}{2}}\right)\right) \quad (34)$$

as $x \rightarrow \infty$, where

$$\begin{aligned} C &= \frac{1}{2\sqrt{2\pi}} \lambda_1^{-\frac{n_1+1}{4}} \delta^{-\frac{n_1-1}{4}} \exp\left\{\frac{s - \sum_{n=1}^{\infty} \delta_n^2}{2\lambda_1} - \frac{\delta}{2}\right\} \\ &\quad \times \prod_{j>n_1}^{\infty} \left(\frac{\lambda_1}{\lambda_1 - \lambda_j}\right)^{\frac{1}{2}} \exp\left\{\frac{\delta_j^2}{2(\lambda_1 - \lambda_j)}\right\}. \end{aligned} \quad (35)$$

(ii) If $\delta = 0$, then

$$p_T(x) = C x^{\frac{n_1-2}{2}} \exp\left\{-\frac{x}{2\lambda_1}\right\} \left(1 + O\left(x^{-1}\right)\right) \quad (36)$$

as $x \rightarrow \infty$, where

$$C = \frac{1}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} \lambda_1^{-\frac{n_1}{2}} \exp\left\{\frac{s - \sum_{n>n_1} \delta_n^2}{2\lambda_1}\right\} \prod_{k>n_1} \left(\frac{\lambda_1}{\lambda_1 - \lambda_k}\right)^{\frac{1}{2}} \exp\left\{\frac{\delta_j^2}{2(\lambda_1 - \lambda_j)}\right\}. \quad (37)$$

In particular, if the process X is centered, then (36) holds with

$$C = \frac{1}{2^{\frac{n_1}{2}} \Gamma\left(\frac{n_1}{2}\right)} \lambda_1^{-\frac{n_1}{2}} \prod_{k>n_1} \left(\frac{\lambda_1}{\lambda_1 - \lambda_k}\right)^{\frac{1}{2}}. \quad (38)$$

Proof. It follows from (10) that $p_T(x) = \frac{1}{\lambda_1} q_T\left(\frac{1}{\lambda_1}(x - \tau)\right)$, where $\tau = s - \sum_{n=1}^{\infty} \delta_n^2$. Now, formula (32) implies that

$$\begin{aligned} p_T(x) &= \frac{A}{2\sqrt{2\pi}} \frac{1}{\lambda_1} \lambda_1^{\frac{n_1-1}{4}} \delta^{-\frac{n_1-1}{4}} \\ &\quad \times \left(\sum_{n=1}^{n_1} \delta_n^2\right)^{-\frac{n_1-1}{4}} \lambda_1^{-\frac{n_1-3}{4}} \exp\left\{\frac{\tau - \sum_{n=1}^{n_1} \delta_n^2}{2\lambda_1}\right\} \\ &\quad \times (x - \tau)^{\frac{n_1-3}{4}} \exp\left\{\sqrt{\frac{\delta(x - \tau)}{\lambda_1}}\right\} \exp\left\{-\frac{x}{2\lambda_1}\right\} \\ &\quad \times \left(1 + O\left(x^{-\frac{1}{2}}\right)\right) \end{aligned} \quad (39)$$

as $x \rightarrow \infty$.

Next, taking into account that

$$(x - \tau)^{\frac{n_1-3}{4}} = x^{\frac{n_1-3}{4}} (1 + O(x^{-1}))$$

and

$$\exp\left\{\sqrt{\frac{\delta(x - \tau)}{\lambda_1}}\right\} = \exp\left\{\sqrt{\frac{\delta}{\lambda_1}} \sqrt{x}\right\} (1 + O(x^{-\frac{1}{2}})),$$

and simplifying the expression on the right-hand side of (39), we obtain (34). The proof of formula (36) is similar, using (33). \square

4. ASSET PRICE ASYMPTOTICS

The model in (1) is described by a linear stochastic differential equation. Therefore, we have

$$S_t = s_0 \exp\left\{rt - \frac{1}{2} \int_0^t X_s^2 ds + \int_0^t |X_s| dW_s\right\}. \quad (40)$$

The previous equality can be derived from the Doléans-Dade formula (see [33]). Since the processes X and W are independent, the following formula holds for the distribution

density D_t of the asset price S_t :

$$D_t(x) = \frac{\sqrt{s_0 e^{rt}}}{\sqrt{2\pi t}} x^{-\frac{3}{2}} \int_0^\infty y^{-1} \exp \left\{ - \left[\frac{\log^2 \frac{x}{s_0 e^{rt}}}{2ty^2} + \frac{ty^2}{8} \right] \right\} \tilde{p}_t(y) dy. \quad (41)$$

In (41), \tilde{p}_t is the distribution density of the random variable $\tilde{Y}_t = \left\{ \frac{1}{t} \int_0^t X_s^2 ds \right\}^{\frac{1}{2}}$. The function \tilde{p}_t is called the mixing density. The proof of formula (41) can be found in [22] (see (3.5) in [22]). It is not hard to see that $\tilde{p}_t(y) = 2ty p_t(ty^2)$, where the symbol p_t stands for the density of the realized volatility $Y_t = \int_0^t X_s^2 ds$.

Suppose first that the volatility process is such that $\delta > 0$. It follows from formula (34) that

$$\tilde{p}_t(y) = \tilde{A} y^{\frac{n_1-1}{2}} \exp \{ \tilde{B} y \} \exp \{ -\tilde{C} y^2 \} \left(1 + O(y^{-1}) \right) \quad (42)$$

as $y \rightarrow \infty$, where

$$\tilde{A} = 2Ct^{\frac{n_1+1}{4}}, \quad \tilde{B} = \sqrt{\frac{\delta t}{\lambda_1}}, \quad \tilde{C} = \frac{t}{2\lambda_1}. \quad (43)$$

In (43), the constant C is defined by (35).

Our next goal is to estimate the function D_t . The asymptotic behavior as $x \rightarrow \infty$ of the integral appearing in (41) was studied in [25] (see also Section 5.3 in [22]). It is explained in [22] how to get an asymptotic formula for the integral in (41) in the case where an asymptotic formula for the mixing density is similar to formula (42). We refer the reader to the derivation of Theorem 6.1 in [22], which is based on formula (5.133) in Section 5.6 of [22] and Theorem 5.5 in [22]. The latter theorem concerns the asymptotic behavior of integrals with lognormal kernels. Having obtained an asymptotic formula for the distribution density of the asset price, we can find a similar asymptotic formula for the call pricing function C at large strikes, and then obtain an asymptotic formula for the implied volatility I (see Section 10.5 in [22]).

Theorem 5.5 in [22] provides an asymptotic formula as $w \rightarrow \infty$ for the integral

$$\int_0^\infty A(y) \exp \left\{ - \left(\frac{w^2}{y^2} + k^2 y^2 \right) \right\} dy,$$

where $k > 0$ is fixed, and it is assumed that

$$A(y) = e^{ly} \zeta(y) (1 + O(b(y)))$$

as $y \rightarrow \infty$. In the previous asymptotic formula, l is a real number, and ζ and b are functions satisfying certain conditions.

Let us fix $T > 0$. Our goal is to use Theorem 5.5 in [22] with

$$A(y) = y^{-1} \tilde{p}_T(y) \exp \{ \tilde{C} y^2 \},$$

$l = \tilde{B}$, $\zeta(y) = \tilde{A}y^{\frac{n_1-3}{2}}$, $b(y) = y^{-1}$, $w = (2T)^{-\frac{1}{2}} \log \frac{x}{s_0 e^{rT}}$, $k = \frac{\sqrt{8\tilde{C}+T}}{2\sqrt{2}}$, and $\gamma = 1$ (see the formulation of Theorem 5.5 in [22] for the meaning of the constant γ). This gives

$$\begin{aligned} \int_0^\infty y^{-1} \exp \left\{ - \left[\frac{\log^2 \frac{x}{s_0 e^{rT}}}{2ty^2} + \frac{Ty^2}{8} \right] \right\} \tilde{p}_T(y) dy &= \frac{\tilde{A} 2^{\frac{n_1-1}{4}} \sqrt{\pi}}{T^{\frac{n_1-3}{8}} (8\tilde{C}+T)^{\frac{n_1+1}{8}}} \left(s_0 e^{rT} \right)^{\frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}}} \\ \exp \left\{ \frac{\tilde{B}^2}{2(8\tilde{C}+T)} \right\} \left(\log \frac{x}{s_0 e^{rT}} \right)^{\frac{n_1-3}{4}} x^{-\frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}}} \exp \left\{ \frac{\tilde{B}\sqrt{2}}{T^{\frac{1}{4}}(8\tilde{C}+T)^{\frac{1}{4}}} \sqrt{\log \frac{x}{s_0 e^{rT}}} \right\} \\ \left(1 + O \left(\left(\log \frac{x}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (44)$$

as $x \rightarrow \infty$. Next, using (41) and (44), we obtain

$$\begin{aligned} D_T(x) &= \frac{\tilde{A} 2^{\frac{n_1-3}{4}}}{T^{\frac{n_1+1}{8}} (8\tilde{C}+T)^{\frac{n_1+1}{8}}} \left(s_0 e^{rT} \right)^{\frac{1}{2} + \frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}}} \exp \left\{ \frac{\tilde{B}^2}{2(8\tilde{C}+T)} \right\} \\ &\quad \left(\log \frac{x}{s_0 e^{rT}} \right)^{\frac{n_1-3}{4}} x^{-\left(\frac{3}{2} + \frac{\sqrt{8\tilde{C}+T}}{2\sqrt{T}} \right)} \exp \left\{ \frac{\tilde{B}\sqrt{2}}{T^{\frac{1}{4}}(8\tilde{C}+T)^{\frac{1}{4}}} \sqrt{\log \frac{x}{s_0 e^{rT}}} \right\} \\ &\quad \left(1 + O \left(\left(\log \frac{x}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (45)$$

as $x \rightarrow \infty$.

