Quantitative Finance

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Refinement by reducing and reusing random numbers of the Hybrid scheme for Brownian semistationary processes

Masaaki Fukasawa & Asuto Hirano

To cite this article: Masaaki Fukasawa & Asuto Hirano (2021): Refinement by reducing and reusing random numbers of the Hybrid scheme for Brownian semistationary processes, Quantitative Finance, DOI: 10.1080/14697688.2020.1866209

To link to this article: https://doi.org/10.1080/14697688.2020.1866209

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 04 Feb 2021.

Submit your article to this journal

Article views: 145

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=rquf20

Quantitative Finance, 2021
https://doi.org/10.1080/14697688.2020.1866209

Refinement by reducing and reusing random numbers of the Hybrid scheme for Brownian semistationary processes

MASAAKI FUKASAWA* and ASUTO HIRANO

Graduate School of Engineering Science, Osaka University, Osaka, Japan

(*Received 1 April 2020; accepted 11 December 2020; published online 4 February 2021*)

We revisit the Hybrid scheme proposed by Bennedsen et al. (2017) for numerical simulations of Brownian semistationary processes, and propose a Refinement by Reducing and Reusing random numbers of the Hybrid scheme (3R Hybrid scheme). The key idea is to reuse random variables through orthogonal projections. An application to the analysis of the rough Bergomi model is also given.

Keywords: Rough volatility; Hybrid scheme; Brownian semistationary process; Monte Carlo methods

1. Introduction

The *Brownian semistationary (BSS)* process introduced by Barndorff-Nielsen and Schmiegel (2009) forms a flexible class of stochastic processes that can reproduce some common features of empirical time series in various fields, including stochastic volatility, roughness, stationarity, and strong dependence. A BSS process is defined as

$$X_t := \int_{-\infty}^{t} g(t-s)\sigma_s \, \mathrm{d}W_s, \quad t \in \mathbb{R}, \tag{1}$$

where σ is a stochastic process adapted to a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}}$ and W is an $\{\mathcal{F}_t\}$ Brownian motion, that is, W is a two-sided Brownian motion with $W_t - W_s$ being $\mathcal{F}_t$ measurable for all (s,t) with $s \le t$. The deterministic function g is called a kernel function. A fractional Ornstein-Uhlenbeck (fOU) process is an example of BSS processes (see Appendix 1). A BSS process X is not a semimartingale when the kernel function g has a power-law behavior near zero; more specifically, when

$$g(x) \sim x^{\alpha}, \quad \text{as } x \to 0$$

for some $\alpha \in (-\frac{1}{2}, \frac{1}{2})\backslash\{0\}$. The trajectories of X then behave locally like those of a *fractional Brownian motion* with the Hurst index $H = \alpha + \frac{1}{2}$.

If σ is a constant as in the case of the fOU process, then X is stationary Gaussian and a sample path of X can be simulated by exact schemes based on the Cholesky decomposition or circulant embedding methods. However, it is 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. Even for the cases where an exact method is available, its computational cost is often too expensive for practical simulations. Therefore in the general case, we must resort to an approximation method. A Fourier-based method is proposed by Benth *et al.* (2014). A state-of-art method which is effective for singular kernel functions is the *Hybrid scheme* proposed by Bennedsen *et al.* (2007). The idea of this scheme is to approximate the kernel function by a power function near its singularity point and by a step function elsewhere. This method is more effective than an Itô-Riemann type approximation of the stochastic integral which approximates the kernel function by a step function everywhere. A drawback of this scheme is that the approximation accuracy becomes worse as α approaches $-\frac{1}{2}$; see Figure 1 for the mean squared error of the Hybrid scheme.

Recently BSS processes with $\alpha \approx -\frac{1}{2}$ has been applied to financial modeling. Gatheral *et al.* (2018) argued that the log-volatility of an asset price process is well modelled by a fractional Brownian motion with the Hurst index $H = \alpha + \frac{1}{2} \approx 0.1$, at any reasonable time scale. A fractional Brownian motion with $H < 1/2$ has a rougher path than a standard Brownian motion does; therefore such a stochastic

*Corresponding author. Email: fukasawa@sigmath.es.osaka-u.ac.jp, fukasawa@math.sci.osaka-u.ac.jp

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

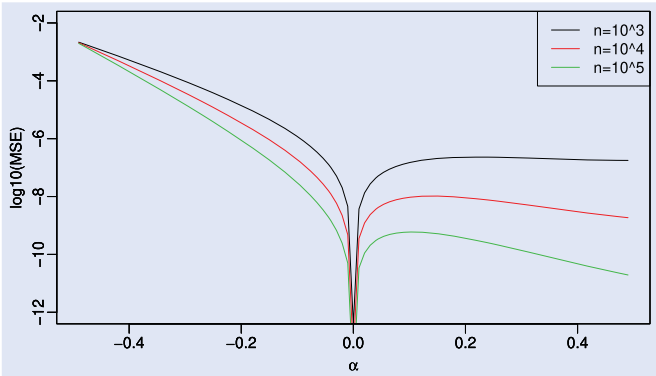

Figure 1. Log mean squared error (MSE) of the Hybrid scheme. g is a fOU kernel (17), $\sigma \equiv 1$, n is the number of discretization steps, and $\kappa + 1 = 3$ is the dimension of i.i.d. Gaussian vectors to generate.

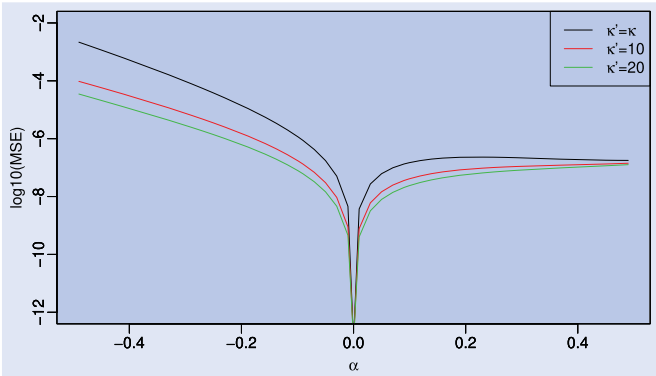

Figure 2. Log MSE of the 3R Hybrid scheme for the same g and $\kappa = 2$ as in Figure 1. $n = 10^3$. $\kappa' - \kappa$ is the number of reuse.

volatility model is now called a '*rough volatility model*'. Fukasawa *et al.* (2019) developed a consistent estimator of H under a fractional volatility model and applied it to stock indices to estimate $H \approx 0.05$. Alòs *et al.* (2007), Fukasawa (2011, 2017, 2021), Bayer *et al.* (2016), Forde and Zhang (2017), Garnier and Solna (2017), El Euch *et al.* (2019) showed that such a model with very small H reproduces an empirically observed power law of the option implied volatility surface. To be consistent further with the VIX option surface, Horvath *et al.* (2020) introduced modulated Volterra stochastic volatility processes, of which the volatility process is a BSS process with σ being independent of W. Since an analytic formula for the option prices under such a model is not available, the option pricing and calibration require intensive Monte Carlo simulations, for which the Hybrid scheme mentioned above, with some variance reduction techniques (McCrickerd and Pakkanen 2018), has been the most popular. See e.g. Jacquier *et al.* (2018).

In this paper, after revisiting the Hybrid scheme, we propose a Refinement by Reducing and Reusing random numbers of the Hybrid scheme (3R Hybrid scheme hereafter). The key idea is to reuse random variables through orthogonal projections. The new simulation method improves the Hybrid scheme, especially when $\alpha \approx -\frac{1}{2}$, or equivalently, $H \approx 0$; the refinement of the approximation accuracy is illustrated in Figure 2. All of the schemes corresponding to the three curves in Figure 2 have almost the same computational cost as that corresponding to the black curve in Figure 1 (the Hybrid scheme with $\kappa = 2$ and $n = 10^3$). The parameter κ' controls the number of reuse of random numbers. Although the rate of convergence with the number of discretization steps is the same as for the Hybrid scheme, that is, $O(n^{-(2\alpha+1)})$, the coefficient of the leading term of MSE is much reduced by choosing κ' enough large.

In Section 2, we describe the framework and assumptions. In Section 3, we review the Hybrid scheme. In Section 4, we propose the 3R Hybrid scheme. In Section 5, we apply the proposed scheme to the numerical analysis of the rough Bergomi model (Bayer *et al.* 2016).

2. Framework

Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\in\mathbb{R}})$ be a filtered probability space satisfying the usual conditions and $W = (W_t)_{t\in\mathbb{R}}$ be a two-sided Brownian motion adapted to $\{\mathcal{F}_t\}$. This means that W is a centered Gaussian process with independent increments and $W_t - W_s \in \mathcal{F}_t$ for all $s \le t$.

2.1. Definitions

Definition 2.1 (Brownian semistationary process) *A stochastic process $X = (X_t)_{t\in\mathbb{R}}$ which admits the representation*

$$X_t = \int_{-\infty}^{t} g(t-s)\sigma_s \, dW_s, \quad t \in \mathbb{R}, \tag{2}$$

is called a Brownian semistationary process (BSS process), where $\sigma = (\sigma_t)_{t\in\mathbb{R}}$ is an $\{\mathcal{F}_t\}$−adapted process with

$$\sup_{t\in\mathbb{R}} \mathbb{E}\left[\sigma_t^2\right] < \infty,$$

and $g : (0,\infty) \to [0,\infty)$ is a Borel-measurable function with

$$\int_0^\infty g(x)^2 \, dx < \infty.$$

Remark 2.2 If σ is covariance-stationary, namely,

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

then X is also covariance-stationary, namely,

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

Definition 2.3 (slowly varying function) *A measurable function $L : (0,1] \to [0,\infty)$ is said to be slowly varying at 0 if for any $t \in \mathbb{R}$,*

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

Definition 2.4 (regular varying function) *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 2.5 (Potter's theorem) Let $L : (0,1] \to (0,\infty)$ be slowly varying at 0 and bounded away from 0 and ∞ on every interval $(u,1]$, $u \in (0,1)$ (i.e. there exist $0 < m_1, m_2 < \infty$ such that $m_1 < L(x) < m_2$ for all $x \in (u,1]$). Then for any $\delta > 0$, there exists a constant $C_\delta > 0$ such that

$$\frac{L(x)}{L(y)} \le C_\delta \max\left\{\left(\frac{x}{y}\right)^{\delta}, \left(\frac{x}{y}\right)^{-\delta}\right\}, \qquad x, y \in (0,1]. \tag{3}$$

In particular, we have

$$\underline{C_\delta} x^{\delta} \le L(x) \le \overline{C_\delta} x^{-\delta}, \quad x \in (0,1]. \tag{4}$$

See (Bingham *et al.* 1989, Theorem 1.5.6) for (3). To get (4), let $x = 1$ and $y = 1$ in (3).

2.2. Assumptions

For their Hybrid scheme for BSS processes, Bennedsen *et al.* (2007) introduced the following conditions (A1) ∼ (A3).

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

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

where $L_g : (0,1] \to [0,\infty)$ is continuously differentiable, slowly varying at 0 and bounded away from 0 (i.e. $^{\exists}m > 0$, $L_g(x) > m$ for all $x \in (0,1]$). Moreover, there exists a constant $C > 0$ such that derivative L_g' of L_g satisfies

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

(A2) The function g is continuously differentiable on $(0,\infty)$, with derivative g' that is ultimately monotonic (i.e.$^{\exists}R \in (0,\infty), R < s < t \Rightarrow g'(s) \lesseqgtr g'(t)$) and also satisfies $\int_1^{\infty} g'(x)^2 \,\mathrm{d}x < \infty$.

