Volatility: A hidden Markov process in financial time series

Zoltán Eisler[*]

Department of Theoretical Physics, Budapest University of Technology and Economics, Budafoki út 8., H-1111, Budapest, Hungary

Josep Perelló[†] and Jaume Masoliver[‡]

Departament de Física Fonamental, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain

Volatility characterizes the amplitude of price return fluctuations. It is a central magnitude in finance closely related to the risk of holding a certain asset. Despite its popularity on trading floors, volatility is unobservable and only the price is known. Diffusion theory has many common points with the research on volatility, the key of the analogy being that volatility is a time-dependent diffusion coefficient of the random walk for the price return. We present a formal procedure to extract volatility from price data by assuming that it is described by a hidden Markov process which together with the price forms a two-dimensional diffusion process. We derive a maximum-likelihood estimate of the volatility path valid for a wide class of two-dimensional diffusion processes. The choice of the exponential Ornstein-Uhlenbeck (expOU) stochastic volatility model performs remarkably well in inferring the hidden state of volatility. The formalism is applied to the Dow Jones index. The main results are that (i) the distribution of estimated volatility is lognormal, which is consistent with the expOU model, (ii) the estimated volatility is related to trading volume by a power law of the form $\sigma \propto V^{0.55}$, and (iii) future returns are proportional to the current volatility, which suggests some degree of predictability for the size of future returns.

I. INTRODUCTION

Volatility measures the amplitude of return fluctuations, and it is one of the central quantities in finance [1]. Investors sometimes place even greater emphasis on the level of volatility than on the market trend itself. The reason for this is mainly that the risk of holding an asset is classically associated with its volatility [2]. The theoretical framework used to quantify aspects of price fluctuations has many common points with areas of physics dealing with noisy signals. The research of random diffusion aims to describe the dynamics of particles in random media and its methods have been applied to a large variety of phenomena in statistical physics and condensed matter [3]. Time series describing solar flares, earthquakes, the human heartbeat, or climate records show strong correlations, multiscaling, non-Gaussian statistics, and self-organized behavior [4–7]. These are properties also observed in financial time series where volatility is considered to play a key role [1,8].

The picture that prices follow a simple diffusion process was first proposed by Bachelier in 1900 [9]. Later in 1959, the physicist Osborne introduced the geometric Brownian motion and suggested that volatility can be viewed as the diffusion coefficient of this random walk [10]. The simplest possible assumption—that it is a time-independent constant—lies at the heart of classical models such as the Black-Scholes option pricing formula [11]. More recently, it has become widely accepted that such an assumption is inadequate to explain the richness of the markets' behavior [1].

Instead, volatility itself should be treated as a random quantity with its own particular dynamics.

Among its most relevant properties [1,8,12–14], volatility is the responsible for the observed clustering in price changes. That is, large fluctuations are commonly followed by other large fluctuations and similarly for small changes [1,8]. Another related feature is that, in clear contrast with price changes which show negligible autocorrelations, volatility autocorrelation is still significant for time lags longer than a year [8,12,13]. Most of these studies introduce a subordinated process which is associated with the volatility in one way or another [15–26].

The main obstacle of an appropriate analysis of volatility is that it is directly unobservable. As we have mentioned, volatility provides important information to traders, but it is very unclear how reliable the estimates of such a hidden process can be. Investors use several proxies to infer the level of current asset volatility. The most common ways are (i) to make it equivalent to the absolute value of return changes, (ii) to assume a proportional law between volatility and market volume [26–29], and (iii) to use the information contained in option prices and obtain the so-called "implied volatility," which, in fact, corresponds to the *market's belief of volatility* [1].

This already raises the following questions: What process is a proper model of volatility and how does one adjust the possible parameters to describe various stocks and markets? Among the possible candidates, multifractals [16–20] and stochastic volatility models [15,21–26] are the most promising. The fact that volatility is directly unobservable makes it even more difficult to get a conclusive answer about the best model. This double challenge has deserved attention from the most diverse disciplines that look at financial markets. All of them have converged on a description of what is technically known as realized volatility and on methodologies for

———————
[*]eisler@maxwell.phy.bme.hu
[†]josep.perello@ub.edu
[‡]jaume.masoliver@ub.edu

reconstructing volatility paths. Several procedures for reconstructing volatility have already been presented being more or less dependent on the volatility model chosen. During the last decade mathematics, econometrics, and finance journals have published a large number of articles devoted to this issue (see, e.g., [15,30–41]).

Our research presents an alternative procedure to estimate volatility from the price dynamics only. Like most of the papers in the literature, price dynamics is represented by a two-dimensional diffusion process: one dimension for price and a second dimension for a volatility described by a hidden Markov process. Our case estimates the volatility subordinated-time series through maximum likelihood, but in contrast with other studies, having fixed in advance the parameters of the model.

We have decided to focus on a particular stochastic volatility model that is able to circumvent both mathematical and computational difficulties: the exponential Ornstein-Uhlenbeck volatility model [21,22], although the same procedure can be used for a larger class of models. As its name indicates, the model assumes that the logarithm of the volatility follows an Ornstein-Uhlenbeck process: that is, a mean reverting process with linear drift. The resulting model is capable of reproducing the statistical properties of the financial markets fairly well [21].

The paper is organized as follows. In Sec. II we outline the general stochastic volatility framework and more specifically the exponential Ornstein-Uhlenbeck model. In Sec. III we present a maximum-likelihood estimator for a wide class of stochastic volatility models. For the case of the exponential Ornstein-Uhlenbeck (expOU) model we present Monte Carlo simulations to show that it performs remarkably well in inferring the hidden state of the volatility process. In Sec. IV the procedure is applied to the Dow Jones Industrial Average. Conclusions are drawn in Sec. V, and some more technical details are left to the Appendixes.

II. STOCHASTIC VOLATILITY MODELS

The geometric Brownian motion (GBM) [10] is the most widely used model in finance. In this setting the asset price $S(t)$ is described through the following Langevin equation (in the Itô sense):

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW_1(t), \tag{1}$$

where σ is the volatility, assumed to be constant, μ is some deterministic drift indicating an eventual trend in the market, and $W_1(t)$ is the Wiener process. We define the zero-mean return $X(t)$ as

$$X(t) = \ln[S(t + t_0)/S(t_0)] - \langle \ln[S(t + t_0)/S(t_0)] \rangle, \tag{2}$$

where the symbol $\langle \cdots \rangle$ designates the average value and t_0 is the initial time which is usually set to be zero. In terms of $X(t)$ the GBM is simply written as

$$dX(t) = \sigma dW_1(t). \tag{3}$$

However, especially after the 1987 market crash, compelling empirical evidence has become available that the assumption of a constant volatility is doubtful [8]. Nor is volatility a deterministic function of time as one might expect on account of the nonstationarity of financial data, but a random quantity [22].