The next assertion can be obtained by using (43) in (45) and simplifying the resulting expressions.

Theorem 9. *If $\delta > 0$, then*

$$\begin{aligned} D_T(x) &= V \left(\log \frac{x}{s_0 e^{rT}} \right)^{\frac{n_1-3}{4}} x^{-\left(\frac{3}{2} + \frac{\sqrt{4+\lambda_1}}{2\sqrt{\lambda_1}} \right)} \exp \left\{ \frac{\sqrt{2\delta}}{\lambda_1^{\frac{1}{4}}(4+\lambda_1)^{\frac{1}{4}}} \sqrt{\log \frac{x}{s_0 e^{rT}}} \right\} \\ &\quad \left(1 + O \left(\left(\log \frac{x}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right) \right) \end{aligned} \quad (46)$$

as $x \rightarrow \infty$, where

$$\begin{aligned} V &= \frac{2^{\frac{n_1-5}{4}}}{\sqrt{\pi} \lambda_1^{\frac{n_1+1}{8}} (4+\lambda_1)^{\frac{n_1+1}{8}}} \delta^{-\frac{n_1-1}{4}} \left(s_0 e^{rT} \right)^{\frac{1}{2} + \frac{\sqrt{4+\lambda_1}}{2\sqrt{\lambda_1}}} \exp \left\{ -\frac{\delta(3+\lambda_1)}{2(4+\lambda_1)} \right\} \\ &\quad \exp \left\{ \frac{s - \sum_{n=1}^\infty \delta_n^2}{2\lambda_1} \right\} \prod_{k>n_1} \left(\frac{\lambda_1}{\lambda_1 - \lambda_k} \right)^{\frac{1}{2}} \exp \left\{ \frac{\delta_k^2}{2(\lambda_1 - \lambda_k)} \right\}. \end{aligned} \quad (47)$$

Formula (46) describes the asymptotic behavior of the asset price density in a Gaussian stochastic volatility model in terms of the Karhunen-Loève parameters, the initial condition s_0 , the interest rate r , and the time horizon T . Note that the Karhunen-Loève parameters depend on T , while the constant V depends on s_0 and r . We will sometimes use the notation $V(s_0, r)$ to emphasize this dependence.

An asymptotic formula similar to that in (46) can be also obtained for $\delta = 0$, using (36) and (37) instead of (34). We will next formulate this asymptotic formula for a special model where the volatility is described by a centered Gaussian process.

Theorem 10. *If the process X is centered, then*

$$D_T(x) = U \left(\log \frac{x}{s_0 e^{rT}} \right)^{\frac{n_1-2}{2}} x^{-\left(\frac{3}{2} + \frac{\sqrt{4+\lambda_1}}{2\sqrt{\lambda_1}}\right)} \left(1 + O \left(\left(\log \frac{x}{s_0 e^{rT}} \right)^{-\frac{1}{2}} \right) \right) \quad (48)$$

as $x \rightarrow \infty$, where

$$U = \frac{1}{\Gamma\left(\frac{n_1}{2}\right) \lambda_1^{\frac{n_1}{4}} (4 + \lambda_1)^{\frac{n_1}{4}}} (s_0 e^{rT})^{\frac{1}{2} + \frac{\sqrt{4+\lambda_1}}{2\sqrt{\lambda_1}}} \prod_{k>n_1}^{\infty} \left(\frac{\lambda_1}{\lambda_1 - \lambda_k} \right)^{\frac{1}{2}}. \quad (49)$$

5. ASYMPTOTICS OF THE IMPLIED VOLATILITY

Taking into account formula (40), we see that the discounted asset price process in a Gaussian stochastic volatility model is given by the following stochastic exponential:

$$\tilde{S}_t = e^{-rt} S_t = s_0 \exp \left\{ -\frac{1}{2} \int_0^t X_s^2 ds + \int_0^t |X_s| dW_s \right\}. \quad (50)$$

The next standard assertion states that Gaussian stochastic volatility models create a risk-neutral environment.

Lemma 11. *Under the restrictions on the volatility process X in (1), the discounted asset price process \tilde{S} is a $\{\mathcal{F}_t\}$ -martingale.*

Proof. Lemma 11 is standard. Using Itô's formula, we first show that the process \tilde{S} in (50) is a positive local martingale. Hence, it is a supermartingale by Fatou's lemma. The conditional distribution of the stochastic integral $\int_0^t |X_s| dW_s$ given $|X|$ is normal with mean zero and variance $\int_0^t X_s^2 ds$. Hence by conditioning on $|X|$ and using the normal MGF, we can prove that $\mathbb{E}[\tilde{S}_t] = s_0$ for all t . However, a supermartingale with a constant expectation is a martingale. This completes the proof of Lemma 11. \square

Let us define the call pricing function in the stochastic volatility model described by (1) by $C(T, K) = e^{-rt} \mathbb{E}[(S_T - K)^+]$, where T is the maturity and K is the strike price, and recall that $S_0 = s_0$ a.s.

If the initial condition for the volatility process X is constant, then the call pricing function C is free of static arbitrage. On the other hand, if the initial condition X_0 is random, then there may be static arbitrage in the function C . We refer the reader to Definition 1.2 in [34] for more details concerning static arbitrage.

Let us fix the maturity T , and consider C as the function $K \mapsto C(K)$ of only the strike price K . The Black-Scholes implied volatility associated with the pricing function C will be denoted by I . More information on the implied volatility can be found in [20, 22].

The asymptotic behavior of the implied volatility for stochastic volatility models, in which the asset price density satisfies

$$D_T(x) = r_1 x^{-r_3} \exp\{r_2 \sqrt{\log x}\} (\log x)^{r_4} (1 + O((\log x)^{-\frac{1}{2}})), \quad x \rightarrow \infty, \quad (51)$$

where $r_1 > 0$, $r_2 \geq 0$, $r_3 > 2$, and $r_4 \in \mathbb{R}$, was characterized in [27]. However, there is an error in the expression for the fourth coefficient in formula (91) in [27]. The correct statement is as follows.

Theorem 12. *Suppose condition (51) holds. Then the following asymptotic formula is valid for the implied volatility:*

$$\begin{aligned}
 I(K) &= \frac{\sqrt{2}}{\sqrt{T}}(\sqrt{r_3-1} - \sqrt{r_3-2})\sqrt{\log \frac{K}{s_0 e^{rT}}} + \frac{r_2}{\sqrt{2T}} \left(\frac{1}{\sqrt{r_3-2}} - \frac{1}{\sqrt{r_3-1}} \right) \\
 &+ \frac{2r_4+1}{2\sqrt{2T}} \left(\frac{1}{\sqrt{r_3-2}} - \frac{1}{\sqrt{r_3-1}} \right) \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} \\
 &+ \left[\frac{1}{\sqrt{2T}} \left(\frac{1}{\sqrt{r_3-1}} - \frac{1}{\sqrt{r_3-2}} \right) \log \frac{\sqrt{r_3-1} - \sqrt{r_3-2}}{2\sqrt{\pi r_1}} + \frac{r_2^2}{4\sqrt{2T}} \left(\frac{1}{(r_3-2)^{\frac{3}{2}}} - \frac{1}{(r_3-1)^{\frac{3}{2}}} \right) \right] \\
 &\times \frac{1}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + \frac{r_2(2r_4+1)}{4\sqrt{2T}} \left(\frac{1}{(r_3-2)^{\frac{3}{2}}} - \frac{1}{(r_3-1)^{\frac{3}{2}}} \right) \frac{\log \log \frac{K}{s_0 e^{rT}}}{\log \frac{K}{s_0 e^{rT}}} + O\left(\frac{1}{\log \frac{K}{s_0 e^{rT}}}\right) \quad (52)
 \end{aligned}$$

as $K \rightarrow \infty$.

The proof of Theorem 12 is exactly the same as that of Theorem 17 in [27].

The next assertions (Theorems 13 and 14) are the main results of the present paper. They provide asymptotic formulas for the implied volatility in the stochastic volatility model given by (1).

Theorem 13. *Suppose $\delta > 0$. Then the following formula holds for the implied volatility $K \mapsto I(K)$:*

$$\begin{aligned}
 I(K) &= M_1 \sqrt{\log \frac{K}{s_0 e^{rT}}} + M_2 + M_3 \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} \\
 &+ M_4 \frac{1}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + M_5 \frac{\log \log \frac{K}{s_0 e^{rT}}}{\log \frac{K}{s_0 e^{rT}}} + O\left(\frac{1}{\log \frac{K}{s_0 e^{rT}}}\right) \quad (53)
 \end{aligned}$$

as $K \rightarrow \infty$, where

$$\begin{aligned}
M_1 &= \frac{\sqrt{2}}{\sqrt{T}} \left(\frac{\sqrt{\lambda_1}}{\sqrt{4 + \lambda_1 + 2}} \right)^{\frac{1}{2}}, & M_2 &= \frac{\sqrt{\delta}}{\sqrt{T}} \left(\frac{\lambda_1}{\sqrt{4 + \lambda_1}(\sqrt{4 + \lambda_1 + 2})} \right)^{\frac{1}{2}}, \\
M_3 &= \frac{n_1 - 1}{4\sqrt{2T}} \left(\frac{\lambda_1^{\frac{3}{2}}}{\sqrt{4 + \lambda_1 + 2}} \right)^{\frac{1}{2}}, \\
M_4 &= -\frac{1}{\sqrt{2T}} \left(\frac{\lambda_1^{\frac{3}{2}}}{\sqrt{4 + \lambda_1 + 2}} \right)^{\frac{1}{2}} \log \left[\frac{1}{2\sqrt{\pi}V(1,0)} \left(\frac{\lambda_1^{\frac{1}{2}}}{\sqrt{4 + \lambda_1 + 2}} \right)^{\frac{1}{2}} \right] \\
&\quad + \frac{\sqrt{2}\delta}{4\sqrt{T}} \left(\frac{\sqrt{\lambda_1}(\sqrt{4 + \lambda_1} - 2)}{4 + \lambda_1} \right)^{\frac{1}{2}} (\sqrt{4 + \lambda_1} + 1), \\
M_5 &= \frac{(n_1 - 1)\sqrt{\delta}}{8\sqrt{T}} \left(\frac{\lambda_1(\sqrt{4 + \lambda_1} - 2)}{\sqrt{4 + \lambda_1}} \right)^{\frac{1}{2}} (\sqrt{4 + \lambda_1} + 1),
\end{aligned} \tag{54}$$

where $V(1,0)$ is the value of V in (47) with $s_0 = 1$ and $r = 0$.