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

$$g(x) = \mathcal{O}\left(x^{\beta}\right), \quad x \to \infty, \quad \text{i.e. } \limsup_{x\to\infty}\left|\frac{g(x)}{x^{\beta}}\right| < \infty.$$

REMARK 2.6 These assumptions ensure that g is square-integrable;

$$\int_0^{\infty} g(x)^2 \,\mathrm{d}x < \infty.$$

In fact, in view of (A3), there exists $M \in \mathbb{R}_+$ satisfying

$$\int_1^{\infty} g(x)^2 \,\mathrm{d}x < M^2 \int_1^{\infty} x^{2\beta} \,\mathrm{d}x < \infty.$$

Further, by (A1) and (4),

$$\int_0^1 g(x)^2 \,\mathrm{d}x = \int_0^1 x^{2\alpha} L_g(x)^2 \,\mathrm{d}x, \quad \alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right)$$
$$\le \overline{C_\delta}^2 \int_0^1 x^{2\alpha - 2\delta} \,\mathrm{d}x.$$

We can take $\delta > 0$ with $2\alpha - 2\delta > -1$.

In this paper, we propose our 3R Hybrid scheme under the same conditions.

3. Hybrid scheme

Here we recall the Hybrid scheme proposed by Bennedsen *et al.* (2007).

3.1. The Hybrid scheme for BSS processes

Let X be a BSS process. Then, by definition, we have

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

If k is small, then due to (A1), we approximate

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

If k is large, or at least $k \ge 2$, then choosing $c_k \in [k-1, k]$ provides an approximation

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

by the continuity of g from (A2). Then we get an approximation

$$X_t \approx \sum_{k=1}^{\infty} \sigma_{t-k/n} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} g(t-s) \,\mathrm{d}W_s \tag{5}$$
$$\approx \sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right) \sigma_{t-k/n} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} (t-s)^{\alpha} \,\mathrm{d}W_s$$
$$+ \sum_{k=\kappa+1}^{\infty} g\left(\frac{c_k}{n}\right) \sigma_{t-k/n} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} \mathrm{d}W_s$$
$$\text{where } \kappa = 0, 1, 2, \ldots \tag{6}$$

When $\kappa = 0$, we require that $c_1 \in (0,1]$. To make numerical implementation feasible, we truncate the second sum on the right-hand side of (5) so that both sums have $N_n \ge \kappa + 1$ terms in total. Thus, we arrive at a discretization scheme for X_t;

$$X_t^n = \dot{X}_t^n + \ddot{X}_t^n$$

where

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

$$\ddot{X}_t^n := \sum_{k=\kappa+1}^{N_n} g\left(\frac{c_k}{n}\right) \sigma_{t-k/n} \left(W_{t-\frac{k}{n}+\frac{1}{n}} - W_{t-\frac{k}{n}}\right).$$

This is called the *Hybrid scheme*. A sequence of real numbers $\mathbf{c} := (c_k)_{k=\kappa+1}^{N_n}$ must satisfy $c_k \in [k-1, k]\backslash\{0\}$ for each $k \ge \kappa + 1$, but otherwise can be chosen freely.

For N_n, we need a condition:
(A4) For some $\gamma > 0$,

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

3.2. Asymptotic behavior of MSE

THEOREM 3.1 (Theorem 2.5 of Bennedsen *et al.* (2007))
Suppose that (A1) ∼ (A4) *hold with*

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

and σ *is covariance-stationary with*

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

for some $\delta > 0$. *Then for all* $t \in \mathbb{R}$,

$$\mathbb{E}\left[|X_t - X_t^n|^2\right] \sim \boldsymbol{J}(\alpha,\kappa,\mathbf{c})\mathbb{E}[\sigma_0^2]n^{-(2\alpha+1)}L_g\left(\frac{1}{n}\right)^2, \quad n \to \infty \tag{7}$$

where

$$\boldsymbol{J}(\alpha,\kappa,\mathbf{c}) := \sum_{k=\kappa+1}^{\infty}\int_{k-1}^{k}(y^\alpha - c_k^\alpha)^2\,dy < \infty.$$

PROPOSITION 3.2 (Proposition 2.8 of Bennedsen *et al.* (2007))

The coefficient of MSE $\boldsymbol{J}(\alpha,\kappa,\mathbf{c})$ *in Theorem* 3.1 *is minimized by the sequence* $\mathbf{c}$ *given by*

$$\mathbf{c} = \mathbf{c}^* = \{c_{\kappa+1}^*, c_{\kappa+2}^*, \ldots, c_{N_n}^*\},$$

where

$$c_k^* = \left(\frac{k^{\alpha+1} - (k-1)^{\alpha+1}}{\alpha+1}\right)^{\frac{1}{\alpha}}, \quad \kappa+1 \le k \le N_n.$$

We set $\mathbf{c} = \mathbf{c}^*$ in the sequel.

3.3. Implementation

Simulating the BSS process X_t 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

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

Provided that we can simulate the Gaussian random variables

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

and a discretized process $\sigma_{i/n}$, $i = -N_n, -N_n+1, \ldots$, we can compute

$$X_{\frac{i}{n}}^n = \underbrace{\sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n}W_{i-k,k}^n}_{=\dot{X}_{\frac{i}{n}}^n} + \underbrace{\sum_{k=\kappa+1}^{N_n} g\left(\frac{c_k}{n}\right)\sigma_{(i-k)/n}W_{i-k}^n}_{=\ddot{X}_{\frac{i}{n}}^n}, \quad i = 0, 1, \ldots, \lfloor nT\rfloor. \tag{8}$$

Note that the FFT can be effectively used in the above convolution. The $\kappa+1$ dimensional random vectors

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

are independent and follow the multivariate Gaussian distribution with mean $\mathbf{0} = (0, \ldots, 0)$ and variance-covariance matrix Σ given by

$$\Sigma_{1,1} = \frac{1}{n}, \tag{9}$$

$$\Sigma_{1,j} = \Sigma_{j,1} = \frac{(j-1)^{\alpha+1} - (j-2)^{\alpha+1}}{(\alpha+1)n^{\alpha+1}}, \tag{10}$$

$$\Sigma_{j,j} = \frac{(j-1)^{2\alpha+1} - (j-2)^{2\alpha+1}}{(2\alpha+1)n^{2\alpha+1}}, \tag{11}$$

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

$$\Sigma_{j,k} = \frac{1}{(\alpha+1)n^{2\alpha+1}} \times \left((j-1)^{\alpha+1}(k-1)^\alpha {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{j-1}{k-1}\right) - (j-2)^{\alpha+1}(k-2)^\alpha {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{j-2}{k-2}\right)\right), \tag{12}$$

for $j, k = 2, \ldots, \kappa+1$ with $j < k$. When $j > k$, set $\Sigma_{j,k} = \Sigma_{k,j}$. Here, ${}_2F_1$ is *Gauss hyper geometric function*, that is

$${}_2F_1(a,b,c,z) := \sum_{n=0}^{\infty}\frac{(a)_n(b)_n z^n}{(c)_n n!},$$

where

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1)\cdots(a+n-1).$$

3.4. Discussion

From Theorem 3.1, we know that the mean squared error of the Hybrid scheme converges to 0 with rate $n^{-(2\alpha+1)}$. This means when α is close to $-\frac{1}{2}$, increasing n does not help

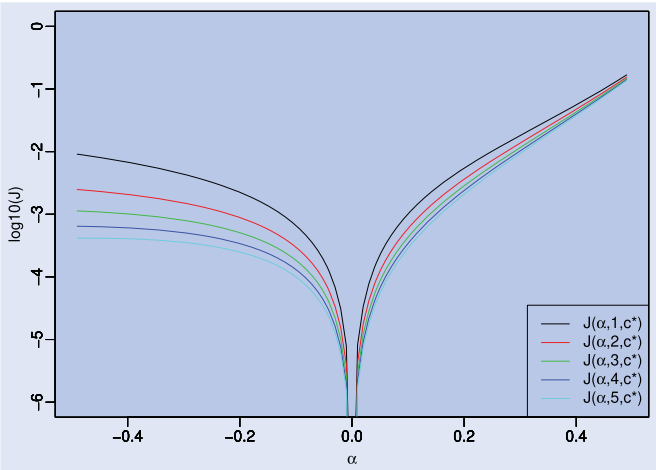

Figure 3. MSE coefficient $\mathbf{J}(\alpha, \kappa, \mathbf{c})$.

much to decrease the error. Nevertheless, the number n has to be large for simulations of rough volatility models (see Section 5). When n is large and $\alpha \approx -\frac{1}{2}$, Theorem 3.1 implies also that the error magnitude is determined by $\mathbf{J}(\alpha, \kappa, \mathbf{c})$. See Figure 3 for the graph of $(\alpha, \log \mathbf{J}(\alpha, \kappa, \mathbf{c}))$ for $\kappa = 1, 2, 3, 4, 5$. From this figure we can observe that increasing κ makes a significant reduction of the mean squared error when $\alpha \approx -\frac{1}{2}$. On the other hand, as seen in the previous subsection, an implementation of the Hybrid scheme requires $N_n + [nT] - 1$ i.i.d. generations of $\kappa + 1$ dimensional Gaussian random variables. This fact hampers the use of large κ in practice. This is because of not only increased computational time but also numerical difficulty to generate the random variable. In fact, as seen in Table 1, the determinant of the covariance matrix Σ is quite small for $\kappa \geq 3$ and it is numerically computed even as a negative value, while of course the determinant of a covariance matrix has to be nonnegative. As a result, the R function chol for computing the Cholesky decomposition fails for Σ with $\kappa \geq 6$ due to the numerical degeneracy. In the next section, we propose an improved scheme which does not increase the dimension of the Gaussian random variables but reduces significantly the mean squared error. The efficiency of our scheme indeed results from the almost degeneracy of Σ.

4. 3R Hybrid scheme

Here we propose a new scheme, the 3R Hybrid scheme, which improves the Hybrid scheme.

4.1. An improved approximation to the kernel function

The key idea of the Hybrid scheme is to use a power function to approximate a singular part of the kernel function. The remained regular part is approximated by a step function. Our idea is to reuse the power function part to improve the approximation to the regular part. This is done by orthogonal projections. Let us consider how to approximate a power function f

$$f(x) := x^{\alpha}, \quad \alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right), x \in \left(0, \frac{\lfloor nT \rfloor}{n}\right]$$

for an arbitrary $T > 0$. Denote

$$f_k^n(x) := \begin{cases} f(x) & x \in \left(\frac{k-1}{n}, \frac{k}{n}\right] \\ 0 & \text{otherwise,} \end{cases}$$

to decompose f as

$$f(x) = \sum_{k=1}^{\lfloor nT \rfloor} f_k^n(x).$$

Our idea is to approximate $f_k^n(x)$ with $k \geq \kappa + 1$ by a linear combination of a constant function and $f_\kappa^n(x - (k-\kappa)/n)$:

$$\begin{aligned} f(x) \approx\, & f_1^n(x) + \cdots + f_\kappa^n(x) \\ & + \left(a_{\kappa+1} + b_{\kappa+1} f_\kappa^n\left(x - \frac{1}{n}\right)\right) \\ & + \cdots + \left(a_{\lfloor nT \rfloor} + b_{\lfloor nT \rfloor} f_\kappa^n\left(x - \frac{\lfloor nT \rfloor - \kappa}{n}\right)\right) \\ = & \sum_{k=1}^{\kappa} f_k^n(x) + \sum_{k=\kappa+1}^{\lfloor nT \rfloor} \left(a_k + b_k f_\kappa^n\left(x - \frac{k-\kappa}{n}\right)\right), \end{aligned}$$

Table 1. $\det(\Sigma)$ for $\kappa \in \{1, \ldots, 6\}$ computed by the R function det. The hypergeometric functions in (11) are by the R function hyperg_2F1 in the package gsl.