In the most general setting one therefore assumes that the volatility σ is a given function of a random process $Y(t)$:

$$\sigma(t) = f(Y(t)). \tag{4}$$

Most stochastic volatility (SV) models assume $Y(t)$ is also a diffusion process that may or may not be correlated with price. The main difference between various models is only the parametrization of this scheme. In a general notation the zero-mean return $X(t)$ defined above is described by the following set of stochastic differential equations:

$$dX(t) = f(Y(t))dW_1(t) \tag{5}$$

$$dY(t) = -g(Y(t))dt + h(Y(t))dW_2(t), \tag{6}$$

where $dX = dS/S - \langle dS/S \rangle$ and f, g, and h are given functions of $Y(t)$. As shown in Eq. (4), $f(Y(t))$ corresponds to the volatility—i.e., the amplitude of return fluctuations. However, since $f(x)$ is usually chosen to be a monotonically increasing function, it is not misleading to think of Y as a measure of volatility. Thus, as far as there is no confusion, we will refer to the process $Y(t)$ as "volatility" as well. On the other hand, the function $g(Y(t))$ describes a reverting force that drives the volatility toward the so-called "normal level." This force brings the volatility process to a stationary regime for long-time horizons, and the normal level is related to the average volatility in that limit. Finally, the subordinated process $Y(t)$ may have a nonconstant diffusion coefficient defined in terms of the function $h(Y(t))$ which is called the volatility-of-volatility ("vol of vol"). The functions g and h fully describe the volatility process. The resulting dynamics is comparable with the one described by a Gaussian particle trapped in a potential well $V(y)$ whose associated force is $-g(y)$, where $g(y) = V'(y)$. In finance one typically proposes convex potentials with only one minimum whose value is related to the normal level of the volatility.

In what follows we will mostly work with one particular SV model, the expOU model, which follows from the substitutions

$$f(x) = me^x, \quad g(x) = \alpha x, \quad h(x) = k,$$

that is,

$$dX(t) = me^{Y(t)}dW_1(t), \tag{7}$$

$$dY(t) = -\alpha Y(t)dt + kdW_2(t). \tag{8}$$

Note that in this model the process $Y(t)$ is precisely the logarithm of the volatility, or "log-volatility" for short. The main statistical properties of the model are thoroughly discussed in Ref. [21]. We simply recall that the stationary distribution of the process $Y(t)$ is a Gaussian (i.e., a lognormal distribution for σ) with zero mean and variance β:

$$p(y) = \frac{1}{\sqrt{2\pi\beta}} \exp(-y^2/2\beta), \qquad (9)$$

where

$$\beta \equiv k^2/2\alpha. \qquad (10)$$

III. VOLATILITY ESTIMATION

A. Wiener measure and volatility estimation

Let **X** and **Y** denote a simultaneous realization of the variables $X(\tau)$ and $Y(\tau)$ in the time interval $t-s \le \tau \le t$. Omitting all **Y**-independent terms, we show in Appendix A that the probability density (likelihood) of such a realization is approximately given by

$$\ln \mathbb{P}(\mathbf{X}, \mathbf{Y}) \simeq -\frac{1}{2} \int_{t-s}^{t} \left[\frac{\dot{X}(\tau)}{f(Y(\tau))} \right]^2 d\tau$$
$$-\frac{1}{2} \int_{t-s}^{t} \left[\frac{\dot{Y} + g(Y(\tau))}{h(Y(\tau))} \right]^2 d\tau + \cdots. \qquad (11)$$

Before proceeding further, we will discuss the meaning of this expression. We first note that Eq. (11) has to be understood in the sense of generalized functions [42] since the Wiener process is only differentiable in this sense and, hence, $\dot{X}(t)$ and $\dot{Y}(t)$ do not exist as ordinary functions and Eq. (11) is just a symbolic expression. Nevertheless, the formula is still valid when the integral and the derivatives therein are discretized with arbitrary small time steps, a requirement that is indeed necessary for numerical computations.

Let us now see some qualitative properties of Eq. (11). The first summand measures the fluctuations of the zero-mean return with respect to the volatility, $[\dot{X}(\tau)/f(Y(\tau))]^2$, and their contribution to the likelihood (probability density) of a given return realization. Note that the higher this contribution is, the lower those "relative" fluctuations are. In the same fashion, the second summand in Eq. (11) measures the fluctuations of the volatility process $Y(t)$ with respect to the vol of vol $h(Y)$, although in this case these fluctuations are gauged with the mean reverting force $-g(Y)$. As before, the lower these fluctuations, the higher their contribution to the log-likelihood (11).

While Eqs. (5) and (6) represent a joint model for return and volatility, the stock market data only include recordings of the return process X. The Y process and, hence, the volatility $f(Y)$ must be inferred indirectly in a Bayesian fashion through Eq. (11). Indeed, the conditional probability density that the realization of the hidden Y process is **Y**, given that the observed return is **X**, reads

$$\ln \mathbb{P}(\mathbf{Y}|\mathbf{X}) = \ln \mathbb{P}(\mathbf{X}, \mathbf{Y}) - \ln \mathbb{P}(\mathbf{X}). \qquad (12)$$

In consequence, *we can find the maximum-likelihood sample path of the (hidden) volatility process by maximizing* Eq. (12) *with respect to* **Y** [recall that **Y** is a realization of $Y(\tau)$ in the interval $t-s \le \tau \le t$]. Since the second summand of Eq. (12)

is independent of **Y**, we can neglect $\mathbb{P}(\mathbf{X})$ in this maximization process. Therefore, the maximization of $\ln \mathbb{P}(\mathbf{Y}|\mathbf{X})$ yields the same result as that of $\ln \mathbb{P}(\mathbf{X}, \mathbf{Y})$.

Note that, besides the specification of the stochastic volatility model (that is, the explicit forms of f, g, and h), the only free parameter is s: the duration of past return data to take into account. After substituting the observed return history as **X**, we will obtain by maximum likelihood the quantity

$$\hat{\mathbf{Y}} = \text{argmax}_{\mathbf{Y}} \ln \mathbb{P}(\mathbf{Y}|\mathbf{X}) = \text{argmax}_{\mathbf{Y}} \ln \mathbb{P}(\mathbf{X}, \mathbf{Y}). \qquad (13)$$

We should mention that similar maximization problems have been studied in the context of hidden Markov models, where this procedure is called "decoding." When the state space (the number of possible X and Y values) is finite, the optimization can be done exactly by the Viterbi algorithm [43], while there has been limited success in the continuous case [44]. A similar technique has been applied to the forecasting of volatility assuming a binomial cascade model, which, unlike stochastic volatility models, has a finite state space (see Ref. [20]). More focused on the stochastic volatility, we should also mention the efforts based on Kalman or particle filtering, Bayesian inference, conditional likelihood, or Fourier methods among other similar techniques (see, e.g., Refs. [15,30–41]). Most of the cited works assume a specific model, although they never focus on the expOU model. In addition, they are mainly worried about intraday (high-frequency) data, while we are here focused on reproducing daily (low-frequency) data. The superposition model of Lévy Ornstein-Uhlenbeck processes for σ^2, which have been crafted to describe high-frequency (intraday) data, deserves special attention since it provides a power-law slow decay for the volatility autocorrelation [15,32]. All the techniques mentioned above provide an efficient way to reproduce the volatility path, but in contrast to our case, the method also serves to estimate the parameters of the model. We have decided to provide the parameters beforehand, thus using an independent way of estimating them (see Sec. IV A). An accurate comparison between our method and others should be left for future research. We may look for a way to implement parameter estimation in our method.

As we have already stated, our main objective is to design a method able to filter the Wiener noise $dW_1(t)$ out of Eq. (5) and thus to obtain a reliable estimate $\hat{Y}(t)$ of the hidden volatility process $Y(t)$. During the rest of the paper we will denote with a hat the estimated value of a certain variable. Thus, in the current case, the hat denotes that we are using an estimation of Y(t), not it's "true" value. The method, an extension of the deconvolution procedure previously presented in [21], has basically the following five steps.

(i) We simulate a random sample path of the Wiener process for $t-s \le t \le \tau$ and consider it to be an estimation of $\widehat{dW_1}(t)$.

