Finance Stoch
DOI 10.1007/s00780-017-0335-5

Hybrid scheme for Brownian semistationary processes

Mikkel Bennedsen[1,2] · **Asger Lunde**[1,2] · **Mikko S. Pakkanen**[2,3]

Received: 21 October 2016 / Accepted: 2 May 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract We introduce a simulation scheme for Brownian semistationary processes, which is based on discretizing the stochastic integral representation of the process in the time domain. We assume that the kernel function of the process is regularly varying at zero. The novel feature of the scheme is to approximate the kernel function by a power function near zero and by a step function elsewhere. The resulting approximation of the process is a combination of Wiener integrals of the power function and a Riemann sum, which is why we call this method a hybrid scheme. Our main theoretical result describes the asymptotics of the mean square error of the hybrid scheme, and we observe that the scheme leads to a substantial improvement of accuracy compared to the ordinary forward Riemann-sum scheme, while having the same computational complexity. We exemplify the use of the hybrid scheme by two numerical experiments, where we examine the finite-sample properties of an estimator of the roughness parameter of a Brownian semistationary process and study Monte Carlo option pricing in the rough Bergomi model of Bayer et al. (Quant. Finance 16:887–904, 2016), respectively.

✉ M.S. Pakkanen
m.pakkanen@imperial.ac.uk

M. Bennedsen
mbennedsen@econ.au.dk

A. Lunde
alunde@econ.au.dk

1 Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark

2 CREATES, Aarhus University, Aarhus, Denmark

3 Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

Published online: 28 June 2017

Keywords Stochastic simulation · Discretization · Brownian semistationary process · Stochastic volatility · Regular variation · Estimation · Option pricing · Rough volatility · Volatility smile

Mathematics Subject Classification (2010) 60G12 · 60G22 · 65C20 · 91G60 · 62M09

JEL Classification C22 · G13 · C13

1 Introduction

We study simulation methods for *Brownian semistationary* ($\mathcal{BSS}$) processes, first introduced by Barndorff-Nielsen and Schmiegel [8, 9], which form a flexible class of stochastic processes that are able to capture some common features of empirical time series, such as stochastic volatility (intermittency), roughness, stationarity, and strong dependence. By now these processes have been applied in various contexts, most notably in the study of turbulence in physics [7, 16] and in finance as models of energy prices [4, 11]. A $\mathcal{BSS}$ process X is defined via the integral representation

$$X_t = \int_{-\infty}^{t} g(t-s)\sigma_s \mathrm{d}W_s, \tag{1.1}$$

where W is a two-sided Brownian motion providing the fundamental noise innovations, the amplitude of which is modulated by a stochastic volatility (intermittency) process σ that may depend on W. This driving noise is then convolved with a deterministic kernel function g that specifies the dependence structure of X. The process X can also be viewed as a *moving average* of volatility-modulated Brownian noise, and setting $\sigma_s = 1$, we see that stationary *Brownian moving averages* are nested in this class of processes.

In the applications mentioned above, the case where X is not a semimartingale is particularly relevant. This situation arises when the kernel function g behaves like a power-law near zero; more specifically, when for some $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$,

$$g(x) \propto x^{\alpha} \quad \text{for small } x > 0. \tag{1.2}$$

Here we write "$\propto$" to indicate proportionality in an *informal* sense, anticipating a rigorous formulation of this relationship given in Sect. 2.2 using the theory of *regular variation* [15], which plays a significant role in our subsequent arguments. The case $\alpha = -\frac{1}{6}$ in (1.2) is important in statistical modeling of turbulence [16] as it gives rise to processes that are compatible with Kolmogorov's scaling law for ideal turbulence. Moreover, processes of similar type with $\alpha \approx -0.4$ have been recently used in the context of option pricing as models of *rough volatility* [1, 10, 18, 19]; see Sects. 2.5 and 3.3 below. The case $\alpha = 0$ would (roughly speaking) lead to a process that is a semimartingale, which is thus excluded.

Under (1.2), the trajectories of X behave locally like the trajectories of a *fractional Brownian motion* with Hurst index $H = \alpha + \frac{1}{2} \in (0, 1) \setminus \{\frac{1}{2}\}$. While the *local*

behavior and roughness of X, measured in terms of Hölder regularity, are determined by the parameter α, the *global* behavior of X (e.g., whether the process has long or short memory) depends on the behavior of $g(x)$ as $x \to \infty$, which can be specified independently of α. This should be contrasted with fractional Brownian motion and related *self-similar* models, which necessarily must conform to a restrictive affine relationship between their Hölder regularity (local behavior and roughness) and Hurst index (global behavior), as elucidated by Gneiting and Schlather [20]. Indeed, in the realm of $\mathcal{BSS}$ processes, local and global behavior are conveniently decoupled, which underlines the flexibility of these processes as a modeling framework.

In connection with practical applications, it is important to be able to simulate the process X. If the volatility process σ is deterministic and constant in time, then X will be strictly stationary and Gaussian. This makes X amenable to exact simulation using the Cholesky factorization or circulant embeddings; see e.g. [2, Chapter XI]. However, it seems difficult, if not impossible, to develop an exact method that is applicable with a stochastic σ, as the process X is then neither Markovian nor Gaussian. Thus in the general case, one must resort to approximative methods. To this end, Benth et al. [13] have recently proposed a Fourier-based method of simulating $\mathcal{BSS}$ processes, and more general *Lévy semistationary* ($\mathcal{LSS}$) processes, which relies on approximating the kernel function g in the frequency domain.

In this paper, we introduce a new discretization scheme for $\mathcal{BSS}$ processes based on approximating the kernel function g in the time domain. Our starting point is the Riemann-sum discretization of (1.1). The Riemann-sum scheme builds on an approximation of g using step functions, which has the pitfall of failing to capture appropriately the steepness of g near zero. In particular, this becomes a serious defect under (1.2) when $\alpha \in (-\frac{1}{2}, 0)$. In our new scheme, we mitigate this problem by approximating g using an appropriate power function near zero and a step function elsewhere. The resulting discretization scheme can be realized as a linear combination of Wiener integrals with respect to the driving Brownian motion W and a Riemann sum, which is why we call it a *hybrid scheme*. The hybrid scheme is only slightly more demanding to implement than the Riemann-sum scheme, and the schemes have the same computational complexity as the number of discretization cells tends to infinity.

Our main theoretical result describes the exact asymptotic behavior of the mean square error (MSE) of the hybrid scheme and, as a special case, that of the Riemann-sum scheme. We observe that switching from the Riemann-sum scheme to the hybrid scheme reduces the asymptotic root mean square error (RMSE) substantially. Using merely the simplest variant the of hybrid scheme, where a power function is used in a single discretization cell, the reduction is at least 50% for all $\alpha \in (0, \frac{1}{2})$ and at least 80% for all $\alpha \in (-\frac{1}{2}, 0)$. The reduction in RMSE is close to 100% as α approaches $-\frac{1}{2}$, which indicates that the hybrid scheme indeed resolves the problem of poor precision that affects the Riemann-sum scheme.

To assess the accuracy of the hybrid scheme in practice, we perform two numerical experiments. Firstly, we examine the finite-sample performance of an estimator of the roughness index α, introduced by Barndorff-Nielsen et al. [6] and Corcuera et al. [16]. This experiment enables us to assess how faithfully the hybrid scheme approximates the fine properties of the $\mathcal{BSS}$ process X. Secondly, we study Monte Carlo option pricing in the rough Bergomi stochastic volatility model of Bayer et al. [10]. We use

the hybrid scheme to simulate the volatility process in this model, and we find that the resulting implied volatility smiles are indistinguishable from those obtained using a method that involves exact simulation of the volatility process. Thus we are able to propose a solution to the problem of finding an efficient simulation scheme for the rough Bergomi model, left open in the paper [10].

The rest of this paper is organized as follows. In Sect. 2, we recall the rigorous definition of a $\mathcal{BSS}$ process and introduce our assumptions. We also introduce the hybrid scheme, state our main theoretical result concerning the asymptotics of the mean square error and discuss an extension of the scheme to a class of *truncated* $\mathcal{BSS}$ processes. Section 3 briefly discusses the implementation of the discretization scheme and presents the numerical experiments mentioned above. Finally, Sect. 4 contains the proofs of the theoretical and technical results given in the paper.

2 The model and theoretical results

2.1 Brownian semistationary process

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}}, \mathbb{P})$ be a filtered probability space, satisfying the usual conditions, supporting a (two-sided) standard Brownian motion $W = (W_t)_{t\in\mathbb{R}}$. We consider a Brownian semistationary process

$$X_t = \int_{-\infty}^{t} g(t-s)\sigma_s \mathrm{d}W_s, \quad t \in \mathbb{R}, \tag{2.1}$$

where $\sigma = (\sigma_t)_{t\in\mathbb{R}}$ is an $(\mathcal{F}_t)_{t\in\mathbb{R}}$-predictable process with locally bounded trajectories, which captures the stochastic volatility (intermittency) of X, and where $g : (0, \infty) \to [0, \infty)$ is a Borel-measurable kernel function.

To ensure that the integral (2.1) is well defined, we assume that the kernel function g is square-integrable, that is, $\int_0^\infty g(x)^2 \mathrm{d}x < \infty$. In fact, we shortly introduce some more specific assumptions on g that imply its square-integrability. **Throughout the paper**, we also assume that the process σ has finite second moments, $\mathbb{E}[\sigma_t^2] < \infty$ for all $t \in \mathbb{R}$, and that the process is covariance-stationary, namely,

$$\mathbb{E}[\sigma_s] = \mathbb{E}[\sigma_t], \quad \mathrm{Cov}(\sigma_s, \sigma_t) = \mathrm{Cov}(\sigma_0, \sigma_{|s-t|}), \quad s, t \in \mathbb{R}.$$

These assumptions imply that also X is covariance-stationary, that is,

$$\mathbb{E}[X_t] = 0, \quad \mathrm{Cov}(X_s, X_t) = \mathbb{E}[\sigma_0^2] \int_0^\infty g(x)g(x + |s-t|)\mathrm{d}x, \quad s, t \in \mathbb{R}.$$

However, the process X need not be *strictly* stationary as the dependence between the volatility process σ and the driving Brownian motion W may be time-varying.

2.2 Kernel function

As mentioned above, we consider a kernel function that satisfies $g(x) \propto x^\alpha$ for some $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$ when $x > 0$ is near zero. To make this idea rigorous and to allow

additional flexibility, we formulate our assumptions on g using the theory of regular variation [15] and, more specifically, slowly varying functions.

To this end, recall that a measurable function $L : (0, 1] \to [0, \infty)$ is *slowly varying* at 0 if for any $t > 0$,

$$\lim_{x \to 0} \frac{L(tx)}{L(x)} = 1.$$

Moreover, a function $f(x) = x^\beta L(x)$, $x \in (0, 1]$, where $\beta \in \mathbb{R}$ and L is slowly varying at 0, is said to be *regularly varying* at 0, with β being the *index of regular variation*.

Remark 2.1 Conventionally, slow and regular variation are defined at ∞ [15, Sects. 1.2.1 and 1.4.1]. However, L is slowly varying (resp. regularly varying) at 0 if and only if $x \mapsto L(1/x)$ is slowly varying (resp. regularly varying) at ∞.

A key feature of slowly varying functions, which will be very important in the sequel, is that they can be sandwiched between polynomial functions as follows. If $\delta > 0$ and L is slowly varying at 0 and bounded away from 0 and ∞ on every interval $(u, 1]$, $u \in (0, 1)$, then there exist constants $\overline{C}_\delta \geq \underline{C}_\delta > 0$ such that

$$\underline{C}_\delta x^\delta \leq L(x) \leq \overline{C}_\delta x^{-\delta}, \quad x \in (0, 1]. \tag{2.2}$$

The inequalities above are an immediate consequence of the so-called *Potter bounds* for slowly varying functions; see [15, Theorem 1.5.6(ii)] and (4.1) below. Making δ very small, we see that slowly varying functions are asymptotically negligible in comparison with polynomially growing/decaying functions. Thus, by multiplying power functions and slowly varying functions, regular variation provides a flexible framework to construct functions that behave asymptotically like power functions.

Our assumptions concerning the kernel function g are as follows:

(A1) For some $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$,

$$g(x) = x^\alpha L_g(x), \quad x \in (0, 1],$$

where $L_g : (0, 1] \to [0, \infty)$ is continuously differentiable, slowly varying at 0 and bounded away from 0. Moreover, there exists a constant $C > 0$ such that the derivative L_g' of L_g satisfies

$$|L_g'(x)| \leq C(1 + x^{-1}), \quad x \in (0, 1].$$

(A2) The function g is continuously differentiable on $(0, \infty)$, with derivative g' that is ultimately monotonic and also satisfies $\int_1^\infty g'(x)^2 \mathrm{d}x$.

(A3) For some $\beta \in (-\infty, -\frac{1}{2})$,

$$g(x) = \mathcal{O}(x^\beta), \quad x \to \infty.$$

(Here, and in the sequel, we use the notation $f(x) = \mathcal{O}(h(x))$, $x \to a$, to indicate that $\limsup_{x\to a} |\frac{f(x)}{h(x)}| < \infty$. Additionally, analogous notation is later used for sequences and computational complexity.) In view of the bound (2.2), these assumptions ensure that g is square-integrable. It is worth pointing out that (A1) accommodates functions L_g with $\lim_{x\to 0} L_g(x) = \infty$, e.g. $L_g(x) = 1 - \log x$.

The assumption (A1) influences the short-term behavior and roughness of the process X. A simple way to assess the roughness of X is to study the behavior of its *variogram* (also called *second-order structure function* in the turbulence literature)

$$V_X(h) := \mathbb{E}[|X_h - X_0|^2], \quad h \geq 0,$$

as $h \to 0$. Note that by covariance-stationarity,

$$V_X(|s-t|) = \mathbb{E}[|X_s - X_t|^2], \quad s, t \in \mathbb{R}.$$

Under our assumptions, we have the following characterization of the behavior of V_X near zero, which generalizes a result of Barndorff-Nielsen [3, p. 9] and implies that X has a locally Hölder-continuous modification. Therein, and in what follows, we write $a(x) \sim b(x)$, $x \to y$, to indicate that $\lim_{x\to y} \frac{a(x)}{b(x)} = 1$. The proof of this result is carried out in Sect. 4.1.