α	1	2	3	4	5	6
−0.49	4.4e+00	2.7e−02	1.0e−07	1.8e−16	9.1e−29	−9.6e−43
−0.4	1.4e−01	3.0e−04	4.5e−10	3.2e−19	6.8e−32	−2.1e−47
−0.3	1.8e−02	1.1e−05	5.0e−12	1.1e−21	7.9e−35	1.9e−49
−0.2	2.6e−03	3.4e−07	3.8e−14	2.1e−24	3.7e−38	−1.7e−52
−0.1	2.4e−04	3.8e−09	5.7e−17	4.4e−28	9.3e−43	2.9e−57
0.1	4.3e−05	1.6e−10	6.3e−19	1.4e−30	1.6e−45	7.7e−61
0.2	7.9e−05	5.4e−10	4.3e−18	1.9e−29	2.1e−44	−2.3e−59
0.3	8.4e−05	5.9e−10	5.0e−18	2.4e−29	3.7e−44	5.9e−60
0.4	7.2e−05	4.0e−10	2.8e−18	1.1e−29	1.9e−44	1.5e−59
0.49	5.7e−05	2.2e−10	1.1e−18	3.0e−30	3.3e−45	−1.4e−59

where a_k, b_k are determined so that the mean squared error (MSE)

$$\begin{aligned}\mathrm{MSE}_k &:= \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(f(x) - \left(a_k + b_k f_\kappa^n\left(x - \frac{k-\kappa}{n}\right)\right)\right)^2 \mathrm{d}x \\ &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(x^\alpha - \left(a_k + b_k\left(x - \frac{k-\kappa}{n}\right)^\alpha\right)\right)^2 \mathrm{d}x \\ &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} x^{2\alpha} + a_k^2 + b_k^2\left(x - \frac{k-\kappa}{n}\right)^{2\alpha} \\ &\quad - 2b_k x^\alpha\left(x - \frac{k-\kappa}{n}\right)^\alpha - 2a_k x^\alpha \\ &\quad + 2a_k b_k\left(x - \frac{k-\kappa}{n}\right)^\alpha \mathrm{d}x \end{aligned} \tag{13}$$

is minimized. It is easy to obtain the minimizer

$$a_k^* = \frac{AE - BC}{AD - B^2}, \quad b_k^* = \frac{CD - BE}{AD - B^2}, \tag{14}$$

where

$$\begin{aligned} A &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(x - \frac{k-\kappa}{n}\right)^{2\alpha} \mathrm{d}x \\ &= \frac{1}{2\alpha+1}\left(\left(\frac{\kappa}{n}\right)^{2\alpha+1} - \left(\frac{\kappa-1}{n}\right)^{2\alpha+1}\right), \\ B &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} \left(x - \frac{k-\kappa}{n}\right)^{\alpha} \mathrm{d}x \\ &= \frac{1}{\alpha+1}\left(\left(\frac{\kappa}{n}\right)^{\alpha+1} - \left(\frac{\kappa-1}{n}\right)^{\alpha+1}\right), \\ C &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} x^\alpha\left(x - \frac{k-\kappa}{n}\right)^{\alpha} \mathrm{d}x \\ &= \frac{1}{(\alpha+1)n^{2\alpha+1}}\left(\kappa^{\alpha+1}k^\alpha {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{\kappa}{k}\right)\right. \\ &\quad \left. -(\kappa-1)^{\alpha+1}(k-1)^\alpha {}_2F_1\left(-\alpha, 1, \alpha+2, \frac{\kappa-1}{k-1}\right)\right), \\ D &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} \mathrm{d}x = \frac{1}{n}, \\ E &= \int_{\frac{k-1}{n}}^{\frac{k}{n}} x^\alpha \mathrm{d}x = \frac{1}{\alpha+1}\left(\left(\frac{k}{n}\right)^{\alpha+1} - \left(\frac{k-1}{n}\right)^{\alpha+1}\right). \end{aligned}$$

Figures 4 and 5 show that our approximation with $\kappa = 2$ is already very precise in spite of the coarse division $n = 10$. To examine the reduction of the MSE with $\kappa = 2$, some numerical examples also can be given; denote

$$\begin{aligned}\mathrm{MSE}^*(\alpha, T, n) &:= \sum_{k=3}^{\lfloor nT\rfloor} \int_{\frac{k-1}{n}}^{\frac{k}{n}} (f(x) - (a_k \\ &\quad + b_k f_2^n\left(x - \frac{k-2}{n}\right))\Big)^2 \mathrm{d}x \end{aligned}$$

$$\mathrm{MSE}^{**}(\alpha, T, n) := \sum_{k=3}^{\lfloor nT\rfloor} \int_{\frac{k-1}{n}}^{\frac{k}{n}} (f(x) - c_k)^2 \mathrm{d}x,$$

where (a_k, b_k) and c_k are the optimized values. When $\alpha = -0.49$, we have

$$\begin{aligned}\mathrm{MSE}^*(-0.49, 1, 10) &= 1.16317 \times 10^{-5}, \\ \mathrm{MSE}^{**}(-0.49, 1, 10) &= 2.27096 \times 10^{-3}, \end{aligned}$$

and when $\alpha = 0.49$,

$$\begin{aligned}\mathrm{MSE}^*(0.49, 1, 10) &= 2.87234 \times 10^{-7}, \\ \mathrm{MSE}^{**}(0.49, 1, 10) &= 3.26240 \times 10^{-4}. \end{aligned}$$

See Figure 6 for plots with different n.

4.2. Implementation

Here we describe how to simulate X_t on the equidistant grid $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{\lfloor nT\rfloor}{n}\}$ for some $T > 0$ using the 3R Hybrid scheme. We introduce a new parameter $\kappa' \in \mathbb{N}$, $\kappa' > \kappa$ and approximate

$$\begin{aligned} g\left(\frac{i}{n} - s\right) &\approx \left(\frac{i}{n} - s\right)^\alpha L_g\left(\frac{k}{n}\right), \\ &\qquad s \in \left[\frac{i-k}{n}, \frac{i-k+1}{n}\right],\ k = 1, \ldots, \kappa' \\ g\left(\frac{i}{n} - s\right) &\approx g\left(\frac{c_k}{n}\right), \\ &\qquad s \in \left[\frac{i-k}{n}, \frac{i-k+1}{n}\right],\ k = \kappa'+1, \ldots, N_n. \end{aligned}$$

As in the Hybrid scheme, we approximate the stochastic integral on the interval $(-\infty, i/n]$ for $X_{i/n}$ by that on the finite interval $((i - N_n)/n, i/n]$. Then,

$$\begin{aligned} X_{\frac{i}{n}} &\approx \sum_{k=1}^{\kappa'} \int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}} g\left(\frac{i}{n} - s\right)\sigma_s \mathrm{d}W_s \\ &\quad + \sum_{k=\kappa'+1}^{N_n} \int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}} g\left(\frac{i}{n} - s\right)\sigma_s \mathrm{d}W_s \\ &\approx \sum_{k=1}^{\kappa'} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n} \int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}} \left(\frac{i}{n} - s\right)^\alpha \mathrm{d}W_s \\ &\quad + \sum_{k=\kappa'+1}^{N_n} g\left(\frac{c_k}{n}\right)\sigma_{(i-k)/n} \int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}} \mathrm{d}W_s. \end{aligned}$$

In the new scheme, the 3R Hybrid scheme, the kernel function for $k \in [\kappa+1, \kappa']$ is approximated using that for $k = \kappa$. We then define

$$\hat{X}^n_{\frac{i}{n}} = \sum_{k=1}^{\kappa} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n} \underbrace{\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}} \left(\frac{i}{n} - s\right)^\alpha \mathrm{d}W_s}_{W^n_{i-k,k}}$$

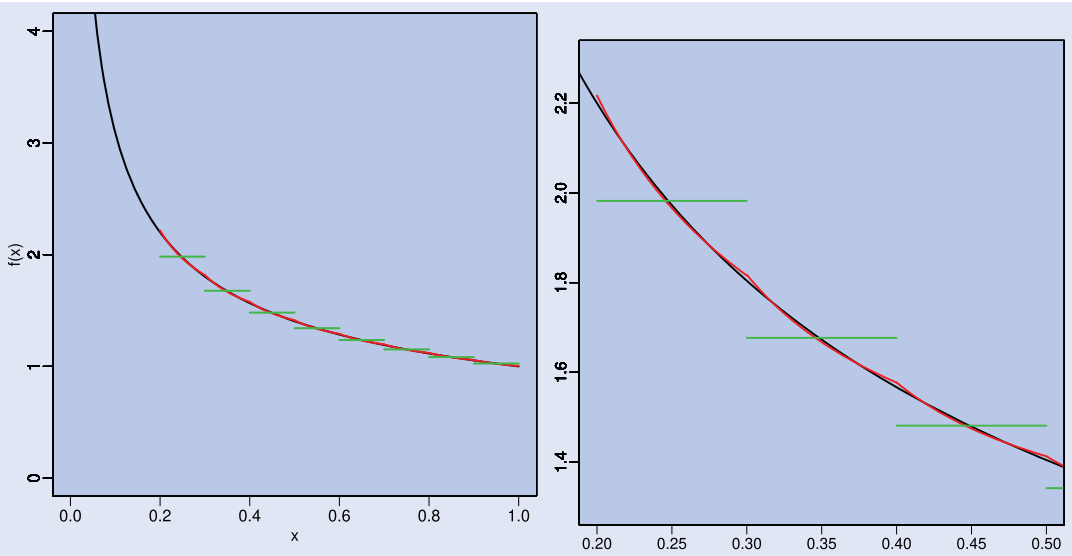

Figure 4. The quality of approximation with the optimal (a_k, b_k) (red) in comparison with a step function approximation with the optimal c_k of the Hybrid scheme (green) ($\alpha = -0.49, T = 1, n = 10, \kappa = 2$). The right figure is a zoom-in of the left one.

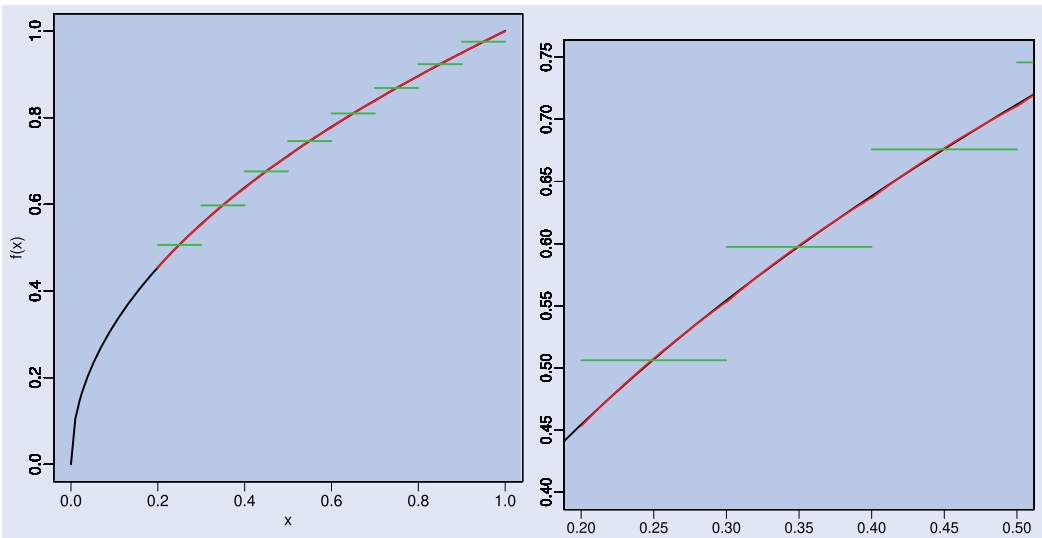

Figure 5. The quality of approximation with the optimal (a_k, b_k) (red) in comparison with a step function approximation with the optimal c_k of the Hybrid scheme (green) ($\alpha = 0.49, T = 1, n = 10, \kappa = 2$). The right figure is a zoom-in of the left one.