(ii) Then a surrogate realization of **Y** is generated as

$$\hat{Y}_s(\tau) = f^{-1}\left(\left|\frac{dX(\tau)}{\widehat{dW_1(\tau)}}\right|\right), \tag{14}$$

where $t - s \leq \tau \leq t$. Note that this equation requires that $f(x)$ be invertible, which implies that $f(x)$ be chosen to be a monotonic function.

(iii) Substitute $\hat{\mathbf{Y}}_s$ and $\mathbf{X}$ into Eq. (11) to calculate the log-likelihood of this realization.

(iv) Iterate (i)–(iii) for I steps, keep the highest likelihood random realization (the conditional median), and assume this to be the proper estimate $\hat{Y}(t)$.

(v) The estimate of the hidden process at time t is then $\hat{Y}(t)$. The estimate of the volatility is given by

$$\hat{\sigma}(t) = f(\hat{Y}(t)). \tag{15}$$

B. Interpretation of the method

Let us further elaborate the meaning of such an estimate. In finance, the volatility is often identified with the absolute value of returns variations. Indeed, as a first approximation, we can replace the noise term in Eq. (5) by its expected value and write

$$\sigma(t) \approx \frac{|dX(t)|}{\langle|dW_1(t)|\rangle}, \tag{16}$$

which shows that the volatility is approximately proportional to the absolute returns. Equation (5) can be thus thought of as a first approximation toward estimating volatility. Our method, based on the maximization of Eq. (12), takes this estimation two steps further. In effect, the first step was taken in Ref. [21] where we replaced the average $\langle|dW_1(t)|\rangle$ by a simulated sample path. We are now taking a second and more refined step in which we are not only replacing the Wiener noise by a random simulation, but in addition, we perform the maximum-likelihood method described by items (i)–(v).

Thus we are basically separating the observed returns $dX(t)$ into two sources: $\sigma(t)$ and $dW_1(t)$. To do this, we have first considered a specific form of stochastic volatility. Second, we have taken the driving Wiener noises $dW_1(t)$ and $dW_2(t)$ appearing in Eqs. (5) and (6) to be uncorrelated. Finally, we have assumed that $\sigma(t)$ is approximately constant over the time step during which we numerically evaluate the derivatives $\dot{Y}(t)$ and $\dot{X}(t)$ appearing in Eq. (11). We incidentally note that if $h(x) = 0$—the vol of vol is equal to zero—then the stationary solution of Eq. (6) is $Y(t) \equiv 0$. Thus the model reduces to the Wiener process in which the volatility is constant and absolute returns are uncorrelated [9].

C. Performance of the estimator

In order to test the performance of the estimator, we simulate the expOU process by using Eqs. (7) and (8) with the realistic parameters obtained in Sec. IV. The relationship between the simulated value of the log-volatility $Y(t)$ and its
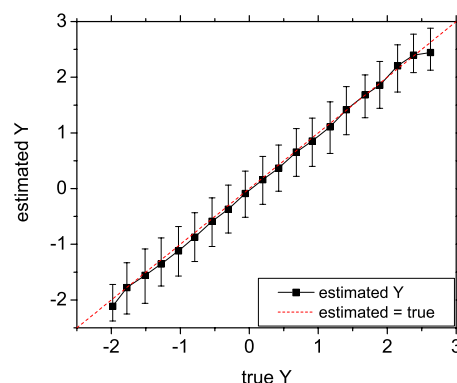

FIG. 1. (Color online) Estimated log-volatility $\hat{Y}$ as a function of the actual log-volatility Y taken from 2×10^5 simulations of the expOU model. Reconstruction used the last $s = 10$ values of returns and $I = 10^5$ iterations. The error bars represent the 25% and 75% quantiles of the distribution estimated volatility.

estimate $\hat{Y}(t)$ is given in Fig. 1. The two quantities agree within error bars, so we may state that

$$Y(t) \approx \hat{Y}(t) \quad \text{(in a mean square sense)}.$$

In what follows we will always use $s = 10$ days of past data and $I = 10^5$ iterations for maximization. The time step for discretization will be $\Delta t = 1$ day. For random optimization, the necessary number of iterations I grows very fast, perhaps exponentially, with $s/\Delta t$. The values of s and Δt cited here were chosen to keep the task computationally feasible. As for the value of I, in this case an increase to $I = 10^6$ does not improve the estimates noticeably. The estimates generated with this parameter set have a negligible bias, and they can efficiently distinguish between low- and high-volatility periods (see Appendix B for a discussion on the robustness and possible bias of the procedure).

An additional verification of the good performance of our estimate is shown in Fig. 2, where we give the actual sample path of Y for a single realization of the expOU model together with the estimated $\hat{Y}$. As we can see, the estimate follows the true log-volatility $Y(t)$ very closely.
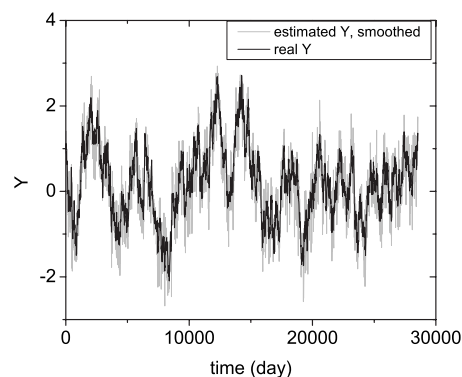

FIG. 2. Estimated and actual volatility for a typical sample path of the expOU model. The estimated values were smoothed by five-neighbor averaging to reduce noise.
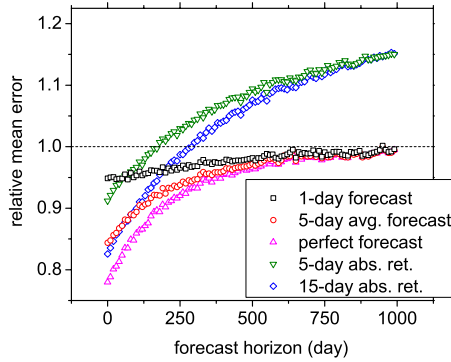
FIG. 3. (Color online) Median forecasting error $E(h)$ of absolute returns calculated according to Eq. (18) by five methods. The errors were normalized by the lowest level of error achievable with the assumption of a time-independent constant volatility (horizontal dashed line). The result was averaged over 150 000 independent realizations of the expOU process; parameters were chosen as outlined in Sec. IV A.

D. Volatility forecasting

Another possible approach to the utility of the expOU model is to evaluate its forecasting performance. One can for example use the mean error of absolute returns $|dX(t+h)|$ for a given Monte Carlo–simulated path:

$$E(h) = \langle ||\widehat{dX}(t+h)| - |dX(t+h)|| \rangle, \qquad (17)$$

where $|\widehat{dX}(t+h)|$ is our estimate for the absolute return $|dX(t+h)|$ given an estimate $\hat{Y}(t)$, while h is the time horizon (number of days) to forecast ahead. We recall that for our estimates we take the conditional median of our Monte Carlo simulations. Thus for the expOU model we have

$$|\widehat{dX}(t+h)| = M[|dW_1(t+h)|]m \exp[\hat{Y}(t)\exp(-\alpha h)], \qquad (18)$$

where $M[|dW_1(t+h)|]$ denotes the median of the absolute value of the Wiener increment. Figure 3 compares five methods to forecast $|dX(t+h)|$ using only information available before time t. For clarity we also give the titles of the five methologies used in parentheses ($\cdots$).

(i) The median of the last 5 absolute returns. This method clearly has predictive power for $h < 150$ days ("5-day abs. ret.").

(ii) The median of the last 15 absolute returns. The longer averaging period brings a substantial improvement for short-term forecasts, but not for long-term ("15-day abs. ret.").