Proof. Set $r_1 = V(1,0)$, $r_2 = \frac{\sqrt{2\delta}}{\lambda_1^{\frac{1}{4}}(4+\lambda_1)^{\frac{1}{4}}}$, $r_3 = \frac{3}{2} + \frac{\sqrt{4+\lambda_1}}{2\sqrt{\lambda_1}}$, and $r_4 = \frac{n_1-3}{4}$. Next, using (46) and (52), and making straightforward simplifications, we get

$$\begin{aligned}
M_1 &= \frac{2\lambda_1^{\frac{1}{4}}}{\sqrt{T} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]}, \\
M_2 &= \frac{\sqrt{2\delta\lambda_1}}{(4 + \lambda_1)^{\frac{1}{4}}\sqrt{T} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]}, \\
M_3 &= \frac{(n_1 - 1)\lambda_1^{\frac{3}{4}}}{4\sqrt{T} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]}, \\
M_4 &= -\frac{\lambda_1^{\frac{3}{4}}}{\sqrt{T} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \\
&\quad \times \log \frac{\lambda_1^{\frac{1}{4}}}{\sqrt{2\pi}V(1,0) \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} \right]} \\
&\quad + \frac{\delta\lambda_1^{\frac{1}{4}}}{8\sqrt{T}(4 + \lambda_1)} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{3}{2}} - (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{3}{2}} \right], \\
M_5 &= \frac{\sqrt{2\lambda_1\delta}(n_1 - 1)}{32\sqrt{T}(4 + \lambda_1)^{\frac{1}{4}}} \left[(\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{3}{2}} - (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{3}{2}} \right].
\end{aligned}$$

Finally, by taking into account the equalities

$$\begin{aligned} (\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} + (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} &= \sqrt{2}(\sqrt{4 + \lambda_1} + 2)^{\frac{1}{2}}, \\ (\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{1}{2}} - (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{1}{2}} &= \sqrt{2}(\sqrt{4 + \lambda_1} - 2)^{\frac{1}{2}}, \\ (\sqrt{4 + \lambda_1} + \sqrt{\lambda_1})^{\frac{3}{2}} - (\sqrt{4 + \lambda_1} - \sqrt{\lambda_1})^{\frac{3}{2}} &= 2^{\frac{3}{2}}(\sqrt{4 + \lambda_1} - 2)^{\frac{1}{2}}(\sqrt{4 + \lambda_1} + 1), \end{aligned}$$

we obtain the formulas for the coefficients in Theorem 13. \square

The constant $V(1, 0)$, given by (47), depends on all the Karhunen-Loève parameters. However, this constant appears for the first time in the fourth term of the asymptotic expansion in (52). By keeping only three terms in (53), we obtain an asymptotic formula for the implied volatility, in which the coefficients do not depend on V . However, now we have the error term of the following form: $O\left(\left(\log \frac{K}{s_0 e^{rT}}\right)^{-\frac{1}{2}}\right)$.

We will next suppose that the volatility is a centered Gaussian process, and study the wing behavior of the implied volatility in such a case. According to formula (48), we can take $r_1 = U(1, 0)$, $r_2 = 0$, $r_3 = \frac{3}{2} + \frac{\sqrt{4 + \lambda_1}}{2\sqrt{\lambda_1}}$, and $r_4 = \frac{n_1 - 2}{2}$. Here $U(1, 0)$ is defined by (49). Then, using Theorems 10 and 12, and reasoning as in the proof of Theorem 13, we obtain the following assertion.

Theorem 14. *Suppose the volatility is modeled by a centered Gaussian process. Then*

$$I(K) = L_1 \sqrt{\log \frac{K}{s_0 e^{rT}}} + L_2 \frac{\log \log \frac{K}{s_0 e^{rT}}}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + L_3 \frac{1}{\sqrt{\log \frac{K}{s_0 e^{rT}}}} + O\left(\frac{1}{\log \frac{K}{s_0 e^{rT}}}\right)$$

as $K \rightarrow \infty$, where

$$\begin{aligned} L_1 &= \frac{\sqrt{2}}{\sqrt{T}} \left(\frac{\sqrt{\lambda_1}}{\sqrt{4 + \lambda_1} + 2} \right)^{\frac{1}{2}}, \quad L_2 = \frac{n_1 - 1}{2\sqrt{2T}} \left(\frac{\lambda_1^{\frac{3}{2}}}{\sqrt{4 + \lambda_1} + 2} \right)^{\frac{1}{2}}, \\ L_3 &= -\frac{1}{\sqrt{2T}} \left(\frac{\lambda_1^{\frac{3}{2}}}{\sqrt{4 + \lambda_1} + 2} \right)^{\frac{1}{2}} \log \left[\frac{1}{2\sqrt{\pi}U(1, 0)} \left(\frac{\lambda_1^{\frac{1}{2}}}{\sqrt{4 + \lambda_1} + 2} \right)^{\frac{1}{2}} \right]. \end{aligned}$$

Remark 15. *Since the processes X and W in (1) are independent, the model in (1) belongs to the class of the so-called symmetric models (see Section 9.8 in [22]). It is known that for a symmetric model,*

$$I(K) = I\left(\frac{(s_0 e^{rT})^2}{K}\right) \quad \text{for all } K > 0. \quad (55)$$

It is clear that, using (55) and Theorem 13, we can characterize the left-wing asymptotic behavior of the implied volatility in the case of a noncentered Gaussian volatility. Similarly, (55) and Theorem 14 can be used in the case of a centered Gaussian volatility.

6. IMPLIED VOLATILITY IN THE UNCORRELATED STEIN-STEIN MODEL

The classical Stein-Stein model is an important special example of a Gaussian stochastic volatility model. The Stein-Stein model was introduced in [36]. The volatility in the uncorrelated Stein-Stein model is the absolute value of an Ornstein-Uhlenbeck process with a constant initial condition m_0 . In this section, we also consider a generalization of the Stein-Stein model, in which the initial condition for the volatility process is a random variable X_0 . Of our interest in the present section is a Gaussian stochastic volatility model with the process X satisfying the equation $dX_t = q(m - X_t)dt + \sigma dZ_t$. Here $q > 0$, $m \geq 0$, and $\sigma > 0$. It will be assumed that the initial condition X_0 is a Gaussian random variable with mean m_0 and variance σ_0^2 , independent of the process Z . It is known that

$$X_t = e^{-qt}X_0 + (1 - e^{-qt})m + \sigma e^{-qt} \int_0^t e^{qu} dZ_u, \quad t \geq 0. \quad (56)$$

If $\sigma_0 = 0$, then the initial condition is equal to the constant m_0 . The mean function of the process X is given by

$$m(t) = e^{-qt}m_0 + (1 - e^{-qt})m, \quad (57)$$

and its covariance function is as follows:

$$Q(t, s) = e^{-q(t+s)} \left\{ \sigma_0^2 + \frac{\sigma^2}{2q} \left(e^{2q \min(t,s)} - 1 \right) \right\}.$$

Therefore, the following formula holds for the variance function:

$$\sigma_t^2 = \frac{\sigma^2}{2q} + e^{-2qt} \left(\sigma_0^2 - \frac{\sigma^2}{2q} \right),$$

and hence, if $\sigma_0^2 = \frac{\sigma^2}{2q}$, then the process $X_t - m(t)$, $t \in [0, T]$, is centered and stationary. In this case, the covariance function is given by

$$Q(t, s) = \frac{\sigma^2}{2q} e^{-q|t-s|}.$$

The Karhunen-Loève expansion of the Ornstein-Uhlenbeck process is known explicitly (see [12]). Denote by w_n the increasingly sorted sequence of the positive solutions to the equation

$$\sigma^2 w \cos(wT) + (q\sigma^2 - w^2\sigma_0^2 - q^2\sigma_0^2) \sin(wT) = 0. \quad (58)$$

If $\sigma_0 = 0$, then the equation in (58) becomes

$$w \cos(wT) + q \sin(wT) = 0. \quad (59)$$

For the OU process in (56) with $\sigma_0 \neq 0$, we have $n_k = 1$ for all $k \geq 1$;

$$\lambda_n = \frac{\sigma^2}{w_n^2 + q^2} \quad (60)$$

for all $n \geq 1$; and

$$e_n(t) = K_n[\sigma_0^2 w_n \cos(w_n t) + (\sigma^2 - q\sigma_0^2) \sin(w_n t)] \quad (61)$$

for all $n \geq 1$ and $t \in [0, T]$. The constant K_n in (61) is determined from

$$\begin{aligned} \frac{1}{K_n^2} &= \frac{1}{2w_n} \sigma_0^2 (\sigma^2 - q\sigma_0^2) (1 - \cos(2w_n T)) + \frac{1}{2} \sigma_0^4 w_n^2 \left(T + \frac{1}{2w_n} \sin(2w_n T) \right) \\ &\quad + \frac{1}{2} (\sigma^2 - q\sigma_0^2)^2 \left(T - \frac{1}{2w_n} \sin(2w_n T) \right) \end{aligned} \quad (62)$$

for all $n \geq 1$. On the other hand, if $\sigma_0 = 0$, then λ_n is given by (60), while the functions e_n are defined by

$$e_n(t) = \frac{1}{\sqrt{\frac{T}{2} - \frac{\sin(2w_n T)}{4w_n}}} \sin(w_n t) \quad (63)$$

for all $n \geq 1$ and $t \in [0, T]$.