Proposition 2.2 *Suppose that* (A1)–(A3) *hold.*

(i) *The variogram of X satisfies*

$$V_X(h) \sim \mathbb{E}[\sigma_0^2]\left(\frac{1}{2\alpha+1} + \int_0^\infty \left((y+1)^\alpha - y^\alpha\right)^2 \mathrm{d}y\right) h^{2\alpha+1} L_g(h)^2, \quad h \to 0,$$

which implies that V_X is regularly varying at zero with index $2\alpha + 1$.

(ii) *The process X has a modification with locally ϕ-Hölder-continuous trajectories for any $\phi \in (0, \alpha + \frac{1}{2})$.*

Motivated by Proposition 2.2, we call α the *roughness index* of the process X. Ignoring the slowly varying factor $L_g(h)^2$ in (2.2), we see that the variogram $V(h)$ behaves like $h^{2\alpha+1}$ for small values of h, which is reminiscent of the scaling property of the increments of a fractional Brownian motion (fBm) with Hurst index $H = \alpha + \frac{1}{2}$. Thus, the process X behaves locally like such an fBm, at least when it comes to second order structure and roughness. (Moreover, the factor $\frac{1}{2\alpha+1} + \int_0^\infty ((y+1)^\alpha - y^\alpha)^2 \mathrm{d}y$ coincides with the normalization coefficient that appears in the Mandelbrot–Van Ness representation [23, Theorem 1.3.1] of an fBm with $H = \alpha + \frac{1}{2}$.)

Let us now look at two examples of a kernel function g that satisfies our assumptions.

Example 2.3 The so-called *gamma kernel*

$$g(x) = x^\alpha e^{-\lambda x}, \quad x \in (0, \infty),$$

with parameters $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$ and $\lambda > 0$, has been used extensively in the literature on $\mathcal{BSS}$ processes. It is particularly important in connection with statistical modeling of turbulence, see Corcuera et al. [16], but it also provides a way to construct generalizations of Ornstein–Uhlenbeck (OU) processes with a roughness that differs from the usual semimartingale case $\alpha = 0$, while mimicking the long-term behavior of an OU process. Moreover, $\mathcal{BSS}$ and $\mathcal{LSS}$ processes defined using the gamma kernel have interesting probabilistic properties; see [24]. An in-depth study of the gamma kernel can be found in [3]. Setting $L_g(x) := e^{-\lambda x}$, which is slowly varying at 0 since $\lim_{x\to 0} L_g(x) = 1$, it is evident that (A1) holds. Since $g(x)$ decays exponentially fast to 0 as $x \to \infty$, it is clear that also (A3) holds. To verify (A2), note that g satisfies

$$g'(x) = \left(\frac{\alpha}{x} - \lambda\right) g(x), \quad g''(x) = \left(\left(\frac{\alpha}{x} - \lambda\right)^2 - \frac{\alpha}{x^2}\right) g(x), \quad x \in (0, \infty),$$

where $\lim_{x\to\infty}((\frac{\alpha}{x} - \lambda)^2 - \frac{\alpha}{x^2}) = \lambda^2 > 0$, so g' is ultimately increasing with

$$g'(x)^2 \le (|\alpha| + \lambda)^2 g(x)^2, \quad x \in [1, \infty).$$

Thus, $\int_1^\infty g'(x)^2 \mathrm{d}x < \infty$ since g is square-integrable.

Example 2.4 Consider the *power-law kernel* function

$$g(x) = x^\alpha (1 + x)^{\beta - \alpha}, \quad x \in (0, \infty),$$

with parameters $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$ and $\beta \in (-\infty, -\frac{1}{2})$. The behavior of this kernel function near zero is similar to that of the gamma kernel, but $g(x)$ decays to zero polynomially as $x \to \infty$, so it can be used to model long memory. In fact, it can be shown that if $\beta \in (-1, -\frac{1}{2})$, then the autocorrelation function of X is not integrable. Clearly, (A1) holds with $L_g(x) := (1 + x)^{\beta - \alpha}$, which is slowly varying at 0 since $\lim_{x\to 0} L_g(x) = 1$. Moreover, note that we can write

$$g(x) = x^\beta K_g(x), \quad x \in (0, \infty),$$

where $K_g(x) := (1 + x^{-1})^{\beta - \alpha}$ satisfies $\lim_{x\to\infty} K_g(x) = 1$. Thus, also (A3) holds. We can check (A2) by computing

$$g'(x) = \frac{\alpha + \beta x}{x(1 + x)} g(x),$$

$$g''(x) = \left(\left(\frac{\alpha + \beta x}{x(1 + x)}\right)^2 + \frac{-\alpha - 2\alpha x - \beta x^2}{x^2(1 + x)^2}\right) g(x), \quad x \in (0, \infty),$$

where $-\alpha - 2\alpha x - \beta x^2 \to \infty$ when $x \to \infty$ (as $\beta < -\frac{1}{2}$), so g' is ultimately increasing. Additionally, we note that

$$g'(x)^2 \le (|\alpha| + |\beta|)^2 g(x)^2, \quad x \in [1, \infty),$$

implying $\int_1^\infty g'(x)^2 \mathrm{d}x < \infty$ since g is square-integrable.

2.3 Hybrid scheme

Let $t \in \mathbb{R}$ and consider discretizing X_t based on its integral representation (2.1) on the grid $\mathcal{G}_t^n := \{t, t - \frac{1}{n}, t - \frac{2}{n}, \ldots\}$ for $n \in \mathbb{N}$. To derive our discretization scheme, let us first note that if the volatility process σ does not vary too much, then it is reasonable to use the approximation

$$X_t = \sum_{k=1}^{\infty} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} g(t-s)\sigma_s \mathrm{d}W_s \approx \sum_{k=1}^{\infty} \sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} g(t-s)\mathrm{d}W_s, \tag{2.3}$$

that is, we keep σ constant in each discretization cell. (Here, and in the sequel, "$\approx$" stands for an informal approximation used for purely heuristic purposes.) If k is "small", then due to (A1), we may approximate

$$g(t-s) \approx (t-s)^{\alpha} L_g\left(\frac{k}{n}\right), \quad t-s \in \left[\frac{k-1}{n}, \frac{k}{n}\right] \setminus \{0\}, \tag{2.4}$$

as the slowly varying function L_g varies "less" than the power function $y \mapsto y^{\alpha}$ near zero; cf. (2.2). If k is "large", or at least $k \geq 2$, then choosing $b_k \in [k-1, k]$ provides an adequate approximation

$$g(t-s) \approx g\left(\frac{b_k}{n}\right), \quad t-s \in \left[\frac{k-1}{n}, \frac{k}{n}\right], \tag{2.5}$$

by (A2). Applying (2.4) to the first κ terms, where $\kappa = 1, 2, \ldots$, and (2.5) to the remaining terms in the approximating series in (2.3) yields

$$\begin{aligned}\sum_{k=1}^{\infty} \sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} g(t-s)\mathrm{d}W_s \approx \sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right)\sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} (t-s)^{\alpha}\mathrm{d}W_s \\ + \sum_{k=\kappa+1}^{\infty} g\left(\frac{b_k}{n}\right)\sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} \mathrm{d}W_s.\end{aligned} \tag{2.6}$$

For completeness, we also allow $\kappa = 0$, in which case we require that $b_1 \in (0, 1]$ and interpret the first sum on the right-hand side of (2.6) as zero. To make numerical implementation feasible, we truncate the second sum on the right-hand side of (2.6) so that both sums have $N_n \geq \kappa + 1$ terms in total. Thus, we arrive at a discretization scheme for X_t, which we call a *hybrid scheme*, given by

$$X_t^n := \check{X}_t^n + \hat{X}_t^n,$$

where

$$\check{X}_t^n := \sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right)\sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} (t-s)^{\alpha}\mathrm{d}W_s, \tag{2.7}$$

$$\hat{X}_t^n := \sum_{k=\kappa+1}^{N_n} g\left(\frac{b_k}{n}\right)\sigma_{t-\frac{k}{n}}\left(W_{t-\frac{k}{n}+\frac{1}{n}} - W_{t-\frac{k}{n}}\right), \tag{2.8}$$

and $\mathbf{b} := (b_k)_{k=\kappa+1}^{\infty}$ is a sequence of real numbers, evaluation points, that must satisfy $b_k \in [k-1, k] \setminus \{0\}$ for each $k \geq \kappa + 1$, but otherwise can be chosen freely.

As it stands, the discretization grid $\mathcal{G}_t^n$ depends on the time t, which may seem cumbersome with regard to sampling X_t^n simultaneously for different times t. However, note that whenever times t and t' are separated by a multiple of $\frac{1}{n}$, the corresponding grids $\mathcal{G}_t^n$ and $\mathcal{G}_{t'}^n$ will intersect. In fact, the hybrid scheme defined by (2.7) and (2.8) can be implemented efficiently, as we shall see in Sect. 3.1, below. Since

$$g\left(\frac{b_k}{n}\right) = g\left(t - \left(t - \frac{b_k}{n}\right)\right),$$

the degenerate case $\kappa = 0$ with $b_k = k$ for all $k \geq 1$ corresponds to the usual Riemann-sum discretization scheme of X_t with (Itô type) forward sums from (2.8). Henceforth, we denote the associated sequence $(k)_{k=\kappa+1}^{\infty}$ by $\mathbf{b}_{\mathrm{FWD}}$, where the subscript "FWD" alludes to forward sums. However, including terms involving Wiener integrals of a power function given by (2.7), that is, having $\kappa \geq 1$, improves the accuracy of the discretization considerably, as we shall see. Having the leeway to select b_k within the interval $[k-1, k] \setminus \{0\}$, so that the function $g(t - \cdot)$ is evaluated at a point that does not necessarily belong to $\mathcal{G}_t^n$, leads additionally to a moderate improvement.

The truncation in the sum (2.8) entails that the stochastic integral (2.1) defining X_t is truncated at $t - \frac{N_n}{n}$. In practice, the value of the parameter N_n should be large enough to mitigate the effect of truncation. To ensure that the truncation point $t - \frac{N_n}{n}$ tends to $-\infty$ as $n \to \infty$ in our asymptotic results, we introduce the following assumption:

(A4) For some $\gamma > 0$,

$$N_n \sim n^{\gamma+1}, \quad n \to \infty.$$

2.4 Asymptotic behavior of mean square error

We are now ready to state our main theoretical result, which gives a sharp description of the asymptotic behavior of the mean square error (MSE) of the hybrid scheme as $n \to \infty$. We defer the proof of this result to Sect. 4.2.

Theorem 2.5 *Suppose that* (A1)–(A4) *hold, so that*

$$\gamma > -\frac{2\alpha + 1}{2\beta + 1}, \tag{2.9}$$

and that for some $\delta > 0$,

$$\mathbb{E}[|\sigma_s - \sigma_0|^2] = \mathcal{O}(s^{2\alpha+1+\delta}), \quad s \downarrow 0. \tag{2.10}$$

Then for all $t \in \mathbb{R}$,

$$\mathbb{E}[|X_t - X_t^n|^2] \sim J(\alpha, \kappa, \mathbf{b})\mathbb{E}[\sigma_0^2]n^{-(2\alpha+1)}L_g(1/n)^2, \quad n \to \infty, \tag{2.11}$$

where

$$J(\alpha, \kappa, \mathbf{b}) := \sum_{k=\kappa+1}^{\infty} \int_{k-1}^{k} (y^{\alpha} - b_k^{\alpha})^2 \mathrm{d}y < \infty. \tag{2.12}$$

Remark 2.6 Note that if $\alpha \in (-\frac{1}{2}, 0)$, then having

$$\mathbb{E}[|\sigma_s - \sigma_0|^2] = \mathcal{O}(s^{\theta}), \quad s \downarrow 0,$$

for all $\theta \in (0, 1)$, ensures that (2.10) holds. (Take, say, $\delta := \frac{1}{2}(1 - (2\alpha + 1)) > 0$ and $\theta := 2\alpha + 1 + \delta = \alpha + 1 \in (0, 1)$.)

When the hybrid scheme is used to simulate the $\mathcal{BSS}$ process X on an equidistant grid $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{\lfloor nT \rfloor}{n}\}$ for some $T > 0$ (see Sect. 3.1 on the details of the implementation), the following consequence of Theorem 2.5 ensures that the covariance structure of the simulated process approximates that of the actual process X.

Corollary 2.7 *Suppose that the assumptions of Theorem 2.5 hold. Then for any s, $t \in \mathbb{R}$ and $\varepsilon > 0$,*

$$|\mathbb{E}[X_t^n X_s^n] - \mathbb{E}[X_t X_s]| = \mathcal{O}(n^{-(\alpha+\frac{1}{2})+\varepsilon}), \quad n \to \infty.$$

Proof Let $s, t \in \mathbb{R}$. Applying the Cauchy–Schwarz inequality, we get

$$\begin{aligned} |\mathbb{E}[X_t^n X_s^n] - \mathbb{E}[X_t X_s]| \leq{}& \mathbb{E}[(X_t^n)^2]^{1/2} \mathbb{E}[|X_s - X_s^n|^2]^{1/2} \\ &+ \mathbb{E}[X_s^2]^{1/2} \mathbb{E}[|X_t - X_t^n|^2]^{1/2}. \end{aligned}$$

We have $\sup_{n\in\mathbb{N}} \mathbb{E}[(X_t^n)^2]^{1/2} < \infty$ since $\mathbb{E}[(X_t^n)^2] \to \mathbb{E}[X_t^2] < \infty$ as $n \to \infty$, by Theorem 2.5. Moreover, Theorem 2.5 and the bound (2.2) imply that we also have $\mathbb{E}[|X_s - X_s^n|^2]^{1/2} = \mathcal{O}(n^{-(\alpha+\frac{1}{2})+\varepsilon})$ and $\mathbb{E}[|X_t - X_t^n|^2]^{1/2} = \mathcal{O}(n^{-(\alpha+\frac{1}{2})+\varepsilon})$ for any $\varepsilon > 0$. □

In Theorem 2.5, the asymptotics of the MSE (2.11) are determined by the behavior of the kernel function g near zero, as specified in (A1). The condition (2.9) ensures that the error from approximating g near zero is asymptotically larger than the error induced by the truncation of the stochastic integral (2.1) at $t - \frac{N_n}{n}$. In fact, a different kind of asymptotics of the MSE, where truncation error becomes dominant, could be derived when (2.9) does not hold, under some additional assumptions, but we do not pursue this direction in the present paper.