(ii) Equation (18) with $\hat{Y}(t) = Y(t)$, which is the true value in the underlying simulation of the expOU model. This gives a substantial decrease in forecasting error, which persists up to $h \sim 500$ days. This curve is also a theoretical lower bound for error achievable by the expOU model in any time series: Here the underlying data are perfectly described by the expOU model, and the parameters and $Y(t)$ are known "perfectly." Neither of these usually happen in real data, and so there one expects a worse performance ("perfect").

(iv) Equation (18) with $\hat{Y}(t)$ estimated by the reconstruction procedure of Sec. III. This estimator does not perform well, due to noise in the optimization procedure ("1-day forecast").

(v) Equation (18) with $\hat{Y}(t)$ estimated by the reconstruction procedure of Sec. III, then averaged for the last 5 days. Such averaging greatly decreases noise, and the accuracy of the forecast is improved for short times (cf. Fig. 2) ("5-day forecast").

Finally, note that for $h \sim 500$ days all estimators based on the expOU model lose any information included in $\hat{Y}(t)$ and converge to the same error level. This is consistent with real data (cf. Sec. IV D).

IV. APPLICATION TO STOCK MARKET DATA

In this section we present an application of the method to actual stock market data. We analyze the Dow Jones Industrial Average (DJIA) index in the period 1 Jan. 1900–26 May 2006, a total of 29 038 days. In order to work with zero-mean returns, the mean return was subtracted from the actual data. Trading volumes for the index are only available for the period 1 April 1993–26 May 2006, a total of 3 375 days.

A. Parameter estimation

We recall that in the estimation procedure presented here one necessarily needs to assume a theoretical model for the volatility. Having done this, the next step is to estimate the parameters involved in the model chosen. For the expOU model, Eqs. (7) and (8), these parameters are m, k, and α.

Before proceeding further we remark that time increments in real data have a finite size since the market always works on discrete times (for daily data the minimum time increment is 1 day). Thus, in practice, the (infinitesimal) return variation $dX(t) = X(t+dt) - X(t)$ corresponds to a (finite) return increment $\Delta X(t) = X(t+\Delta t) - X(t)$ where Δt is the time step between two consecutive ticks. Also, the Wiener differentials $dW(t)$ correspond, in a mean-square sense [45], to the increments

$$\Delta W(t) \approx \varepsilon(t)\sqrt{\Delta t}, \qquad (19)$$

where $\varepsilon(t)$ is a Gaussian process with zero mean and unit variance [45]. In the present case our time step has a fixed width and is equal to $\Delta t = 1$ day.

Coming back to the estimation of parameters, we show in Appendix C that

$$\ln m \approx (\gamma + \ln 2)/2 + \langle \ln(|\Delta X|/\sqrt{\Delta t}) \rangle, \qquad (20)$$

where $\gamma = 0.5772\ldots$ is the Euler constant. Taking into account that the third summand can be evaluated from data, we see that Eq. (20) provides a direct estimation of m.

On the other hand, if the expOU model is appropriate, then the empirical estimate $\hat{Y}(t)$ of the hidden volatility $Y(t)$ should also be a Gaussian process with a stationary distribution of zero mean and variance given by $\beta = k^2/2\alpha$ [see Eq. (10)]. As shown in Fig. 4, if one takes $\beta = 0.61 \pm 0.05$, the distribution of $\hat{Y}(t)$ is Gaussian and coincides with the theoretical distribution of $Y(t)$ given by Eq. (9). The assumption
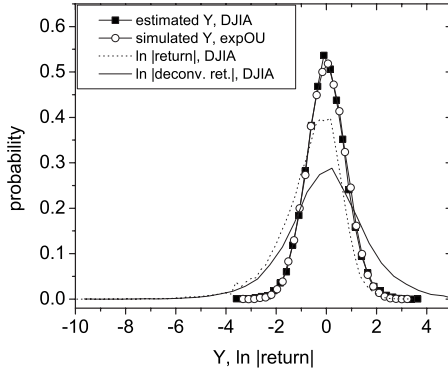
FIG. 4. A comparison of estimates of log-volatility for Dow Jones. Black boxes ("estimated Y, DJIA") represent our maximum-likelihood method applied to empirical data. Open circles ("simulated Y, expOU") represent the distribution of the simulated sample path of the log-volatility, assuming that $Y(t)$ follows the expOU model. We also plot the empirical distributions of two estimates: the dashed line ("ln [return], DJIA") was obtained through Eq. (16) and the solid line ("ln [deconv. ret], DJIA") via Eq. (21).

of a Gaussian distribution for our estimate is robust and it holds for a wide range of parameters.

To fully specify the model we have to obtain the parameter α. We have chosen the value found in Ref. [21] which was obtained in order to capture the long-range correlations of the volatility, at least up to 500 days. The model is able to provide the appropriate long-range behavior with an infinite sum of exponentials, but it is also true that it does not provide a pure power-law decay like the models in Refs. [15,17]. In any case, we have seen in Ref. [21] that at least for daily data our approach is satisfactory. Our final set of parameter estimates is thus

$$m = (7.5 \pm 0.5) \times 10^{-3} \ \text{days}^{-1/2},$$

$$\alpha = (1.82 \pm 0.03) \times 10^{-3} \ \text{days}^{-1},$$

and

$$k = (4.7 \pm 0.3) \times 10^{-2} \ \text{days}^{-1/2}.$$

The errors were determined based on Fig. 4, similarly to the error of β. In this parameter range the distributions of the estimated $\hat{Y}(t)$ and simulated $Y(t)$ series agree well. The results reported throughout this paper are insensitive to the misspecification of these parameters, even beyond these error bars (see also Appendix B).

The distributions of log-volatility are compared in Fig. 4 for four cases: our maximum-likelihood procedure applied to Dow Jones, a simulation of the expOU model, the simple estimate of Eq. (16), and the deconvolution procedure introduced in Ref. [21] which can be written as

$$\hat{Y}_{\text{decon}}(t) = \ln \frac{1}{m} \left| \frac{dX(t)}{\widehat{dW_1}(t)} \right|, \tag{21}$$

where $W_1(t)$ is a simulation of the Wiener process. Note that this deconvoluted log-return estimator has indeed a Gaussian distribution, but with a larger variance as hinted in Fig. 4 in
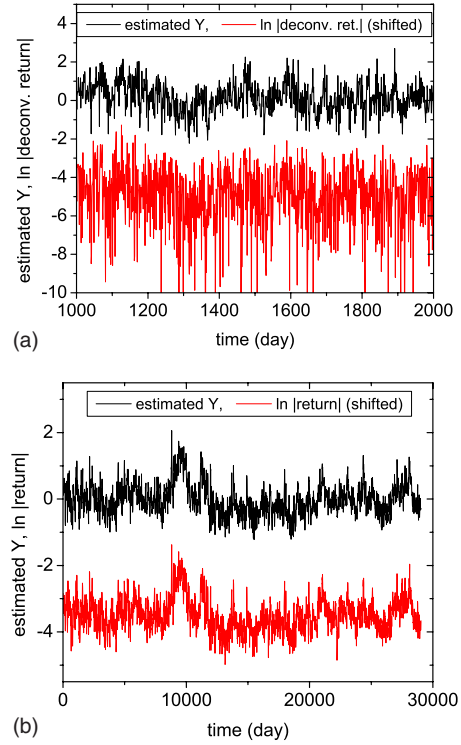

(a)

(b)

FIG. 5. (Color online) (Top) A comparison between the estimate $\hat{Y}(t)$ and $\hat{Y}_{\text{decon}}(t)$ for a typical 1000-day period of Dow Jones. These curves were not smoothed in order to show the substantial reduction of both the noise level and the asymmetry in $\hat{Y}(t)$ compared to $\hat{Y}_{\text{decon}}(t)$ with a random approximation of the noise term dW. (Bottom) The estimate $\hat{Y}(t)$ and the logarithm of absolute return variations for the whole sample of Dow Jones.

view of its wider density. We finally mention that the estimate for the log-volatility $\ln(\sigma/m) \approx \ln|dX|$ given by Eq. (16) shows a non-Gaussian and biased distribution as was also reported in Ref. [21]. This suggests that $\hat{Y}_{\text{decon}}$ is an appropriate "null model" to contrast with $\hat{Y}$. Both quantities are generated by dividing $dX(t)$ by the increments of a realization of the Wiener process and then taking the logarithm of the absolute value of this ratio. The difference lays in the fact that $\hat{Y}$ takes the realizations that satisfy a maximum-likelihood requirement while $\hat{Y}_{\text{decon}}$ takes a Wiener process realization that is purely random. In such a way, our method keeps the divisor correlated with dX and, as we will see next, it seems to conserve clustering and memory effects in the log-volatility time series.