By the Karhunen-Loève theorem, the Ornstein-Uhlenbeck process X in (56) can be represented as follows:

$$X_t = e^{-qt} m_0 + (1 - e^{-qt}) m + \sum_{n=1}^{\infty} \sqrt{\lambda_n} e_n(t) Z_n$$

where $\{Z_n\}_{n \geq 1}$ is an i.i.d. sequence of standard normal variables. The eigenvalues λ_n , $n \geq 1$, and the eigenfunctions e_n , $n \geq 1$, are given by (60) and (61) if $\sigma_0 \neq 0$, and by (60) and (63) if $\sigma_0 = 0$. Recall that the numbers w_n , $n \geq 1$, in (60) are solutions to the equation in (58) if $\sigma_0 \neq 0$, and to the equation in (59) if $\sigma_0 = 0$. We refer the interested reader to [12] for more details.

Our next goal is to discuss the constants in the asymptotic formulas for the implied volatility at extreme strikes in the Stein-Stein model. Since $n_1 = 1$ for any OU process, the third and fifth terms in the expansion of Theorem 13 are zero, and with the exception of the term $V(1, 0)$ in M_4 , the only parameters needed to compute the above-mentioned constants are λ_1 and δ_1 . If $\sigma_0 \neq 0$, then we have

$$\lambda_1 = \frac{\sigma^2}{w_1^2 + q^2}, \quad (64)$$

where w_1 is the smallest strictly positive solution to the equation in (58).

The next assertion provides explicit formulas for the number $\delta_1 = \int_0^T m(t) e_1(t) dt$.

Lemma 16. (i) For the generalized uncorrelated Stein-Stein model with $\sigma_0 \neq 0$,

$$\begin{aligned} \delta_1 &= \frac{K_1 m (\sigma^2 - q\sigma_0^2) (1 - \cos(w_1 T))}{w_1} + K_1 \sigma_0^2 \sin(w_1 T) [(m_0 - m) e^{-qT} + m] \\ &\quad + K_1 \sigma^2 (m_0 - m) \frac{w_1 [1 - e^{-qT} \cos(w_1 T)] - q e^{-qT} \sin(w_1 T)}{q^2 + w_1^2}, \end{aligned} \quad (65)$$

where the constant K_1 is determined from (62) with $n = 1$. The symbol w_1 in (65) stands for the smallest strictly positive solution to (58).

(ii) For the uncorrelated Stein-Stein model with $X_0 = m_0$ \mathbb{P} -almost surely,

$$\delta_1 = \frac{mq^2(1 - \cos(w_1T)) + w_1^2(m_0 - m \cos(w_1T))}{w_1(q^2 + w_1^2)\sqrt{\frac{T}{2} - \frac{\sin(2w_1T)}{4w_1}}}. \quad (66)$$

Proof. Taking into account (57) and (61), we see that

$$\begin{aligned} \delta_1 &= b_1 \int_0^T \cos(w_1t)dt + b_2 \int_0^T e^{-qt} \cos(w_1t)dt \\ &\quad + b_3 \int_0^T \sin(w_1t)dt + b_4 \int_0^T e^{-qt} \sin(w_1t)dt, \end{aligned} \quad (67)$$

where

$$\begin{aligned} b_1 &= mK_1\sigma_0^2w_1, \quad b_2 = (m_0 - m)K_1\sigma_0^2w_1, \\ b_3 &= mK_1(\sigma^2 - q\sigma_0^2), \quad \text{and} \quad b_4 = (m_0 - m)K_1(\sigma^2 - q\sigma_0^2). \end{aligned} \quad (68)$$

It remains to evaluate the integrals in (67). We have

$$\int_0^T \cos(w_1t)dt = \frac{\sin(w_1T)}{w_1}, \quad (69)$$

$$\int_0^T e^{-qt} \cos(w_1t)dt = \frac{q[1 - e^{-qT} \cos(w_1T)] + w_1e^{-qT} \sin(w_1T)}{q^2 + w_1^2}, \quad (70)$$

$$\int_0^T \sin(w_1t)dt = \frac{1 - \cos(w_1T)}{w_1}, \quad (71)$$

and

$$\int_0^T e^{-qt} \sin(w_1t)dt = \frac{w_1[1 - e^{-qT} \cos(w_1T)] - qe^{-qT} \sin(w_1T)}{q^2 + w_1^2}. \quad (72)$$

In the proof of (70) and (72), we use the integration by parts formula twice. Now, taking into account formulas (67-72) and making simplifications, we establish formula (65).

Next, suppose $\sigma_0 = 0$. Then (65) implies that

$$\begin{aligned} \delta_1 &= \frac{m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1T)}{4w_1}}} \frac{1 - \cos(w_1T)}{w_1} \\ &\quad + \frac{m_0 - m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1T)}{4w_1}}} \frac{w_1[1 - e^{-qT} \cos(w_1T)] - qe^{-qT} \sin(w_1T)}{q^2 + w_1^2}, \end{aligned}$$

where w_1 denotes the smallest strictly positive solution to (59). It is not hard to see, using the equality $w_1 \cos(w_1T) + q \sin(w_1T) = 0$, that

$$\delta_1 = \frac{m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1T)}{4w_1}}} \frac{1 - \cos(w_1T)}{w_1} + \frac{m_0 - m}{\sqrt{\frac{T}{2} - \frac{\sin(2w_1T)}{4w_1}}} \frac{w_1}{q^2 + w_1^2}, \quad (73)$$

and it is clear that (73) and (66) are equivalent. This completes the proof of Lemma 16. \square

Remark 17. *Since for the generalized Stein-Stein model with a random initial condition we have $n_1 = 1$, one can use the asymptotic formulas in Theorem 13 with $M_3 = M_5 = 0$ to characterize the wing-behavior of the implied volatility. The dependence of the parameters λ_1 and δ_1 , appearing in those formulas, on the model parameters is described in (64), (65), and (66). Originally, sharp asymptotic formulas for the implied volatility at extreme strikes in the uncorrelated Stein-Stein model with $X_0 = m_0$ were obtained in [25] (see also Section 10.5 in [22]). However, explicit expressions, obtained in [25] and [22] for the coefficients in the asymptotic formulas for the implied volatility in the Stein-Stein model, are significantly more complicated than those found in the present paper.*

7. NUMERICAL ILLUSTRATION

A basic calibration strategy when presented with asymptotic results such as those given in this paper is to assume one can place oneself in the corresponding regime, and then determine model parameters by reading asymptotic coefficient off of market option prices. We now illustrate how this strategy can produce positive results, and discuss its limitations, when the top of the KL spectrum is simple ($n_1 = 1$). As noted in the introduction, in this case, the third and fifth terms in the expansion are null. The idea is to ignore the big O term in the asymptotic (54), and calibrate parameters to the remaining coefficients. Denoting the discounted log-moneyness $\log(S_0 e^{rT}/K)$ by k for compactness of notation, we thus have, for $|k|$ sufficiently large,

$$I(k) \simeq M_1 \sqrt{|k|} + M_2 + M_4 \frac{1}{\sqrt{|k|}}, \quad (74)$$

for three constants M_1 , M_2 , and M_4 , which can, in principle, be read off of market data. By the explicit expressions for the first two constants in (54) in terms of λ_1 and δ_1 , we then express the latter in terms of M_1 and M_2 as

$$\begin{aligned} \lambda_1 &= \frac{64T^2 M_1^4}{(4 - T^2 M_1^4)^2}, \\ \delta_1 &= \frac{4\sqrt{2T} M_2 \sqrt{4 + T^2 M_1^4}}{4 - T^2 M_1^4}. \end{aligned} \quad (75)$$

Here we use (74). One notices that, conveniently, λ_1 can be calibrated using only the coefficient M_1 , while given M_1 , δ_1 is then proportional to M_2 .

At this stage, one may simply conclude that the extreme strike asymptotics given in the market are consistent with any Gaussian volatility model whose top of eigenstructure is represented by the values computed in the above expressions for λ_1 and δ_1 . However, practitioners will prefer to determine a more specific model, perhaps by choosing a classical parametric one, and using other non-asymptotic-calibration techniques for estimating some of its parameters. The expressions in (75) can then be used to pin down other parameters by calibration, as long as one can relate the model's parameters to the pair (λ_1, δ_1) from the top of its KL spectrum, whether analytically or numerically. The expression for M_4 , given in (47) and (54), may be too complex to provide a reliable method for calibrating parameters beyond the pair (λ_1, δ_1) , but we will see below that the existence of the

corresponding term in the expansion, combined with a truncation of the formula for M_4 , is very helpful for implementing the calibration based on (75).

We provide illustrations of this strategy in two cases: the stationary Stein-Stein model, where the KL expansion is known semi-explicitly, and the Stein-Stein model's long-memory version, where the volatility is also known as the fractional Ornstein-Uhlenbeck (fOU) model, and the KL expansion is computed numerically. The data we use is also generated numerically: for each model, we compute option prices and their corresponding implied volatilities, by classical Monte-Carlo, given that the underlying pair of stochastic processes is readily simulated. Specifically, in the Stein-Stein (standard OU) case, 10^6 paths are generated via Euler's method based on discretizing the stochastic differential equation satisfied by X started from a r.v. sampled from X 's stationary distribution, and the explicit expression for $\log S$ given X , also approximated via Euler with the same time steps; 10^3 time steps are used in $[0, T]$ for the various values of T we illustrate below (1, 2, 3 and 6 months, measured in years). Option prices are derived by computing average pay-offs over the 10^6 paths. The details are well known, and are omitted. In the fOU case, the exact same methodology is used, except that one must specify the technique used to simulate increments of the fBm process which drives X : we used the circulant method, which is based on fBm's spectral properties, and was proposed by A.T. Dieker in a 2002 thesis: see [16, 17].

