

Detecting Rough Volatility

Damian, Camilla; Frey, Rüdiger

Published: 24/02/2023

[Link to publication](#)

Citation for published version (APA):

Damian, C., & Frey, R. (2023). *Detecting Rough Volatility: A Filtering Approach*.
<http://arxiv.org/abs/2302.12612v1>

DETECTING ROUGH VOLATILITY: A FILTERING APPROACH

CAMILLA DAMIAN* AND RÜDIGER FREY†

ABSTRACT. In this paper, we focus on the estimation of historical volatility of asset prices from high-frequency data. Stochastic volatility models pose a major statistical challenge: since in reality historical volatility is not observable, its current level and, possibly, the parameters governing its dynamics have to be estimated from the observable time series of asset prices. To complicate matters further, recent research has analyzed the rough behavior of volatility time series to challenge the common assumption that the volatility process is a Brownian semimartingale. In order to tackle the arising inferential task efficiently in this setting, we use the fact that a fractional Brownian motion can be represented as a superposition of Markovian semimartingales (Ornstein-Uhlenbeck processes) and we solve the filtering (and parameter estimation) problem by resorting to more ‘standard’ techniques, such as particle methods.

Keywords: high-frequency data; rough volatility; nested particle filter.

1. INTRODUCTION

It is well-known that, from a statistical point of view, one of the main challenges in the context of *stochastic volatility* models is that of inference: since in reality historical volatility is not observable, its current level and, possibly, the parameters governing its dynamics have to be estimated from the observable time series of asset prices. In particular, the unobservability of volatility represents a crucial issue when modelling high-frequency data, i.e. data recorded on an irregular and remarkably small time scale.

To complicate matters further, recent research has analyzed the *rough* behavior of volatility time series to challenge the assumption, common to most stochastic volatility models, that the volatility process is a Brownian semimartingale. For instance, Gatheral et al. (2018) use historical volatility proxy data to estimate the smoothness of the volatility process with a regression approach. Their empirical finding is that increments of log-volatility of several assets enjoy a monofractal scaling property: this, together with the well-established stylized fact that the distribution of such increments is approximately Gaussian, means log-volatility behaves essentially as a fractional Brownian motion. In particular, given their estimates of the smoothness parameter, these authors claim that log-volatility can be modeled as a fractional Brownian motion with Hurst exponent $H \approx 0.1$. This is in contrast to other fractional stochastic volatility models in the literature (see Comte and Renault (1998)) – typically assuming an Hurst index $H > 1/2$ – and, more generally, to the quite widespread belief that volatility is a long-memory process.

*INSTITUTE OF STATISTICS AND MATHEMATICAL METHODS IN ECONOMICS, TU WIEN

†INSTITUTE FOR STATISTICS AND MATHEMATICS, VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

E-mail addresses: camilla.damian@tuwien.ac.at, ruediger.frey@wu.ac.at.

Date: February 27, 2023.

The first author was in part supported by the Austrian Science Fund (FWF, grant ZK 35 and grant Y 1235).

A related work is that of Bennedsen et al. (2021), who focus on intraday volatility and, by means of an extensive empirical study, find it to be rough *and* persistent. However, as explained in Gneiting and Schlather (2004), models based on processes with the self-similarity property, such as fractional Brownian motion, cannot account for both effects. Therefore, Bennedsen et al. (2021) propose to use the so-called Brownian semistationary processes (Barndorff-Nielsen and Schmiegel (2009)) to model volatility instead, as this choice allows to decouple the short- and the long-term behavior of volatility. If we focus on their estimation of the roughness parameter, we find similarities with Gatheral et al. (2018) both in the linear regression approach¹ and in the results. In fact, the estimates for H , computed for a variety of assets, are qualitatively in agreement with those found by Gatheral et al. (2018).

Fukasawa et al. (2019), however, apply the regression methodology of Gatheral et al. (2018) using the 5-minute realized volatility from simulated price paths and find that, given this sampling frequency, this approach gives a remarkable fit with $H \approx 0.1$ *regardless* of the true value of H used in their simulation. They attribute this finding to the use of a volatility proxy, claiming that it is the error in approximating spot volatility that results in an illusory scaling property. Prompted by these findings, these authors take such ‘proxy error’ explicitly into account and develop a novel quasi-likelihood estimator of the Hurst exponent. In any case, when applying their estimator to real data, Fukasawa et al. (2019) find that volatility is indeed rough: remarkably, their estimates for H for various stock indices are not only smaller than 0.5, but even smaller than those of Gatheral et al. (2018). On the other hand, further follow-up studies to Gatheral et al. (2018), such as Rogers (2023) and – more recently – Cont and Das (2022), find that ‘simpler’ (meaning driven by a Brownian diffusion) volatility models can explain market features just as well as fractional ones. Cont and Das (2022), in particular, identify the cause of apparent roughness in realized volatility with microstructure noise rather than with an inherent, ‘true’ roughness in (spot) volatility.

In this paper, the recent literature regarding *rough* volatility, as well as its intrinsic unobservability, will be taken into account. As we will explain in the next section, we assume that the asset price process changes at discrete times (essentially, the times when new orders arrive to the market), which we model as jump times of a doubly-stochastic Poisson process. We take the intensity of the jump process to be driven by an unobservable signal and we explain how such intensity relates to asset price volatility. To account for roughness in volatility, the central idea is to consider the dynamics of the unobservable process to be driven by a fractional Brownian motion (fBM) with a Hurst index $H < 1/2$. This results in the volatility process being neither a semimartingale nor a Markov process, posing considerable challenges in the analysis of the model, in particular for what concerns the filtering and parameter estimation tasks. Even in the flexible context of a particle filtering approach, the non-Markovianity of the signal implies that sampling a particle at a given time instance requires calculations involving the entire trajectory up to that point. Clearly, the practical consequence is a great increase in computational cost and in memory allocation requirements over time.

¹Although Bennedsen et al. (2021) also present an estimation procedure based on a non-linear regression: this alternative can mitigate the bias arising from using a volatility proxy, which can affect the OLS estimate of the roughness parameter. However, in practice, the two estimators behave in a qualitatively similar fashion.

Here, we circumvent this issue by employing a suitable approximation of the underlying non-Markovian volatility process. The starting point is the work of Carmona and Coutin (1998) and Carmona et al. (2000): they have shown that one can represent an fBM with $H < 1/2$ as a superposition of infinitely many Ornstein-Uhlenbeck processes. We can exploit this idea to approximate the unobservable state process by means of a finite number of such processes, each of them being a Markovian semimartingale: this allows us to use a standard (recursive) particle filter to retrieve the signal; moreover, we can infer about model parameters – in particular, the Hurst index – by adapting the *nested particle filter algorithm* of Crisan and Miguez (2018) to our specific framework. In this way, we can reformulate the complex problem of determining the roughness of a hidden path in terms of a – more easily approachable – parameter estimation task. Simulation experiments indicate that this approach yields satisfactory results, both in terms of filtering and of parameter estimation, while still fully taking into account the intrinsic unobservability of the volatility process. Moreover, in the spirit of Rogers (2023) and Cont and Das (2022), we apply our methodology to synthetic data generated in a setting where (spot) volatility is driven by a Brownian diffusion and assess whether the resulting estimates for H are more consistent with such a setup, rather than mistakenly reflecting a spurious strong roughness effect.

The remainder of this paper is as follows. Section 2 introduces the model: motivated by a specific continuous-time modelling framework involving a fractional resp. Liouville Brownian motion with $H < 1/2$, we explain how such framework can be discretized in time and how such processes can be approximated to obtain a modelling framework suitable for the algorithm of Crisan and Miguez (2018), which is presented in Section 3. The numerical experiments are presented in Section 4; in particular, we include an application of the proposed methodology to the “alternative” models of Cont and Das (2022) and Rogers (2023). Section 5 concludes and illustrates possible extensions.

2. MODEL

In this section, we discuss key ingredients of our model in detail. We start by explaining the motivation behind our specific model design choices for observation and state process in continuous time; then, we proceed to illustrate a suitable approximation of the underlying non-Markovian state process that will make it possible to tackle filtering and parameter estimation in our context. We conclude this section with the description of the discrete counterpart of our modelling framework, which we will need in order to apply the algorithm in Section 3.