B. Clustering and the estimated volatility

To support our claim that the technique presented is powerful enough in filtering the noise dW_1 out of returns we will give a visual comparison based on the following qualitative experiments.

Figure 5 (left) displays a comparison over a 1000-day time interval. One can observe there that the noise level in $\hat{Y}$ is substantially smaller than in $\hat{Y}_{\text{decon}}$, as also inferred from Fig. 4.
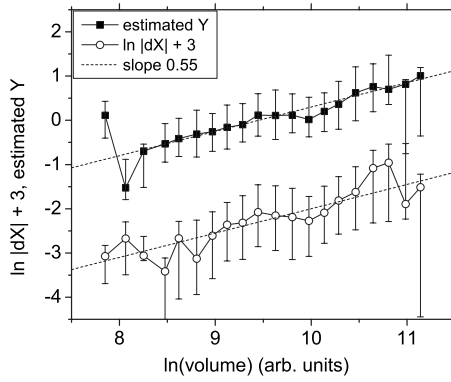
FIG. 6. Logarithm of daily absolute return and estimated log-volatility $\hat{Y}$ as a function of the daily volume. Days with similar volumes were binned for better visibility. The symbols represent the medians and the error bars the 25%–75% quantiles in the bins.

In order to show that such a correlation is responsible for suppressing large fluctuations in the ratio, we can perform a second experiment. Thus in Fig. 5 (right) we see a comparison between the logarithm of absolute returns variations, $\ln|dX|$, and the estimated volatility $\hat{Y}$. Note that the proper clustering of volatility becomes clearly visible.

C. Comparison with the trading volume

The hidden nature of the volatility process has been addressed by several authors [26–28]. For instance, Ref. [26] suggests that, instead of the volatility, a good estimate would be the square root of the daily trading volume—that is,

$$\mathrm{M}[\sigma(t)] \propto V(t)^{\alpha}, \qquad (22)$$

where $\alpha=0.5$ and $\mathrm{M}[\cdot]$ denotes the median. In Fig. 6 we show evidence that supports this assumption. Again, the first estimation for the volatility is $|dX(t)|$. Therefore, we regress $\ln|dX(t)|$ versus $\ln V(t)$ as shown in Fig. 6. In the same figure we also present the regression between the maximum-likelihood estimate $\hat{Y}(t)$ and $\ln V(t)$ which appears to be less noisy than the former regression in accordance with the smaller variance of $\hat{Y}(t)$ compared to that of $\ln|dX(t)|$. Nevertheless, the exponent $\alpha \approx 0.55$ is the same for both regressions. There have been similar [46], albeit controversial [47,48], findings for the price impact of single transactions. However, Eq. (22) does not yet imply that volatility is proportional to volume, only that *its typical value is* (i.e., the median). Fluctuations around the average behavior due to changes in liquidity might have a key role in the process [48].

D. Predictive power of volatility

From Eqs. (7) and (8), we know that for the expOU model a simple relationship can be given between $\ln|dX(t)|$ and $Y(t)$:

$$\ln|dX(t)| = \ln[m|dW_1(t)|] + Y(t).$$

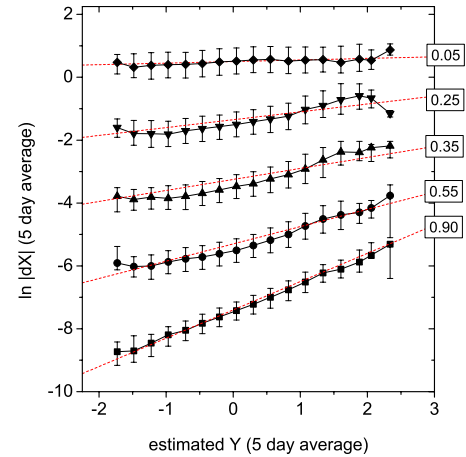Therefore, the conditional median of $\ln|dX(t)|$ given $Y(t)$ is

FIG. 7. (Color online) The proportionality between the estimated volatility and of the logarithm of absolute return variations. In order to decrease noise, five-day moving averages have been used. Numbers on the right indicate the slopes of the corresponding regression lines. Time shifts from bottom to top: $h=0$ days ($\blacksquare$), 5 days ($\bullet$), 20 days ($\blacktriangle$), 100 days ($\blacktriangledown$), and 1000 days ($\blacklozenge$). Days with similar absolute returns were binned for better visibility. The symbols represent the medians and the error bars the 25%–75% quantiles in the bins.

$$M[\ln|dX(t)| \, |Y(t)\,] = \mathrm{const} + Y(t). \qquad (23)$$

We point out that this relationship implies some degree of predictability of the absolute changes in return through their median if one knows the current value of the log-volatility $Y(t)$. We test Eq. (23) for real data and with $Y(t)$ replaced by its estimate $\hat{Y}(t)$. As shown by the bottom curve of Fig. 7, the slope of the linear regression between $M[\ln|dX(t)| \, |Y\,]$ and $\hat{Y}(t)$ is not equal to 1—as would have been implied by Eq. (23)—but 0.9, which still suggests strong predictive power.

Recall that the minimum time step of the empirical data used is 1 day. Hence, Eq. (23) implies the prediction of tomorrow's return based on today's volatility and return. We now want to extend the prediction horizon. To this end we generalize Eq. (23) and propose the following ansatz:

$$M[\ln|dX(t+h)| \, |\hat{Y}(t)\,] = \mathrm{const} + \gamma(h)\hat{Y}(t), \qquad (24)$$

where $h=0,1,2,\ldots$. In Fig. 7 we test this ansatz for several values of the horizon: $h=0$, 5, 20, 100, and 1000 days. We find that the slope $\gamma(h)$ is a decreasing function from the value $\gamma(0)=0.9$ to practically zero when $h=1000$ days which means a complete loss of memory. Note that, when $h=100$ trading days, we have $\gamma=0.25$, still implying a slight degree of prediction after approximately 5 months, which is of the same order of magnitude than the DJIA characteristic time scale, $1/\alpha \sim 500$ days, for the relaxation of the volatility [21].

V. CONCLUSIONS

The volatility is a crucial quantity for financial markets since it provides a measure of the amplitude of price fluctuations. Traders try to follow carefully the level of volatility

because it gives the perception of the risk associated with any asset. Although volatility was originally conceived to be the diffusion coefficient of the price return random walk, there is compelling evidence not to consider it a constant, but a subordinated random process. The framework is analogous to that of random diffusion processes which have been applied to a large variety of phenomena in statistical mechanics and condensed matter physics.