While the rate of convergence in (2.11) is fully determined by the roughness index α, which may seem discouraging at first, it turns out that the quantity $J(\alpha, \kappa, \mathbf{b})$, which we call the *asymptotic* MSE, can vary a lot, depending on how we choose κ and $\mathbf{b}$, and can have a substantial impact on the precision of the approximation of X. It is immediate from (2.12) that increasing κ will decrease $J(\alpha, \kappa, \mathbf{b})$. Moreover, for given α and κ, it is straightforward to choose $\mathbf{b}$ so that $J(\alpha, \kappa, \mathbf{b})$ is minimized, as shown in the following result.

Proposition 2.8 *Let $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$ and $\kappa \geq 0$. Among all sequences $\mathbf{b} = (b_k)_{k=\kappa+1}^{\infty}$ with $b_k \in [k-1, k] \setminus \{0\}$ for $k \geq \kappa + 1$, the function $J(\alpha, \kappa, \mathbf{b})$, and consequently the asymptotic MSE induced by the discretization, is minimized by the sequence $\mathbf{b}^*$ given by*

$$b_k^* = \left(\frac{k^{\alpha+1} - (k-1)^{\alpha+1}}{\alpha + 1} \right)^{1/\alpha}, \quad k \geq \kappa + 1.$$

Proof Clearly, a sequence $\mathbf{b} = (b_k)_{k=\kappa+1}^{\infty}$ minimizes the function $J(\alpha, \kappa, \mathbf{b})$ if and only if b_k minimizes $\int_{k-1}^{k} (y^\alpha - b_k^\alpha)^2 \mathrm{d}y$ for any $k \geq \kappa + 1$. By standard L^2-space theory, $c \in \mathbb{R}$ minimizes the integral $\int_{k-1}^{k} (y^\alpha - c)^2 \mathrm{d}y$ if and only if the function $y \mapsto y^\alpha - c$ is orthogonal in L^2 to all constant functions. This is tantamount to

$$\int_{k-1}^{k} (y^\alpha - c) \mathrm{d}y = 0,$$

and computing the integral and solving for c yields

$$c = \frac{k^{\alpha+1} - (k-1)^{\alpha+1}}{\alpha + 1}.$$

Setting $b_k^* := c^{1/\alpha} \in (k-1, k)$ completes the proof. □

To understand how much increasing κ and using the optimal sequence $\mathbf{b}^*$ from Proposition 2.8 improves the approximation, we study numerically the asymptotic *root mean square error* (RMSE) $\sqrt{J(\alpha, \kappa, \mathbf{b})}$. In particular, we assess how much the asymptotic RMSE decreases relative to the RMSE of the forward Riemann-sum scheme ($\kappa = 0$ and $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$) by using the quantity

$$\text{reduction in asymptotic RMSE} = -\frac{\sqrt{J(\alpha, \kappa, \mathbf{b})} - \sqrt{J(\alpha, 0, \mathbf{b}_{\mathrm{FWD}})}}{\sqrt{J(\alpha, 0, \mathbf{b}_{\mathrm{FWD}})}} \cdot 100\%. \tag{2.13}$$

The results are presented in Fig. 1. We find that employing the hybrid scheme with $\kappa \geq 1$ leads to a substantial reduction in the asymptotic RMSE relative to the forward Riemann-sum scheme when $\alpha \in (-\frac{1}{2}, 0)$. Indeed, when $\kappa \geq 1$, the asymptotic RMSE, as a function of α, does not blow up as $\alpha \to -\frac{1}{2}$, while with $\kappa = 0$ it does. This explains why the reduction in the asymptotic RMSE approaches 100% as $\alpha \to -\frac{1}{2}$. When $\alpha \in (0, \frac{1}{2})$, the improvement achieved using the hybrid scheme is more modest, but still considerable. Figure 1 also highlights the importance of using the optimal sequence $\mathbf{b}^*$, instead of $\mathbf{b}_{\mathrm{FWD}}$, as evaluation points in the scheme, in particular when $\alpha \in (0, \frac{1}{2})$. Finally, we observe that increasing κ beyond 2 does not appear to lead to a significant further reduction. Indeed, in our numerical experiments, reported in Sects. 3.2 and 3.3 below, we observe that using $\kappa = 1, 2$ already leads to good results.

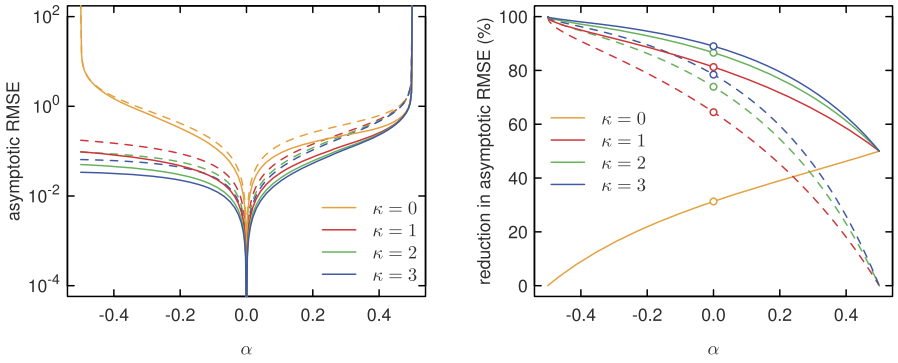

Fig. 1 (*Left*) The asymptotic RMSE given by $\sqrt{J(\alpha,\kappa,\mathbf{b})}$ as a function of $\alpha \in (-\frac{1}{2},\frac{1}{2}) \setminus \{0\}$ for $\kappa = 0,1,2,3$ using $\mathbf{b} = \mathbf{b}^*$ of Proposition 2.8 (*solid lines*) and $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$ (*dashed lines*). (*Right*) Reduction in the asymptotic RMSE relative to the forward Riemann-sum scheme ($\kappa = 0$ and $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$) given by the formula (2.13), plotted as a function of $\alpha \in (-\frac{1}{2},\frac{1}{2}) \setminus \{0\}$ for $\kappa = 0,1,2,3$ using $\mathbf{b} = \mathbf{b}^*$ (*solid lines*) and for $\kappa = 1,2,3$ using $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$ (*dashed lines*). In all computations, we have used the approximations outlined in Remark 2.9 with $N = 1000000$

Remark 2.9 It is non-trivial to evaluate the quantity $J(\alpha,\kappa,\mathbf{b})$ numerically. Computing the integral in (2.12) explicitly, we can approximate $J(\alpha,\kappa,\mathbf{b})$ by

$$J_N(\alpha,\kappa,\mathbf{b}) := \sum_{k=\kappa+1}^{N} \left(\frac{k^{2\alpha+1} - (k-1)^{2\alpha+1}}{2\alpha+1} - \frac{2b_k^{\alpha}(k^{\alpha+1} - (k-1)^{\alpha+1})}{\alpha+1} + b_k^{2\alpha} \right)$$

with some large $N \in \mathbb{N}$. This approximation is adequate when $\alpha \in (-\frac{1}{2}, 0)$, but its accuracy deteriorates when $\alpha \to \frac{1}{2}$. In particular, the singularity of the function $\alpha \mapsto J(\alpha,\kappa,\mathbf{b})$ at $\frac{1}{2}$ is difficult to capture using $J_N(\alpha,\kappa,\mathbf{b})$ with numerically feasible values of N. To overcome this numerical problem, we introduce a correction term in the case $\alpha \in (0, \frac{1}{2})$. The correction term can be derived informally as follows. By the mean value theorem, and since $b_k^* \approx k - \frac{1}{2}$ for large k, we have

$$(y^{\alpha} - b_k^{\alpha})^2 = \alpha^2 \xi^{2\alpha-2}(y - b_k)^2 \approx \begin{cases} \alpha^2 k^{2\alpha-2}(y-k)^2, & \mathbf{b} = \mathbf{b}_{\mathrm{FWD}}, \\ \alpha^2 k^{2\alpha-2}(y-k+\frac{1}{2})^2, & \mathbf{b} = \mathbf{b}^*, \end{cases}$$

where $\xi = \xi(y, b_k) \in [k-1, k]$, for large k. Thus, for large N, we obtain

$$\begin{aligned} &J(\alpha,\kappa,\mathbf{b}) - J_N(\alpha,\kappa,\mathbf{b}) \\ &= \sum_{k=N+1}^{\infty} \int_{k-1}^{k} (y^{\alpha} - b_k^{\alpha})^2 \mathrm{d}y \\ &\approx \begin{cases} \alpha^2 \sum_{k=N+1}^{\infty} k^{2\alpha-2} \int_{k-1}^{k} (y-k)^2 \mathrm{d}y, & \mathbf{b} = \mathbf{b}_{\mathrm{FWD}}, \\ \alpha^2 \sum_{k=N+1}^{\infty} k^{2\alpha-2} \int_{k-1}^{k} (y-k+\frac{1}{2})^2 \mathrm{d}y, & \mathbf{b} = \mathbf{b}^*, \end{cases} \\ &= \begin{cases} \frac{\alpha^2}{3} \zeta(2-2\alpha, N+1), & \mathbf{b} = \mathbf{b}_{\mathrm{FWD}}, \\ \frac{\alpha^2}{12} \zeta(2-2\alpha, N+1), & \mathbf{b} = \mathbf{b}^*, \end{cases} \end{aligned}$$

where $\zeta(x,s) := \sum_{k=0}^{\infty} \frac{1}{(k+s)^x}$, $x > 1$, $s > 0$, is the *Hurwitz zeta function*, which can be evaluated using accurate numerical algorithms.

Remark 2.10 Unlike the Fourier-based method of Benth et al. [13], the hybrid scheme does not require truncating the singularity of the kernel function g when $\alpha \in (-\frac{1}{2}, 0)$, which is beneficial to maintaining the accuracy of the scheme when α is near $-\frac{1}{2}$. Let us briefly analyze the effect on the approximation error of truncating the singularity of g; cf. [13, pp. 75–76]. Consider for any $\varepsilon > 0$ the modified $\mathcal{BSS}$ process

$$\tilde{X}_t^{\varepsilon} := \int_{-\infty}^{t} g_{\varepsilon}(t-s)\sigma_s \mathrm{d}W_s, \quad t \in \mathbb{R},$$

defined using the truncated kernel function

$$g_{\varepsilon}(x) := \begin{cases} g(\varepsilon), & x \in (0, \varepsilon], \\ g(x), & x \in (\varepsilon, \infty). \end{cases}$$

Adapting the proof of Theorem 2.5 in a straightforward manner, it is possible to show that under (A1) and (A3),

$$\begin{aligned} \mathbb{E}[|X_t - \tilde{X}_t^{\varepsilon}|^2] &= \mathbb{E}[\sigma_0^2] \int_0^{\varepsilon} \big(g(s) - g(\varepsilon)\big)^2 \mathrm{d}s \\ &\sim \underbrace{\left(\frac{1}{2\alpha+1} - \frac{2}{\alpha+1} + 1\right)}_{=:\tilde{J}(\alpha)} \mathbb{E}[\sigma_0^2] \varepsilon^{2\alpha+1} L_g(\varepsilon)^2, \quad \varepsilon \downarrow 0, \end{aligned}$$

for any $t \in \mathbb{R}$. While the rate of convergence, as $\varepsilon \downarrow 0$, of the MSE that arises from replacing g with g_{ε} is analogous to the rate of convergence of the hybrid scheme, it is important to note that the factor $\tilde{J}(\alpha)$ blows up as $\alpha \downarrow -\frac{1}{2}$. In fact, $\tilde{J}(\alpha)$ is equal to the first term in the series that defines $J(\alpha, 0, \mathbf{b}_{\mathrm{FWD}})$ and

$$\tilde{J}(\alpha) \sim J(\alpha, 0, \mathbf{b}_{\mathrm{FWD}}), \quad \alpha \downarrow -\frac{1}{2},$$

which indicates that the effect of truncating the singularity, in terms of MSE, is similar to the effect of using the forward Riemann-sum scheme to discretize the process when α is near $-\frac{1}{2}$. In particular, the truncation threshold ε would then have to be very small in order to keep the truncation error in check.

2.5 Extension to truncated Brownian semistationary processes

It is useful to extend the hybrid scheme to a class of non-stationary processes that are closely related to $\mathcal{BSS}$ processes. This extension is important in connection with an application to the so-called rough Bergomi model which we discuss in Sect. 3.3 below. More precisely, we consider processes of the form

$$Y_t = \int_0^t g(t-s)\sigma_s \mathrm{d}W_s, \quad t \geq 0, \tag{2.14}$$

where the kernel function g, volatility process σ and driving Brownian motion W are as before. We call Y a *truncated Brownian semistationary* ($\mathcal{TBSS}$) process, as Y is obtained from the $\mathcal{BSS}$ process X by truncating the stochastic integral in (2.1) at 0. Of the preceding assumptions, only (A1) and (A2) are needed to ensure that the stochastic integral in (2.14) exists—in fact, of (A2), only the requirement that g is differentiable on $(0, \infty)$ comes into play.

The $\mathcal{TBSS}$ process Y does not have covariance-stationary increments, so we define its (time-dependent) variogram as

$$V_Y(h,t) := \mathbb{E}[|Y_{t+h} - Y_t|^2], \quad h, t \geq 0.$$

Extending Proposition 2.2, we can describe the behavior of $h \mapsto V_Y(h,t)$ near zero as follows. The existence of a locally Hölder-continuous modification is then a straightforward consequence. We omit the proof of this result, as it would be a straightforward adaptation of the proof of Proposition 2.2.

Proposition 2.11 *Suppose that* (A1) *and* (A2) *hold.*

(i) *The variogram of Y satisfies for any $t \geq 0$ that as $h \to 0$,*

$$V_Y(h,t) \sim \mathbb{E}[\sigma_0^2]\left(\frac{1}{2\alpha+1} + \mathbf{1}_{(0,\infty)}(t)\int_0^\infty \left((y+1)^\alpha - y^\alpha\right)^2 \mathrm{d}y\right) h^{2\alpha+1} L_g(h)^2,$$

which implies that $h \mapsto V_Y(h,t)$ is regularly varying at zero with index $2\alpha+1$.