2.1. Motivation and Continuous-Time Framework. As mentioned in the introduction, our goal is to investigate the tasks of volatility estimation and inference in the context of a model which is suitable for high-frequency data and which accounts for *roughness* and *unobservability* of volatility dynamics. We can include these features in our setup by means of specific design choices concerning *observation* and *state* processes, respectively.

2.1.1. Observations. The inspiration for the modelling framework considered in this paper lies in Frey and Runggaldier (2001), who also consider the problem of volatility estimation when observations consist of high-frequency data. In their model, a (logarithmic) asset price is assumed to change – and to be observed – only at random, discrete points in time in order to

better mimic the characteristics of high-frequency data. These random times are considered to be the jump times of a marked point process whose intensity and jump-size distribution depend on the level of a hidden state process, closely related to asset price volatility; thus, volatility estimation amounts to a filtering problem with marked point process observations. This represents a departure from more commonly used diffusion models, such as the Heston model, and it makes for a more plausible modelling framework for asset prices observed on a very small time-scale. In particular, the dependence of the intensity on a hidden state variable introduces randomness in market activity over time.

Therefore, with the application to high-frequency observations in mind, we follow Frey and Runggaldier (2001) in that we model observed prices with a jump process, rather than with a continuous one. However, note that in this paper the plan is to consider *event* data (that is, all transactions). As prices recorded at such ultra high frequency often move only by one or few ticks, the size of the jumps is rather uninformative; thus, we can simplify the filtering task by assuming that the unobservable signal influences only the intensity of the point process modelling the observed jump times. We will specify the form of the intensity in more detail in Equation (1) (continuous-time framework) resp. (12) (discrete-time framework); moreover, we will explain the relationship between our model and a continuous model with stochastic volatility.

Another important difference between our setup and the original one from Frey and Runggaldier (2001) is that they assume the hidden state process to be a time-homogeneous Markov process, a setting which is not suitable for fractional or Liouville Brownian motion. Following the approach of Kushner (1977), they obtain a recursive approximation to the optimal filter. In Cvitanić et al. (2006) these results are extended and, using a Bayesian approach, a recursive, closed-form optimal filter is obtained. Introducing *rough* volatility amounts to considering the state process dynamics to be driven by a non-Markovian, non-semimartingale process, which prevents us from exploiting these results directly. The details on the specific form of the state process and on how to tackle the corresponding filtering problem will be the topic of the following sections.

2.1.2. *Hidden State.* As mentioned previously, it is crucial in this paper to choose a model which can account for volatility *roughness*, in alignment with results in recent literature. In our framework, this can be achieved by considering the dynamics of the hidden state process to be driven by a fractional Brownian motion with Hurst index $H < 1/2$, or by a process ‘close’ to it.

Note that, in such a rough framework, it is not possible to use ‘classical’ filtering results directly because the state process is neither a semimartingale nor a Markov process. Some theoretical results regarding nonlinear filtering for fractional Brownian motion (both in the state and in the observation process) have been obtained for example by Decreusefond and Üstünel (1998) and Coutin and Decreusefond (1999), but they are restricted to the case $H \geq 1/2$ and thus not appropriate for the *rough* volatility setting considered here. However, this issue can be bypassed by employing a suitable approximation of the underlying non-Markovian volatility process, as we will describe in detail in Section 2.2.

2.1.3. *Continuous-Time Modelling Framework.* Now that we have introduced the main ingredients of the model, we can describe the corresponding framework. We will generically denote by X the hidden state process, defined on some underlying filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ satisfies the usual conditions. Later, in Section 2.2, we will explain how X is constructed to approximate a fractional resp. Liouville Brownian motion with Hurst index $H < 1/2$; in particular, we will see how X is in fact a functional of a finite number of Ornstein-Uhlenbeck process driven by the same Brownian motion.

Then, in continuous time, our observation process is a Cox process denoted by D for which we assume an intensity of the form

$$\lambda_t := \lambda(X_t) = b \cdot \exp(X_t), \quad (1)$$

where b is a positive constant.

2.1.4. *Relationship to Stochastic Volatility.* Now that the continuous-time model has been introduced, we shall detail the relationship between the intensity of D and stochastic volatility. To this, we assume that a stock price S is given by the following doubly-stochastic compound Poisson process

$$S_t = S_0 + \sum_{i=1}^{D_t} \nu_i, \quad \{\nu_i\}_{i \in \mathbb{N}} \text{ i.i.d.}, \quad \mathbb{E}\{\nu_i\} = 0, \quad \text{Var}(\nu_i) = \sigma^2, \quad (2)$$

where $\sigma > 0$ and D is a doubly-stochastic Poisson process with intensity $\lambda(X_t) = b \cdot \exp(X_t)$, for a Gaussian process X . Moreover, D and $\{\nu_i\}_{i \in \mathbb{N}}$ are independent.

Note that, when the observations consist of event data, the $\{\nu_i\}_{i \in \mathbb{N}}$ are small and bounded as prices move only by one or few ticks at a time; at the same time, there are many jumps (that is, there are frequent trades), so that $b > 0$ should be large. Hence, one expects that the model (2) is close to a stochastic volatility model with instantaneous variance given by $\sigma_t^2 = \sigma^2 \lambda(X_t)$. This concept is formalized in Proposition A.1 in Appendix A.

Now we shall detail how X can be chosen such that the modelling framework accounts for roughness in volatility dynamics, while being fairly tractable in the sense that the filtering and parameter estimation tasks for X can be performed via more ‘standard’ methods.

2.2. **Approximation of fBM.** In this section, we will explain how we can use some results from Carmona and Coutin (1998) and Carmona et al. (2000) to approximate fractional Brownian motion (resp. Liouville Brownian motion), as well as how this is helpful in our context.

More specifically, one considers the Mandelbrot-Van Ness representation of a fractional Brownian motion W^H , which is given by

$$W_t^H = c_H \int_0^t (t-s)^{H-\frac{1}{2}} dB_s + c_H \int_{-\infty}^0 \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right) dB_s \quad (3)$$

where B is a two-sided Brownian motion and c_H a constant depending on H . In particular, the choice

$$c_H = \sqrt{\frac{\pi H(2H-1)}{\Gamma(2-2H)\Gamma(H+1/2)^2 \sin(\pi(H-1/2))}} \quad (4)$$

ensures that the autocovariance function of W^H is given by

$$\mathbb{E}\{W_t^H W_s^H\} = \frac{1}{2} \{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}. \quad (5)$$

Similarly, here we introduce the so-called Liouville Brownian motion V^H as

$$V_t^H = c_H \int_0^t (t-s)^{H-\frac{1}{2}} dB_s. \quad (6)$$

Now, if $H < 1/2$, it has been shown in Carmona and Coutin (1998) and Carmona et al. (2000) that expressing $u \mapsto u^{H-\frac{1}{2}}$ as a Laplace transform

$$u^{H-\frac{1}{2}} \propto \int_0^\infty e^{-xu} x^{-H-\frac{1}{2}} dx$$

and applying stochastic Fubini theorem results in a Markovian representation. In particular, we can express (3) as

$$W_t^H = \int_0^\infty \int_0^t e^{-x(t-s)} dB_s \mu(dx) + \int_0^\infty (e^{-xt} - 1) \int_{-\infty}^0 e^{sx} dB_s \mu(dx),$$

where

$$\mu(dx) = c_H \frac{x^{-H-1/2}}{\Gamma(1/2-H)} dx. \quad (7)$$

Denote by Z^x the Ornstein-Uhlenbeck process $\int_0^t e^{-x(t-s)} dB_s$, starting at zero, and set $Q_0^x := \int_{-\infty}^0 e^{xs} dB_s$, so that

$$W_t^H = \int_0^\infty Z_t^x \mu(dx) + \int_0^\infty (e^{-xt} - 1) Q_0^x \mu(dx)$$

and

$$V_t^H = \int_0^\infty Z_t^x \mu(dx).$$

Spatial Discretization. This idea can be exploited in the context of this paper as follows. For a fixed Hurst parameter $H < 1/2$, we will approximate the corresponding fractional resp. Liouville Brownian motion by a finite sum of Ornstein-Uhlenbeck processes (all driven by the same Brownian motion). This, in turn, will allow us to perform filtering for the unobservable state process by employing more ‘standard’ techniques, such as particle filtering. The core idea is to approximate the measure μ in (7) by a finite sum of Dirac measures; that is, for some $J \in \mathbb{N}$, $\mu \approx \sum_{j=1}^J c_j \delta_{\kappa_j}$ for positive coefficients $(c_j)_{j=1,\dots,J}$ and positive mean-reversion speeds $(\kappa_j)_{j=1,\dots,J}$.