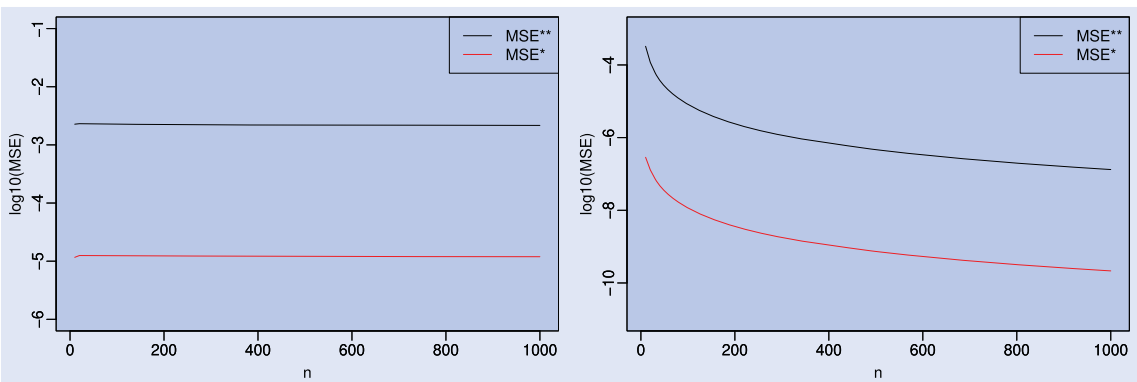

Figure 6. Comparison of $\mathrm{MSE}^*(\alpha, 1, n)$ and $\mathrm{MSE}^{**}(\alpha, 1, n)$. (Left) $\alpha = -0.49$. (Right) $\alpha = 0.49$.

$$
\begin{aligned}
&+\sum_{k=\kappa+1}^{\kappa'} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n}\\
&\times\Bigg(a_k\underbrace{\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}\mathrm{d}W_s}_{W^n_{i-k}}\\
&+b_k\underbrace{\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}\left(\frac{i+\kappa-k}{n}-s\right)^{\alpha}\mathrm{d}W_s}_{W^n_{i-k,\kappa}}\Bigg)\\
&+\sum_{k=\kappa'+1}^{N_n} g\left(\frac{c_k}{n}\right)\sigma_{(i-k)/n}\underbrace{\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}\mathrm{d}W_s}_{W^n_{i-k}},
\end{aligned}
\tag{15}
$$

where a_k and b_k are determined so that

$$
\begin{aligned}
\mathrm{MSE}_k &= \mathbb{E}\Bigg[\Bigg\{\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}\left(\frac{i}{n}-s\right)^{\alpha}\mathrm{d}W_s\\
&\quad-\Bigg(a_k\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}dW_s\\
&\quad+b_k\int_{\frac{i-k}{n}}^{\frac{i-k+1}{n}}\left(\frac{i+\kappa-k}{n}-s\right)^{\alpha}\mathrm{d}W_s\Bigg)\Bigg\}^2\Bigg]\\
&=\mathbb{E}\left[\left\{W^n_{i-k,k}-\left(a_kW^n_{i-k}+b_kW^n_{i-k,\kappa}\right)\right\}^2\right],\\
&k\in[\kappa+1,\kappa']
\end{aligned}
$$

is minimized. It is easy to check this MSE coincides with (13). The minimizer is therefore given by (14). The $\kappa'+1$ dimensional random vectors

$$
\mathbf{W}^n_{i-k}:=(W^n_{i-k},W^n_{i-k,1},\dots,W^n_{i-k,\kappa'})
$$

has the expected value $\mathbf{0}=(0,\dots,0)$ and the variance-covariance matrix Σ given by (9)$\sim$(11) for $j=2,\dots,\kappa'+1$ and (12) for $j,k=2,\dots,\kappa'+1$ with $j<k$. We can rewrite MSE_k using Σ;

$$
\begin{aligned}
\mathrm{MSE}_k &= \Sigma_{1,1}\left\{a_k+\frac{\Sigma_{1,\kappa+1}}{\Sigma_{1,1}}b_k-\frac{\Sigma_{1,k+1}}{\Sigma_{1,1}}\right\}^2\\
&\quad+\left(\Sigma_{\kappa+1,\kappa+1}-\frac{\Sigma^2_{1,\kappa+1}}{\Sigma_{1,1}}\right)\\
&\quad\times\left\{b_k-\frac{\Sigma_{1,1}\Sigma_{\kappa+1,k+1}-\Sigma_{1,\kappa+1}\Sigma_{1,k+1}}{\Sigma_{1,1}\Sigma_{\kappa+1,\kappa+1}-\Sigma^2_{1,\kappa+1}}\right\}^2+D_k,
\end{aligned}
$$

where

$$
\begin{aligned}
D_k &= -\left(\Sigma_{\kappa+1,\kappa+1}-\frac{\Sigma^2_{1,\kappa+1}}{\Sigma_{1,1}}\right)\\
&\quad\times\left(\frac{\Sigma_{1,1}\Sigma_{\kappa+1,k+1}-\Sigma_{1,\kappa+1}\Sigma_{1,k+1}}{\Sigma_{1,1}\Sigma_{\kappa+1,\kappa+1}-\Sigma^2_{1,\kappa+1}}\right)^2\\
&\quad+\Sigma_{k+1,k+1}-\frac{\Sigma^2_{1,k+1}}{\Sigma_{1,1}}.
\end{aligned}
\tag{16}
$$

The minimizer (a^*_k,b^*_k) is then expressed as

$$
\begin{aligned}
a^*_k &= \frac{\Sigma_{1,k+1}\Sigma_{\kappa+1,\kappa+1}-\Sigma_{1,\kappa+1}\Sigma_{\kappa+1,k+1}}{\Sigma_{1,1}\Sigma_{\kappa+1,\kappa+1}-\Sigma^2_{1,\kappa+1}},\\
b^*_k &= \frac{\Sigma_{1,1}\Sigma_{\kappa+1,k+1}-\Sigma_{1,\kappa+1}\Sigma_{1,k+1}}{\Sigma_{1,1}\Sigma_{\kappa+1,\kappa+1}-\Sigma^2_{1,\kappa+1}}.
\end{aligned}
$$

Hereafter, we assume $\mathbf{a}=\mathbf{a}^*=\{a^*_k\}_{k=\kappa+1\cdots,\kappa'}$, $\mathbf{b}=\mathbf{b}^*=\{b_k\}_{k=\kappa+1\cdots,\kappa'}$, $\mathbf{c}=\mathbf{c}^*$.

The 3R Hybrid scheme and the Hybrid scheme require the same size of random number generation for the same value of κ: $N_n+[nT]-1$ i.i.d samples of a $\kappa+1$ dimensional multivariate normal random variable. An additional computation is required for

$$
\Sigma_{1,k+1},\quad \Sigma_{\kappa+1,k+1},\quad k\in[\kappa+1,\kappa']
$$

to determine a^*_k,b^*_k in the 3R Hybrid scheme; the matrix Σ is however not random and so, can be computed within negligible time at the beginning of simulations. Therefore the computational cost for the 3R Hybrid scheme is almost the same as the Hybrid scheme with the same value of κ. A numerical experiment to confirm this is reported in Appendix 2.

4.3. Asymptotic behavior of MSE

Here, we state a theorem describing the approximate accuracy of the 3R Hybrid scheme. We omit the proof because it is essentially a repetition of the proof of Theorem 3.1 given in Bennedsen *et al.* (2007).

THEOREM 4.1 *Under the same assumptions as in Theorem* 3.1, *for any t on a grid* $\{1/n,2/n,\dots\}$,

$$
\begin{aligned}
&\mathbb{E}\left[|X_t-\hat{X}^n_t|^2\right]\\
&\quad\sim\left(\sum_{k=\kappa+1}^{\kappa'}E_k+\boldsymbol{J}(\alpha,\kappa',\mathbf{c})\right)\mathbb{E}[\sigma_0^2]n^{-(2\alpha+1)}L_g\\
&\quad\times\left(\frac{1}{n}\right)^2,\quad n\to\infty,
\end{aligned}
\tag{17}
$$

where $E_k=D_kn^{2\alpha+1}$ *and* D_k *is given by* (16).

REMARK 4.2 E_k does not depend on n.

4.4. The choice of κ

Here we argue about the choice of κ for the 3R Hybrid scheme. We propose to choose $\kappa=2$ by the following two reasons : (1) it enables a fast computation due to its mild requirement of only $\kappa+1=3$ dimensional i.i.d. Gaussian random number generations, and (2) it achieves a relatively

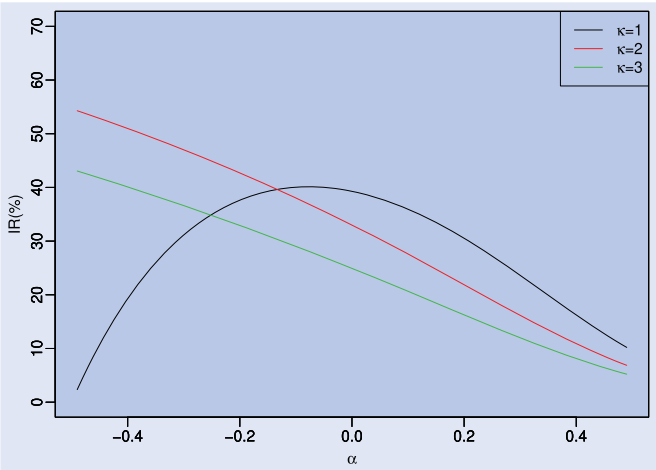

Figure 7. The variation of $IR_1(\alpha, \kappa, \mathbf{c})$.

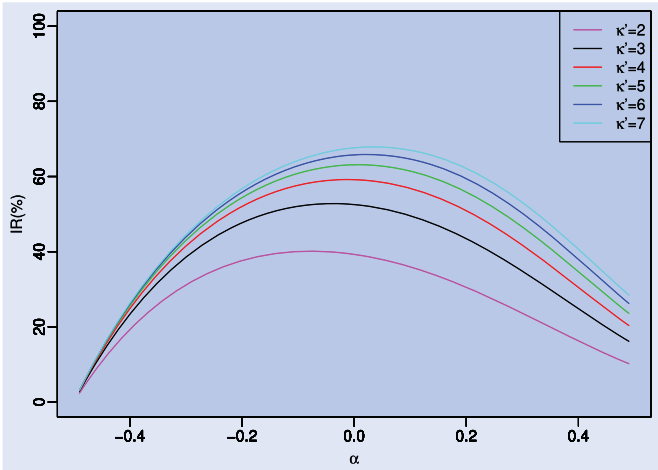

Figure 8. The improvement ratio $IR_{3R}(\alpha, 1, \kappa', \mathbf{c})$.

significant reduction of the mean squared error. To make the second point precise, based on Theorem 4.1, we look at the improvement ratio (IR) of the 3R Hybrid scheme with $\kappa' = \kappa + 1$ compared with the Hybrid scheme with the same κ;

$$IR_1(\alpha, \kappa, \mathbf{c}) := \left(1 - \frac{E_{\kappa+1} + \mathbf{J}(\alpha, \kappa+1, \mathbf{c})}{\mathbf{J}(\alpha, \kappa, \mathbf{c})}\right) \times 100\ (\%).$$

Figure 7 shows the IR for $\kappa = 1, 2, 3$ for various values of α. We see that the IR for $\kappa = 2$ is higher than $\kappa = 3$ for any α. The improvement ratio of the 3R Hybrid scheme with general $\kappa' > \kappa$ is defined as

$$IR_{3R}(\alpha, \kappa, \kappa', \mathbf{c}) := \left(1 - \frac{\displaystyle\sum_{k=\kappa+1}^{\kappa'} E_k + \mathbf{J}(\alpha, \kappa', \mathbf{c})}{\mathbf{J}(\alpha, \kappa, \mathbf{c})}\right) \times 100\ (\%).$$

Figures 8 and 9 show the values for $\kappa = 1$ and $\kappa = 2$ respectively. We see that the approximate precision is not much improved when $\kappa = 1$ for $\alpha \approx -\frac{1}{2}$. On the other hand, when $\kappa = 2$, the approximate precision is significantly improved by increasing κ' for $\alpha \in (-\frac{1}{2}, 0)$.