The main obstacle to get a better knowledge of the volatility's nature is that it is not directly observed. In fact, this is precisely the motivation behind the present research. Our main objective has been to develop a tool which visualizes the sample path of volatility. The procedure derives a maximum-likelihood estimate assuming that the volatility is a hidden Markov process. To do so, one needs also to assume a specific model for the volatility. We have chosen a class of two-dimensional diffusions commonly known as stochastic volatility models, where the volatility acts as a diffusion particle trapped in a potential well. We have focused on the expOU model and obtained promising results, especially for three reasons: (i) the model is computationally feasible, (ii) its parameters can be easily obtained and fit the data reasonably well, and (iii) the distribution of the estimated volatility is log-normal, which is consistent with the assumed expOU model.

We have shown for the Dow Jones index daily data that the sample path of our estimated volatility improves other estimates. We have compared our estimation with a rather typical one which identifies volatility with absolute return changes. Our estimation is able to remove the existing bias in the stationary distribution of volatility while still preserving the clustering in volatility time series. We have also studied the estimate of volatility that deconvolutes the return by the simulation of a random Wiener path [21]. This last procedure also provides a Gaussian distribution for the log-volatility, albeit the distribution has too fat tails and pays the price of losing clustering and memory in the volatility time series. Our new procedure is in fact a more sophisticated variant of this estimate since it filters out Wiener realizations via maximum likelihood. The estimate drastically reduces the noise in the volatility path, thus preserving data clustering. In this way we have thus proposed an alternative method to those already provided by the mathematical finance literature (see, e.g., [15,30–41]). An accurate comparison between our method and others should, however, be left for future research. We may even look for a way to implement parameter estimation in our method.

The median of the estimated volatility has also been related to the trading volume by the power-law expression $M[\sigma] \propto V^{0.55}$. A link between volatility and trading volume has been previously mentioned in different studies; however, our estimate is again capable of providing a less noisy regression. We must, indeed, stress the fact that this does not imply that volatility is proportional to a power of the volume, but only that its typical value is and that fluctuations around the average might play an important role.

We have also seen that current returns are proportional to the estimated volatility, as otherwise expected. However, the main novelty is that we have observed how future returns are proportional to current volatility and their predictability di-minishes monotonically with the number of time steps ahead. This last finding implies that our estimation method can be applied to predict the size of future returns with knowledge of current volatility.

As a final remark, we stress the fact that the technique herein presented can be applied to a variety of physical phenomena besides finance. One typical problem of this sort is provided by the Brownian motion inside a field of force in which inertial effects are not negligible [49]. In this situation the dynamics of the particle is described by a two-dimensional diffusion process $(X(t), V(t))$ representing the position and the velocity of the Brownian particle. The maximum-likelihood technique might provide a reliable estimate of the velocity in the case that, for instance, the only accessible experimental measures are the positions of the particle at wide time steps, so that a measure of the velocity—which implies the knowledge of two very close positions—is too noisy and unreliable.

APPENDIX A: DERIVATION OF THE LIKELIHOOD FUNCTION

In order to make the notation more compact, in this appendix the time dependence of the stochastic processes is mostly indicated as a lower index. A generic stochastic volatility model is defined as

$$dX_t = f(Y_t)dW_1(t), \tag{A1}$$

$$dY_t = -g(Y_t)dt + h(Y_t)dW_2(t). \tag{A2}$$

To explain the procedure it is more convenient to work with the discrete time version of the model. To this end, suppose that Δt is a small time step and that the driving noises in Eqs. (A1) and (A2) can be approximated by [cf. Eq. (19)]

$$dW_i(t) \approx \varepsilon_i(t)\sqrt{\Delta t} \quad (i=1,2), \tag{A3}$$

where $\varepsilon_i(t)$ are independent standard Gaussian processes with zero mean and unit variance. We remark that the approximation (A3) has to be understood in a mean-square sense [45]. The discrete time equations of the model thus read

$$X_t - X_{t-\Delta t} = f(Y_{t-\Delta t})\varepsilon_1(t-\Delta t)\sqrt{\Delta t}, \tag{A4}$$

$$Y_t - Y_{t-\Delta t} = -g(Y_{t-\Delta t})\Delta t + h(Y_{t-\Delta t})\varepsilon_2(t-\Delta t)\sqrt{\Delta t}, \tag{A5}$$

from which we get

$$\varepsilon_1(t-\Delta t) = \frac{X_t - X_{t-\Delta t}}{f(Y_{t-\Delta t})\sqrt{\Delta t}}, \tag{A6}$$

$$\varepsilon_2(t - \Delta t) = \frac{Y_t - Y_{t-\Delta t} + g(Y_{t-\Delta t})\Delta t}{h(Y_{t-\Delta t})\sqrt{\Delta t}}. \quad \text{(A7)}$$

For a given number of realizations, the probability of the set $\{X_\tau, Y_\tau\}$ ($\tau = t - \Delta t, t - 2\Delta t, \ldots, t - s$) can be easily obtained, as we will see next.

Let us denote the set of realizations as $\{\mathbf{X}, \mathbf{Y}\}$. Then the Markov property of the process ensures that one can decompose the joint probability density function (PDF) of this set as a chain of products between conditional probability densities. In consequence, the PDF of the whole sample path can be written as

$$\mathbb{P}(\{\mathbf{X}, \mathbf{Y}\}) = \mathbb{P}(X_{t-s}, Y_{t-s}) \prod_{\tau = t-s}^{t-\Delta t} \mathbb{P}(X_\tau, Y_\tau | X_{\tau-\Delta t}, Y_{\tau-\Delta t}), \quad \text{(A8)}$$

where the first term $\mathbb{P}(X_{t-s}, Y_{t-s})$ corresponds to the initial realizations of X and Y $s/\Delta t$ time steps far from the present time t; all the remaining terms of the form $\mathbb{P}(X_\tau, Y_\tau | X_{\tau-\Delta t}, Y_{\tau-\Delta t})$ are the conditional PDFs for transitions between consecutive steps:

$$(X_{\tau-\Delta t}, Y_{\tau-\Delta t}) \rightarrow (X_\tau, Y_\tau).$$

The logarithm of Eq. (A8) is

$$\ln \mathbb{P}(\{\mathbf{X}, \mathbf{Y}\}) = \ln \mathbb{P}(X_{t-s}, Y_{t-s}) + \sum_{\tau = t-s}^{t-\Delta t} \ln \mathbb{P}(X_\tau, Y_\tau | X_{\tau-\Delta t}, Y_{\tau-\Delta t}). \quad \text{(A9)}$$

On the other hand, from Eqs. (A4) and (A5) we realize that

$$\mathbb{P}(X_\tau, Y_\tau | X_{\tau-\Delta t}, Y_{\tau-\Delta t}) = |J| \mathbb{P}(\varepsilon_1(\tau - \Delta t), \varepsilon_2(\tau - \Delta t)),$$

where $|J|$ is the Jacobian of the transformation $(X_\tau, Y_\tau) \rightarrow (\varepsilon_1(\tau - \Delta t), \varepsilon_2(\tau - \Delta t))$ defined by Eqs. (A4) and (A5)—that is,

$$|J| = \frac{1}{f(Y_{\tau-\Delta t}) h(Y_{\tau-\Delta t}) \Delta t}.$$

But ε_1 and ε_2 are independent standard Gaussians, hence

$$\mathbb{P}(\varepsilon_1, \varepsilon_2) = (1/2\pi) \exp[-(\varepsilon_1^2 + \varepsilon_2^2)/2],$$

whence

$$\mathbb{P}(X_\tau, Y_\tau | X_{\tau-\Delta t}, Y_{\tau-\Delta t}) = \frac{1/(2\pi\Delta t)}{f(Y_{\tau-\Delta t}) h(Y_{\tau-\Delta t})}$$
$$\times \exp\left[-\frac{\varepsilon_1^2(\tau - \Delta t) + \varepsilon_2^2(\tau - \Delta t)}{2} \right]. \quad \text{(A10)}$$