Given this simulated data, before embarking on the task of calibrating parameters, to ensure that our methodology is relevant in practice, it is important to discuss liquidity issues. It is known that the out-of-the-money call options market is poorly liquid, implying that the large strike asymptotics for call and IV prices are typically not visible in the data. We concentrate instead on small strike asymptotics. There, depending on the market segment, options with three-month maturity can be liquid with small bid-ask spread for log moneyness k as far down as -1 or even a bit further. Options with six-month maturity with very small bid-ask spread can be liquid as far down as -1.5 . Convincing visual evidence of this can be found in Figures 3 and 4 in [21] which report 2011 data for SPX options. We will also consider examples with one-month and two-month maturity, where liquidity will be assumed to exist down to $k = -0.8$, based on corresponding evidence in the same figures. We will illustrate calibration using intervals of relatively short length which start on the left side within these observed liquidity ranges. Beyond these lower bounds, liquidity is insufficient to measure IV. In these ranges of k , the constant term M_2 and the expressions $\sqrt{-k}$ and $1/\sqrt{-k}$ are of similar magnitude, which may call into question whether the expansion can be of any use in the range where liquidity exists. However, one may expect that the KL expansion converges fast enough that the three terms $M_1\sqrt{-k}$, M_2 , and $M_4/\sqrt{-k}$ are of different orders because the three constants are. This turns out to be the case in the two example classes we consider, so that our three-term expansion allows us to calibrate λ_1 and δ_1 to M_1 and M_2 as in (75). This works very well in practice, as our examples below now show.

We begin with the stationary uncorrelated Stein-Stein model with constant mean-reversion level m , rate of mean reversion q , and so called vol-vol parameter σ . Referring to the notation in Section 6, since now X is stationary, we have $m_0 = m$ and $\sigma_0^2 = \sigma^2 / (2q)$, and the constant K_1 , which is determined from equation (62), will play an important role for

us. The systems of equations needed to perform calibration here have a somewhat triangular structure. According to Section 6, if one were to calibrate q , access to δ_1 would be needed, if one were to rely on independent knowledge of the level of mean reversion m . Specifically, one would solve the following system

$$\begin{aligned} q \sin(wT) + w \cos(wT) &= 0 \\ C_1 \left(\sin(wT) + \frac{q}{w}(1 - \cos(wT)) \right) &= \frac{\delta_1}{m} \end{aligned} \quad (76)$$

where $C_1 = K_1 \sigma_0^2$. As noted via (62), unfortunately the constant C_1 also depends on (q, w) in the following non-trivial way:

$$\frac{1}{C_1^2} = \frac{q}{2}(1 - \cos(2wT)) + \frac{w^2}{2} \left(T + \frac{\sin(2wT)}{2w} \right) + \frac{q^2}{2} \left(T - \frac{\sin(2wT)}{2w} \right). \quad (77)$$

When q is not fixed, the task of determining which value of w represents the minimal solution of the first equation above, given the large number of solutions to the above system, is difficult. We did not pursue this avenue further for this reason. Instead, we will assume that q , which determines the rate of mean reversion, is known, and we will calibrate the pair (m, σ) .

The equations for finding (σ, m) given prior knowledge of q , and given measurement of M_1 and M_2 which imply values of (λ_1, δ_1) via (75), are much simpler. Indeed, since q is assumed given, the base frequency w is computed easily as the smallest positive solution of (59). Then according to equation (64) and part (ii) of Lemma 16, with C_1 given by (77), we obtain immediately

$$\sigma^2 = \lambda_1 (w^2 + q^2); \quad (78)$$

$$m = \frac{\delta_1}{C_1 \left(\sin(wT) + \frac{q}{w}(1 - \cos(wT)) \right)}. \quad (79)$$

Any fitting method can in principle be used to estimate the coefficients M_1 , M_2 , and M_4 when working from a data-based IV curve. However, it turns out that, in the ranges of liquidity which we described above, any estimation will contain a certain amount of instability. We now give the details of an iterative technique which increases the stability of the method dramatically by exploiting the fact that M_4 is significantly smaller than M_1 and M_2 .

We use simulated IV data for the call option with $m = 0.2$ (signifying a typical mean level of volatility of 20%), $q = 7$ (fast mean reversion, every eight weeks or so), and $\sigma = 1.2$ (high level of volatility uncertainty). How to estimate M_1 from the data is not unambiguous. We adopt a least-squares method, on an interval of k -values of fixed length; after experimentation, as a rule of thumb, an interval of length 0.10 or 0.20 provides a good balance between providing a local estimate and drawing on enough datapoints. One should start the interval as far to the left as possible while avoiding any range with insufficient liquidity in practice. As a guide to assess this liquidity, we use the study reported in [21, Section 5.4], which depends heavily on the option maturity, as we mentioned in this section. The following are intervals employed.

Maturity T	1/12 (1 mo.)	1/12 (1 mo.)	1/6 (2 mos.)	1/6 (2 mos.)
Interval used	$[-0.8, -0.6]$	$[-0.7, -0.6]$	$[-0.8, -0.6]$	$[-0.7, -0.6]$
Maturity T	0.25 (3 mos.)	0.25 (3 mos.)	0.5 (6 mos.)	0.5 (6 mos.)
Interval used	$[-1.1, -0.9]$	$[-1.0, -0.9]$	$[-1.4, -1.2]$	$[-1.3, -1.2]$

Graphs of the data versus the asymptotic curve in (74), showing excellent agreement, are given in Fig. 1a thru 1d, though a case-by-case need for an analysis of the trade-off between this agreement and the liquidity-dictated calibration intervals, is apparent as one considers various possible maturities (note the difference in ranges for log-moneyness k on the horizontal axes).

Our stabilized calibration method starts with a least-squares measurement of M_1 and M_2 based on the asymptotic curve with only the first two terms. The value of M_1 is used to calibrate σ . A guess is expressed for m to initiate the procedure; in our examples we use $m = 0.22$, to signify an educated guess which misses the mark by 10%, as would be reasonable to expect when using a proxy such as the VIX to visually estimate this so-called mean reversion level. The next step uses the values of σ and m previously determined, along with the known value q , to compute a large number of terms in the KL expansion of the OU process (we use 500 terms), and uses those terms to compute M_4 via the expressions in (47) and (54). The value of M_4 just obtained is also used to refine the non-linear least-squares estimation of M_1 and M_2 based on the three-term asymptotic function in (74) where the term $M_4/\sqrt{|k|}$ is assumed known. The third step then calibrates σ and m based on the new values of M_1 and M_2 , and then recomputes M_4 using the same procedure as in the second step, which allows a new estimation of M_1 and M_2 using the full asymptotics including the just-updated term $M_4/\sqrt{|k|}$. One then repeats the third step iteratively, until one notices a stabilization. In the examples we report, the method either stabilizes on a single set of values for the pair (σ, m) , or loops between two very close sets of values; this occurs after 6 or 7 steps. We think this needed number of repeats, and the precision obtained in the end, are typical, because they are functions of the small magnitude of M_4 compared to M_1 (order of 2% to 10% for our maturities from one month to six months), this M_4 being considered as a nuisance term whose rough estimation helps sharpen the estimation of the other two constants significantly. Summarizing the procedure, we have:

- (0) Assume q is known. Compute w as smallest frequency solving (76).
- (1) Use two-term asymptotics to estimate M_1 and M_2 , calibrate σ to M_1 via (75) and (79). Initialize m using a good guess for rate of mean reversion.
- (2) Use σ and m from step 1 (and q from step 0) to compute a large number (e.g. 500) of terms in the KL expansion of X . Use truncated theoretical formula in (47) and (54) to compute K_4 from this expansion. Re-estimate M_1 and M_2 by using full three-term asymptotics (74) assuming $M_4/\sqrt{|k|}$ is known.
- (3) Calibrate σ from the new M_1 and m from the new pair (M_1, M_2) via (75), (79), and (78). Recompute the KL expansion of X based on the new (σ, m) , and recompute K_4 using the new KL expansion in the theoretical formula. Re-estimate M_1 and M_2 by using full three-term asymptotics (74) assuming $M_4/\sqrt{|k|}$ is known using the new M_4 .

(4) Repeat step 3 iteratively until stabilization of (σ, m) occurs.

We report our findings for the calibration of (σ, m) in our 8 examples of interest in the following tables. The "true values" of M_1 , M_2 , and M_4 in these tables are those which are computed from the Stein-Stein model with $(\sigma, m, q) = (1.2, 0.2, 7)$ via its KL elements; as explained above, only the first-order KL eigen-elements are needed for M_1 , M_2 , while for M_4 , we use the full theoretical formula in which we ignore the eigen-elements after rank 500.