4.5. Relative error

Here we argue that when choosing $\kappa = 2$, the 3R Hybrid scheme achieves almost the same accuracy as the Hybrid scheme with $\kappa = \kappa' > 2$, while its computational cost remains the same as the latter with $\kappa = 2$. Figure 10 shows the magnitude of E_k, $k = 3, 4, \ldots$ with $\kappa = 2$. We see that E_k is much smaller than J shown in Figure 3.

The relative errors R_X and $R_{\hat{X}}$ of the Hybrid and the 3R Hybrid schemes respectively are defined as

$$R_X = \frac{\mathbb{E}[|X_t - X_t^n|^2]}{\mathbb{E}[X_t^2]}, \quad R_{\hat{X}} = \frac{\mathbb{E}[|X_t - \hat{X}_t^n|^2]}{\mathbb{E}[X_t^2]}.$$

Here, we take $\sigma \equiv 1$ and a fOU kernel as g:

$$g(x) = x^\alpha - e^{-x} \int_0^x s^\alpha e^s \, \mathrm{d}s. \tag{18}$$

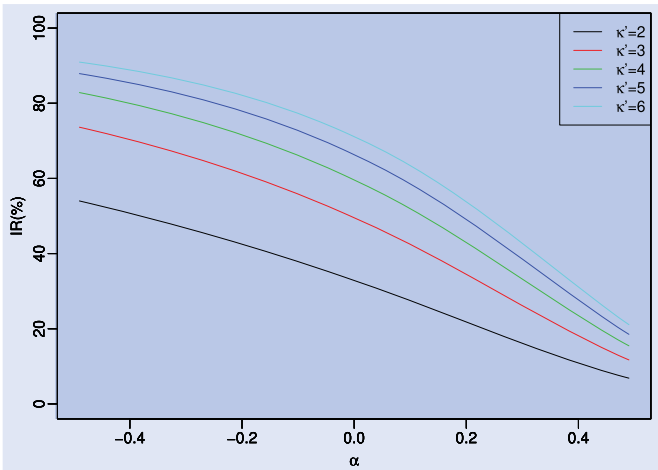

Figure 9. The improvement ratio $IR_{3R}(\alpha, 2, \kappa', \mathbf{c})$.

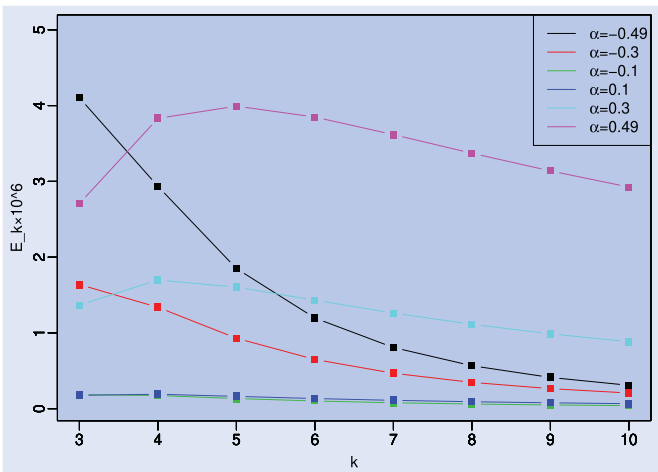

Figure 10. The variation of the error coefficient E_k with $\kappa = 2$.

See Appendix 1 for details of fOU processes. The denominator is then computed as

$$\mathbb{E}[X_t^2] = \mathbb{E}\left[\left(\int_{-\infty}^t g(t-s)\, \mathrm{d}W_s\right)^2\right] = \int_0^\infty g(x)^2 \, \mathrm{d}x$$

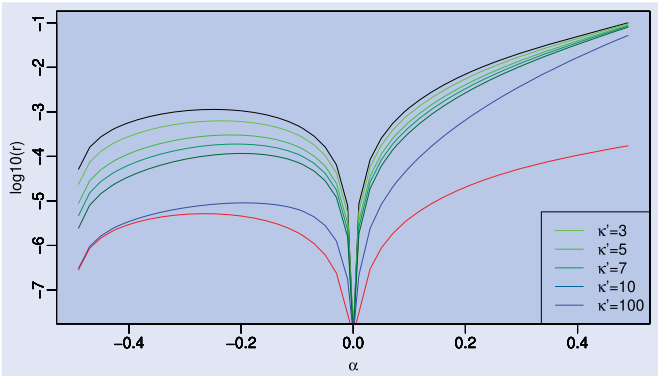

Figure 11. $r_{\hat{X}}(2,\kappa')$ as a function of α.

by the Itô isometry. The numerators are represented

$$\mathbb{E}\left[|X_t - X_t^n|^2\right] \sim \boldsymbol{J}(\alpha,\kappa,\mathbf{c})n^{-(2\alpha+1)}L_g\left(\frac{1}{n}\right)^2, \quad n\to\infty,$$

$$\mathbb{E}\left[|X_t - \hat{X}_t^n|^2\right] \sim \left(\sum_{k=\kappa+1}^{\kappa'} E_k + \boldsymbol{J}(\alpha,\kappa',\mathbf{c})\right) n^{-(2\alpha+1)}L_g\left(\frac{1}{n}\right)^2, \quad n\to\infty$$

by (7) and (17). Since the convergence rates are the same, it is reasonable to compare

$$r_X(\kappa) := \frac{\boldsymbol{J}(\alpha,\kappa,\mathbf{c})}{\int_0^\infty g(x)^2\,\mathrm{d}x}, \quad r_{\hat{X}}(\kappa,\kappa') := \frac{\sum_{k=\kappa+1}^{\kappa'} E_k + \boldsymbol{J}(\alpha,\kappa',\mathbf{c})}{\int_0^\infty g(x)^2\,\mathrm{d}x}.$$

In Figure 11, the black line is for $r_X(\kappa) = r_{\hat{X}}(\kappa,\kappa)$ with $\kappa = 2$, and the red line is for

$$r^* = \frac{\sum_{k=\kappa+1}^{N} E_k}{\int_0^\infty g(x)^2\,\mathrm{d}x}.$$

with $\kappa = 2$ and $N = 10^4$, which is essentially $\lim_{\kappa'\to\infty} r_{\hat{X}}(2,\kappa')$. We see that r^* is quite small especially for $\alpha < 0$, which means that $r_{\hat{X}}(2,\kappa') \approx r_X(\kappa')$, that is, the 3R Hybrid scheme achieves almost the same accuracy as the Hybrid scheme with $\kappa = \kappa' > 2$, while its computational cost (the size of random number generations) remains the same as the latter with $\kappa = 2$.

5. The 3R Hybrid scheme for rough volatility models

Here we apply the 3R Hybrid scheme to the analysis of the rough Bergomi model (Bayer *et al.* 2016).

5.1. Practical implementation

Definition 5.1 (tBSS) *Let Y be of the form*

$$Y_t = \int_0^t g(t-s)\sigma_s\,dW_s, \quad t \ge 0$$

where the volatility process σ and the driving Brownian motion W are as in Definition 2.1. The kernel function k is assumed to be locally square integrable. Following Bennedsen et al. *(2007), we call Y_t a truncated Brownian semistationary(tBSS) process.*

A tBSS process is a BSS process with $\sigma_s = 0$ for $s < 0$. But as a result of the truncation, it can be defined without the square integrability of the kernel function g. Similarly to BSS processes, tBSS process Y_t can be approximated as

$$\begin{aligned}\hat{Y}^n_{\frac{i}{n}} \approx &\sum_{k=1}^{\min\{i,\kappa\}} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n}W^n_{i-k,k} \\ &+ \sum_{k=\kappa+1}^{\min\{i,\kappa'\}} L_g\left(\frac{k}{n}\right)\sigma_{(i-k)/n}\left(a_k W^n_{i-k} + b_k W^n_{i-k,\kappa}\right) \\ &+ \sum_{k=\kappa'+1}^{i} g\left(\frac{c_k}{n}\right)\sigma_{(i-k)/n}W^n_{i-k}\end{aligned}$$

for $i = 0,1,\ldots,\lfloor nT\rfloor$. In the rough Bergomi model, an asset price process S is defined by

$$S_t = S_0\exp\left\{\int_0^t \sqrt{V_s}[\rho\,\mathrm{d}W_s + \sqrt{1-\rho^2}\,\mathrm{d}W_s^\perp] - \frac{1}{2}\int_0^t V_s\,\mathrm{d}s\right\},$$

$$V_t = \xi(t)\exp\left\{\int_0^t g(t-s)\,\mathrm{d}W_s - \frac{1}{2}\int_0^t g(t-s)^2\,\mathrm{d}s\right\}$$

with $g(x) = \eta x^\alpha$, $\eta > 0$, $\alpha \in (-\frac{1}{2}, 0)$, where $W^\perp$ is a Brownian motion independent of W, $\rho \in [-1,1]$, and ξ is a deterministic continuous function, called the forward variance curve due to $E[V_t] = \xi(t)$. The process $V_{\frac{i}{n}}$ can be simulated by using the above 3R Hybrid scheme for tBBS processes. The log asset process $U_{\frac{i}{n}} = \log S_{\frac{i}{n}}$ is computed as

$$U_{\frac{i+1}{n}} = U_{\frac{i}{n}} - \frac{1}{2n}V_{\frac{i}{n}} + \sqrt{V_{\frac{i}{n}}}\left(Z_{\frac{i+1}{n}} - Z_{\frac{i}{n}}\right),$$
$$i = 0,1,\ldots,\lfloor nT\rfloor - 1,$$

where

$$Z_{\frac{i+1}{n}} - Z_{\frac{i}{n}} = \rho W_i^n + \sqrt{1-\rho^2}W_i^{\perp n}$$

and $W_i^{\perp n} \sim N(0,\frac{1}{n})$ independent of $(W_i^n, W_{i,1}^n, \cdots W_{i,\kappa}^n)$. If the purpose of the simulation is the Monte-Carlo pricing of an option, one should apply the conditional pricing formula by Romano and Touzi (1997) that allows to avoid $W_i^{\perp n}$ and to reduce the variance of the Monte Carlo significantly. See also McCrickerd and Pakkanen (2018) for other variance reduction techniques.

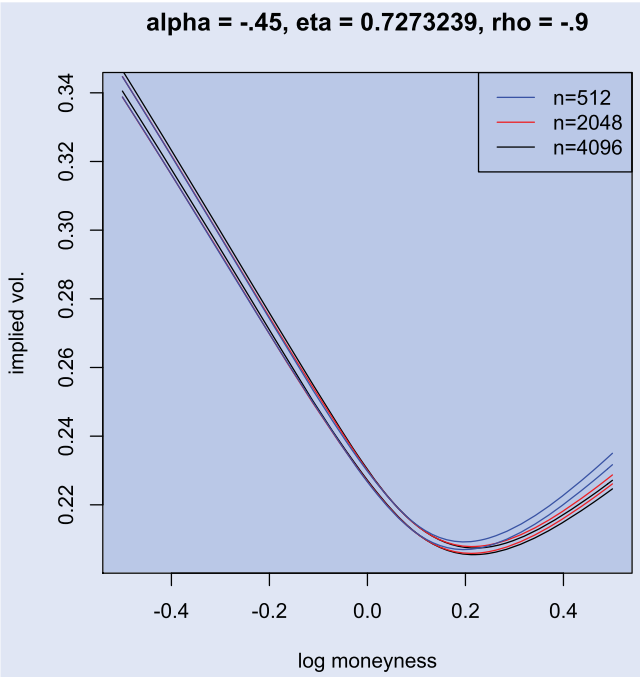

Figure 12. A slow convergence of weak approximation.