In particular, similarly to Carmona and Coutin (1998) and Carmona et al. (2000), we perform the following spatial discretization. Given $H \in (0, 1/2)$, fix $J \in \mathbb{N}$ and consider $[\xi_0, \xi_J]$, a compact subset of $(0, \infty)$. Split this interval into subintervals by auxiliary terms $0 < \xi_0 < \xi_1 < \dots < \xi_J < \infty$ and compute, for $j = 1, \dots, J$,

$$c_j = \int_{\xi_{j-1}}^{\xi_j} \mu(dx) \quad \text{and} \quad \kappa_j = \frac{1}{c_j} \int_{\xi_{j-1}}^{\xi_j} x \mu(dx).$$

With a slight abuse of notation, now and in what follows we write Z^j and Q_0^j in place of Z^{κ_j} resp. $Q_0^{\kappa_j}$.

An approximation of fractional Brownian motion is then given by:

$$W_t^H \approx \sum_{j=1}^J c_j \left(Z_t^j + (e^{-\kappa_j t} - 1) Q_0^j \right), \quad (8)$$

where $\left(Q_0^j \right)_{j=1}^J$ can be easily simulated, since $(Q_0^x)_{x>0}$ is a centered Gaussian process with covariance function $\Gamma(x, y) = \frac{1}{x+y}$.

In this paper, we will instead consider the following approximation of the Liouville Brownian motion V^H of Equation (6). This approximation is denoted by X , that is

$$V_t^H \approx X_t := \sum_{j=1}^J c_j Z_t^j, \quad (9)$$

as this eases the exposition and notation in what follows.

Quality of the Approximation. Several papers study the accuracy and the convergence of the approximation of fractional Brownian motion via a finite number of OU processes: other than Carmona and Coutin (1998) and Carmona et al. (2000), we refer the readers to Harms (2019) for an approach involving quadrature rules and to Coutin and Pontier (2007) for further results concerning also the temporal approximation.

A standard choice for auxiliary terms is a geometric partition of ratio $r \in (1, 2)$. The quality of the approximation obviously depends on the chosen J and r . Numerical experiments, in agreement with theoretical results, suggest that for large J one should choose a value of r close to 1. However, if in practical applications we choose to keep the dimension of the approximation manageable, numerical experiments indicate that we should increase the value of r . In particular, for a given natural number $J > 16$, here we fix a compact $[\xi_0, \xi_J]$ and split it into subintervals using a geometric partition, with ratio $\left(\frac{\xi_J}{\xi_0} \right)^{1/J}$. Typically, one would want to choose a smaller value of ξ_0 for H close to 0.5, and a larger value of ξ_J for H close to 0. Writing $\alpha = H + 1/2$, we can, for example, choose the values $\xi_0 = J^{-2\alpha}$ and $\xi_J = J^{4-2\alpha}$, so that $\left(\frac{\xi_J}{\xi_0} \right)^{1/J} = J^{4/J}$. Thus, with careful choices for ξ_0 and ξ_J , we can fix J and construct a partition which is in line with what observed above and performs reasonably well numerically also for modest $J > 16$.

Remark 2.1. *In practice, note that the choice of J will also depend on the fineness of the time scale on which we want to approximate W^H resp. V^H . Fix a time horizon $T > 0$ and consider the partition $0 = t_0 < t_1 < \dots < t_N = T$ of $[0, T]$ with a (for simplicity assumed constant) step size $\Delta = T/N$. Then, for smaller value of Δ resp. higher values of N , we need to increase J to still obtain a similarly satisfactory approximation accuracy as on a coarser time grid. Coutin and Pontier (2007) study the accuracy and the convergence of the approximation of fractional Brownian motion via a finite number of OU processes taking also into account the effect of the temporal approximation. With their results in mind, we will make the following choices concerning the number of OU processes building the approximations (8) resp. (9).*

(i) Given N and $H < 1/2$, we take

$$J := J(N, H) = \lfloor 2 \cdot N^\zeta \cdot \log(N) \rfloor, \quad \zeta = \log(1 + H), \quad (10)$$

so that $\zeta < H$ as $H \in (0, 1/2)$.

(ii) If we do not want the dimension of the approximation to depend on H ,²

$$J := J(N) = \lfloor 2 \cdot N^\zeta \cdot \log(N) \rfloor, \quad \zeta = \log(1 + 0.25). \quad (11)$$

2.3. Discrete-Time Modelling Framework. In order to solve the estimation and inference problems resulting from our modelling setup, one possibility is to use the *nested particle filtering algorithm* introduced by Crisan and Miguez (2018). In order to fit our setup within their framework, we must consider a time-discretized version of our model. To this, we fix a time horizon $T > 0$ and consider the partition $0 = t_0 < t_1 < \dots < t_N = T$ of $[0, T]$ with a (for simplicity assumed constant) step size $\Delta = T/N$. We only consider changes in the unobservable state process at these discrete time points, which implies that the intensity has a piecewise-constant form. In particular, we assume

$$\lambda_u = b \cdot \exp(X_{t_{n-1}}) \quad (12)$$

for all $u \in (t_{n-1}, t_n]$, $n = 1, \dots, N$. As long as Δ is small enough, this assumption is not too restrictive as the choice (12) still allows to mimic true market behavior reasonably well.

We will use the following notation for the discretized model. We write $X_n = X_{t_{n-1}}$ for the random sequence corresponding to the underlying state process, λ_n for the corresponding intensity and, for what concerns observations, we consider the sequence given by $y_n = D_{t_n} - D_{t_{n-1}}$, $n = 1, \dots, N$. Note that the piecewise-constant form of the intensity implies that y_n is Poisson distributed; in particular we have

$$y_n \sim \text{Pois}(\lambda_n \Delta). \quad (13)$$

Updating Scheme. To conclude this section, note that the particle filtering algorithms of Section 3 require us to ‘update’ the value of J Ornstein-Uhlenbeck processes building the approximation (9); that is, given their values at some time t_{n-1} , we need to calculate their values at some later time t_n . Since each of these processes starts at and reverts to zero and has unit variance, the dynamics of each Z^j , $j = 1, \dots, J$ are given by

$$dZ_t^j = -\kappa_j Z_t^j dt + dB_t.$$

As explained in Gillespie (1996), an Euler-Maruyama discretization of these dynamics is accurate only for a suitably small time-discretization step Δ . In particular, Δ should be much smaller than the reciprocal of the mean reversion speed, which is clearly problematic in our context, given the high values of some of the κ_j ’s in the approximation (9). Gillespie (1996) explains how to circumvent this problem by using so-called *exact updating formulae*, which allow us to calculate the value of $Z_{t_n}^j$ given $z := Z_{t_{n-1}}^j$ as

$$Z_{t_n}^j = ze^{-\kappa_j \Delta} + \sqrt{\frac{1 - e^{-2\kappa_j \Delta}}{2\kappa_j}} v \quad (14)$$

²When implementing the nested particle filter of of Section 3.2, in order to be able to update the particles in the recursive step, the dimension of the OU approximation corresponding to a jittered parameter has to be the same as the dimension of the OU approximation corresponding to the ‘previous’,unjittered parameter. This works quite well as, with our choice of jittering kernel, jittered parameters are obtained as small perturbations of the existing particles in the parameter space; however, it might be more consistent to keep the dimension fixed from the beginning and for all particles. This also allows us to work with tridimensional arrays in R.

where v is sampled from a standard normal distribution.