$T = 1/12$ (1 mo.) Calibration over the interval $[-0.8, -0.6]$

	M_1	M_2	M_4	σ	m
True values	0.7117	0.0706	0.0188	1.2	0.2
Step 1	0.6875	0.1113		1.1196	0.22
Step 2	0.6875	0.0777	0.0188		
Step 3	0.7145	0.0661	0.0183	1.2096	0.1873
Step 4	0.7138	0.0673	0.0184	1.2072	0.1907
Step 5	0.7140	0.0671	0.0184	1.2077	0.1900

$T = 1/12$ (1 mo.) Calibration over the interval $[-0.7, -0.6]$

	M_1	M_2	M_4	σ	m
True values	0.7117	0.0706	0.0188	1.2	0.2
Step 1	0.6859	0.1126		1.1142	0.22
Step 2	0.6859	0.0777	0.0187	1.2102	0.1872
Step 3	0.7147	0.0661	0.0183	1.2102	0.1872
Step 4	0.7141	0.0671	0.0184	1.2081	0.1901
Step 5	0.7143	0.0668	0.0184	1.2086	0.1894

$T = 1/6$ (2 mos.) Calibration over the interval $[-0.8, -0.6]$

	M_1	M_2	M_4	σ	m
True values	0.5743	0.0704	0.0245	1.2	0.2
Step 1	0.5370	0.1309		1.0490	0.22
Step 2	0.5370	0.0775	0.0232		
Step 3	0.5705	0.0752	0.0251	1.1839	0.2134
Step 4	0.5732	0.0706	0.0245	1.1953	0.2005
Step 5	0.5723	0.0720	0.0247	1.1917	0.2046
Step 6	0.5726	0.0716	0.0246	1.1929	0.2032
Step 7	0.5725	0.0718	0.0246	1.1923	0.2039

$T = 1/6$ (2 mos.) Calibration over the interval $[-0.7, -0.6]$

	M_1	M_2	M_4	σ	m
True values	0.5743	0.0704	0.0245	1.2	0.2
Step 1	0.5354	0.1322		1.0428	0.22
Step 2	0.5354	0.0775	0.0231		
Step 3	0.5711	0.0748	0.0251	1.1866	0.2123
Step 4	0.5742	0.0698	0.0244	1.1995	0.1982
Step 5	0.5731	0.0715	0.0246	1.2011	0.1997
Step 6	0.5734	0.0710	0.0246	1.1962	0.2017

$T = 1/4$ (3 mos.) Calibration over the interval $[-1.1, -0.9]$

	M_1	M_2	M_4	σ	m
True values	0.5001	0.0702	0.0295	1.2	0.2
Step 1	0.4699	0.1299		1.0591	0.22
Step 2	0.4699	0.0773	0.0279		
Step 3	0.4980	0.0740	0.0300	1.1896	0.2107
Step 4	0.5001	0.0698	0.0294	1.1997	0.1987
Step 5	0.4995	0.0710	0.0295	1.1969	0.2021
Step 6	0.4996	0.0708	0.0295	1.1973	0.2016

$T = 1/4$ (3 mos.) Calibration over the interval $[-1.0, -0.9]$

	M_1	M_2	M_4	σ	m
True values	0.5001	0.0702	0.0295	1.2	0.2
Step 1	0.4655	0.1342		1.0396	0.22
Step 2	0.4655	0.0773	0.0275		
Step 3	0.4945	0.0777	0.0305	1.1732	0.2214
Step 4	0.4977	0.0716	0.0295	1.1883	0.2039
Step 5	0.4966	0.0736	0.0298	1.1832	0.2097
Step 6	0.4969	0.0730	0.0297	1.1847	0.2080
Step 7	0.4968	0.0732	0.0297	1.1842	0.2086

$T = 1/2$ (6 mos.) Calibration over the interval $[-1.4, -1.2]$

	M_1	M_2	M_4	σ	m
True values	0.3838	0.0695	0.0428	1.2	0.2
Step 1	0.3521	0.1432		1.0094	0.22
Step 2	0.3521	0.0765	0.0385		
Step 3	0.3817	0.0757	0.0442	1.1869	0.2178
Step 4	0.3861	0.0657	0.0423	1.2144	0.1890
Step 5	0.3846	0.0690	0.0429	1.2052	0.1986
Step 6	0.3851	0.0679	0.0427	1.2081	0.1956
Step 7	0.3849	0.0683	0.0428	1.2071	0.1966
Step 8	0.3850	0.0681	0.0427	1.2076	0.1961

$T = 1/2$ (6 mos.) Calibration over the interval $[-1.3, -1.2]$

	M_1	M_2	M_4	σ	m
True values	0.3838	0.0695	0.0428	1.2	0.2
Step 1	0.3493	0.1464		0.9934	0.22
Step 2	0.3493	0.0765	0.0380		
Step 3	0.3797	0.0783	0.0446	1.1745	0.2255
Step 4	0.3850	0.0665	0.0423	1.2075	0.1915
Step 5	0.3832	0.0706	0.0430	1.1959	0.2034
Step 6	0.3837	0.0694	0.0428	1.1994	0.1998
Step 7	0.3836	0.0697	0.0429	1.1984	0.2008
Step 8	0.3836	0.0696	0.0428	1.1989	0.2003

We obtain excellent agreement of the calibration with the true values, with errors lower than 1% after 5 to 8 steps. Other calibrations, not reported here because of their similarity with these examples, show that calibration accuracy would increase with more liquid options since these allow being able to use intervals further to the left, ensuring a better match with the asymptotic regime (74). The examples reported above in full correspond to realistic liquidity assumptions.

We now propose a calibration method to estimate the memory parameter in the fOU volatility model. This model was introduced in [9] as a way to model long-range dependence in non-linear functionals of asset returns, while preserving the uncorrelated semi-martingale structure at the level of returns themselves. This is the model for X in (2) where the process Z is a fractional Brownian motion, i.e. the continuous Gaussian process started at 0 with covariance determined by $\mathbf{E}[(Z_t - Z_s)^2] = |t - s|^{2H}$, with ‘‘Hurst’’ parameter $H \in (0.5, 1)$. In [8], it was shown empirically that standard statistical methods for long-memory data are inadequate for estimating H . This difficulty can be attributed to the fact that the volatility process X can have non-stationary increments. In addition, some of the classical methods use path regularity or self-similarity as a proxy for long memory, which cannot be exploited in practice since there is a lower limit to how frequent observations can be made without running into microstructure noise. To make matters worse, the process X is not directly observed; in such a partial observation case, a general theoretical result was given in [35], by which the rate of convergence of any estimator of H cannot exceed an optimal H -dependent rate which is always slower than $N^{-1/4}$, where N is the number of observations. Given the non-stationarity of the parameter H on a monthly scale, a realistic time series at the highest observation frequency where microstructure noise can be ignored (e.g. one stock observation every 5 minutes) would not permit even the optimal estimators described in [35] from pinning down a value of H with any acceptable confidence level. The work in [8] proposes a calibration technique based on a straightforward comparison of simulated and market option prices to determine H . Our strategy herein is similar, but based on implied volatility.

Our goal is to calibrate the fOU model described above with the following parameters: $T = 1/4$, $m = 0.2$, $q = 7$, $\sigma = 1.2$, with different values of the Hurst parameter H , namely

$$H \in \{0.51, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85\} \quad (80)$$

As mentioned above, our simulated option prices use standard Monte Carlo, where the fOU process is produced by A.T. Dieker's circulant method. Since the values of λ_1 for each $H > 0.5$ are not known explicitly or semi-explicitly, we resorted to the method developed in by S. Corlay in [10] for optimal quantization: there, the infinite-dimensional eigenvalue problem is converted to a matrix eigenvalue problem which uses a low-order quadrature rule for approximating integrals (a trapezoidal rule is recommended), after which a Richardson-Romberg extrapolation is used to improve accuracy. We repeat this procedure for the fOU process with the above parameters, for each value of H from 0.50 to 0.99, with increments of 0.01. The corresponding values we obtain for λ_1 in each case are collected in the following table:

$H =$	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
$\lambda_1 =$	0.0157	0.0155	0.0152	0.0150	0.0148	0.0146	0.0144	0.0142	0.0140	0.0138
$H =$	0.60	0.61	0.62	0.63	0.64	0.65	0.66	0.67	0.68	0.69
$\lambda_1 =$	0.0136	0.0134	0.0132	0.0130	0.0128	0.0126	0.0124	0.0122	0.0120	0.0118
$H =$	0.70	0.71	0.72	0.73	0.74	0.75	0.76	0.77	0.78	0.79
$\lambda_1 =$	0.0116	0.0115	0.0113	0.0111	0.0109	0.0108	0.0106	0.0104	0.0103	0.0101
$H =$	0.80	0.81	0.82	0.83	0.84	0.85	0.86	0.87	0.88	0.89
$\lambda_1 =$	0.0100	0.0098	0.0097	0.0095	0.0094	0.0092	0.0091	0.0089	0.0088	0.0087
$H =$	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
$\lambda_1 =$	0.0085	0.0084	0.0083	0.0082	0.0080	0.0079	0.0078	0.0077	0.0076	0.0075

Our illustration of the calibration method then consists of starting with simulated IV data for a fOU model with a fixed H from the set in (80), then, similarly to what we did for the Stein-Stein model, calibrate the value of λ_1 to the first term of the simulated IV curve over an interval of length 0.1. For our choice of $T = 1/4$ we use $k \in [-1.0, -0.9]$ to determine λ_1 , which is realistic in terms of liquidity constraints. We then match that value of λ_1 to the closest value in the above table, thereby concluding that the simulated data is consistent with the corresponding value of H in the table. Because of the instability in determining M_4 in (74) by curve fitting, as noted for the standard Stein-Stein model, rather than using the iterative technique described above, we fit our simulated data curve to the first two terms in this expansion only, resulting in a robust estimate for M_1 in all cases, from which our calibrated λ_1 results via (75). This is more efficient since we are only calibrating the single parameter H . The results of this method are summarized here.

$T = 1/4$ (3 mos.)	Calibration of H via λ_1 over the interval $[-1.0, -0.9]$							
True H	0.51	0.55	0.60	0.65	0.70	0.75	0.80	0.85
True λ_1	0.0155	0.0146	0.0136	0.0126	0.0116	0.0108	0.0100	0.00923
calibrated λ_1	0.0152	0.0147	0.0134	0.0127	0.0115	0.0109	0.0101	0.00937
calibrated H	0.52	0.55	0.61	0.64	0.71	0.74	0.79	0.84

Our method shows a good level of accuracy. One notes a bias between the curve $M_1\sqrt{-k} + M_2$ and the simulated IV data, as illustrated in Figures 2a to 2h, which appears to shift downward as H increases. Since we are only calibrating H via λ_1 which is inferred from M_1 , this bias has no influence on the calibration. At the cost of computing M_4 as we did for the Stein-Stein model, which would be more onerous in the fOU

case because one would need to push Corlay's method much further for estimating KL eigenlements, we could obtain the 3-term expansion in (74), resulting in curves which would have much less of a bias than in Figures 2a to 2h, but this would not improve the calibration of λ_1 and H .