(ii) *The process Y has a modification with locally ϕ-Hölder-continuous trajectories for any $\phi \in (0, \alpha + \frac{1}{2})$.*

Note that while the increments of Y are not covariance-stationary, the asymptotic behavior of $V_Y(h,t)$ is the same as that of $V_X(h)$ as $h \to 0$ (cf. Proposition 2.2) for any $t > 0$. Thus, the increments of Y (apart from increments starting at time 0) are locally like the increments of X.

We define the hybrid scheme to discretize Y_t, for any $t \geq 0$, as

$$Y_t^n := \check{Y}_t^n + \hat{Y}_t^n, \tag{2.15}$$

where

$$\check{Y}_t^n := \sum_{k=1}^{\min\{\lfloor nt \rfloor, \kappa\}} L_g\left(\frac{k}{n}\right)\sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} (t-s)^\alpha \mathrm{d}W_s,$$

$$\hat{Y}_t^n := \sum_{k=\kappa+1}^{\lfloor nt \rfloor} g\left(\frac{b_k}{n}\right)\sigma_{t-\frac{k}{n}}\left(W_{t-\frac{k}{n}+\frac{1}{n}} - W_{t-\frac{k}{n}}\right).$$

In effect, we simply drop the summands in (2.7) and (2.8) that correspond to integrals and increments on the negative real line. We make remarks on the implementation of this scheme in Sect. 3.1 below.

The MSE of the hybrid scheme for the $\mathcal{TBSS}$ process Y has the following asymptotic behavior as $n \to \infty$, which is in fact identical to the asymptotic behavior of the

MSE of the hybrid scheme for $\mathcal{BSS}$ processes. We omit the proof of this result, which would be a simple modification of the proof of Theorem 2.5.

Theorem 2.12 *Suppose that* (A1) *and* (A2) *hold, and that for some* $\delta > 0$,

$$\mathbb{E}[|\sigma_s - \sigma_0|^2] = \mathcal{O}(s^{2\alpha+1+\delta}), \quad s \downarrow 0.$$

Then for all $t > 0$,

$$\mathbb{E}[|Y_t - Y_t^n|^2] \sim J(\alpha, \kappa, \mathbf{b})\mathbb{E}[\sigma_0^2]n^{-(2\alpha+1)}L_g(1/n)^2, \quad n \to \infty,$$

where $J(\alpha, \kappa, \mathbf{b})$ *is as in Theorem* 2.5.

Remark 2.13 Under the assumptions of Theorem 2.12, the conclusion of Corollary 2.7 holds mutatis mutandis. In particular, the covariance structure of the discretized $\mathcal{TBSS}$ process approaches that of Y when $n \to \infty$.

3 Implementation and numerical experiments

3.1 Practical implementation

Simulating the $\mathcal{BSS}$ process X on the equidistant grid $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{\lfloor nT \rfloor}{n}\}$ for some $T > 0$ using the hybrid scheme entails generating

$$X_{\frac{i}{n}}^n, \quad i = 0, 1, \ldots, \lfloor nT \rfloor. \tag{3.1}$$

Provided that we can simulate the random variables

$$W_{i,j}^n := \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(\frac{i+j}{n} - s\right)^\alpha \mathrm{d}W_s, \qquad \begin{matrix} i = -N_n, -N_n+1, \ldots, \lfloor nT \rfloor - 1, \\ j = 1, \ldots, \kappa, \end{matrix} \tag{3.2}$$

$$W_i^n := \int_{\frac{i}{n}}^{\frac{i+1}{n}} \mathrm{d}W_s, \qquad i = -N_n, -N_n+1, \ldots, \lfloor nT \rfloor - 1, \tag{3.3}$$

$$\sigma_i^n := \sigma_{\frac{i}{n}}, \qquad i = -N_n, -N_n+1, \ldots, \lfloor nT \rfloor - 1,$$

we can compute (3.1) via the formula

$$X_{\frac{i}{n}}^n = \underbrace{\sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right)\sigma_{i-k}^n W_{i-k,k}^n}_{=\check{X}_{\frac{i}{n}}^n} + \underbrace{\sum_{k=\kappa+1}^{N_n} g\left(\frac{b_k^*}{n}\right)\sigma_{i-k}^n W_{i-k}^n}_{=\hat{X}_{\frac{i}{n}}^n}. \tag{3.4}$$

In order to simulate (3.2) and (3.3), it is instrumental to note that the $\kappa + 1$-dimensional random vectors

$$\mathbf{W}_i^n := (W_i^n, W_{i,1}^n, \ldots, W_{i,\kappa}^n), \quad i = -N_n, -N_n+1, \ldots, \lfloor nT \rfloor - 1,$$

are i.i.d. according to a multivariate Gaussian distribution with mean zero and covariance matrix Σ given by

$$\Sigma_{1,1} = \frac{1}{n}, \quad \Sigma_{1,j} = \Sigma_{j,1} = \frac{(j-1)^{\alpha+1} - (j-2)^{\alpha+1}}{(\alpha+1)n^{\alpha+1}},$$
$$\Sigma_{j,j} = \frac{(j-1)^{2\alpha+1} - (j-2)^{2\alpha+1}}{(2\alpha+1)n^{2\alpha+1}},$$

for $j = 2, \dots, \kappa+1$, and

$$\Sigma_{j,k} = \frac{1}{(\alpha+1)n^{2\alpha+1}}\Big((j-1)^{\alpha+1}(k-1)^{\alpha}{}_2F_1\Big(-\alpha, 1, \alpha+2, \frac{j-1}{k-1}\Big) - (j-2)^{\alpha+1}(k-2)^{\alpha}{}_2F_1\Big(-\alpha, 1, \alpha+2, \frac{j-2}{k-2}\Big)\Big), \quad (3.5)$$

for $j, k = 2, \dots, \kappa+1$ such that $j < k$, where ${}_2F_1$ stands for the *Gauss hypergeometric function*; see e.g. [17, Sect. 2.1.1] for the definition. (When $k < j$, set $\Sigma_{j,k} = \Sigma_{k,j}$.) For the convenience of the reader, we provide a proof of (3.5) in Sect. 4.3.

Thus, $(\mathbf{W}_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$ can be generated by taking independent draws from the multivariate Gaussian distribution $N_{\kappa+1}(\mathbf{0}, \Sigma)$. If the volatility process σ is independent of W, then $(\sigma_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$ can be generated separately, possibly using exact methods. (Exact methods are available e.g. for Gaussian processes, as mentioned in the introduction, and diffusions; see [14].) In the case where σ depends on W, simulating $(\mathbf{W}_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$ and $(\sigma_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$ is less straightforward. That said, if σ is driven by a standard Brownian motion Z, correlated with W, say, one could rely on a factor decomposition

$$Z_t := \rho W_t + \sqrt{1-\rho^2}\, W_t^{\perp}, \quad t \in \mathbb{R}, \quad (3.6)$$

where $\rho \in [-1, 1]$ is the correlation parameter and $(W_t^{\perp})_{t\in[0,T]}$ is a standard Brownian motion independent of W. Then one would first generate $(\mathbf{W}_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$, use (3.6) to generate $(Z_{\frac{i+1}{n}} - Z_{\frac{i}{n}})_{i=-N_n}^{\lfloor nT\rfloor-1}$, and employ some appropriate approximate method to produce $(\sigma_i^n)_{i=-N_n}^{\lfloor nT\rfloor-1}$ thereafter. This approach has, however, the caveat that it induces an additional approximation error, not quantified in Theorem 2.5.

Remark 3.1 In the case of the $\mathcal{TBSS}$ process Y introduced in Sect. 2.5, the observations $Y_{\frac{i}{n}}^n$, $i = 0, 1, \dots, \lfloor nT\rfloor$, given by the hybrid scheme (2.15) can be computed via

$$Y_{\frac{i}{n}}^n = \sum_{k=1}^{\min\{i,\kappa\}} L_g\Big(\frac{k}{n}\Big)\sigma_{i-k}^n W_{i-k,k}^n + \sum_{k=\kappa+1}^{i} g\Big(\frac{b_k^*}{n}\Big)\sigma_{i-k}^n W_{i-k}^n, \quad (3.7)$$

using the random vectors $\mathbf{W}_i^n$, $i = 0, 1, \dots, \lfloor nT\rfloor - 1$, and random variables σ_i^n, $i = 0, 1, \dots, \lfloor nT\rfloor - 1$.

In the hybrid scheme, it typically suffices to take κ to be at most 3. Thus, in (3.4), the first sum $\check{X}^n_{\frac{i}{n}}$ requires only a negligible computational effort. By contrast, the number of terms in the second sum $\hat{X}^n_{\frac{i}{n}}$ increases as $n \to \infty$. It is then useful to note that

$$\hat{X}^n_{\frac{i}{n}} = \sum_{k=1}^{N_n} \Gamma_k \Xi_{i-k} = (\Gamma \star \Xi)_i,$$

where

$$\Gamma_k := \begin{cases} 0, & k = 1, \dots, \kappa, \\ g(\frac{b^*_k}{n}), & k = \kappa+1, \kappa+2, \dots, N_n, \end{cases}$$
$$\Xi_k := \sigma^n_k W^n_k, \qquad k = -N_n, -N_n+1, \dots, \lfloor nT \rfloor - 1,$$

and $\Gamma \star \Xi$ stands for the discrete convolution of the sequences Γ and Ξ. It is well known that the discrete convolution can be evaluated efficiently using a fast Fourier transform (FFT). The computational complexity of simultaneously evaluating $(\Gamma \star \Xi)_i$ for *all* $i = 0, 1, \dots, \lfloor nT \rfloor$ using an FFT is $\mathcal{O}(N_n \log N_n)$, see [22, Sect. 3.3.4], which under (A4) translates to $\mathcal{O}(n^{\gamma+1} \log n)$. The computational complexity of the entire hybrid scheme is then $\mathcal{O}(n^{\gamma+1} \log n)$, provided that $(\sigma^n_i)_{i=-N_n}^{\lfloor nT \rfloor - 1}$ is generated using a scheme with complexity not exceeding $\mathcal{O}(n^{\gamma+1} \log n)$. As a comparison, we mention that the complexity of an exact simulation of a stationary Gaussian process using circulant embeddings is $\mathcal{O}(n \log n)$ [2, Chapter XI, Sect. 3], whereas the complexity of the Cholesky factorization is $\mathcal{O}(n^3)$ [2, Chapter XI, Sect. 2].

Remark 3.2 With $\mathcal{TBSS}$ processes, the computational complexity of the hybrid scheme via (3.7) is $\mathcal{O}(n \log n)$.

Figure 2 presents examples of trajectories of the $\mathcal{BSS}$ process X using the hybrid scheme with $\kappa = 1, 2$ and $\mathbf{b} = \mathbf{b}^*$. We choose the kernel function g to be the gamma kernel (Example 2.3) with $\lambda = 1$. We also discretize X using the Riemann-sum scheme, $\kappa = 0$ with $\mathbf{b} \in \{\mathbf{b}_{\mathrm{FWD}}, \mathbf{b}^*\}$ (that is, the forward Riemann-sum scheme and its counterpart with optimally chosen evaluation points). We can make two observations. Firstly, we see how the roughness parameter α controls the regularity properties of the trajectories of X—as we decrease α, the trajectories of X become increasingly rough. Secondly, and more importantly, we see how the simulated trajectories coming from the Riemann-sum and hybrid schemes can be rather different, even though we use the same innovations for the driving Brownian motion. In fact, the two variants of the hybrid scheme ($\kappa = 1, 2$) yield almost identical trajectories, while the Riemann-sum scheme ($\kappa = 0$) produces trajectories that are comparatively smoother, this difference becoming more apparent as α approaches $-\frac{1}{2}$. Indeed, in the extreme case with $\alpha = -0.499$, both variants of the Riemann-sum scheme break down and yield anomalous trajectories with very little variation, while the hybrid scheme continues to produce accurate results. The fact that the hybrid scheme is

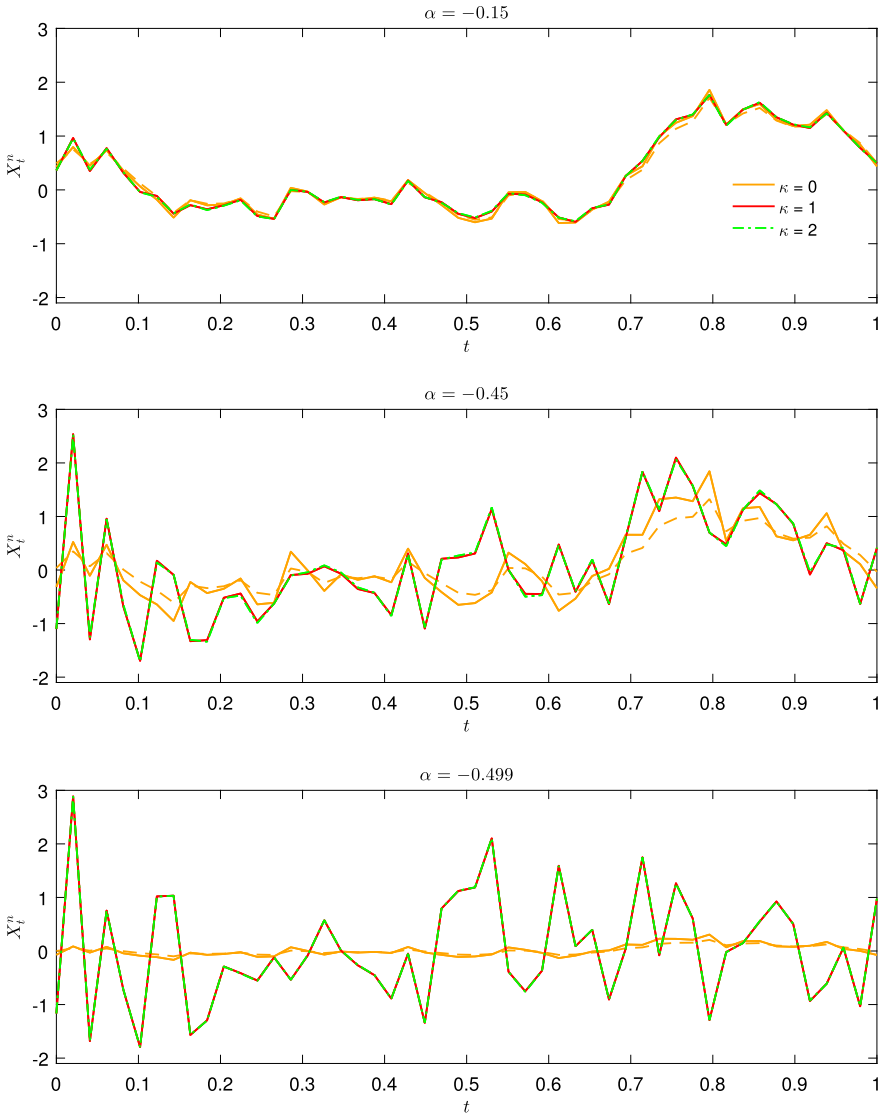

Fig. 2 Discretized trajectories of a $\mathcal{BSS}$ process, where g is the gamma kernel (Example 2.3), $\lambda = 1$ and $\sigma_t = 1$ for all $t \in \mathbb{R}$. Trajectories consisting of $n = 50$ observations on $[0, 1]$ were generated with the hybrid scheme ($\kappa = 1, 2$ and $\mathbf{b} = \mathbf{b}^*$) and Riemann-sum scheme ($\kappa = 0$ and $\mathbf{b} = \mathbf{b}^*$ (*solid lines*), $\mathbf{b} = \mathbf{b}_{\text{FWD}}$ (*dashed lines*)), using the same innovations for the driving Brownian motion in all cases and $N_n = \lfloor 50^{1.5} \rfloor = 353$. The simulated processes were normalized to have unit (stationary) variance

able to reproduce the fine properties of rough $\mathcal{BSS}$ processes, even for values of α very close to $-\frac{1}{2}$, is backed up by a further experiment reported in the following section.