It is important to note that, since all the OU processes in (9) are driven by the same Brownian motion, we only need to generate a single standard normal random variable to update all of their values from time t_{n-1} to time t_n , making for an efficient implementation of the particle filtering algorithms of Section 3.

Once again, we will stay within the discretized framework; therefore, in what follows we will consider the random sequence indexed by $n = 1, \dots, N$ which we obtain by setting $Z_n^j = Z_{t_{n-1}}^j$.

3. NESTED PARTICLE FILTER

The goal of this paper is not only to perform the filtering task, but also to estimate model parameters (in particular, H). To this, we rely on a *nested particle filter* approach, as described in Crisan and Miguez (2018): the authors introduce a nested structure, employing two layers of particle filters, for the approximation of the joint posterior distribution of the signal and the unknown (static) parameters.

Main Assumptions for the Applicability of the Algorithm. To determine the applicability of the algorithm in a specific framework, two main assumptions identified in Crisan and Miguez (2018) must be satisfied. The first assumption states that the parameter space must be a compact set and that the conditional p.d.f. of the observations must be well-behaved (positive and upper bounded) uniformly over that set. We will assume that the parameter space of our parameter of interest, the Hurst index H , is a compact subset of $(0, 1/2)$, so that this assumption is satisfied.

Secondly and most importantly, the optimal filter for the model of interest must be continuous with respect to the parameter; that is, small changes in the parameter should lead only to small changes in the posterior of the state given the observations. This assumption is crucial to obtain a recursive algorithm. In fact, at a given time instance, the key quantity required to approximate the posterior measure of the parameter is its marginal likelihood. At time n , consider the sampled parameter $\bar{\theta}_n$ obtained, for instance, as a random mutation of θ_{n-1} through jittering: one could approximate the likelihood of $\bar{\theta}_n$ by running a standard bootstrap filter conditional on $\Theta = \bar{\theta}_n$ from time 0 to time n . However, this non-recursive approach would dramatically increase the computational cost with time. Due to the continuity assumption, one can instead use the particle approximation of the filter corresponding to θ_{n-1} as a particle approximation of the filter corresponding to $\bar{\theta}_n$. This is crucial to approximate both the predictive measure and the likelihood of $\bar{\theta}_n$ by exploiting the ‘mismatched’ parameter value θ_{n-1} instead, and it ensures the recursive property of the algorithm. We discuss this continuity property in our setting in Appendix A (see, in particular, Proposition A.2).

3.1. ‘Inner’ (Bootstrap) Filter. To fix notation and describe how the approximation of fBM can be exploited in the given context, in this section we review the standard particle filter (so-called *bootstrap* filter); this represents the “inner” filter of the two nested layers of particle filters described in Crisan and Miguez (2018) and it yields, at each time point n , an approximation π_n of the posterior measure for the state, conditional on the observations up to the time point n and the given (known or sampled) parameters.

We apply a bootstrap filter with M particles in the context of the discrete-time version of our model, as described at the end of Section 2.1. Regarding observations, we record the values of the observed realization of the counting process D on a grid of step size $\Delta = T/N$ to then obtain the observation sequence $(y_n)_{1 \leq n \leq N}$. Here, we consider both the constant b and the Hurst index $H < 1/2$ to be known; then, using (10) we compute the number $J = J(H, N)$ of OU processes needed for the approximation. Finally, note that in the following we write $\mathbf{Z}_n := (Z_n^1, \dots, Z_n^J)^T$.

(1) **Initialization** ($n = 0$)

- Compute coefficients and mean-reversion speeds for the approximation (9).
- For each $m = 1, \dots, M$, initialize: sample \mathbf{Z}_0^m from a suitable prior distribution.³

(2) **Recursive step** (from $n - 1$ to n)

For $m = 1, \dots, M$, let $\{\mathbf{Z}_{n-1}^m\}_{1 \leq m \leq M}$ be the particles (Monte Carlo samples) available at time $n - 1$. At time n :

- (a) Update: for $m = 1, \dots, M$, generate $v^m \sim \mathcal{N}(0, 1)$ and draw $\bar{\mathbf{Z}}_n^m$ conditional on \mathbf{Z}_{n-1}^m , see also (14).
- (b) Compute normalized weights: for $m = 1, \dots, M$, compute (see (9))

$$\bar{X}_n^m = \sum_{j=1}^J c_j \bar{Z}_n^{j,m},$$

the corresponding intensity λ_n^m and the corresponding likelihood weight (see also (13))

$$w_{t+\Delta}^m \propto (\lambda_n^m \Delta)^{y_n} \cdot \exp(-\lambda_n^m \Delta).$$

Normalize the weights to obtain $\bar{w}_{t+\Delta}^m = w_{t+\Delta}^m / \sum_{m=1}^M w_{t+\Delta}^m$.

- (c) Resample: For $m = 1, \dots, M$, let $\mathbf{Z}_n^m = \bar{\mathbf{Z}}_n^q$ with probability $\bar{w}_{t+\Delta}^q$, $q \in \{1, \dots, M\}$.

3.2. Nested Particle Filter and Parameter Estimation. Consider the setting of the previous section, but let Θ be the vector of model parameters to be estimated. In our case, we aim at estimating $\Theta = H$. The estimation of b will not be addressed in the specific context of this model, as it can typically be chosen to match the average number of price movements observed in one (continuous-time) unit. The nested particle filter for the approximation, at each discrete time point n , of the posterior distribution of the parameter is given below.

Other than the time step size $\Delta = T/N$ and the recorded observations $(y_n)_{1 \leq n \leq N}$, the required inputs are the value of constant b and the number of particles $K \cdot M$. For the parameter, an appropriate prior distribution has to be chosen: we use a uniform distribution over $(0, 1/2)$.

(1) **Initialization** ($n = 0$)

³In our simulation studies, we will typically choose a centered J -dimensional normal distribution with a small variance.

- Draw K i.i.d. samples θ_0^k , $k = 1, \dots, K$ from the prior distribution $\pi_0^\theta(d\theta)$. Then given the total number of time steps N , compute the corresponding $J^k = J(N, \theta_0^k)$, $k = 1, \dots, K$.
- For each $m = 1, \dots, M$ and for each $k = 1, \dots, K$, initialize: sample $\mathbf{Z}_0^{(k,m)}$ from a suitable prior distribution.

(2) **Recursive step** (from $n - 1$ to n)

For $k = 1, \dots, K$, let $\left(\theta^k, \left\{\mathbf{Z}_n^{(k,m)}\right\}_{1 \leq m \leq M}\right)$ be the particle set available at time $n - 1$. At time n :

(a) Jittering and state propagation: for each $k = 1, \dots, K$, perform the following steps.

- Draw $\bar{\theta}_n^k$ from $\kappa_K^{\theta_0^k}(d\theta)$. This corresponds to *jittering* the samples in the parameter space: for possible ways to perform such step, we refer to Section 4.2 of Crisan and Miguez (2018).
- For each $\bar{\theta}_n^k$, $k = 1, \dots, K$, compute the corresponding coefficients and mean-reversion speeds for the approximation (9).
- Given $\bar{\theta}_n^k$, perform Step 2(a) of the standard particle filter described in the previous section to obtain $\bar{\mathbf{Z}}_n^{(k,m)}$ (and $\bar{X}_n^{(k,m)}$), $m = 1, \dots, M$.
- Compute the approximate likelihood of the parameter $\bar{\theta}_n^k$, given by

$$u_n^M(\bar{\theta}_n^k) = \frac{1}{M} \sum_{m=1}^M L_{\bar{\theta}_n^k}(\bar{\mathbf{Z}}_n^{(k,m)}),$$

where $L_{\bar{\theta}_n^k}(\bar{\mathbf{Z}}_n^{(k,m)})$ correspond to the (Poisson) likelihood given $\bar{\theta}_n^k$.

- Perform steps 2(b) and 2(c) of the standard particle filter described in the previous section to update and obtain $\tilde{\mathbf{Z}}_n^{(k,m)}$, $m = 1, \dots, M$.