Substituting Eqs. (A6) and (A7) into this equation and the result into Eq. (A9) we finally get

$$\ln \mathbb{P}(\{\mathbf{X}, \mathbf{Y}\}) = -\frac{s \ln(2\pi\Delta t)}{\Delta t} - \sum_{\tau = t-s}^{t-\Delta t} [\ln f(Y_{\tau-\Delta t}) + \ln h(Y_{\tau-\Delta t})]$$
$$+ \ln \mathbb{P}(X_{t-s}, Y_{t-s}) - \frac{1}{2} \sum_{\tau = t-s}^{t-\Delta t} \left[\frac{X_\tau - X_{\tau-\Delta t}}{f(Y_{\tau-\Delta t})\Delta t} \right]^2 \Delta t$$
$$- \frac{1}{2} \sum_{\tau = t-s}^{t-\Delta t} \left[\frac{Y_\tau - Y_{\tau-\Delta t}}{h(Y_{\tau-\Delta t})\Delta t} + \frac{g(Y_{\tau-\Delta t})}{h(Y_{\tau-\Delta t})} \right]^2 \Delta t. \quad \text{(A11)}$$

Let us briefly explain the origin of some of these contributions. The first summand comes from the normalization constant of the Gaussian distribution (A10). It appears in every conditional probability density, and this is the reason for the factor $s/\Delta t$, which is the number of time steps between $t - s$ and t. The resulting term does not depend on the realization, so that we can neglect it for a maximization with respect to $\mathbf{Y}$. The term also goes to $-\infty$ in the $\Delta t \rightarrow 0$ limit, which means that any individual realization has a probability measure zero.

The second summand is mostly the sum of the Jacobian transformations of each transition probability and depends on Y. Stochastic volatility models typically assume that these f and g [cf. Eqs. (A1) and (A2)] are continuous and monotonically increasing functions or even constants. For instance, in the expOU model we have $f(x) = m \exp(x)$ and $g(x) = k$. Because of this, we will also neglect this term in the maximization procedure. The contribution shifts the maximum at the excessive cost of adding more noise to the numerical computations. We can, however, look at the situation from another point of view. Ignoring this term is equivalent to omitting the Jacobian transformation between the two probability density measures [cf. Eq. (A10)]. In this way, we are stating that what we are really going to maximize is the probability of the realization of $\varepsilon_1(t)$ and $\varepsilon_2(t)$—instead of $Y(t)$—in terms of the past history of the process expressed by $\{\mathbf{X}, \mathbf{Y}\}$.

The term $\ln \mathbb{P}(X_{t-s}, Y_{t-s})$ is fixed by the initial conditions of the process. If we assume a known initial return X—which can be set to zero—and take a random Y_{t-s} following the stationary distribution $\mathbb{P}_{st}(Y_{t-s})$ given by Eq. (9), then $\mathbb{P}(X_{t-s}, Y_{t-s}) = \delta(X_{t-s} - X) P_{st}(Y_{t-s})$ and hence

$$\ln \mathbb{P}(X_{t-s}, Y_{t-s}) = \ln \mathbb{P}_{st}(Y_{t-s}) + \ln \delta(X_{t-s} - X). \quad \text{(A12)}$$

Had we taken another initial condition, the technique would have given equivalent results (we have checked this by using several initial distributions). For this reason and in order to improve the convergence of the maximum-likelihood estimate we have neglected also this contribution.

We therefore write

$$\ln \mathbb{P}(\{\mathbf{X}, \mathbf{Y}\}) \simeq -\frac{1}{2} \sum_{\tau = t-s}^{t-\Delta t} \left[\frac{X_\tau - X_{\tau-\Delta t}}{f(Y_{\tau-\Delta t})\Delta t} \right]^2 \Delta t$$
$$- \frac{1}{2} \sum_{\tau = t-s}^{t-\Delta t} \left[\frac{Y_\tau - Y_{\tau-\Delta t}}{h(Y_{\tau-\Delta t})\Delta t} + \frac{g(Y_{\tau-\Delta t})}{h(Y_{\tau-\Delta t})} \right]^2 \Delta t + \cdots. \quad \text{(A13)}$$

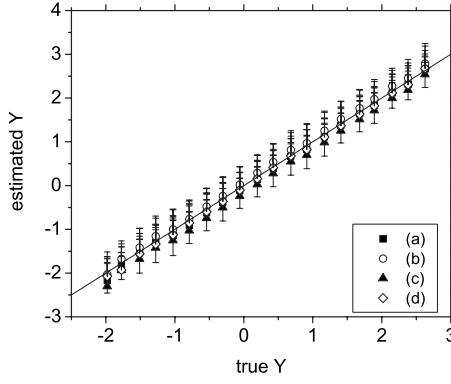We can represent this equation in the continuous-time frame-

FIG. 8. Estimated log-volatility $\hat{Y}$ as a function of the actual log-volatility Y taken from 29 038 simulations of the expOU model in four ways. Reconstruction uses the last $s=10$ values of returns and $I=10^5$ iterations. The error bars represent the 25% and 75% quantiles of the distribution estimated volatility.

work if Δt is sufficiently small and if $f(x)$, $h(x)$, and $g(x)$ are continuous. In such a case, Eq. (A13) yields the result given in Eq. (11).

APPENDIX B: ROBUSTNESS OF THE PROCEDURE

In order to show that our volatility estimation procedure is robust, we carry out four measurements along the lines of Sec. III C. We generate artificial time series of the same length as the DJIA data set, with the same parameters as therein: $m=(7.5\pm0.5)\times10^{-3}$ days$^{-1/2}$, $\alpha=(1.82\pm0.03)\times10^{-3}$ days^{-1}, and $k=(4.7\pm0.3)\times10^{-2}$ days$^{-1/2}$.

Then we reproduce Fig. 1 in several variations:

(a) In a way identical to the paper as shown in Fig. 1.

(b) Data generated as in (a), but for reconstructing $\hat{Y}$ we use the above parameters *minus* twice the error specified above.

(c) Data generated as in (a), but for reconstructing $\hat{Y}$ we use the above parameters *plus* twice the error specified above.

(d) In a way identical to (a), but the data is detrended by the Monte Carlo sample mean μ. We recall, however, that the process is driftless as given by Eq. (3). In this way we want to check whether detrending causes a systematic bias in the estimation.

The results are given in Fig. 8. The plot shows that either option gives the same width of the $\hat{Y}$ distribution and the bias introduced by either specification error is small. We have also run similar simulations, but with very different parameters values, and the procedure still provides consistent results similar to those given by Fig. 8.