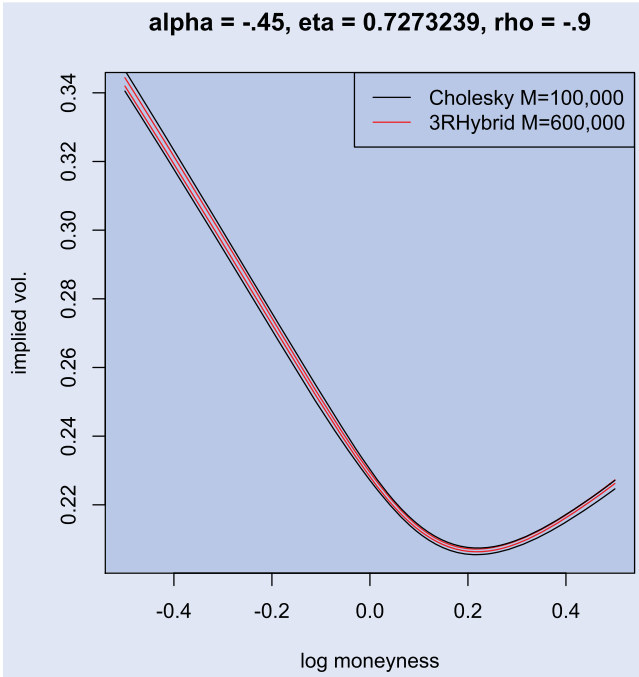

Figure 13. 99.7% confidence bound 3R Hybrid.

5.2. *The 3R Hybrid scheme v.s. the Cholesky decomposition*

For the Monte-Carlo pricing under the rough Bergomi model, we have to take care of three sources of error; (1) the Monte-Carlo error, (2) the discretization error and (3) the approximation error to a BSS process, where (2) is due to the Euler-Maruyama type approximation to the SDE outlined above, and (3) is due to the use of the Hybrid or 3R Hybrid scheme to approximate the tBSS process. Although the last one can be avoided by an exact simulation method using the Cholesky decomposition, here we argue that it is unfortunately not practical. Figure 12 shows the implied volatility smiles under the discretized rough Bergomi model with $\alpha = -0.45$, $\rho = -0.9$, $\eta = 2.3\sqrt{2\alpha+1}$, $\xi \equiv 0.04$ and maturity $T = 1$ (a set of calibrated parameters in Bayer *et al.* (2016)), with various choices of n. For each n, there are two curves that correspond to the Monte Carlo mean $\pm$ three standard deviations (99.7 % confidence bounds). Both call-option-implied and put-option implied volatilities were computed and we took one with smaller standard deviation for each log moneyness. The tBSS process, that is, the log volatility process was simulated using the Cholesky decomposition of the covariance matrix of $(\log V_{1/n}, \ldots, \log V_1, W_1^n, \ldots, W_n^n)$ that is of size $2n$. From the figure, one would hesitate to declare a convergence even with $n = 4,096$. The figure indicates a very slow convergence rate of weak approximation when $\alpha \approx -1/2$. As is well-known, when $\alpha = 0$, the weak approximation error is of $O(n^{-1})$, and we actually observed fast convergences when $\alpha \approx 0$ in our experiments. Although the weak approximation rate under rough volatility models is an open question and it is beyond our scope here, our numerical experiments suggest it depends on α, and we should warn here that n has to be far larger for $\alpha \approx -1/2$ than for $\alpha = 0$ (the classical situation). Then, the problem is that the

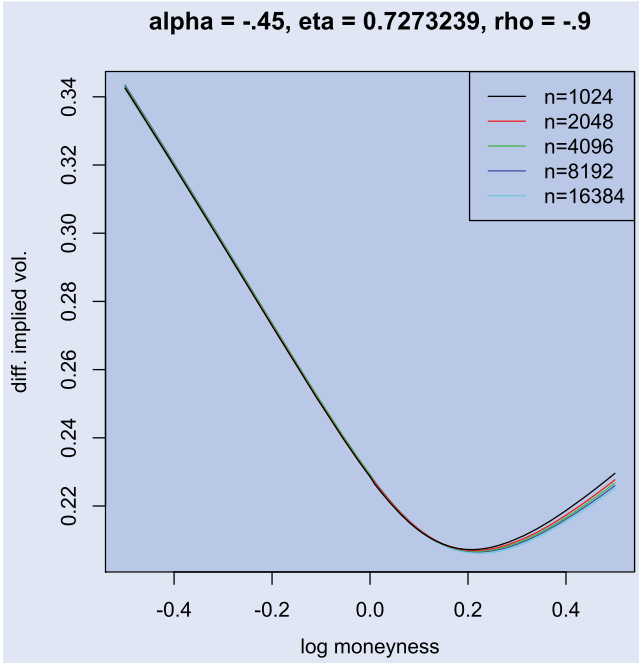

Figure 14. The convergence of the volatility smile.

Cholesky decomposition method is quite time-consuming for large n. For 100,000 Monte Carlo samples to get curves in Figure 12, with R programming, our Mac mini (2018) computer with 3.2 GHz 6 Cores Intel Core i7 took 345 minutes for $n = 2,048$ (the red curves) and 1,574 minutes for $n = 4,096$ (the black curves). The bottleneck of the computation is the time-consuming multiplication of the Cholesky-decomposed matrix and the standard normal random variable of size $2n$ in every path generations. Roughly, as n increases, with the size of the Monte-Carlo fixed, the computational cost for the

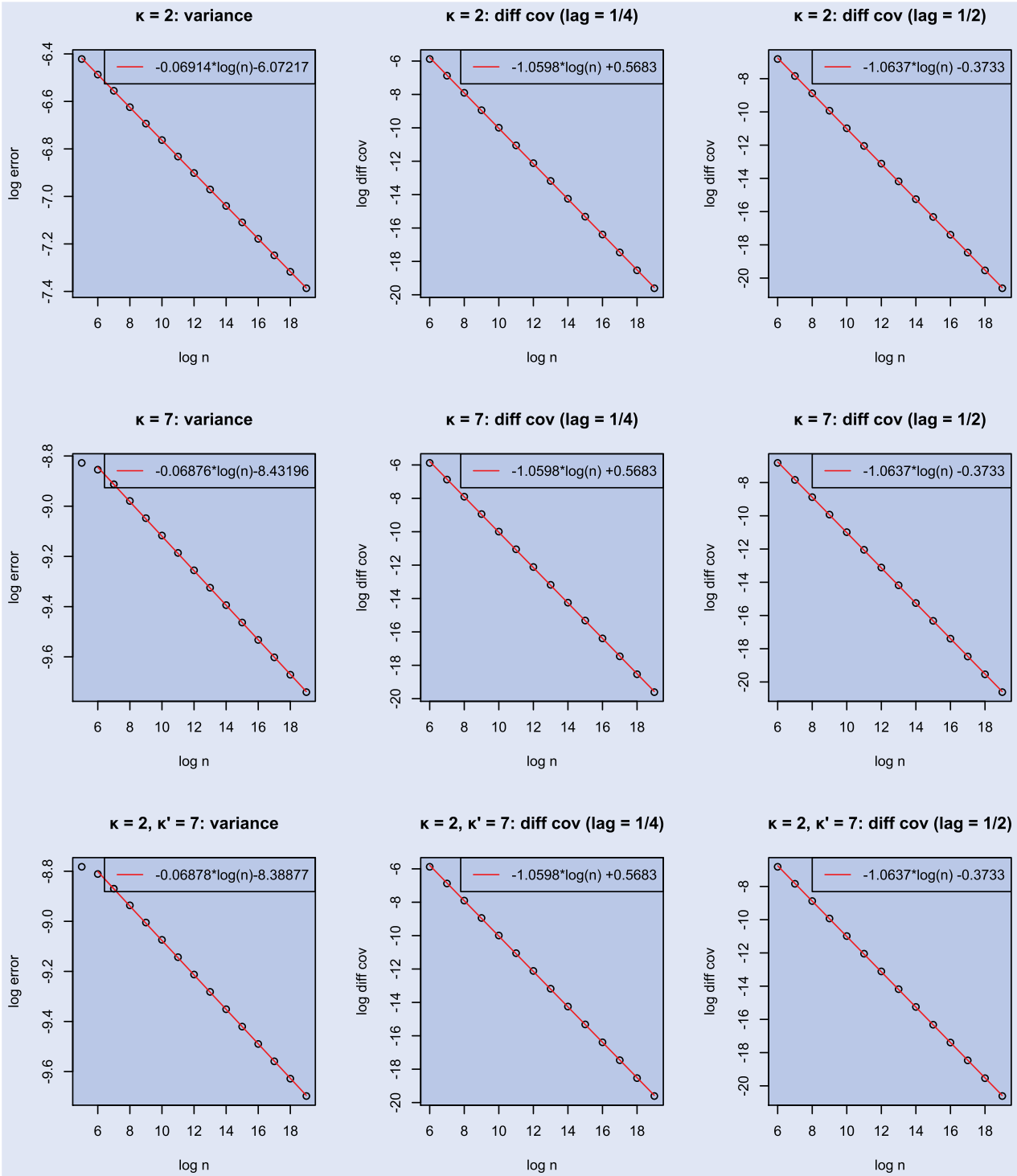

Figure 15. Errors in approximating the autocovariance.

Cholesky method grows with order n^2 (set aside a preliminary computation of the Cholesky decomposition of order n^3), while those for the Hybrid and the 3R Hybrid schemes do with order $n \log n$ thanks to the FFT convolution.

For a fixed number n, say, $n = 4,096$, the 3R Hybrid scheme achieves much faster the same error tolerance level as the Cholesky method does. The red curves in Figure 13 represent a 99.7% confidence band based on 600,000 Monte Carlo samples by the 3R Hybrid scheme with $(\kappa, \kappa') = (2, 10)$, computed in only 19 minutes by the same R programming and the same computer. As seen in the figure, it is within the 99.7% confidence band by the Cholesky method (the black curves, the same as the ones in Figure 12). Although the R programming is not the best choice for the Monte-Carlo pricing in terms of computation time, the superiority of the 3R Hybrid scheme to the Cholesky method is clear.

To discuss how n has to be large for the discretization error to be negligible, Figure 14 shows how the volatility smile converges as $n \to \infty$ for the same parameters as in Figures 12 and 13. The 3R Hybrid scheme with $(\kappa, \kappa') = (2, 10)$ was used because the Cholesky method is too slow. For this level of

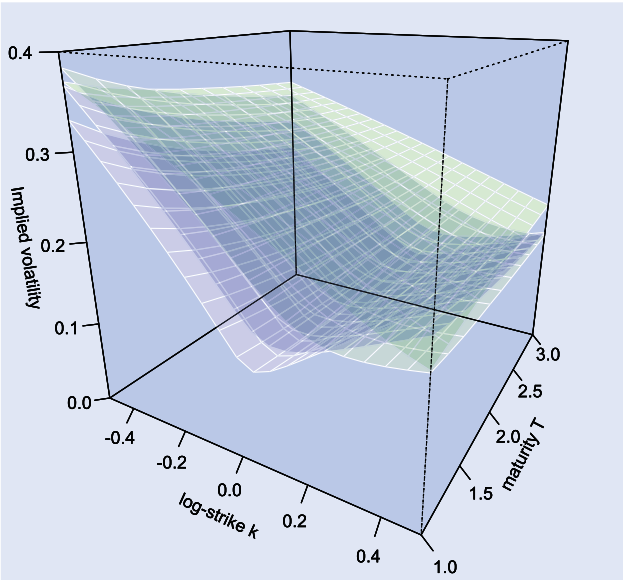

Figure 16. IV surfaces with various η: η =0.5 (green), 1, 1.5, and 2 (blue). $\alpha = -0.4$, $S_0 = 1$, $\rho = -0.9$, $\xi \equiv 0.3^2$.