(b) Compute normalized weights: for each $k \in \{1, \dots, K\}$, compute normalized weights $\bar{w}_n^k \propto u_n^M(\bar{\theta}_n^k)$.

(c) Resample: For $k = 1, \dots, K$, let

$$\left(\theta_n^k, \left\{\mathbf{Z}_n^{(k,m)}\right\}_{1 \leq m \leq M}\right) = \left(\bar{\theta}_n^l, \left\{\tilde{\mathbf{Z}}_n^{(l,m)}\right\}_{1 \leq m \leq M}\right)$$

with probability \bar{w}_n^l , $l \in \{1, \dots, K\}$. Furthermore, for each resampled θ_n^k , $k = 1, \dots, K$, we take the corresponding J^k .

(d) Approximate the posterior measure by $\mu_n \approx \frac{1}{K} \sum_{k=1}^K \delta_{\theta_n^k}$.

4. NUMERICAL RESULTS

In this section, we present the results of a simulation study aimed at testing the accuracy of the presented algorithms in the context of our model, as well as an application of the proposed methodology to the “alternative” models of Cont and Das (2022) and Rogers (2023).

4.1. Simulation. For the simulation experiments of this subsection, we fix a time horizon $T = 1$ and a small time step of size $\Delta = 1/960$; that is, we consider half-minute intervals in a trading day of 8 hours. Except when analyzing the impact of the ‘informativeness’ of the observation process, we set $b = 8000$ as the average amount of price changes in a given day – a relatively conservative choice for a liquid stock.

Bootstrap Filter for Known H . Here we consider the Hurst index H to be known and thus we implement the bootstrap filter of Section 3.1 with $M = 600$ particles. We consider two cases for the true Hurst index, $H = 0.1$ and $H = 0.4$; given the total number of time steps $N = 960$, the number of OU processes building the approximation (9) is given by Equation (10) as $J(960, 0.1) = 26$ and $J(960, 0.4) = 138$, respectively.

Figure 1 shows simulated trajectories of Liouville Brownian motion (gray lines) with Hurst index $H = 0.1$ (upper panel) resp. $H = 0.4$ (lower panel), as well as the corresponding filtered estimates (black lines). We observe that the filtered trajectory tracks the true one reasonably well in both cases, particularly if we take into account the small time scale considered.

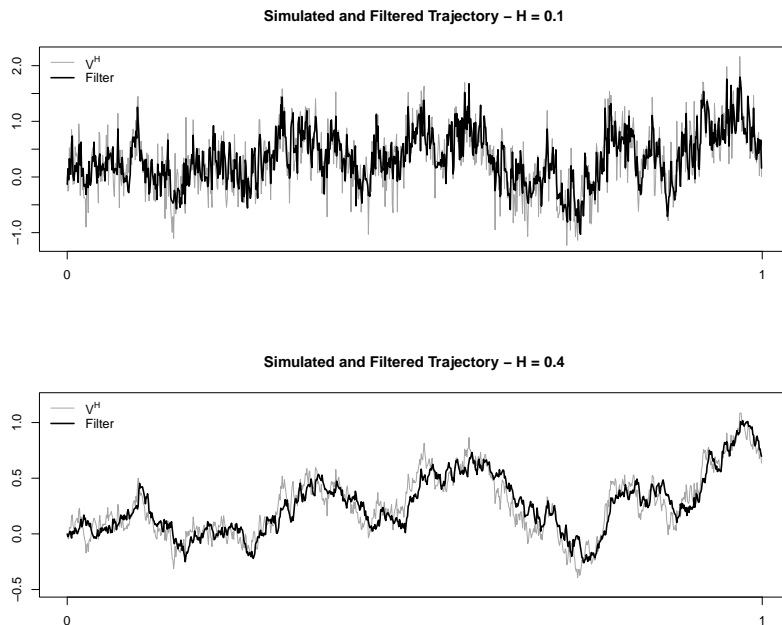


FIGURE 1. True (gray) and filtered (black) state trajectory of V^H .

Nested Particle Filter for Unknown H . Here we consider the Hurst index H to be unknown and thus we estimate it accordingly by implementing the nested particle filter of Section 3.2 with $K \cdot M = 300^2$ particles. We consider two cases for the true Hurst index, $H = 0.1$ and $H = 0.4$; however, unlike in the filtering experiment, we keep the number of OU processes building the approximation (9) equal in both cases. This number depends only on the total number of time steps $N = 960$, and it is given by Equation (11) as $J(960) = 63$.

Figure 2 shows – on a log scale – the average, over 50 independent simulations, of the relative error $|\hat{H}_n - H|/H$, $n = 1, \dots, N$, in the case in which the true Hurst index H is given by $H = 0.1$ (upper panel) resp. $H = 0.4$ (lower panel). We denote by \hat{H}_n the mean of the posterior

distribution μ_n obtained in a given run at the n^{th} time step. Overall, the behavior of parameter estimates over time is quite satisfactory, as the estimation problem at hand is relatively complex (particularly when considering such a fine time grid and only one continuous-time unit).

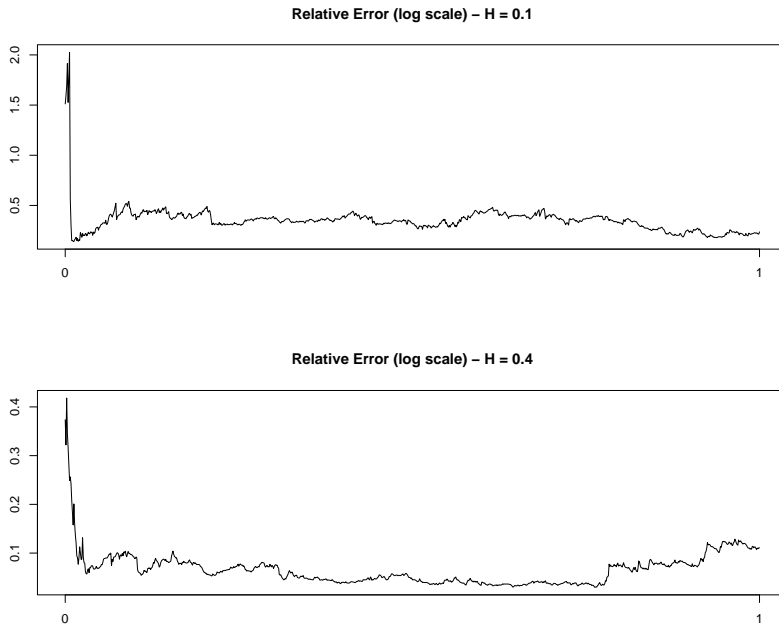


FIGURE 2. Mean relative error (log scale) in parameter estimation.

Sensitivity w.r.t. Value of b . The biggest influence on the accuracy of the algorithms is that of the parameter b , which is why we focus on it next. Obviously, the number of particles (M resp. $K \cdot M$) has also an effect, in the sense that the larger the number of particles, the better the accuracy of the algorithm. However, any reasonable, albeit quite conservative, choice would produce satisfactory filtering results.

Here we assume that the true Hurst index is $H = 0.3$ and we analyze two cases, one in which the observation process is less informative (i.e., $b = 3000$) and one in which it is more informative (i.e., $b = 10000$). First, we consider the Hurst index to be known and thus we implement the bootstrap filter of Section 3.1 with $M = 300$ particles and – as specified by Equation (10) – we use $J(0.3, 960) = 83$ OU processes to build the approximation (9). Then, we consider H to be unknown and we estimate it by implementing the nested particle filter of Section 3.2 with $K \cdot M = 300^2$ particles, where the dimension of the approximation depends only on the total number of time steps N and it is given by Equation (11) as $J(960) = 63$.

The upper panel of Figure 3 shows the unobserved trajectory of Liouville Brownian motion with Hurst index $H = 0.3$ (black solid line) and the area comprised between the 1%- and 99%-quantile of the posterior distribution of the state process. The lower panel of Figure 3 shows the posterior distribution, at final time, of the Hurst index in each case; the vertical lines indicate the respective means. In both plots, the lighter color corresponds to the less informative case, the darker one to the more informative one. In the lower panel, the parameter prior is also depicted (dotted line). We can observe how a higher value of b improves the accuracy in

both filtering and parameter estimation; however, results are still quite satisfactory in the less informative case. Note that, in order to make the modeling framework consistent with the large number of transactions characterizing high-frequency trading of liquid stocks, b would typically take a large value in most real-world applications.