3.2 Estimation of the roughness parameter

Suppose that we have observations $X_{\frac{i}{m}}$, $i = 0, 1, \dots, m$, of the $\mathcal{BSS}$ process X given by (2.1), for some $m \in \mathbb{N}$. Barndorff-Nielsen et al. [6] and Corcuera et al. [16] discuss how the roughness index α can be estimated consistently as $m \to \infty$. The method is based on the *change-of-frequency* (COF) statistics

$$\mathrm{COF}(X, m) = \frac{\sum_{k=5}^{m} |X_{\frac{k}{m}} - 2X_{\frac{k-2}{m}} + X_{\frac{k-4}{m}}|^2}{\sum_{k=3}^{m} |X_{\frac{k}{m}} - 2X_{\frac{k-1}{m}} + X_{\frac{k-2}{m}}|^2}, \quad m \geq 5,$$

which compare the realized quadratic variations of X, using second-order increments, with two different lag lengths. Corcuera et al. [16] have shown that under some assumptions on the process X, which are similar to (A1)–(A3) albeit slightly more restrictive, it holds that

$$\hat{\alpha}(X, m) := \frac{\log \mathrm{COF}(X, m)}{2 \log 2} - \frac{1}{2} \xrightarrow{\mathbb{P}} \alpha, \quad m \to \infty. \tag{3.8}$$

An in-depth study of the finite sample performance of this COF estimator can be found in [12].

To examine how well the hybrid scheme reproduces the fine properties of the $\mathcal{BSS}$ process in terms of regularity/roughness, we apply the COF estimator to discretized trajectories of X, where the kernel function g is again the gamma kernel (Example 2.3) with $\lambda = 1$, generated using the hybrid scheme with $\kappa = 1, 2, 3$ and $\mathbf{b} = \mathbf{b}^*$. We consider the case where the volatility process satisfies $\sigma_t = 1$, that is, the process X is Gaussian. This allows us to quantify and control the intrinsic bias and noisiness, measured in terms of standard deviation, of the estimation method itself, by initially applying the estimator to trajectories that have been simulated using an exact method based on the Cholesky factorization. We then study the behavior of the estimator when applied to a discretized trajectory, while decreasing the step size of the discretization scheme. More precisely, we simulate $\hat{\alpha}(X^n, m)$, where $m = 500$ and X_n is the hybrid scheme for X with $n = ms$ and $s \in \{1, 2, 5\}$. This means that we compute $\hat{\alpha}(X^n, m)$ using m observations obtained by subsampling every sth observation in the sequence $X^n_{\frac{i}{n}}$, $i = 0, 1, \dots, n$. As a comparison, we repeat these simulations substituting the hybrid scheme with the Riemann-sum scheme, using $\kappa = 0$ with $\mathbf{b} \in \{\mathbf{b}_{\mathrm{FWD}}, \mathbf{b}^*\}$.

The results are presented in Fig. 3. We observe that the intrinsic bias of the estimator with $m = 500$ observations is negligible, and hence the bias of the estimates computed from discretized trajectories is then attributable to approximation error arising from the respective discretization scheme, where positive (resp. negative) bias indicates that the simulated trajectories are smoother (resp. rougher) than those of the process X. Concentrating first on the baseline case $s = 1$, we note that the hybrid scheme produces essentially unbiased results when $\alpha \in (-\frac{1}{2}, 0)$, while there is moderate bias when $\alpha \in (0, \frac{1}{2})$, which disappears when passing from $\kappa = 1$ to $\kappa = 3$, even for values of α very close to $\frac{1}{2}$. (The largest value of α considered in our simulations is $\alpha = 0.49$; one would expect the performance to weaken as α approaches $\frac{1}{2}$,

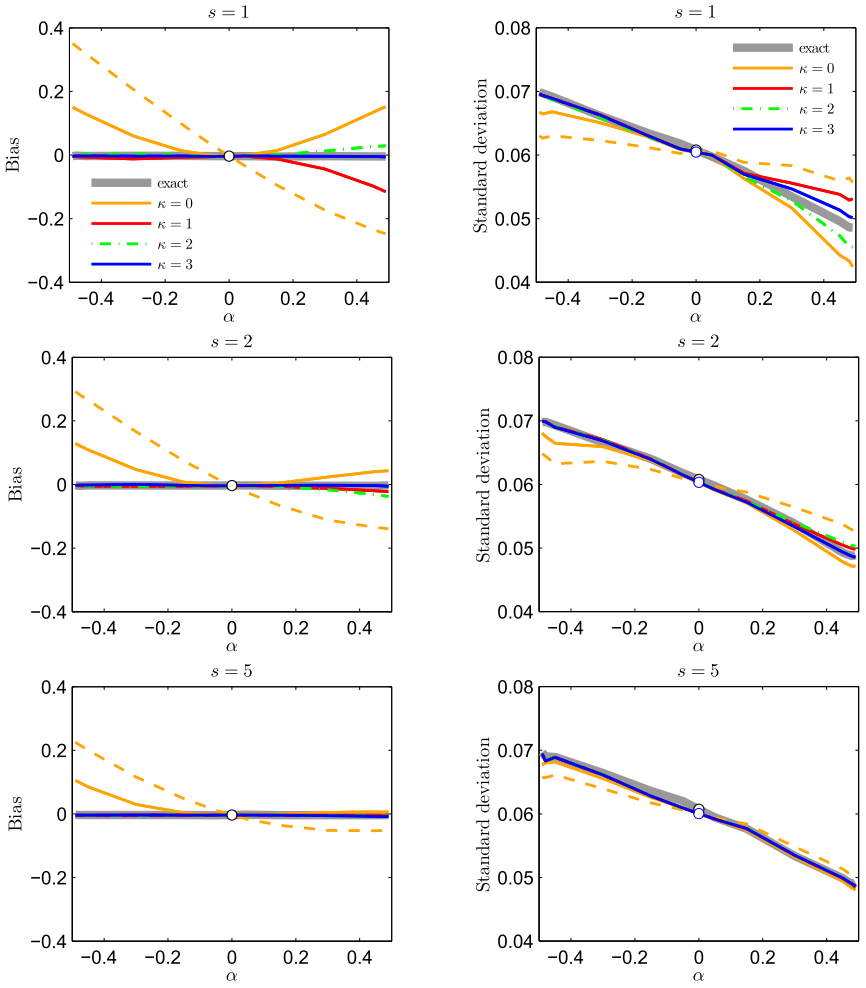

Fig. 3 Bias and standard deviation of the COF estimator (3.8) of the roughness index α, when applied to discretized trajectories of a $\mathcal{BSS}$ process with the gamma kernel (Example 2.3), $\lambda = 1$ and $\sigma_t = 1$ for all $t \in \mathbb{R}$. Trajectories were generated using an exact method based on the Cholesky factorization, the hybrid scheme ($\kappa = 1, 2, 3$ and $\mathbf{b} = \mathbf{b}^*$) and the Riemann-sum scheme ($\kappa = 0$ and $\mathbf{b} = \mathbf{b}^*$ (*solid lines*), $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$ (*dashed lines*)). In the experiment, $n = ms$ observations were generated on $[0, 1]$, where $m = 500$ and $s \in \{1, 2, 5\}$, using $N_n = \lfloor n^{1.5} \rfloor$. Every sth observation was then subsampled, resulting in $m = 500$ observations that were used to compute the estimate $\hat{\alpha}(X^n, m)$ of the roughness index α. The number of Monte Carlo replications is 10000

cf. Fig. 1, but this range of parameter values seems to be of limited practical interest.) The standard deviations exhibit a similar pattern. The corresponding results for the Riemann-sum scheme are clearly inferior, exhibiting significant bias, while using optimal evaluation points ($\mathbf{b} = \mathbf{b}^*$) improves the situation slightly. In particular, the bias in the case $\alpha \in (-\frac{1}{2}, 0)$ is positive, indicating too smooth discretized trajectories, which is connected with the failure of the Riemann-sum scheme with α near $-\frac{1}{2}$, il-

lustrated in Fig. 2. With $s = 2$ and $s = 5$, the results improve with both schemes. Notably, in the case $s = 5$, the performance of the hybrid scheme even with $\kappa = 1$ is on a par with the exact method. However, the improvements with the Riemann-sum scheme are more meager, as considerable bias persists when α is near $-\frac{1}{2}$.

3.3 Option pricing under rough volatility

As another experiment, we study Monte Carlo option pricing in the *rough Bergomi* (rBergomi) model of Bayer et al. [10]. In the rBergomi model, the logarithmic spot variance of the price of the underlying is modeled by a rough Gaussian process, which is a special case of (2.14). By virtue of the rough volatility process, the model fits well to observed implied volatility smiles [10, Sect. 5].

More precisely, the price of the underlying in the rBergomi model with time horizon $T > 0$ is defined, under an equivalent martingale measure identified with $\mathbb{P}$, as

$$S_t := S_0 \exp\left(\int_0^t \sqrt{v_s} \mathrm{d}Z_s - \frac{1}{2} \int_0^t v_s \mathrm{d}s \right), \quad t \in [0, T],$$

using the spot variance process

$$v_t := \xi_t^0 \exp\left(\underbrace{\eta \sqrt{2\alpha + 1} \int_0^t (t - s)^\alpha \mathrm{d}W_s}_{=:Y_t} - \frac{\eta^2}{2} t^{2\alpha+1} \right), \quad t \in [0, T].$$

Above, $S_0 > 0$, $\eta > 0$ and $\alpha \in (-\frac{1}{2}, 0)$ are deterministic parameters, and Z is a standard Brownian motion given by

$$Z_t := \rho W_t + \sqrt{1 - \rho^2} W_t^\perp, \quad t \in [0, T], \tag{3.9}$$

where $\rho \in (-1, 1)$ is the correlation parameter and $(W_t^\perp)_{t \in [0,T]}$ is a standard Brownian motion independent of W. The process $(\xi_t^0)_{t \in [0,T]}$ is the so-called *forward variance curve* [10, Sect. 3], which we assume here to be flat, $\xi_t^0 = \xi > 0$ for all $t \in [0, T]$.

We aim to compute by using Monte Carlo simulation the price of a European call option struck at $K > 0$ with maturity T, which is given by

$$\mathrm{C}(S_0, K, T) := \mathbb{E}[(S_T - K)^+]. \tag{3.10}$$

The approach suggested by Bayer et al. [10] involves sampling the Gaussian processes Z and Y on a discrete time-grid using exact simulation and then approximating S and v using Euler discretization. We modify this approach by using the hybrid scheme to simulate Y, instead of the computationally more costly exact simulation. As the hybrid scheme involves simulating increments of the Brownian motion W driving Y, we can conveniently simulate the increments of Z, needed for the Euler discretization of S, by using the representation (3.9).

We map the option price $\mathrm{C}(S_0, K, T)$ to the corresponding Black–Scholes implied volatility $\mathrm{IV}(S_0, K, T)$. Reparametrizing the implied volatility using the log-strike $k := \log(K/S_0)$ allows us to drop the dependence on the initial price, so we abuse

Table 1 Parameter values used in the rBergomi model

S_0	ξ	η	α	ρ
1	0.235^2	1.9	−0.43	−0.9

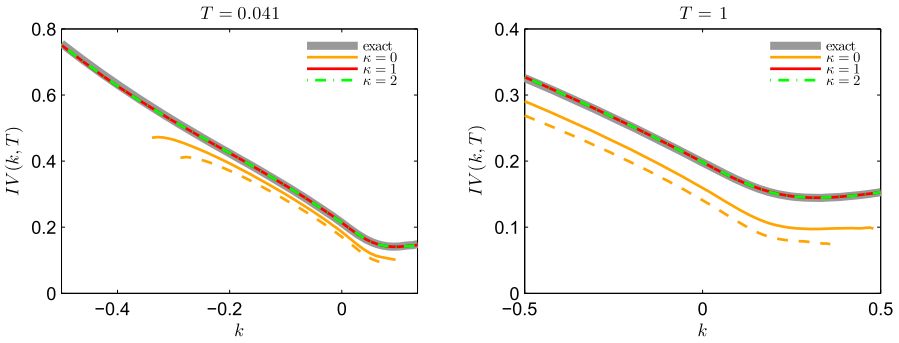

Fig. 4 Implied volatility smiles corresponding to the option price (3.10), computed using Monte Carlo simulation (500 time steps, 1000000 replications), with two maturities: $T = 0.041$ (*left*) and $T = 1$ (*right*). The spot variance process v was simulated using an exact method, the hybrid scheme ($\kappa = 1, 2$ and $\mathbf{b} = \mathbf{b}^*$) and the Riemann-sum scheme ($\kappa = 0$ and $\mathbf{b} = \mathbf{b}^*$ (*solid lines*), $\mathbf{b} = \mathbf{b}_{\mathrm{FWD}}$ (*dashed lines*)). The parameter values used in the rBergomi model are given in Table 1

notation slightly and write $\mathrm{IV}(k, T)$ for the corresponding implied volatility. Figure 4 displays implied volatility smiles obtained from the rBergomi model using the hybrid and Riemann-sum schemes to simulate Y, as discussed above, and compares these to the smiles obtained using an exact simulation of Y via Cholesky factorization. The parameter values are given in Table 1. They have been adopted from Bayer et al. [10], who demonstrate that they result in realistic volatility smiles. We consider two different maturities: "short", $T = 0.041$, and "long", $T = 1$.