REFERENCES

- [1] M. Abramovitz and I. A. Stegun (Eds.). *Handbook of Mathematical Functions*. Applied Mathematics Series 55, National Bureau of Standards, 1972.
- [2] R. J. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. IMS Lecture Notes Monogr. Ser. **12**, 1990.
- [3] A. Alexanderian. A brief note on the Karhunen-Loève expansion, technical note, available on users.ices.utexas.edu/~alen/articles/KL.pdf
- [4] S. Benaim and P. K. Friz. Regular variation and smile asymptotics. *Mathematical Finance* **19** (2009), 1-12.
- [5] S. Benaim and P. K. Friz. Smile asymptotics, II: models with known moment generating function. *Journal of Applied Probability* **45** (2008), 16-32.
- [6] S. Benaim, P. K. Friz, and R. Lee. On black-Scholes implied volatility at extreme strikes. In: R. Cont (Ed.), *Frontiers in Quantitative Finance: Volatility and Credit Risk Modeling*, Wiley, Hoboken (2009), 19-45.
- [7] R. Beran. The probabilities of noncentral quadratic forms. *The Annals of Statistics* **3** (1975), 969-974.
- [8] A. Chronopoulou and F. Viens. Stochastic volatility models with long-memory in discrete and continuous time. *Quantitative Finance*, **12** no. 4 (2012), 635-649.
- [9] F. Comte and E. Renault. Long Memory in Continuous-time Stochastic Volatility Models. *Mathematical Finance*, **8** (1998), 291-323.
- [10] S. Corlay. Quelques aspects de la quantification optimale, et applications en finance (in English, with French summary). Ph.D. Thesis. *Université Pierre et Marie Curie*, 2011.
- [11] S. Corlay. *Properties of the Ornstein-Uhlenbeck bridge*. Preprint, 2014, available at <https://hal.archives-ouvertes.fr/hal-00875342v4>
- [12] S. Corlay and G. Pagès. Functional quantization-based stratified sampling methods. Preprint, 2014, available on <https://hal.archives-ouvertes.fr/hal-00464088v3>
- [13] P. Deheuvels and G. Martynov. A Karhunen-Loeve decomposition of a Gaussian process generated by independent pairs of exponential random variables. *J. Funct. Anal.*, **255** (2008), 23263-2394.
- [14] J.-D. Deuschel, P. K. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility I: Theoretical foundations. *Communications on Pure and Applied Mathematics* **67** (2014), 40-82.
- [15] J.-D. Deuschel, P. K. Friz, A. Jacquier and S. Violante. Marginal density expansions for diffusions and stochastic volatility I: Applications. *Communications on Pure and Applied Mathematics* **67** (2014), 321-350.
- [16] A. B. Dieker. *Simulation of fractional Brownian motion*. Master Thesis (2002), published by Universiteit Twente, 2004. Available at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.8880>.
- [17] A. B. Dieker and M. Mandjes. On spectral simulation of fractional Brownian motion, *Probab. Engrg. Inform. Sci.*, **17** (2003), 417-434.
- [18] K. Gao and R. Lee. Asymptotics of Implied Volatility to Arbitrary Order. To appear in *Finance and Stochastics*, 2015. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1768383.
- [19] J. Gatheral. A parsimonious arbitrage-free implied volatility parametrization with application to the valuation of the volatility derivatives. In: *Global Derivatives and Risk Management*, Madrid, May 26, 2004.
- [20] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley New York 2006.
- [21] J. Gatheral and A. Jacquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance* **14** (1) (2014), 59-71.
- [22] A. Gulisashvili. *Analytically Tractable Stochastic Stock Price Models*, Springer-Verlag Berlin Heidelberg 2012.

- [23] A. Gulisashvili. Asymptotic formulas with error estimates for call pricing functions and the implied volatility at extreme strikes. *SIAM Journal on Financial Mathematics* **1** (2010), 609-641.
- [24] A. Gulisashvili. Asymptotic equivalence in Lee's moment formulas for the implied volatility, asset price models without moment explosions, and Piterbarg's conjecture. *International Journal of Theoretical and Applied Finance* **15** (2012), 1250020.
- [25] A. Gulisashvili and E. M. Stein. Asymptotic behavior of the stock price distribution density and implied volatility in stochastic volatility models. *Applied Mathematics and Optimization* **61** (2010), 287-315.
- [26] A. Gulisashvili, F. Viens, and X. Zhang. *Small-time asymptotics for Gaussian self-similar stochastic volatility models*. Preprint, 2015, available at <http://arxiv.org/abs/1505.05256>.
- [27] A. Gulisashvili and J. Vives. Asymptotic analysis of stock price densities and implied volatilities in mixed stochastic models. *SIAM J. Finan. Math.* **6** (2015), 158-188.
- [28] W. Hoeffding. On a theorem of V. M. Zolotarev. *Theor. Probab. Appl.* **9** (1964), 89-92.
- [29] I. A. Ibragimov and Y. A. Rozanov. *Gaussian Random Processes*, Springer-Verlag New York Heidelberg Berlin 1978.
- [30] R. Lee. The moment formula for implied volatility at extreme strikes. *Math. Finance* **14** (2004), 469-480.
- [31] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. Cambridge U.P., 2012.
- [32] E. Renault and N. Touzi. Option Hedging and Implicit Volatilities, *Mathematical Finance* **6** (1996), 279-302.
- [33] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, Springer-Verlag Berlin 2004.
- [34] M. Roper. Arbitrage free implied volatility surfaces, preprint, 2010.
- [35] M. Rosenbaum. Estimation of the volatility persistence in a discretely observed diffusion model. *Stochastic Processes and their Applications*, **118** (8) (2008), 1434-1462.
- [36] E. Stein and J. Stein. Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financ. Stud.* **4** (1991), 727-752.
- [37] A. M. Yaglom. *Correlation Theory of Stationary and Related Random functions, Vol. I*, Springer-Verlag New York 1987.
- [38] V. M. Zolotarev. Concerning a certain probability problem. *Theor. Probab. Appl.* **6** (1961), 201-204.

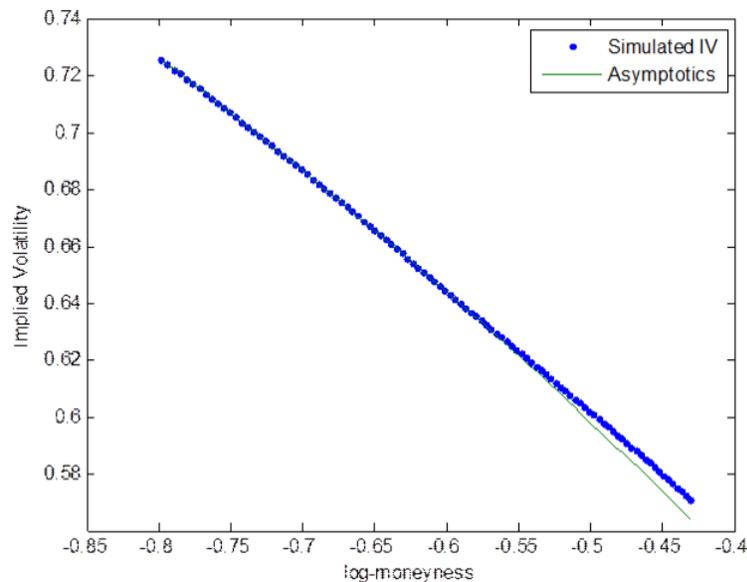


FIGURE 1. Figure 1a. One-month IV for Stein-Stein model with parameters $m = 0.2, q = 7, \sigma = 1.2$

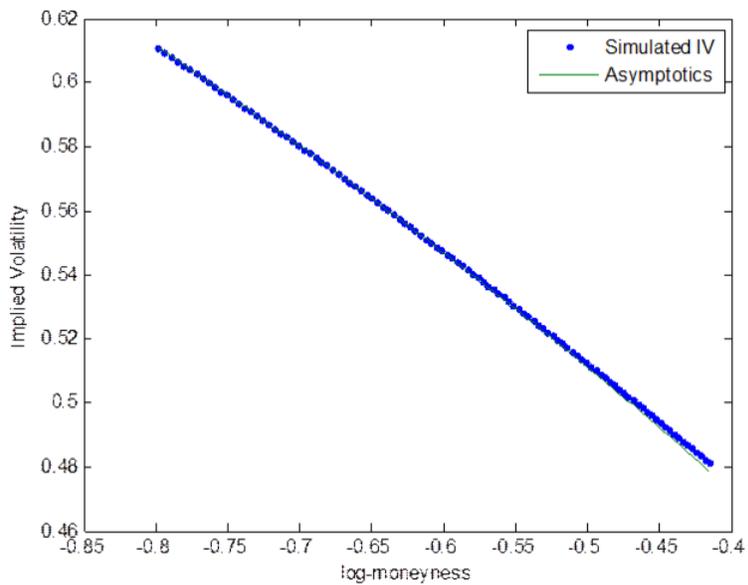


FIGURE 2. Figure 1b. Two-month IV for Stein-Stein model with parameters $m = 0.2, q = 7, \sigma = 1.2$