Finally, a more detailed test along the lines of point (d) above can be performed (that is, a test on whether the detrending causes a bias in our estimation procedure). We thus have repeated the test (d) also including errors 10μ and 100μ. The procedure tolerates up to 10 times more detrending error than expected in the real data set. Only, when we
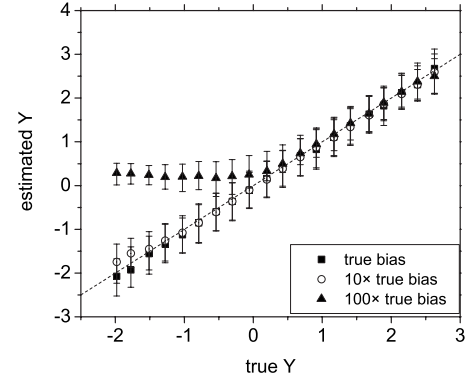

FIG. 9. Estimated log-volatility $\hat{Y}$ as a function of the actual log-volatility Y taken from 29 038 simulations of the expOU model in three different levels of "detrending error." Reconstruction used the last $s=10$ values of returns and $I=10^5$ iterations. The error bars represent the 25% and 75% quantiles of the distribution estimated volatility. Note that there only exists a clear deviation when the error is 100 times the true bias.

have about 100 times the error does the reconstruction show a strong upward bias for low-volatility periods as shown in Fig. 9. All these results imply that possible errors in the detrending procedure could affect our procedure only when any wrong specification of the drift is far outside the error domain involved in our DJIA data set.

APPENDIX C: DERIVATION OF Eq. (20)

We start from Eq. (7) which we write in the approximate form

$$\Delta X(t) \simeq me^{Y(t)}\Delta W(t);$$

thus, $\ln|\Delta X(t)| = m + Y(t) + \ln|\Delta W_1(t)|$ and, taking into account that $\langle Y(t)\rangle = 0$, we have

$$\langle\ln|\Delta X(t)|\rangle \simeq \ln m + \langle\ln|\Delta W_1(t)|\rangle. \tag{C1}$$

On the other hand, we know that $\Delta W_1(t) \approx \varepsilon\sqrt{\Delta t}$, where ε is a standard Gaussian variable [cf. Eq. (19)]. Hence

$$\langle\ln|\Delta W_1(t)|\rangle \approx \langle\ln|\varepsilon|\rangle + (\ln\Delta t)/2.$$

But

$$\langle\ln|\varepsilon|\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\varepsilon^2/2}\ln|\varepsilon|d\varepsilon,$$

which, after a simple change of variables inside the integral, yields [50,51]

$$\langle\ln|\varepsilon|\rangle = \frac{1}{2\sqrt{2\pi}}\int_{0}^{\infty}x^{-1/2}e^{-x/2}\ln xdx = -\sqrt{\pi/2}(\gamma + \ln 2),$$

where $\gamma = 0.5772\ldots$ is the Euler constant. Therefore,

$$\langle\ln|\Delta W_1(t)|\rangle \approx (\ln\Delta t)/2 - (\gamma + \ln 2)/2. \tag{C2}$$

Substituting Eq. (C2) into Eq. (C1) proves Eq. (20).

[1] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk* (Cambridge University Press, Cambridge, England, 2000).

[2] H. Markowitz, J. Financ. **7**, 77 (1952).

[3] D. Ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, 2000).

[4] M. Paczuski, S. Boettcher, and M. Baiesi, Phys. Rev. Lett. **95**, 181102 (2005).

[5] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).

[6] A. Bunde, S. Havlin, J. W. Kantelhardt, T. Penzel, J.-H. Peter, and K. Voigt, Phys. Rev. Lett. **85**, 3736 (2000).

[7] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).

[8] R. Cont, Quant. Finance **1**, 223 (2001).

[9] L. Bachelier, Ann. Sci. Ec. Normale Super. **III-17**, 21 (1900).

[10] M. Osborne, Oper. Res. **7**, 145 (1959).

[11] F. Black and M. Scholes, J. Polit. Econ. **81**, 637 (1973).

[12] A. Lo, Econometrica **59**, 1279 (1991).

[13] Z. Ding, C. W. J. Granger, and R. F. Engle, J. Empirical Finance **1**, 83 (1993).

[14] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Phys. Rev. E **60**, 1390 (1999).

[15] O. E. Barndorff-Nielsen, and N. Shephard, J. R. Stat. Soc. Ser. B (Stat. Methodol.) **63**, 167 (2001).

[16] A. Saichev and D. Sornette, Phys. Rev. E **74**, 011111 (2006).

[17] E. Bacry, J. Delour, and J. F. Muzy, Phys. Rev. E **64**, 026103 (2001).

[18] L. Calvet and A. Fisher, Rev. Econ. Stat. **84**, 381 (2002).

[19] T. Lux, Journal of Business and Economics Statistics (to be published).

[20] L. E. Calvet and A. J. Fisher, J. Financ. Econometr. **2**, 49 (2004).

[21] J. Masoliver and J. Perelló, Quant. Finance **6**, 423 (2006).

[22] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar, *Derivatives in Financial Markets with Stochastic Volatility* (Cambridge University Press, Cambridge, England, 2000).

[23] J. Perelló and J. Masoliver, Phys. Rev. E **67**, 037102 (2003).

[24] J. Perelló, J. Masoliver, and N. Anento, Physica A **344**, 134 (2004).

[25] J. Masoliver and J. Perelló, Int. J. Theor. Appl. Finance **5**, 541 (2002).

[26] A. A. Drăgulescu and V. M. Yakovenko, Quant. Finance **2**, 443 (2002).

[27] T. G. Andersen and J. Lund, J. Econometr. **77**, 343 (1997).

[28] W. A. Brock and B. D. LeBaron, Rev. Econ. Stat. **78**, 94 (1996).

[29] C. Lindberg, Math. Finance **16**, 549 (2006).

[30] V. Genon-Catalot, T. Jeantheau, and C. Larédo, Bernoulli **6**, 1051 (2000).

[31] O. Elerian, S. Chib, and N. Shephard, Econometrica **69**, 959 (2001).

[32] O. E. Barndorff-Nielsen and N. Shephard, J. R. Stat. Soc. Ser. B (Stat. Methodol.) **64**, 253 (2002).

[33] V. Genon-Catalot, T. Jeantheau, and C. Larédo, Scand. J. Stat. **30**, 297 (2003).

[34] F. Comte and V. Genon-Catalot, Scand. J. Stat. **33**, 875 (2006).

[35] J. E. Griffin and M. F. J. Steel, J. Econometr. **134**, 605 (2006).

[36] G. Jongbloed and F. H. V. der Meulen, Scand. J. Stat. **33**, 825 (2006).

[37] K. Morimune, Jpn. Econ. Rev. **58**, 1 (2007).

[38] T. Andersen and T. Bollerslev, J. Financ. **53**, 219 (1998).

[39] E. Barucci and R. Renò, Econ. Lett. **74**, 371 (2002).

[40] Y. Omori, S. Chib, N. Shephard, and J. Nakajima, J. Econometr. **140**, 425 (2006).

[41] R. Renò, Econometric Theory (to be published).

[42] I. M. Gelfand and G. E. Shilov, *Generalized Functions* (Academic Press, New York, 1964).

[43] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice Hall, Upper Saddle River, NJ, 2000).

[44] C. R. Champlin and D. Morrell, in *Proceedings of the 31st Conference on Decision and Control*, edited by Control Systems Society Staff and Institute of Electrical and Electronics Engineers, Inc. Staff(Tuscon, AZ, 2002), p. 3707.

[45] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1997).

[46] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. Stanley, Nature (London) **423**, 267 (2003).

[47] Z. Eisler and J. Kertész, Eur. Phys. J. B **51**, 145 (2006).

[48] J. Doyne Farmer, L. Gillemot, F. Lillo, S. Mike, and A. Sen, Quant. Finance **4**, 383 (2004).

[49] J. Masoliver, Phys. Rev. E **48**, 121 (1993).

[50] I. S. Gradstein and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, San Diego, 1994).

[51] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of the Mathematical Physics* (Springer-Verlag, Berlin, 1966).