We observe that the Riemann-sum scheme ($\kappa = 0$, $\mathbf{b} \in \{\mathbf{b}_{\mathrm{FWD}}, \mathbf{b}^*\}$) is able to capture the shape of the implied volatility smile, but not its level. But the method even breaks down with more extreme log-strikes (the prices are so low that the root-finding algorithm used to compute the implied volatility would return zero). In contrast, the hybrid scheme with $\kappa = 1, 2$ and $\mathbf{b} = \mathbf{b}^*$ yields implied volatility smiles that are indistinguishable from the benchmark smiles obtained using exact simulation. Further, there is no discernible difference between the smiles obtained using $\kappa = 1$ and $\kappa = 2$. As in the previous section, we observe that the hybrid scheme is indeed capable of producing very accurate trajectories of $\mathcal{TBSS}$ processes, in particular in the case $\alpha \in (-\frac{1}{2}, 0)$, even when $\kappa = 1$.

4 Proofs

Throughout the proofs below, we rely on two useful inequalities. The first one is the Potter bound for slow variation at 0, which follows immediately from the corresponding result for slow variation at ∞ [15, Theorem 1.5.6]. Namely, if $L : (0, 1] \to (0, \infty)$

is slowly varying at 0 and bounded away from 0 and ∞ on every interval $(u, 1]$, $u \in (0, 1)$, then for any $\delta > 0$, there exists a constant $C_\delta > 0$ such that

$$\frac{L(x)}{L(y)} \leq C_\delta \max\left\{\left(\frac{x}{y}\right)^{\delta}, \left(\frac{x}{y}\right)^{-\delta}\right\}, \quad x, y \in (0, 1]. \tag{4.1}$$

The second one is the elementary inequality

$$|x^\alpha - y^\alpha| \leq |\alpha|(\min\{x, y\})^{\alpha-1}|x - y|, \quad x, y \in (0, \infty), \quad \alpha \in (-\infty, 1), \tag{4.2}$$

which can be easily shown by using the mean value theorem. Additionally, we use the following variant of Karamata's theorem for regular variation at 0. Its proof is similar to the one of the usual Karamata theorem for regular variation at ∞ [15, Proposition 1.5.10].

Lemma 4.1 *If $\alpha \in (-1, \infty)$ and $L : (0, 1] \to [0, \infty)$ is slowly varying at 0, then*

$$\int_0^y x^\alpha L(x)\mathrm{d}x \sim \frac{1}{\alpha + 1} y^{\alpha+1} L(y), \quad y \to 0.$$

4.1 Proof of Proposition 2.2

(i) By the covariance-stationarity of the volatility process σ, we may express the variogram $V(h)$ for any $h \geq 0$ as

$$\begin{aligned} V(h) &= \mathbb{E}[|X_h - X_0|^2] \\ &= \int_{-\infty}^{h} \big(g(h-u) - g(-u)\mathbf{1}_{(-\infty,0)}(u)\big)^2 \mathbb{E}[\sigma_u^2]\mathrm{d}u \\ &= \mathbb{E}[\sigma_0^2]\bigg(\int_0^h g(x)^2\mathrm{d}x + \int_0^\infty \big(g(x+h) - g(x)\big)^2\mathrm{d}x\bigg). \end{aligned} \tag{4.3}$$

Invoking (A1) and Lemma 4.1, we find that

$$\int_0^h g(x)^2\mathrm{d}x \sim \frac{1}{2\alpha + 1} h^{2\alpha+1} L_g(h)^2, \quad h \to 0. \tag{4.4}$$

We may clearly assume that $h < 1$, which allows us to work with the decomposition

$$\int_0^\infty \big(g(x+h) - g(x)\big)^2\mathrm{d}x = A_h + A_h',$$

where

$$A_h := \int_0^{1-h} \big(g(x+h) - g(x)\big)^2\mathrm{d}x, \quad A_h' := \int_{1-h}^\infty \big(g(x+h) - g(x)\big)^2\mathrm{d}x.$$

According to (A2), there exists $M > 1$ such that $x \mapsto |g'(x)|$ is nonincreasing on $[M, \infty)$. Thus, using the mean value theorem, we deduce that

$$|g(x+h) - g(x)| = |g'(\xi)|h \leq \begin{cases} \sup_{y \in (1-h, M]} |g'(y)|h, & x \in (1-h, M), \\ |g'(x)|h, & x \in [M, \infty), \end{cases}$$

where $\xi = \xi(x, h) \in [x, x+h]$. It follows then that

$$\limsup_{h \to 0} \frac{A_h'}{h^2} \leq (M-1) \sup_{y \in [1,M]} g'(y)^2 + \int_1^\infty g'(x)^2 \mathrm{d}x < \infty,$$

which in turn implies that

$$A_h' = \mathcal{O}(h^2), \quad h \to 0. \tag{4.5}$$

Making the substitution $y = \frac{x}{h}$, we next obtain

$$\begin{aligned} A_h = \int_0^{1-h} \big(g(x+h) - g(x)\big)^2 \mathrm{d}x &= h \int_0^{1/h-1} \Big(g\big(h(y+1)\big) - g(hy)\Big)^2 \mathrm{d}y \\ &= h^{2\alpha+1} L_g(h)^2 \int_0^\infty G_h(y) \mathrm{d}y, \end{aligned}$$

where

$$G_h(y) := \left((y+1)^\alpha \frac{L_g(h(y+1))}{L_g(h)} - y^\alpha \frac{L_g(hy)}{L_g(h)}\right)^2 \mathbf{1}_{(0,1/h-1)}(y), \quad y \in (0, \infty).$$

By the definition of slow variation at 0,

$$\lim_{h \to 0} G_h(y) = \big((y+1)^\alpha - y^\alpha\big)^2, \quad y \in (0, \infty).$$

We show below that the functions G_h, $h \in (0,1)$, have an integrable majorant. Thus, by the dominated convergence theorem,

$$A_h \sim h^{2\alpha+1} L_g(h)^2 \int_0^\infty \big((y+1)^\alpha - y^\alpha\big)^2 \mathrm{d}y, \quad h \to 0. \tag{4.6}$$

Since $\alpha < \frac{1}{2}$, we have $\lim_{h \to 0} \frac{A_h'}{h^{2\alpha+1} L_g(h)^2} = 0$ by (2.2) and (4.5), so we get from (4.4) and (4.6) that

$$\begin{aligned} &\int_0^h g(x)^2 \mathrm{d}x + \int_0^\infty \big(g(x+h) - g(x)\big)^2 \mathrm{d}x \\ &\quad \sim \left(\frac{1}{2\alpha+1} + \int_0^\infty \big((y+1)^\alpha - y^\alpha\big)^2 \mathrm{d}y\right) h^{2\alpha+1} L_g(h)^2, \quad h \to 0, \end{aligned}$$

which together with (4.3) implies the assertion.

It remains to justify the use of the dominated convergence theorem to deduce (4.6). For any $y \in (0, 1]$, we have by the Potter bound (4.1) and the elementary inequality $(u+v)^2 \le 2u^2 + 2v^2$ that

$$\begin{aligned} G_h(y) &\le 2(y+1)^{2\alpha}\left(\frac{L_g(h(y+1))}{L_g(h)}\right)^2 + 2y^{2\alpha}\left(\frac{L_g(hy)}{L_g(h)}\right)^2 \\ &\le 2C_{\delta_1}^2\big((y+1)^{2(\alpha+\delta_1)} + y^{2(\alpha-\delta_1)}\big), \end{aligned}$$

where we choose $\delta_1 \in (0, \alpha + \frac{1}{2})$ to ensure that $2(\alpha - \delta_1) > -1$. Consider then $y \in [1, \infty)$. By adding and subtracting the term $(y+1)^\alpha \frac{L_g(hy)}{L_g(h)}$ and using again the inequality $(u+v)^2 \le 2u^2 + 2v^2$, we get

$$\begin{aligned} G_h(y) &= \left((y+1)^\alpha \frac{L_g(h(y+1))}{L_g(h)} - (y+1)^\alpha \frac{L_g(hy)}{L_g(h)} \right. \\ &\qquad \left. + (y+1)^\alpha \frac{L_g(hy)}{L_g(h)} - y^\alpha \frac{L_g(hy)}{L_g(h)}\right)^2 \mathbf{1}_{(0,1/h-1)}(y) \\ &\le 2(y+1)^{2\alpha}\left(\frac{L_g(h(y+1)) - L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right)^2 \\ &\qquad + 2\big((y+1)^\alpha - y^\alpha\big)^2\left(\frac{L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right)^2. \end{aligned}$$

We recall that $\underline{L}_g := \inf_{x\in(0,1]} L_g(x) > 0$ by (A1), so

$$\begin{aligned} &\left|\frac{L_g(h(y+1)) - L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right| \\ &\quad \le \frac{1}{\underline{L}_g}\big|L_g\big(h(y+1)\big) - L_g(hy)\big|\mathbf{1}_{(0,1/h-1)}(y). \end{aligned}$$

Using the mean value theorem and the bound for the derivative of L_g from (A1), we observe that

$$\big|L_g\big(h(y+1)\big) - L_g(hy)\big| = |L_g'(\xi)||h(y+1) - hy| \le hC\left(1 + \frac{1}{\xi}\right) \le C\left(h + \frac{1}{y}\right),$$

where $\xi = \xi(y,h) \in [hy, h(y+1)]$. Noting that the constraint $y < \frac{1}{h} - 1$ is equivalent to $h < \frac{1}{y+1}$, we obtain further

$$\begin{aligned} \left|\frac{L_g(h(y+1)) - L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right| &\le \frac{C}{\underline{L}_g}\left(h + \frac{1}{y}\right)\mathbf{1}_{(0,1/h-1)}(y) \\ &\le \frac{C}{\underline{L}_g}\left(\frac{1}{y+1} + \frac{1}{y}\right) \le \frac{C}{\underline{L}_g}\frac{3}{y+1} \end{aligned}$$

as $y \geq 1$, which we then use to deduce that

$$2(y+1)^{2\alpha}\left(\frac{L_g(h(y+1)) - L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right)^2 \leq \frac{18C^2}{\underline{L}_g^2}(y+1)^{2(\alpha-1)}.$$

Additionally, we observe that by (4.1) and (4.2),

$$2\big((y+1)^\alpha - y^\alpha\big)^2\left(\frac{L_g(hy)}{L_g(h)}\mathbf{1}_{(0,1/h-1)}(y)\right)^2 \leq 2C_{\delta_2}^2\alpha^2 y^{2(\alpha-1+\delta_2)},$$

where we choose $\delta_2 \in (0, \frac{1}{2} - \alpha)$, ensuring that $2(\alpha - 1 + \delta_2) < -1$. We may finally define the function

$$G(y) := \begin{cases} 2C_{\delta_1}^2((y+1)^{2(\alpha+\delta_1)} + y^{2(\alpha-\delta_1)}), & y \in (0,1], \\ \frac{18C^2}{\underline{L}_g^2}(y+1)^{2(\alpha-1)} + 2C_{\delta_2}^2\alpha^2 y^{2(\alpha-1+\delta_2)}, & y \in (1,\infty), \end{cases}$$

which satisfies $0 \leq G_h(y) \leq G(y)$ for any $y \in (0,\infty)$ and $h \in (0,1)$, and is integrable on $(0,\infty)$ with the aforementioned choices of δ_1 and δ_2.

(ii) To show existence of the modification, we need a localization procedure that involves the ancillary process

$$F_t := \int_1^\infty g'(s)^2\sigma_{t-s}^2 \mathrm{d}s, \quad t \in \mathbb{R}.$$

We first check that F is locally bounded under (A1) and (A2), which is essential for localization. To this end, let $T \in (0,\infty)$ and write, for any $t \in [-T, T]$,

$$F_t = F_t^\flat + F_t^\sharp,$$

where

$$F_t^\flat := \int_1^{M+2T} g'(s)^2\sigma_{t-s}^2 \mathrm{d}s, \quad F_t^\sharp := \int_{M+2T}^\infty g'(s)^2\sigma_{t-s}^2 \mathrm{d}s,$$

and $M > 1$ is such that $x \mapsto |g'(x)|$ is nonincreasing on $[M,\infty)$, as in the proof of (i). Since g' is continuous on $(0,\infty)$ and σ is locally bounded, we have for any $t \in [-T, T]$ that

$$0 \leq F_t^\flat \leq (M+2T-1)\sup_{y\in[1,M+2T]} g'(y)^2 \sup_{u\in[-M-3T,T-1]} \sigma_u^2 < \infty.$$

Further, when $t \in [-T, T]$,

$$F_t^\sharp = \int_{-\infty}^{t-(M+2T)} g'(t-u)^2\sigma_u^2 \mathrm{d}u,$$

where $g'(t-u)^2 \leq g'(-T-u)^2$ since the arguments satisfy

$$t - u \geq -T - u \geq -T - \big(t - (M+2T)\big) \geq M.$$

Thus,

$$0 \le F_t^\sharp \le \int_{-\infty}^{-(M+T)} g'(-T-u)^2 \sigma_u^2 \mathrm{d}u \le \int_1^\infty g'(s)^2 \sigma_{-T-s}^2 \mathrm{d}s < \infty$$

for any $t \in [-T, T]$ almost surely, as we have

$$\mathbb{E}\left[\int_1^\infty g'(s)^2 \sigma_{-T-s}^2 \mathrm{d}s\right] = \int_1^\infty g'(s)^2 \mathbb{E}[\sigma_{-T-s}^2]\mathrm{d}s = \mathbb{E}[\sigma_0^2]\int_1^\infty g'(s)^2 \mathrm{d}s < \infty,$$

where we change the order of expectation and integration relying on Tonelli's theorem and where the final equality follows from the covariance-stationarity of σ. So we can conclude that F is indeed locally bounded.