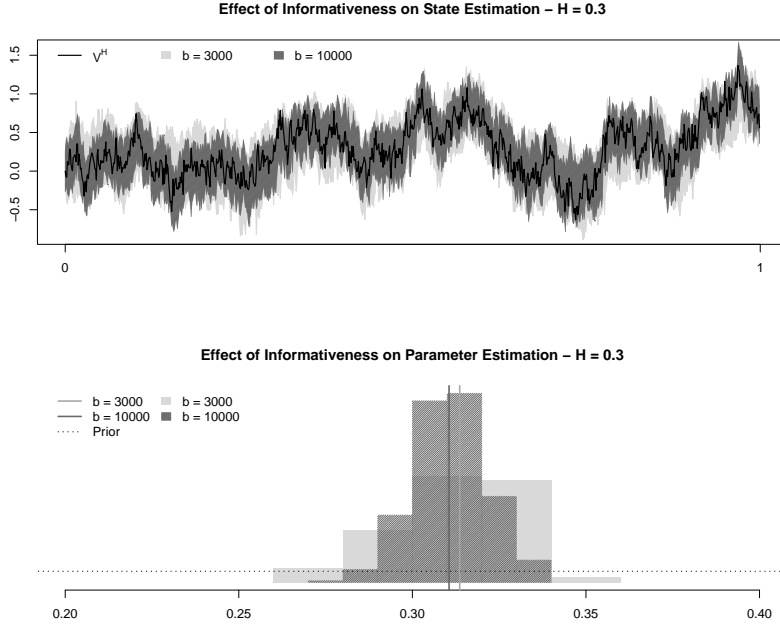


FIGURE 3. True and filtered state trajectory of V^H with $H = 0.3$ (upper panel) and posterior distribution of H at final time (lower panel) in the cases $b = 10000$ resp. $b = 3000$.

4.2. Experiments in the Context of Non-Rough Stochastic Volatility Models. In this section, in the spirit of Rogers (2023) and Cont and Das (2022), we apply our methodology to synthetic data generated in a setting where (spot) volatility is driven by standard Brownian diffusions and assess whether the resulting estimates for H are more consistent with such a setup, rather than mistakenly reflecting a spurious strong roughness effect. This amounts to modify the continuous-time model of Section 2.1, discretize it in the same fashion as in Section 2.3 and estimate the unobservable volatility trajectory and the Hurst index H using the nested particle filter algorithm of Section 3.2 using $K \cdot M = 300^2$ particles.

Volatility as Modulus of Brownian Motion. To be consistent with Section 4.1 of Cont and Das (2022), here we assume our observation process D to have an intensity of the form (to be compared with (1)):

$$\lambda_t := \lambda(W_t) = b \cdot |W_t|^2, \quad (15)$$

where b is a positive constant and W is a standard Brownian motion, i.e. here $H = 1/2$. However, we consider the Hurst index H to be unknown and thus we estimate it accordingly by implementing the nested particle filter of Section 3.2. We fix a time step of size $\Delta = 1/480$ (that is, we consider one-minute intervals in a trading day of 8 hours) and a time horizon of $T = 5$ days, so that $N = 2400$. Given the total number of time steps N , Equation (11) gives $J(2400) = 88$ for the number of OU processes building the approximation (9).

It is very important to stress that this experiment aims at assessing whether the resulting estimate for H reflects a spurious strong roughness effect, attributable for instance to microstructure noise, rather than at estimating H correctly. In fact, the value $H = 1/2$ can *never* be estimated correctly without modifying the methodology and the algorithm substantially, as the approximations (8) and (9) are valid for $H < 1/2$ and the parameter space of nested particle filtering algorithm of Section 3.2 is assumed to be a compact subset of $(0, 1/2)$.

The upper panel of Figure 4 shows a simulated trajectory of such $|W|$ (gray line) and its corresponding filtered estimate (black line). The lower panel shows the mean (dark gray line) and the 1%- resp. 99%- quantiles (light gray lines) of the estimated posterior μ_n , $n = 1, \dots, N$, of the unknown parameter H , when the underlying unobservable process is (the modulus of) a standard Brownian motion, i.e. $H = 1/2$.

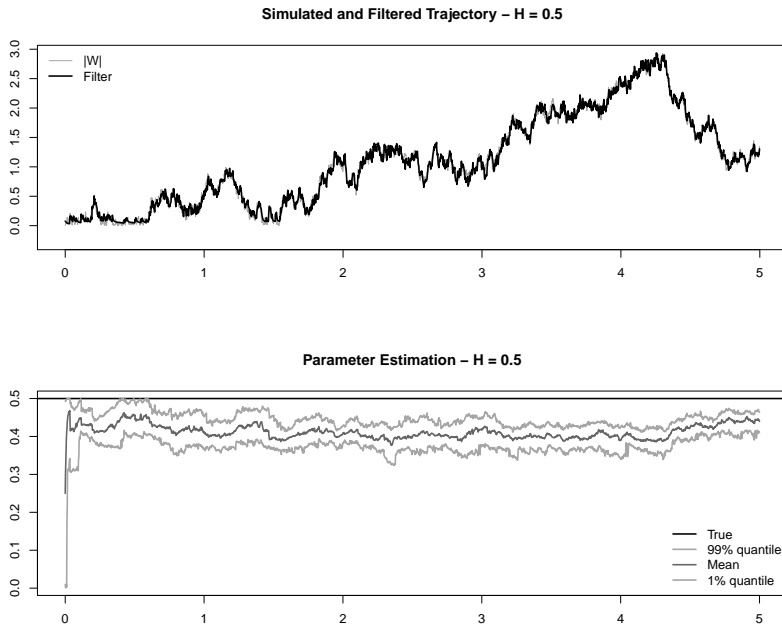


FIGURE 4. True and filtered state trajectory of $|W|$ and behavior of parameter estimates over time.

Overall, considering that $H = 1/2$ can never be estimated in our current setup, the results are quite satisfactory, both in terms of filtered trajectory and – especially – in terms of parameter estimate, since the latter is not falsely identified to be close to zero, but rather it nears $1/2$.

OU-OU Model for Volatility. To be consistent with Section 4 of Rogers (2023), here we assume our observation process D to have an intensity of the form (to be compared with (1)):

$$\lambda_t := \lambda(V_t) = b \cdot V_t^2, \quad (16)$$

where V is given by the so-called OU-OU model:

$$\begin{aligned} dR_t &= -\beta R_t dt + \sigma_R dW'_t, \\ dV_t &= \kappa(R_t - V_t) dt + \sigma_V dW_t. \end{aligned} \quad (17)$$

To simulate model and observations, we choose the same parameters as in the original paper; that is, $\sigma_V^2 = 20$, $\sigma_R^2 = 0.625$, $\kappa = 210$ and $\beta = 2.5$. However, when implementing our estimation methodology, we suppose that the unobservable process V is not of its ‘true’ form (17), but rather a Liouville Brownian motion with unknown Hurst index H , which we estimate using the nested particle filter of Section 3.2. We fix a time horizon of $T = 5$ days and make two distinct choices for the time step Δ to analyze its impact on the estimation procedure. In particular, we consider $\Delta = 1/240$ (corresponding to two-minute intervals in a trading day of 8 hours) and $\Delta = 1/960$ (corresponding to thirty-second intervals in a trading day of 8 hours); thus, the total number of steps is $N = 1200$ resp. $N = 4800$ and Equation (11) gives $J(1200) = 68$ resp. $J(4800) = 112$ for the number of OU processes building the approximation (9).