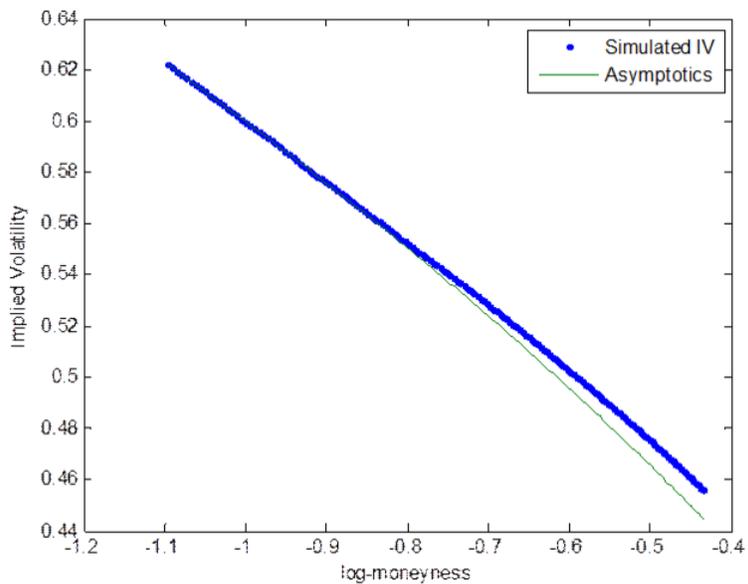


FIGURE 3. Figure 1c. Three-month IV for Stein-Stein model with parameters $m = 0.2, q = 7, \sigma = 1.2$

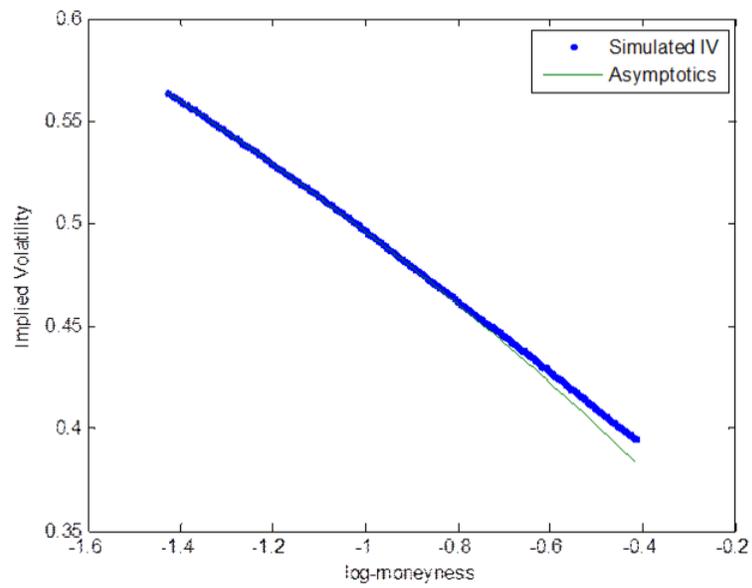


FIGURE 4. Figure 1d. Six-month IV for Stein-Stein model with parameters $m = 0.2, q = 7, \sigma = 1.2$

DEPARTMENT OF MATHEMATICS, OHIO UNIVERSITY, ATHENS OH 45701, GULISASH@OHIO.EDU

DEPARTMENT OF STATISTICS AND PROBABILITY, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824, VIENS@MSU.EDU

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907.

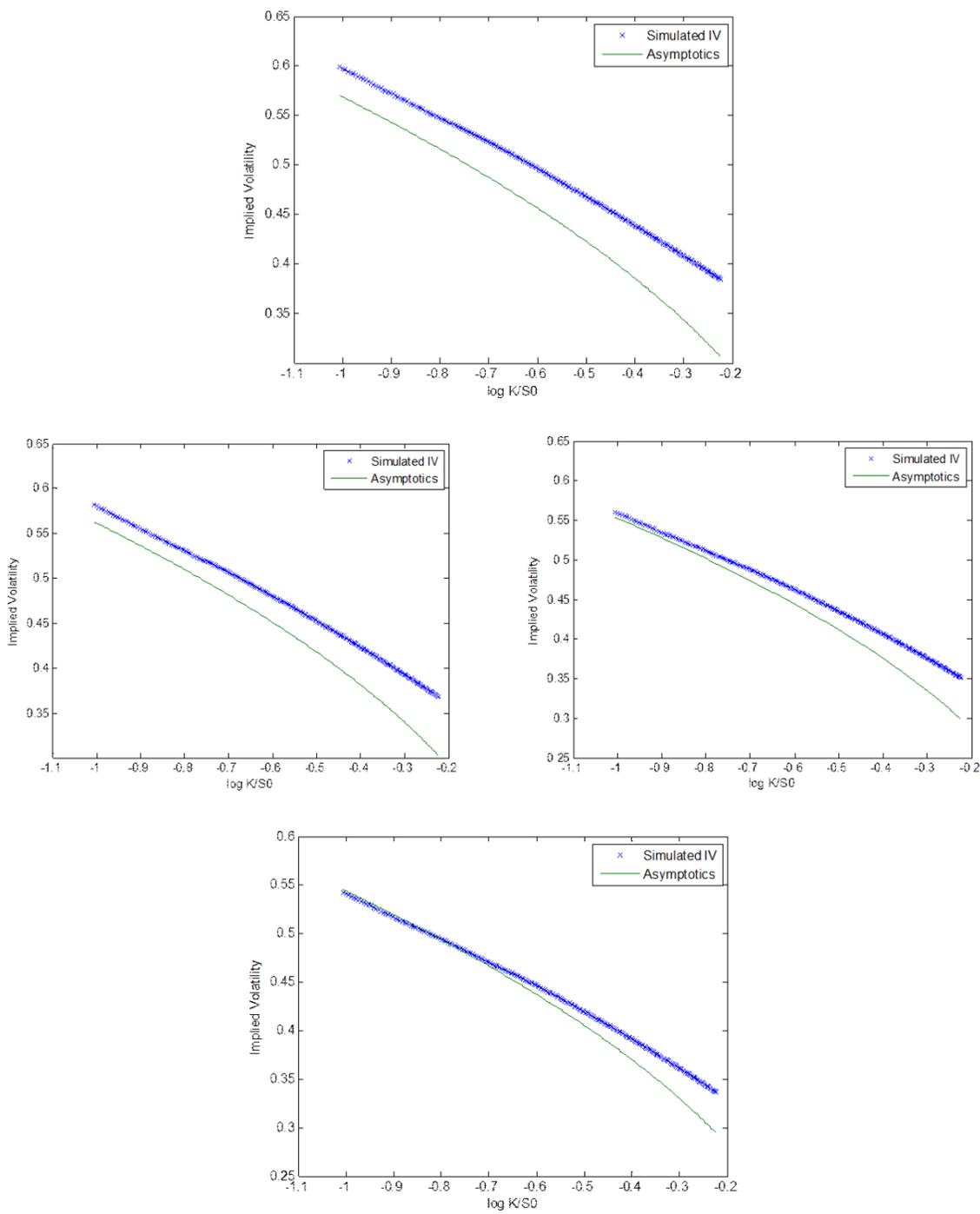


FIGURE 5. Figures 2a, 2b, 2c, 2d. IV for fOU model with $H = 0.51, H = 0.55, H = 0.60, H = 0.65$

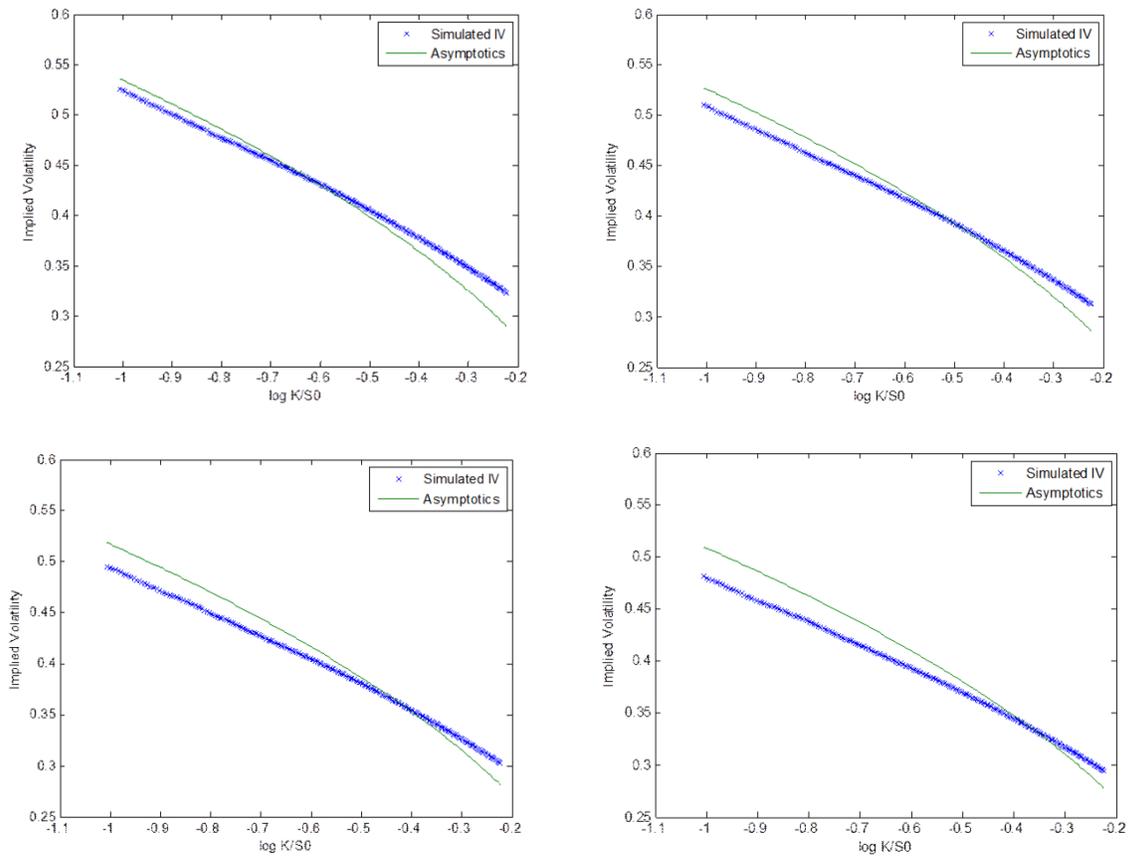


FIGURE 6. Figures 2e, 2f, 2g, 2h. IV for fOU model with $H = 0.70$, $H = 0.75$, $H = 0.80$, $H = 0.85$