Let now $m \in \mathbb{N}$ and, for localization, define the sequence of stopping times

$$\tau_{m,n} := \inf\{t \in [-m, \infty) : F_t > n \text{ or } |\sigma_t| > n\}, \quad n \in \mathbb{N},$$

that satisfies $\tau_{m,n} \uparrow \infty$ almost surely as $n \to \infty$ since both F and σ are locally bounded. (We follow the usual convention that $\inf \varnothing = \infty$.) Consider now the modified $\mathcal{BSS}$ process

$$X_t^{m,n} := \int_{-\infty}^t g(t-s)\sigma_{\min\{s,\tau_{m,n}\}} \mathrm{d}W_s, \quad t \in [-m, \infty),$$

that coincides with X on the stochastic interval $[\![-m, \tau_{m,n}]\!]$. The process $X^{m,n}$ satisfies the assumptions of [5, Lemma 1]; so for any $p > 0$, there exists a constant $\hat{C}_p > 0$ such that

$$\mathbb{E}[|X_s^{m,n} - X_t^{m,n}|^p] \le \hat{C}_p V(|s-t|)^{p/2}, \quad s, t \in [-m, \infty). \tag{4.7}$$

Using the upper bound in (2.2), we can deduce from (i) that for any $\delta > 0$, there are constants $\tilde{C}_\delta > 0$ and $\underline{h}_\delta > 0$ such that

$$V(h) \le \tilde{C}_\delta h^{2\alpha+1-\delta}, \quad h \in (0, \underline{h}_\delta). \tag{4.8}$$

Applying (4.8) to (4.7), we get for $s, t \in [-m, \infty)$, $|s-t| < \underline{h}_\delta$ that

$$\mathbb{E}[|X_s^{m,n} - X_t^{m,n}|^p] \le \hat{C}_p \tilde{C}_\delta^{p/2} |s-t|^{1+p(\alpha+\frac{1}{2}-\frac{\delta}{2}-\frac{1}{p})}.$$

We may note that $p(\alpha + \frac{1}{2} - \frac{\delta}{2} - \frac{1}{p}) > 0$ for small enough δ and large enough p, and in particular,

$$\frac{p(\alpha + \frac{1}{2} - \frac{\delta}{2} - \frac{1}{p})}{p} \uparrow \alpha + \frac{1}{2}$$

as $\delta \downarrow 0$ and $p \uparrow \infty$. Thus it follows from the Kolmogorov–Chentsov theorem [21, Theorem 3.22] that $X^{m,n}$ has a modification with locally ϕ-Hölder-continuous trajectories for any $\phi \in (0, \alpha + \frac{1}{2})$. Moreover, a modification of X on $\mathbb{R}$, having locally ϕ-Hölder-continuous trajectories for any $\phi \in (0, \alpha + \frac{1}{2})$, can then be constructed from these modifications of $X^{m,n}$, $m \in \mathbb{N}$, $n \in \mathbb{N}$, by letting first $n \to \infty$ and then $m \to \infty$. □

4.2 Proof of Theorem 2.5

As a preparation, we first establish an auxiliary result that deals with the asymptotic behavior of certain integrals of regularly varying functions.

Lemma 4.2 *Suppose that $L : (0, 1] \to [0, \infty)$ is bounded away from 0 and ∞ on any set of the form $(u, 1]$, $u \in (0, 1)$, and slowly varying at 0. Moreover, let $\alpha \in (-\frac{1}{2}, \infty)$ and $k \geq 1$. If $b \in [k-1, k] \setminus \{0\}$, then*

$$\text{(i)} \lim_{n\to\infty} \int_{k-1}^{k} \left(x^\alpha \frac{L(x/n)}{L(1/n)} - b^\alpha \frac{L(b/n)}{L(1/n)} \right)^2 \mathrm{d}x = \int_{k-1}^{k} (x^\alpha - b^\alpha)^2 \mathrm{d}x < \infty.$$

$$\text{(ii)} \lim_{n\to\infty} \int_{k-1}^{k} x^{2\alpha} \left(\frac{L(x/n)}{L(1/n)} - \frac{L(b/n)}{L(1/n)} \right)^2 \mathrm{d}x = 0.$$

Proof We only prove (i) as (ii) can be shown similarly. By the definition of slow variation at 0, the function

$$f_n(x) := \left(x^\alpha \frac{L(x/n)}{L(1/n)} - b^\alpha \frac{L(b/n)}{L(1/n)} \right)^2, \quad x \in [k-1, k] \setminus \{0\},$$

satisfies $\lim_{n\to\infty} f_n(x) = (x^\alpha - b^\alpha)^2$ for any $x \in [k-1, k] \setminus \{0\}$. In view of the dominated convergence theorem, it suffices to find an integrable majorant for the functions f_n, $n \in \mathbb{N}$. The construction of the majorant is quite similar to the one seen in the proof of Proposition 2.2, but we provide the details for the convenience of the reader.

Using the Potter bound (4.1) and the inequality $(u+v)^2 \leq 2u^2 + 2v^2$, we find that for any $x \in [k-1, k] \setminus \{0\}$,

$$\begin{aligned} 0 \leq f_n(x) &\leq 2x^{2\alpha} \left(\frac{L(x/n)}{L(1/n)} \right)^2 + 2b^{2\alpha} \left(\frac{L(b/n)}{L(1/n)} \right)^2 \\ &\leq 2C_\delta^2 \left(x^{2\alpha} \max\{x^\delta, x^{-\delta}\}^2 + b^{2\alpha} \max\{b^\delta, b^{-\delta}\}^2 \right) =: f(x), \end{aligned}$$

where we choose $\delta \in (0, \alpha + \frac{1}{2})$. When $k \geq 2$, we have $x \geq 1$ and $b \geq 1$, so

$$f(x) = 2C_\delta^2 \left(x^{2(\alpha+\delta)} + b^{2(\alpha+\delta)} \right)$$

is a bounded function of x on $[k-1, k]$. When $k = 1$, we have $x \leq 1$ and $b \leq 1$, implying that

$$f(x) = 2C_\delta^2 \left(x^{2(\alpha-\delta)} + b^{2(\alpha-\delta)} \right),$$

where $2(\alpha - \delta) > -1$ with our choice of δ, so f is an integrable function on $(0, 1]$. □

Proof of Theorem 2.5 Let $t \in \mathbb{R}$ be fixed. It will be convenient to write X_t^n as

$$X_t^n = \sum_{k=1}^{\kappa} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} (t-s)^\alpha L_g\left(\frac{k}{n}\right) \sigma_{t-\frac{k}{n}} \mathrm{d}W_s + \sum_{k=\kappa+1}^{N_n} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} g\left(\frac{b_k}{n}\right) \sigma_{t-\frac{k}{n}} \mathrm{d}W_s.$$

Moreover, we introduce an ancillary approximation of X_t, namely

$$\bar{X}_t^n = \sum_{k=1}^{N_n} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} g(t-s)\sigma_{t-\frac{k}{n}} \mathrm{d}W_s + \int_{-\infty}^{t-\frac{N_n}{n}} g(t-s)\sigma_s \mathrm{d}W_s.$$

By Minkowski's inequality, we have

$$\begin{aligned}
\mathbb{E}[|X_t^n - X_t|^2]^{\frac{1}{2}} &\geq \mathbb{E}[|X_t^n - \bar{X}_t^n|^2]^{\frac{1}{2}} - \mathbb{E}[|\bar{X}_t^n - X_t|^2]^{\frac{1}{2}}, \\
\mathbb{E}[|X_t^n - X_t|^2]^{\frac{1}{2}} &\leq \mathbb{E}[|X_t^n - \bar{X}_t^n|^2]^{\frac{1}{2}} + \mathbb{E}[|\bar{X}_t^n - X_t|^2]^{\frac{1}{2}},
\end{aligned}$$

which together, after taking squares, imply that

$$E_n\left(1 - 2\sqrt{\frac{E_n'}{E_n}} + \frac{E_n'}{E_n}\right) \leq \mathbb{E}[|X_t^n - X_t|^2] \leq E_n\left(1 + 2\sqrt{\frac{E_n'}{E_n}} + \frac{E_n'}{E_n}\right), \tag{4.9}$$

where

$$E_n := \mathbb{E}[|X_t^n - \bar{X}_t^n|^2], \quad E_n' := \mathbb{E}[|X_t - \bar{X}_t^n|^2].$$

Using the Itô isometry and recalling that σ is covariance-stationary, we obtain

$$\begin{aligned}
E_n' &= \sum_{k=1}^{N_n} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} g(t-s)^2 \mathbb{E}[(\sigma_{t-\frac{k}{n}} - \sigma_s)^2] \mathrm{d}s \\
&\leq \sup_{u \in (0, \frac{1}{n}]} \mathbb{E}[|\sigma_u - \sigma_0|^2] \int_0^\infty g(s)^2 \mathrm{d}s
\end{aligned}$$

and

$$\begin{aligned}
E_n &= \sum_{k=1}^{\kappa} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} \left((t-s)^\alpha L_g\left(\frac{k}{n}\right) - g(t-s)\right)^2 \mathbb{E}[\sigma^2_{t-\frac{k}{n}}] \mathrm{d}s \\
&\quad + \sum_{k=\kappa+1}^{n} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} \left(g\left(\frac{b_k}{n}\right) - g(t-s)\right)^2 \mathbb{E}[\sigma^2_{t-\frac{k}{n}}] \mathrm{d}s \\
&\quad + \sum_{k=n+1}^{N_n} \int_{t-\frac{k}{n}}^{t-\frac{k-1}{n}} \left(g\left(\frac{b_k}{n}\right) - g(t-s)\right)^2 \mathbb{E}[\sigma^2_{t-\frac{k}{n}}] \mathrm{d}s \\
&\quad + \int_{-\infty}^{t-\frac{N_n}{n}} g(t-s)^2 \mathbb{E}[\sigma_s^2] \mathrm{d}s \\
&= \mathbb{E}[\sigma_0^2](D_n + D_n' + D_n'' + D_n'''),
\end{aligned}$$

where

$$D_n := \sum_{k=1}^{\kappa} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(s^\alpha L_g\Big(\frac{k}{n}\Big) - g(s) \right)^2 \mathrm{d}s,$$

$$D_n' := \sum_{k=\kappa+1}^{n} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(g\Big(\frac{b_k}{n}\Big) - g(s) \right)^2 \mathrm{d}s,$$

$$D_n'' := \sum_{k=n+1}^{N_n} \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(g\Big(\frac{b_k}{n}\Big) - g(s) \right)^2 \mathrm{d}s,$$

$$D_n''' := \int_{\frac{N_n}{n}}^{\infty} g(s)^2 \mathrm{d}s.$$

(We may assume without loss of generality that $N_n > n > \kappa$, as this will be the case for large enough n.) In what follows, we study the asymptotic behavior of the terms D_n, D_n', D_n'' and D_n''' separately, showing that D_n, D_n'' and D_n''' are negligible in comparison with D_n', and that D_n' gives rise to the convergence rate given in Theorem 2.5.

Let us analyze the terms D_n''', D_n'' and D_n first. By (A3) and (A4), we have

$$D_n''' = \mathcal{O}\left(\Big(\frac{N_n}{n}\Big)^{2\beta+1} \right) = \mathcal{O}(n^{\gamma(2\beta+1)}), \quad n \to \infty. \tag{4.10}$$

Regarding the term D_n'', recall that by (A2), there is $M > 1$ such that $x \mapsto |g'(x)|$ is nonincreasing on $[M, \infty)$. So we have by the mean value theorem that

$$\left| g\Big(\frac{b_k}{n}\Big) - g(s) \right| = |g'(\xi)| \left| \frac{b_k}{n} - s \right| \le \begin{cases} \frac{1}{n} \sup_{y \in [1,M]} |g'(y)|, & \frac{k-1}{n} < M, \\ \frac{1}{n} |g'(\frac{k-1}{n})|, & \frac{k-1}{n} \ge M, \end{cases}$$

where $\xi = \xi(\frac{b_k}{n}, s) \in [\frac{k-1}{n}, \frac{k}{n}]$. Thus,

$$\limsup_{n\to\infty} n^2 D_n'' \le (M-1) \sup_{y \in [1,M]} g'(y)^2 + \int_1^\infty g'(s)^2 \mathrm{d}s < \infty,$$

which implies

$$D_n'' = \mathcal{O}(n^{-2}), \quad n \to \infty. \tag{4.11}$$

To analyze the behavior of D_n, we substitute $y = ns$ and invoke (A1), yielding

$$\begin{aligned} D_n &= \sum_{k=1}^{\kappa} \int_{k-1}^{k} \left(\Big(\frac{y}{n}\Big)^\alpha L_g\Big(\frac{k}{n}\Big) - g\Big(\frac{y}{n}\Big) \right)^2 \frac{\mathrm{d}y}{n} \\ &= n^{-(2\alpha+1)} L_g(1/n)^2 \sum_{k=1}^{\kappa} \int_{k-1}^{k} y^{2\alpha} \left(\frac{L_g(k/n)}{L_g(1/n)} - \frac{L_g(y/n)}{L_g(1/n)} \right)^2 \mathrm{d}y, \end{aligned}$$

where by Lemma 4.2(ii), we have

$$\lim_{n\to\infty}\int_{k-1}^{k} y^{2\alpha}\left(\frac{L_g(k/n)}{L_g(1/n)}-\frac{L_g(y/n)}{L_g(1/n)}\right)^2\mathrm{d}y=0$$

for any $k=1,\dots,\kappa$. Thus we find that

$$\lim_{n\to\infty}\frac{D_n}{n^{-(2\alpha+1)}L_g(1/n)^2}=0. \tag{4.12}$$

The asymptotic behavior of the term D_n' is more delicate to analyze. By (A1) and substituting $y=ns$, we can write

$$\begin{aligned}
D_n' &= \sum_{k=\kappa+1}^{n}\int_{k-1}^{k}\left(g\Big(\frac{b_k}{n}\Big)-g\Big(\frac{y}{n}\Big)\right)^2\frac{\mathrm{d}y}{n}\\
&= n^{-(2\alpha+1)}\sum_{k=\kappa+1}^{n}\int_{k-1}^{k}\left(b_k^{\alpha}L_g\Big(\frac{b_k}{n}\Big)-y^{\alpha}L_g\Big(\frac{y}{n}\Big)\right)^2\mathrm{d}y\\
&= n^{-(2\alpha+1)}L_g(1/n)^2\sum_{k=\kappa+1}^{n}A_{n,k},
\end{aligned}$$

where

$$A_{n,k}:=\int_{k-1}^{k}\left(y^{\alpha}\frac{L_g(y/n)}{L_g(1/n)}-b_k^{\alpha}\frac{L_g(b_k/n)}{L_g(1/n)}\right)^2\mathrm{d}y.$$