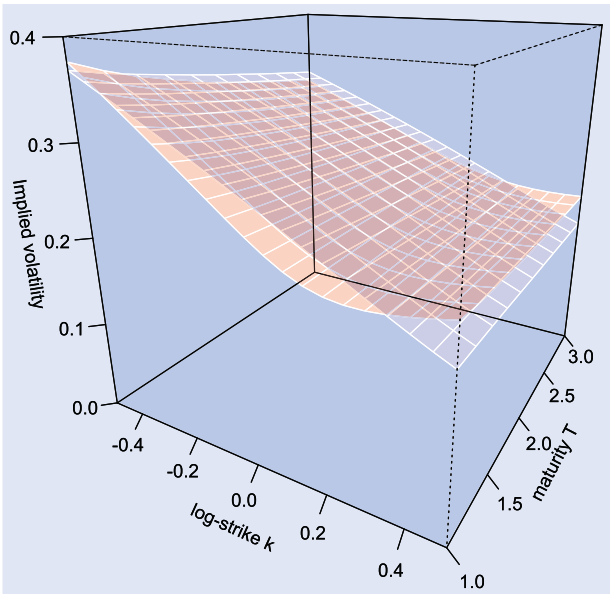

Figure 17. IV surfaces with various α: $\alpha = -0.45$ (red), and -0.1 (blue). $\eta = 0.8$, $S_0 = 1$, $\rho = -0.9$, $\xi \equiv 0.3^2$.

$\alpha \approx -0.45$, taking the confidence bounds also into account, we can give a rough criterion that $n = 2^{13} = 8,192$ gives an accuracy of order 10^{-3}.

5.3. The 3R Hybrid scheme v.s. the Hybrid scheme

For the Monte Carlo pricing, pathwise errors are not directly relevant and it is enough that the law of $(\log V_{1/n}, \ldots, \log V_1, W_1^n, \ldots, W_n^n)$ is well approximated. Therefore in this context, a more relevant approximation error to the tBSS process than the mean squared error studied in Theorems 3.1 and 4.1 is the difference between the autocovariance functions of the exact and approximating Gaussian sequences. Here we examine it for the Hybrid and 3R

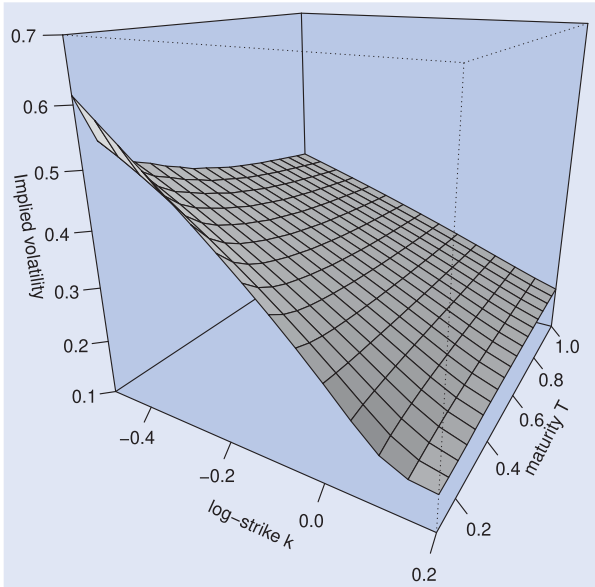

Figure 18. 1. rough Bergomi.

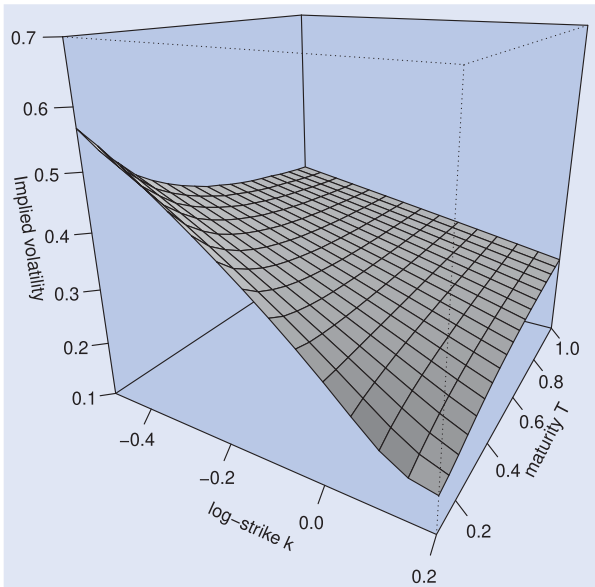

Figure 19. 2. fOU ($\lambda = 3$).

Hybrid schemes by some numerical examples. The left column of Figure 15 presents log-log plots of $E[Y_1^2] - E[|Y_1^n|^2]$, where

$$Y_1 = \int_0^1 g(1-s)\,\mathrm{d}W_s, \quad g(t) = 2\sqrt{2\alpha+1}\,t^{\alpha}, \quad \alpha = -0.45$$

and Y_1^n is its approximation by the Hybrid scheme with $\kappa = 2$ for the first row, the one by the Hybrid scheme with $\kappa = 7$ for the second row, and the one by the 3R Hybrid scheme with $(\kappa, \kappa') = (2, 7)$ for the third row. The logarithmic scale of x-axis is of base 2. First we notice its slow convergence with rate being approximately $n^{-0.07}$ that is not too far from $n^{-(2\alpha+1)}$. The values of intercept in the linear fitting of the log-log plots are approximately -8.4 for both the Hybrid scheme with $\kappa = 7$ and the 3R Hybrid scheme with $(\kappa, \kappa') = (2, 7)$.

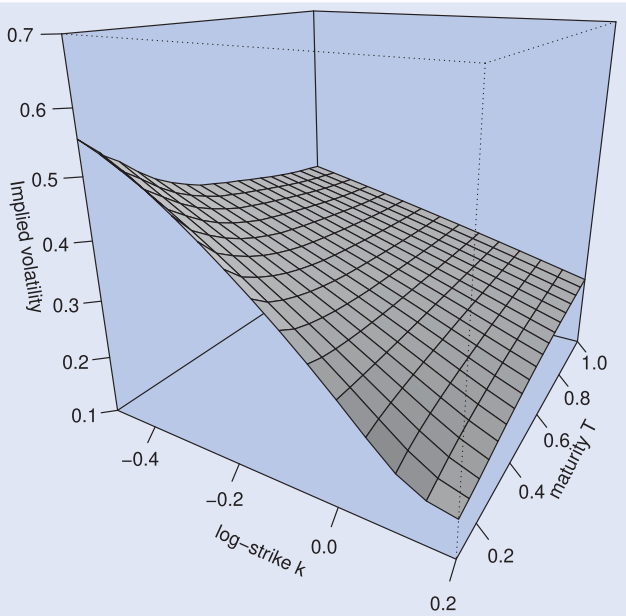

Figure 20. 3. gamma ($\lambda = 3$).

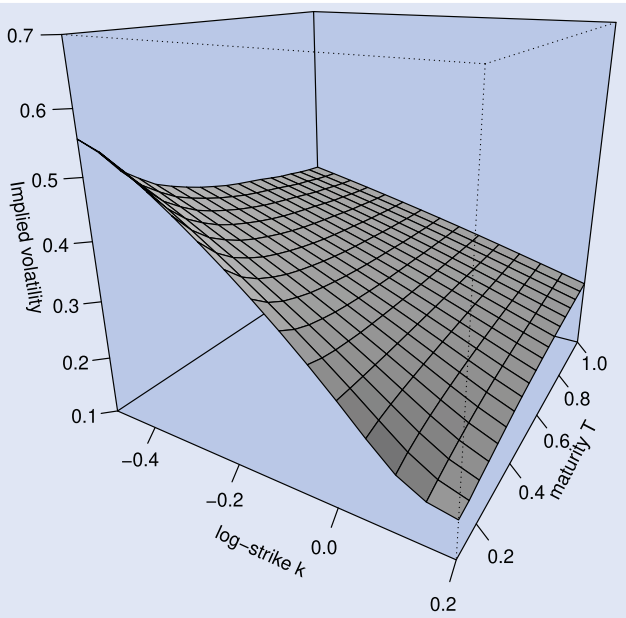

Figure 21. 4. power-low ($\beta = -3$).

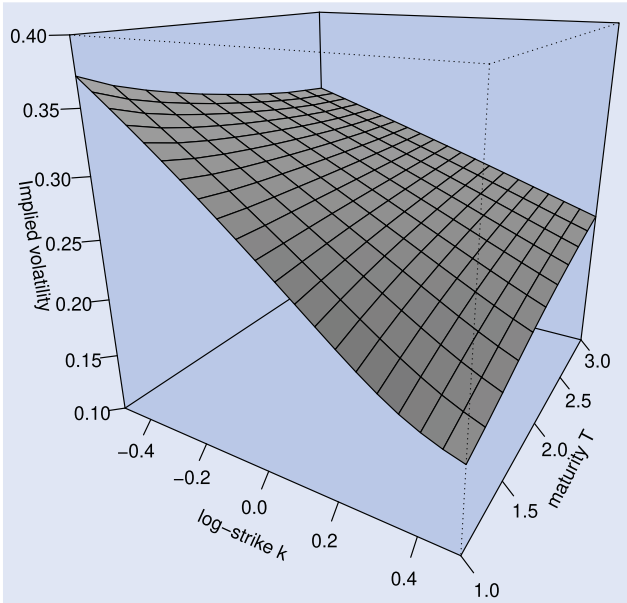

Figure 22. The fOU kernel ($\lambda = 1$).

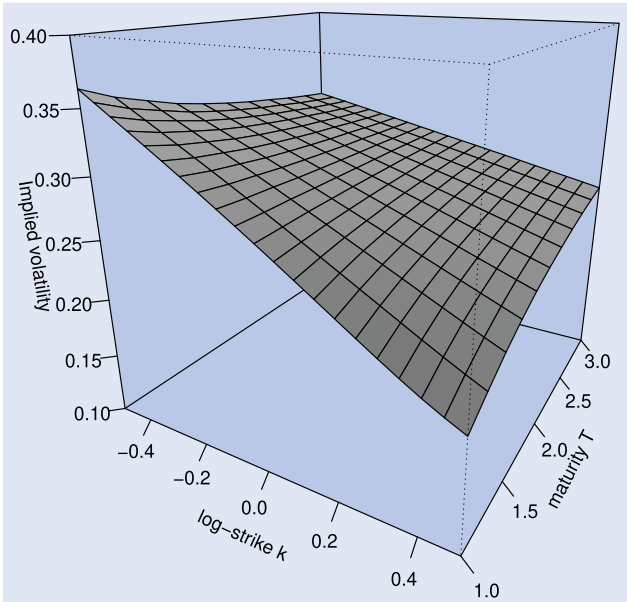

Figure 23. The fOU kernel ($\lambda = 2$).

This is consistent to our discussion in the previous sections that the accuracy of the 3R Hybrid scheme with (κ, κ') is essentially the same as that of the Hybrid scheme with $\kappa = \kappa'$, and better than that of the Hybrid scheme with the same value of κ. The middle and right columns of Figure 15 give log-log plots of $E_n - E_{n/2}$, where $E_n = E[Y_1 Y_\ell] - E[Y_1^n Y_\ell^n]$, for $\ell = 3/4$ and $\ell = 1/2$ respectively. Again the first, the second and the third rows are respectively for the Hybrid scheme with $\kappa = 2$, that with $\kappa = 7$ and the 3R Hybrid scheme with $(\kappa, \kappa') = (2, 7)$. Interestingly, each row gives indistinguishable results, and indicates that $E_n = O(n^{-1})$, a rate much faster than $O(n^{-(2\alpha+1)})$. Recall that the Hybrid scheme with $\kappa \geq 6$ is actually not feasible due to the almost degeneracy of the matrix Σ (see Section 3.4), and that the Hybrid and the 3R Hybrid schemes have almost the same computational complexity for the same value of κ. Consequently, the 3R Hybrid scheme has a certain advantage to the Hybrid scheme also in this context of the Monte Carlo pricing.