Experiments show that here the estimates for H are more consistent with a rough volatility model, meaning that the proposed OU-OU model seem to mimic that type of behavior. However, if we simulate the model on a finer time grid and apply the corresponding algorithm, the estimate for the Hurst index increases. The upper panel of Figure 5, in particular, shows the posterior distribution, at final time, of the Hurst index estimated from the OU-OU model in the case $\Delta = 1/240$ and in the case $\Delta = 1/960$; the vertical lines indicate the respective means. The lower panel shows the corresponding posterior distributions for a model of intensity of the form (16) above, i.e. $\lambda_t := \lambda(V_t^H) = b \cdot (V_t^H)^2$, where the driving process V^H is of the form specified in Equation (6) with $H = 0.2$ (roughly corresponding to the average of the estimates obtained for the OU-OU model on the two different time grids). In both plots, the lighter color corresponds to the first case, the darker one to the more finely discretized one and the dotted line to the parameter prior. We can observe how Hurst index estimates resulting from applying the algorithm to observations generated from the OU-OU model vary strongly depending on the fineness of the chosen time scale (in particular, estimates increases – and thus mimic less roughness – the finer the time grid is), while this is not the case when observations are generated from the ‘true’ model.

5. CONCLUSIONS AND OUTLOOK

In this paper, we discuss filtering and parameter estimation in a rough volatility model for high-frequency data. We consider a time-discretized version of a continuous-time framework where observations are given by a trajectory of a Cox process whose intensity is driven by a ‘rough’ process, such as fractional resp. Liouville Brownian motion with Hurst parameter $H < 1/2$. Basing ourselves on a representation of such signal as a superposition of Ornstein Uhlenbeck processes (as introduced by Carmona and Coutin (1998)), we detail how it can be filtered using standard particle filtering techniques. Moreover, we explain how to estimate H using the *nested particle filtering* algorithm of Crisan and Miguez (2018). We run a comprehensive simulation study to test the accuracy of the algorithms and their sensitivity with respect to specific modelling choices. We find that the results are satisfactory; in particular, we are able to estimate the Hurst index with reasonable accuracy, while still taking the unobservability of volatility fully into account. Moreover, the parameter estimation methodology seems to be able to adequately distinguish between ‘true’ rough dynamics in spot volatility and spurious roughness effects arising from microstructure noise, as documented e.g. in Cont and Das (2022).

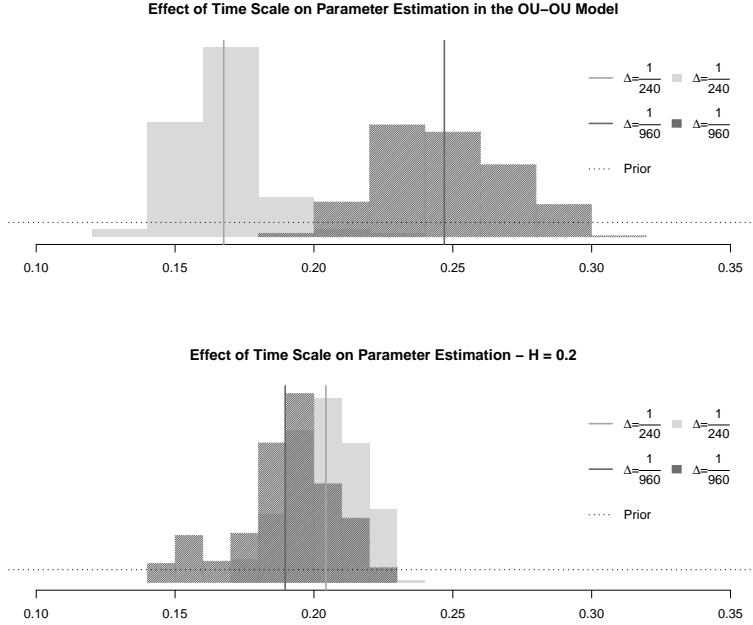


FIGURE 5. Posterior distribution of H at final time in the cases $\Delta = 1/240$ and $\Delta = 1/960$.

From an implementation point of view, a possible extension involves modifying the algorithms so that they do not perform resampling at every step as they currently do, but only when the diversity in the particles is reduced beyond a certain threshold. This is typically done by estimating the so-called Effective Sample Size (ESS), although in the setting of the *nested particle filtering* algorithm of Crisan and Miguez (2018) the usual estimator of ESS should be modified to accommodate the special characteristics of the algorithm and to avoid becoming uninformative (for further reading about the *normalized* ESS estimator devised by Crisan and Miguez (2018), we refer the reader to Section 5.6 of that paper).

Moreover, the proposed methodology should be applied to real data. It would be particularly interesting to compare estimates of H obtained using this methodology to those obtained from Gatheral et al. (2018), Bennedsen et al. (2021), Fukasawa et al. (2019) and Cont and Das (2022), among others. However, as briefly touched upon in Section 1, the *long-memory* feature of volatility has been investigated in several empirical studies (see, for instance, Andersen et al. (2001)) and it is often considered a stylized fact. In the modelling framework considered so far, it is not possible to capture *persistence*: as explained in Gneiting and Schlather (2004), models based on processes with the self-similarity property (such as fractional Brownian motion) cannot account for both roughness and long-run dependence. Moreover, neither fractional nor Liouville Brownian motion are stationary, which represents a problem when one seeks to study the long-term properties of such processes. Therefore, a possible extension of the present work is to find an alternative model that allows for rough paths and for a polynomial decay of the autocorrelation function $\rho(s)$ for $s \rightarrow \infty$ (that is, for long memory). This is very much in the spirit of Bennedsen et al. (2021), who advocate the use of the so-called Brownian semistationary processes (see Barndorff-Nielsen and Schmiegel (2009)) to model volatility. Preliminary results

indicate that it is possible to modify the previously-considered modelling framework in a way such that it accommodates both effects, while preserving a structure suitable for the filtering and estimation strategies presented in Section 3; however, a detailed analysis and a comprehensive simulation study for this alternative model are topics of future research.

Finally, it is important to note that models where underlying random factors are driven by fractional Brownian motion do not only find their applications in financial settings: among several examples, we highlight the recent work by Alòs et al. (2020), which focuses on the COVID-19 pandemic, and that by Leppanen et al. (2021), which concerns the analysis of functional magnetic resonance imaging (fMRI) scans. Therefore, the methodology presented in this paper might be applied to estimate fractional models in a variety of contexts beyond the present one.

APPENDIX A. FURTHER RESULTS

A.1. Relationship to Stochastic Volatility.

Proposition A.1. *Consider a sequence of models indexed by p and given by $S_t^p = S_0 + \sum_{i=1}^{D_t^p} \nu_i^p$, where for all p*

- (i) D^p is a doubly-stochastic Poisson process with intensity $\lambda^p(X_t)$;
- (ii) $\{\nu_i\}_{i \in \mathbb{N}}$ are i.i.d. with zero mean and variance given by $(\sigma^p)^2$;
- (iii) $(\nu_i^p)^2 \leq \bar{c}_p$ for a sequence $\bar{c}_p \rightarrow 0$;
- (iv) $(\sigma^p)^2 \lambda^p(x) \rightarrow \sigma^2 \lambda(x)$, uniformly on compacts.

Then the pair (S^p, X) converges in distribution to (S, X) , where S solves

$$dS_t = \sigma \sqrt{\lambda(X_t)} dW_t$$

for a standard Brownian motion W .

Sketch of proof. We first condition on $\mathcal{F}_\infty^X = \sigma(X_s, s \geq 0)$, so that D^p is a Poisson process with time-dependent intensity. Thus, consider the filtration $\{\mathcal{F}_t^S \vee \mathcal{F}_\infty^X\}_{t \geq 0}$. Note that, for each p , S^p is a martingale on this filtration. Moreover, it holds that

$$d(S_t^p)^2 = 2S_{t-}^p dS_t^p + d[S^p]_t,$$

where $[S^p]_t = \sum_{i=1}^{D_t^p} (\nu_i^p)^2$. That is, $[S^p]_t - \int_0^t (\sigma^p)^2 \lambda^p(X_s) ds$ is a martingale. It follows that $(S_t^p)^2 - \int_0^t (\sigma^p)^2 \lambda^p(X_s) ds$ is also a martingale.