Let us study the asymptotic behavior of the sum $\sum_{k=\kappa+1}^{n}A_{n,k}$ as $n\to\infty$. By Lemma 4.2, we have for any $k\in\mathbb{N}$ that

$$\lim_{n\to\infty}A_{n,k}=\int_{k-1}^{k}(y^{\alpha}-b_k^{\alpha})^2\mathrm{d}y<\infty.$$

To be able to then deduce, using the dominated convergence theorem, that

$$\lim_{n\to\infty}\sum_{k=\kappa+1}^{n}A_{n,k}=\sum_{k=1}^{\infty}\int_{k-1}^{k}(y^{\alpha}-b_k^{\alpha})^2\mathrm{d}y=J(\alpha,\kappa,\mathbf{b})<\infty, \tag{4.13}$$

we seek a sequence $(A_k)_{k=\kappa+1}^{\infty}\subseteq[0,\infty)$ such that

$$0\le A_{n,k}\le A_k,\quad k=\kappa+1,\dots,n,\quad n\in\mathbb{N},$$

and such that $\sum_{k=\kappa+1}^{\infty}A_k<\infty$. Let us assume, without loss of generality, that $\kappa=0$. Clearly, we may set $A_1:=\sup_{n\in\mathbb{N}}A_{n,1}<\infty$. Consider now $k\ge 2$. The construction of A_k in this case parallels some arguments seen in the proof of Proposition 2.2, but we provide the details for the sake of clarity. By adding and subtracting $b_k^{\alpha}\frac{L_g(y/n)}{L_g(1/n)}$

and using the inequality $(u+v)^2 \le 2u^2 + 2v^2$, we get

$$\begin{aligned} A_{n,k} &= \left(y^\alpha \frac{L_g(y/n)}{L_g(1/n)} - b_k^\alpha \frac{L_g(y/n)}{L_g(1/n)} + b_k^\alpha \frac{L_g(y/n)}{L_g(1/n)} - b_k^\alpha \frac{L_g(b_k/n)}{L_g(1/n)} \right)^2 \\ &\le 2 \int_{k-1}^{k} (y^\alpha - b_k^\alpha)^2 \left(\frac{L_g(y/n)}{L_g(1/n)} \right)^2 \mathrm{d}y + 2 b_k^{2\alpha} \int_{k-1}^{k} \left(\frac{L_g(y/n) - L_g(b_k/n)}{L_g(1/n)} \right)^2 \mathrm{d}y \\ &=: I_{n,k} + I'_{n,k}. \end{aligned} \tag{4.14}$$

Recall that $\underline{L}_g := \inf_{x \in (0,1]} L_g(x) > 0$; so by the estimates

$$b_k^{2\alpha} \le \max\{k^{2\alpha}, (k-1)^{2\alpha}\} \le 2(k-1)^{2\alpha},$$

valid when $\alpha < \frac{1}{2}$, we obtain

$$I'_{n,k} \le \frac{4}{\underline{L}_g^2} (k-1)^{2\alpha} \int_{k-1}^{k} \big(L_g(y/n) - L_g(b_k/n) \big)^2 \mathrm{d}y.$$

Note that thanks to (A1) and the mean value theorem,

$$|L_g(y/n) - L_g(b_k/n)| = |L'_g(\xi)| \left| \frac{y}{n} - \frac{b_k}{n} \right| \le \frac{C(1+\xi^{-1})}{n} \le \frac{C}{n} + \frac{C}{k-1} \le \frac{2C}{k-1},$$

where $\xi = \xi(y/n, b_k/n) \in [\frac{k-1}{n}, \frac{k}{n}]$ and where the final inequality follows since $k - 1 < n$. Thus,

$$I'_{n,k} \le \frac{16C^2}{\underline{L}_g^2} (k-1)^{2(\alpha-1)}. \tag{4.15}$$

Moreover, the Potter bound (4.1) and inequality (4.2) imply

$$I_{n,k} \le 2\alpha^2 C_\delta^2 \int_{k-1}^{k} (\min\{y, b_k\})^{2(\alpha-1)} y^{2\delta} \mathrm{d}y \le 2^{1+2\delta} \alpha^2 C_\delta^2 (k-1)^{2(\alpha-1+\delta)}, \tag{4.16}$$

where we choose $\delta \in (0, \frac{1}{2} - \alpha)$. Applying the bounds (4.15) and (4.16) to (4.14) shows that

$$A_{n,k} \le 2^{1+2\delta} \alpha^2 C_\delta^2 (k-1)^{2(\alpha-1+\delta)} + \frac{16C^2}{\underline{L}_g^2} (k-1)^{2(\alpha-1)} =: A_k,$$

where $2(\alpha-1) < -1$ and $2(\alpha - 1 + \delta) < -1$ with our choice of δ, so that we get $\sum_{k=1}^{\infty} A_k < \infty$. Thus we have shown (4.13), which in turn implies that

$$D'_n \sim J\big(\alpha, \kappa, (b_k)_{k=\kappa+1}^{\infty}\big) n^{-(2\alpha+1)} L_g(1/n)^2, \quad n \to \infty. \tag{4.17}$$

We now use the obtained asymptotic relations, (4.10)–(4.12) and (4.17), to complete the proof. To this end, it is convenient to introduce the relation $x_n \gg y_n$ for any sequences $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ of positive real numbers satisfying $\lim_{n\to\infty} \frac{x_n}{y_n} = \infty$.

By (4.12), we have $D_n' \gg D_n$. Since $2\alpha + 1 < 2$, we find that also $D_n' \gg D_n''$, in view of (4.11). The assumption $\gamma > -\frac{2\alpha+1}{2\beta+1}$ is equivalent to $-(2\alpha+1) > \gamma(2\beta+1)$; so by the estimate (2.2) for slowly varying functions, we have $D_n' \gg D_n'''$. It then follows that $E_n \sim \mathbb{E}[\sigma_0^2] D_n'$ as $n \to \infty$. Further, the condition (2.10) implies that $E_n \gg E_n'$. In view of (4.9), we finally find that $\mathbb{E}[|X_t^n - X_t|^2] \sim \mathbb{E}[\sigma_0^2] D_n'$ as $n \to \infty$, which completes the proof. □

4.3 Proof of equation (3.5)

In order to prove (3.5), we rely on the following integral identity for the Gauss hypergeometric function ${}_2F_1$.

Lemma 4.3 *For all $\alpha \in (-1, \infty)$ and $0 \le a < b$,*

$$\int_0^a (a-x)^\alpha (b-x)^\alpha \mathrm{d}x = \frac{a^{\alpha+1} b^\alpha}{\alpha+1} {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{a}{b}\right).$$

Proof In the case $a = 0$, the asserted identity becomes trivial, so we may assume that $a > 0$. Substituting $y = \frac{x}{a}$, we get

$$\int_0^a (a-x)^\alpha (b-x)^\alpha \mathrm{d}x = a^{\alpha+1} b^\alpha \int_0^1 (1-y)^\alpha \left(1 - \frac{a}{b} y\right)^\alpha \mathrm{d}y.$$

Using Euler's formula [17, Sect. 2.1.3, Eq. (10)], we find that

$$\int_0^1 (1-y)^\alpha \left(1 - \frac{a}{b} y\right)^\alpha \mathrm{d}y = \frac{\Gamma(1)\Gamma(\alpha+1)}{\Gamma(\alpha+2)} {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{a}{b}\right),$$

where Γ denotes the gamma function; note that this is valid since $\alpha + 2 > 1 > 0$ and $0 < \frac{a}{b} < 1$ under our assumptions. By the connection between the gamma function Γ and the beta function B (see e.g. [17, Sect. 1.5, Eq. (5)]), we obtain further

$$\frac{\Gamma(1)\Gamma(\alpha+1)}{\Gamma(\alpha+2)} = \mathrm{B}(1, \alpha+1) = \int_0^1 (1-x)^\alpha \mathrm{d}x = \frac{1}{\alpha+1},$$

concluding the proof. □

Proof of equation (3.5) Let $\alpha \in (-\frac{1}{2}, \frac{1}{2}) \setminus \{0\}$ and let $j, k = 2, \dots, \kappa + 1$ be such that $j < k$. By the Itô isometry, we have

$$\begin{aligned}\Sigma_{j,k} = \mathbb{E}[W_{0,j-1}^n W_{0,k-1}^n] &= \int_0^{\frac{1}{n}} \left(\frac{j-1}{n} - s\right)^\alpha \left(\frac{k-1}{n} - s\right)^\alpha \mathrm{d}s \\ &= \frac{1}{n^{2\alpha+1}} \int_0^1 (j-1-x)^\alpha (k-1-x)^\alpha \mathrm{d}x,\end{aligned}$$

where we have substituted $x = ns$. The second integral above can now be expressed as

$$\begin{aligned}
&\int_0^1 (j-1-x)^\alpha (k-1-x)^\alpha \mathrm{d}x \\
&\quad = \int_0^{j-1} (j-1-x)^\alpha (k-1-x)^\alpha \mathrm{d}x - \int_0^{j-2} (j-2-x)^\alpha (k-2-x)^\alpha \mathrm{d}x \\
&\quad = \frac{1}{\alpha+1}\Bigg((j-1)^{\alpha+1}(k-1)^\alpha {}_2F_1\Big(-\alpha, 1, \alpha+2, \frac{j-1}{k-1}\Big) \\
&\qquad\qquad - (j-2)^{\alpha+1}(k-2)^\alpha {}_2F_1\Big(-\alpha, 1, \alpha+2, \frac{j-2}{k-2}\Big)\Bigg),
\end{aligned}$$

where the second equality follows from Lemma 4.3, which is applicable to both integrals on the second line as $0 < j-1 < k-1$ and $0 \le j-2 < k-2$. □

Acknowledgements We should like to thank Heidar Eyjolfsson and Emil Hedevang for useful discussions regarding the simulation of $\mathcal{BSS}$ processes and Ulises Márquez for assistance with symbolic computation. Our research has been supported by CREATES (DNRF78), funded by the Danish National Research Foundation, by Aarhus University Research Foundation (project "Stochastic and Econometric Analysis of Commodity Markets") and by the Academy of Finland (project 258042).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Alòs, E., León, J.A., Vives, J.: On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. Finance Stoch. **11**, 571–589 (2007)
2. Asmussen, S., Glynn, P.W.: Stochastic Simulation: Algorithms and Analysis. Springer, New York (2007)
3. Barndorff-Nielsen, O.E.: Notes on the gamma kernel (2012). Thiele Centre Research Report, No. 03, May 2012. Available online at http://data.math.au.dk/publications/thiele/2012/math-thiele-2012-03.pdf
4. Barndorff-Nielsen, O.E., Benth, F.E., Veraart, A.E.D.: Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. Bernoulli **19**, 803–845 (2013)
5. Barndorff-Nielsen, O.E., Corcuera, J.M., Podolskij, M.: Multipower variation for Brownian semistationary processes. Bernoulli **17**, 1159–1194 (2011)
6. Barndorff-Nielsen, O.E., Corcuera, J.M., Podolskij, M.: Limit theorems for functionals of higher order differences of Brownian semistationary processes. In: Shiryaev, A.N., et al. (eds.) Prokhorov and Contemporary Probability Theory, pp. 69–96. Springer, Berlin (2013)
7. Barndorff-Nielsen, O.E., Pakkanen, M.S., Schmiegel, J.: Assessing relative volatility/intermittency/energy dissipation. Electron. J. Stat. **8**, 1996–2021 (2014)
8. Barndorff-Nielsen, O.E., Schmiegel, J.: Ambit processes: with applications to turbulence and tumour growth. In: Benth, F.E., et al. (eds.) Stochastic Analysis and Applications, Abel Symp., vol. 2, pp. 93–124. Springer, Berlin (2007)
9. Barndorff-Nielsen, O.E., Schmiegel, J.: Brownian semistationary processes and volatility/intermittency. In: Albrecher, H., et al. (eds.) Advanced Financial Modelling. Radon Ser. Comput. Appl. Math., vol. 8, pp. 1–25. Walter de Gruyter, Berlin (2009)

10. Bayer, C., Friz, P.K., Gatheral, J.: Pricing under rough volatility. Quant. Finance **16**, 887–904 (2016)
11. Bennedsen, M.: A rough multi-factor model of electricity spot prices. Energy Econ. **63**, 301–313 (2017)
12. Bennedsen, M., Lunde, A., Pakkanen, M.S.: Discretization of Lévy semistationary processes with application to estimation (2014). Working paper. Available online at http://arxiv.org/abs/1407.2754
13. Benth, F.E., Eyjolfsson, H., Veraart, A.E.D.: Approximating Lévy semistationary processes via Fourier methods in the context of power markets. SIAM J. Financ. Math. **5**, 71–98 (2014)
14. Beskos, A., Roberts, G.O.: Exact simulation of diffusions. Ann. Appl. Probab. **15**, 2422–2444 (2005)
15. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Cambridge University Press, Cambridge (1989)
16. Corcuera, J.M., Hedevang, E., Pakkanen, M.S., Podolskij, M.: Asymptotic theory for Brownian semistationary processes with application to turbulence. Stoch. Process. Appl. **123**, 2552–2574 (2013)
17. Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F.G.: Higher Transcendental Functions, vol. I. McGraw-Hill, New York (1953)
18. Fukasawa, M.: Short-time at-the-money skew and rough fractional volatility. Quant. Finance **17**, 189–198 (2017)
19. Gatheral, J., Jaisson, T., Rosenbaum, M.: Volatility is rough (2014). Working paper. Available online at http://arxiv.org/abs/1410.3394
20. Gneiting, T., Schlather, M.: Stochastic models that separate fractal dimension and the Hurst effect. SIAM Rev. **46**, 269–282 (2004)
21. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002)
22. Mallat, S.: A Wavelet Tour of Signal Processing, 3rd edn. Elsevier, Amsterdam (2009)
23. Mishura, Y.S.: Stochastic Calculus for Fractional Brownian Motion and Related Processes. Springer, Berlin (2008)
24. Pedersen, J., Sauri, O.: On Lévy semistationary processes with a gamma kernel. In: Mena, R.H., et al. (eds.) XI Symposium on Probability and Stochastic Processes, pp. 217–239. Birkhäuser, Basel (2015)