By (iv), one has that $\int_0^t (\sigma^p)^2 \lambda^p(X_s) ds \rightarrow \int_0^t \sigma^2 \lambda(X_s) ds$. Moreover, for fixed T , (iii) gives

$$\sup_{t \leq T} (S_t^p - S_{t-}^p)^2 = \sup_{i \leq D_T^p} (\nu_i^p)^2 \leq \bar{c}_p \rightarrow 0.$$

Hence, Theorem 7.4.1 in Ethier and Kurtz (1986) gives that the law of S^p given \mathcal{F}_∞^X converges to the law of the process S with $dS_t = \sigma \sqrt{\lambda(X_t)} dW_t$. The claim then follows by integrating with respect to the law of X . \square

Note that above we assume a linear model for the stock price itself, which could theoretically lead to negative prices; however, when considering a short time horizon this possibility can safely be ignored. When analyzing the model on a longer time horizon, the linear model should rather be used for the log price.

A.2. Continuity of the filter in the Hurst index. Recall that, in our setup, the state process is in fact a functional of J (time-discretized) OU processes driven by the same Brownian motion. In what follows we write $\mathbf{Z}_n := (Z_n^1, \dots, Z_n^J)^T$, where for $n = 1, \dots, N$ and $j = 1, \dots, J$ we have set $Z_n^j = Z_{t_{n-1}}^j$ (as described in Section 2.3). For fixed H , we consider mean-reversion speeds $\kappa_j := \kappa_j(H)$ and coefficients $c_j := c_j(H)$, $j = 1, \dots, J$, and we assume that they are continuous in H . Moreover, we let

$$\lambda_n := \lambda(\mathbf{Z}_n, H) = b \exp \left(\sum_{j=1}^J c_j Z_n^j \right).$$

As for the observations, y_n denotes the number of jumps over $(t_{n-1}, t_n]$, $n = 1, \dots, N$. Thus, in our setting, y_1, \dots, y_n are conditionally independent given $\mathbf{Z}_1, \dots, \mathbf{Z}_n$. Then, since for $n = 1, \dots, N$, $y_n \sim \text{Pois}(\lambda(\mathbf{Z}_n, H)\Delta)$, we can define the likelihood

$$g(y|z, H) = \exp(-\lambda(z, H)\Delta) \frac{(\lambda(z, H)\Delta)^y}{y!}. \quad (18)$$

Now we can state the following continuity result, which is weaker than the Lipschitz continuity of Crisan and Miguez (2018), but nonetheless useful in our context.

Proposition A.2. *Consider a sequence $H_h \rightarrow H$, $h \rightarrow \infty$ and fix a discrete time point $n \leq N$. Then, for $f : \mathbb{R}^J \rightarrow \mathbb{R}$ bounded and continuous, we have*

$$\lim_{h \rightarrow \infty} \mathbb{E}[f(\mathbf{Z}_n) | y_1, \dots, y_n; H_h] = \mathbb{E}[f(\mathbf{Z}_n) | y_1, \dots, y_n; H]. \quad (19)$$

Proof. First note that, for generic H , we have the following formula (of Kallianpur-Striebel type):

$$\mathbb{E}[f(\mathbf{Z}_n) | y_1, \dots, y_n; H] = \frac{\mathbb{E}[f(\mathbf{Z}_n) \prod_{i=1}^n g(y_i | \mathbf{Z}_i, H) | H]}{\mathbb{E}[\prod_{i=1}^n g(y_i | \mathbf{Z}_i, H) | H]}.$$

To establish convergence, it suffices to consider the numerator. Define the set

$$\mathcal{Z}^L = \{ \mathbf{Z}_1, \dots, \mathbf{Z}_N : \mathbf{Z}_n \in \mathbb{R}^J, \|\mathbf{Z}_n\| \leq L \quad \forall n = 1, \dots, N \}.$$

As the \mathbf{Z}_n are Gaussian with bounded variance, for fixed y_1, \dots, y_N it holds that

$$\lim_{L \rightarrow \infty} \sup_h \mathbb{E} \left[\prod_{i=1}^n g(y_i | \mathbf{Z}_i, H_h); (\mathcal{Z}^L)^c | H_h \right] = 0.$$

Now, since κ_j , $j = 1, \dots, J$ are continuous in H , the law of $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ given H_h converges to the law of $\mathbf{Z}_1, \dots, \mathbf{Z}_N$ given H . Moreover, for the intensity it holds that $\lambda(z, H_h) \rightarrow \lambda(z, H)$ locally uniformly for $h \rightarrow \infty$. This gives

$$\lim_{h \rightarrow \infty} \mathbb{E} \left[f(\mathbf{Z}_n) \prod_{i=1}^n g(y_i | \mathbf{Z}_i, H_h) | H_h \right] = \mathbb{E} \left[f(\mathbf{Z}_n) \prod_{i=1}^n g(y_i | \mathbf{Z}_i, H) | H \right]$$

by standard arguments. □

REFERENCES

- E. Alòs, M. E. Mancino, R. Merino, and S. Sanfelici. A fractional model for the COVID-19 pandemic: Application to Italian data. *Stochastic Analysis and Applications*, pages 1–19, 2020.
- T. G. Andersen, T. Bollerslev, F. X. Diebold, and H. Ebens. The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76, 2001.
- O. Barndorff-Nielsen and J. Schmiegel. Brownian semistationary processes and volatility/intermittency. *Advanced Financial Modelling*, 8:1–26, 2009.
- M. Bennedsen, A. Lunde, and M. Pakkanen. Decoupling the Short- and Long-Term Behavior of Stochastic Volatility. *Journal of Financial Econometrics*, 2021.
- P. Carmona and L. Coutin. Fractional Brownian motion and the Markov property. *Electronic Communications in Probability*, 3:95–107, 1998.
- P. Carmona, L. Coutin, and G. Montseny. Approximation of some Gaussian processes. *Statistical Inference for Stochastic Processes*, 3(1-2):161–171, 2000.
- F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323, 1998.
- R. Cont and P. Das. Rough volatility: fact or artefact? *arXiv preprint arXiv:2203.13820*, 2022.
- L. Coutin and L. Decreusefond. Abstract nonlinear filtering theory in the presence of fractional Brownian motion. *The Annals of Applied Probability*, 9(4):1058–1090, 1999.
- L. Coutin and M. Pontier. Approximation of the fractional Brownian sheet via Ornstein-Uhlenbeck sheet. *ESAIM: Probability and Statistics*, 11:115–146, 2007.
- D. Crisan and J. Miguez. Nested particle filters for online parameter estimation in discrete-time state-space Markov models. *Bernoulli*, 24(4A):3039–3086, 2018.
- J. Cvitanić, R. Liptser, and B. Rozovskii. A filtering approach to tracking volatility from prices observed at random times. *The Annals of Applied Probability*, pages 1633–1652, 2006.
- L. Decreusefond and A. S. Üstünel. Fractional Brownian motion: theory and applications. In *ESAIM: Proceedings*, volume 5, pages 75–86, 1998.
- S. Ethier and T. Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 1986.
- R. Frey and W. Runggaldier. A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal of Theoretical and Applied Finance*, 4(02):199–210, 2001.
- M. Fukasawa, T. Takabatake, and R. Westphal. Is volatility rough? *arXiv preprint arXiv:1905.04852*, 2019.
- J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.
- D. Gillespie. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Physical Review E*, 54(2):2084, 1996.
- T. Gneiting and M. Schlather. Stochastic models that separate fractal dimension and the hurst effect. *SIAM Review*, 46(2):269–282, 2004.

- P. Harms. Strong convergence rates for markovian representations of fractional brownian motion. *arXiv preprint arXiv:1902.01471*, 2019.
- H. Kushner. *Probability methods for approximations in stochastic control and for elliptic equations*, volume 129. Academic Press, 1977.
- J. Leppanen, H. Stone, D. J. Lythgoe, S. Williams, and B. Horvath. Sailing in rough waters: examining volatility of fMRI noise. *Magnetic Resonance Imaging*, 78:69–79, 2021.
- L. C. G. Rogers. Things we think we know. In *Options—45 years since the Publication of the Black–Scholes–Merton Model: The Gershon Fintech Center Conference*, pages 173–184. World Scientific, 2023.