5.4. *On the choice of (κ, κ') for the 3R Hybrid scheme*

It is not possible to specify (κ, κ') that is optimal uniformly in g, n or other parameters. First, we note that the larger κ' results in the smaller mean squared error as shown in Theorem 4.1. There are however three points to be cared for the use of large κ': the first is the computational time of the matrix Σ defined in Section 4.2 that is of size $\kappa' + 1$ and requires evaluations of the hypergeometric function. This would be usually negligible for $\kappa' \leq 100$. The second is that the asymptotic theory relies

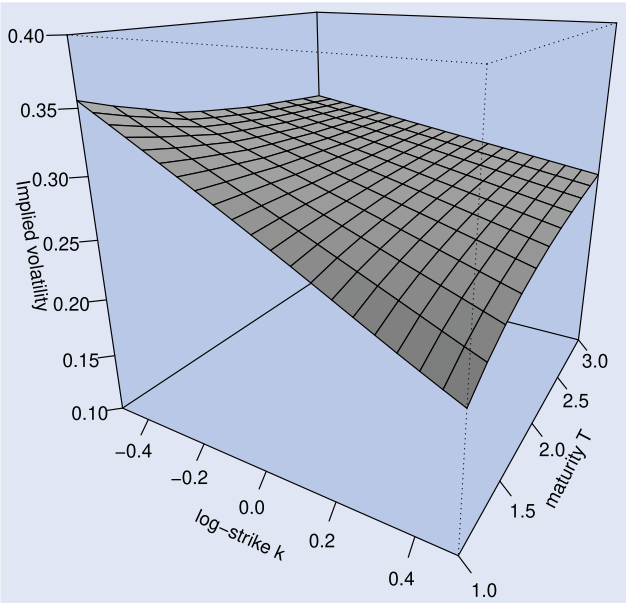

Figure 24. The fOU kernel ($\lambda = 3$).

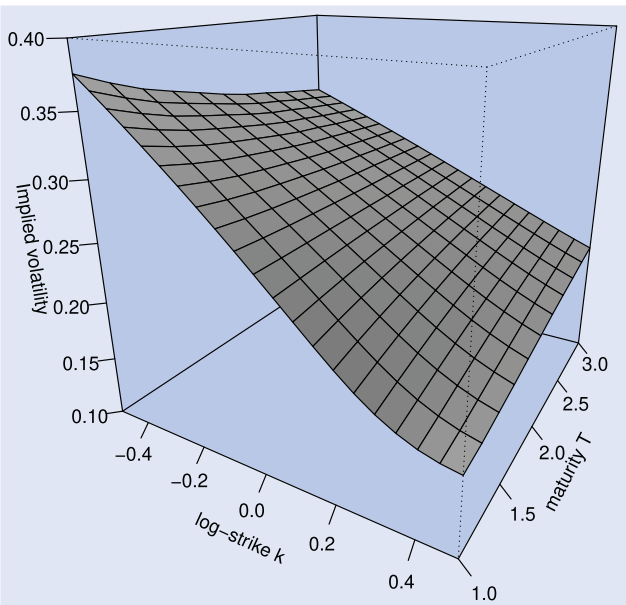

Figure 25. The gamma kernel ($\lambda = 1$).

on that $L_g(x) \approx 1$ for $x \in (0, \kappa'/n)$ and so, we need $\kappa' << n$. The rough Bergomi kernel is however an exception because we have $L_g \equiv 1$. The third is the computation time for the second term of (14). For large κ', although the FFT convolution is applicable, this term results in a nonnegligible difference from the Hybrid scheme in terms of computation time. Note that for small κ' such as $\kappa' = 10$, the FFT convolution would be better avoided for this term.

As seen in Section 4.4, E_k are rather small, so the cheapest choice of $\kappa = 2$ would be enough for the Monte Carlo pricing. Indeed we have already seen in Figure 13 that the choice $\kappa = 2$ was enough to achieve the same error tolerance level as the exact method.

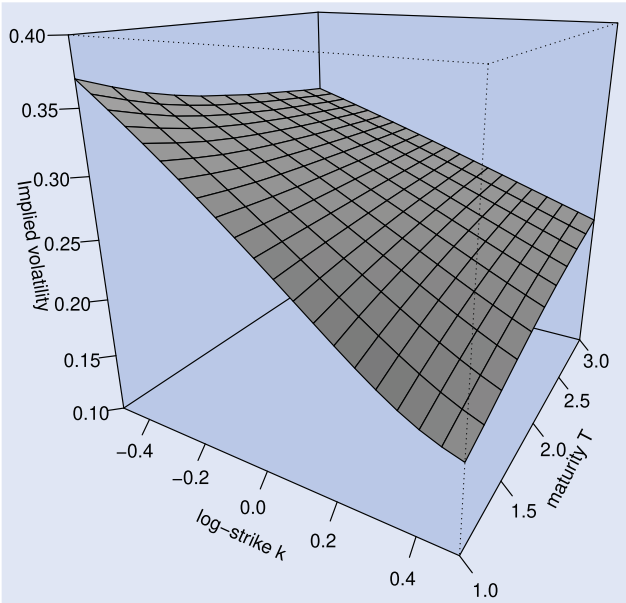

Figure 26. The gamma kernel ($\lambda = 2$).

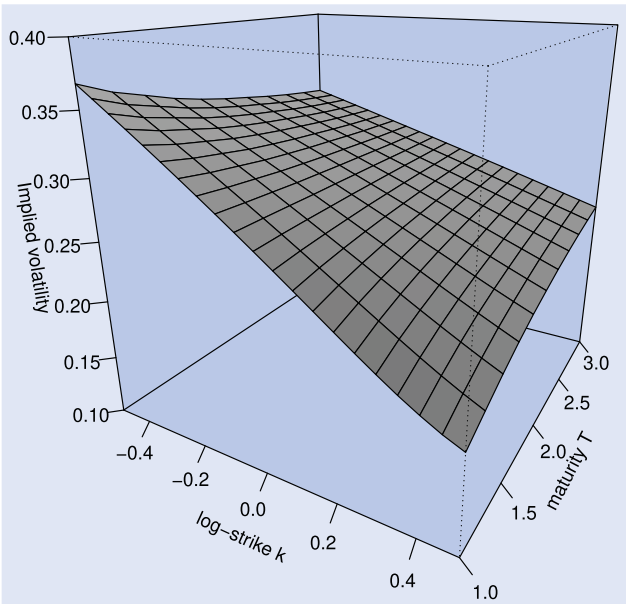

Figure 27. The gamma kernel ($\lambda = 3$).

5.5. The volatility surface

Not only the rough Bergomi kernel $g(x) = \eta x^\alpha$, we can deal with a more general parametric family as the kernel function g which potentially improves fit to the option market prices. Our aim here is to figure out the sensitivity of the implied volatility (IV) surface with respect to the parameters. We consider the following four type kernel functions.

(1) rough Bergomi kernel

$$g(x) = \eta x^\alpha,$$

(2) fOU kernel

$$g(x) = \eta x^\alpha - \eta\lambda e^{-\lambda x} \int_0^x s^\alpha e^{\lambda s}\, \mathrm{d}s, \quad \lambda > 0,$$

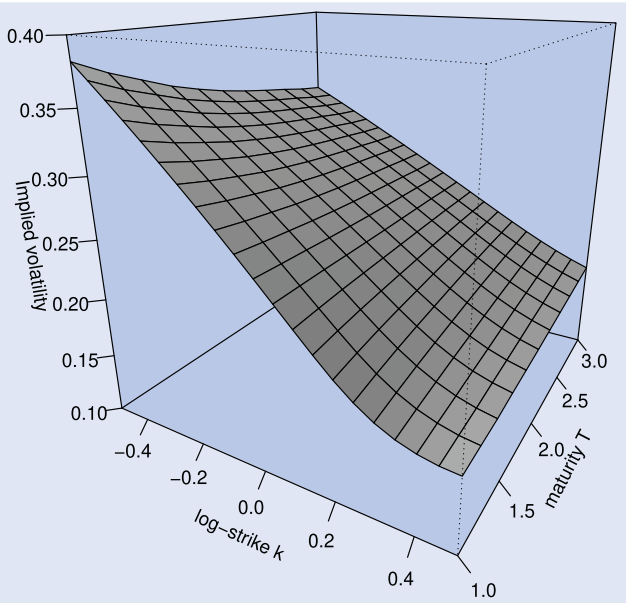

Figure 28. The power-law kernel ($\beta = -1$).

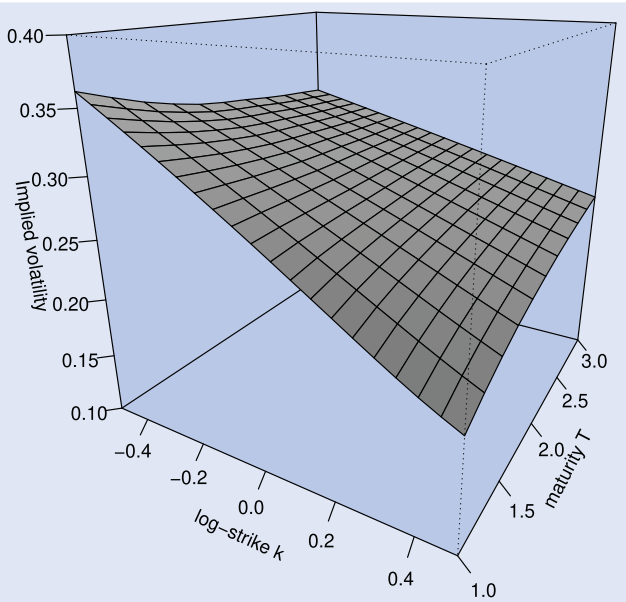

Figure 29. The power-law kernel ($\beta = -5$).

(3) gamma kernel

$$g(x) = \eta x^{\alpha} \exp\{-\lambda x\}, \quad \lambda > 0,$$

(4) power-law kernel

$$g(x) = \eta x^{\alpha} (1+x)^{\beta-\alpha}, \quad \beta < -\frac{1}{2}.$$

First, Figures 16 and 17 illustrate how η and α affect the surface. These surfaces are by the Monte Carlo simulations with the 3R Hybrid scheme under the rough Bergomi kernel. The smaller α, the stronger skew and curvature are observed, and this tendency remains for longer maturities with larger η. The other kernel functions exhibit the same dependence with respect to η and α.

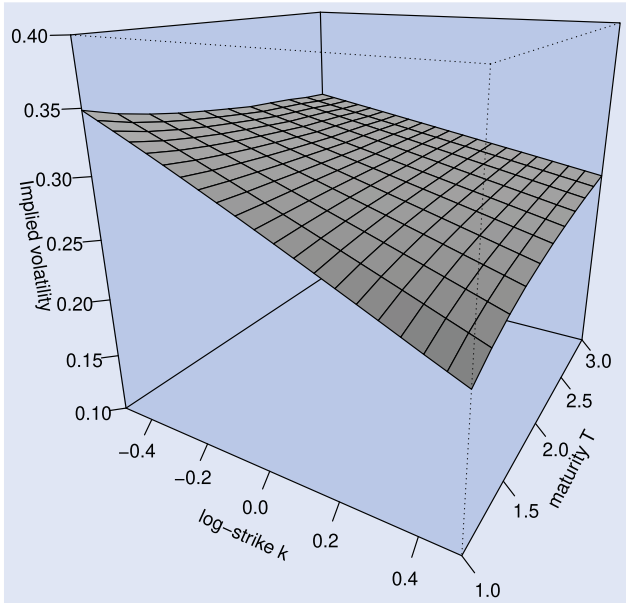

Figure 30. The power-law kernel ($\beta = -10$).

The parameter λ of the fOU kernel controls the strength of mean-reversion (see Appendix 1 for the detail). We therefore expect an averaging effect of volatility with large λ, which will result in a flat IV smile for sufficiently large maturities. Comparing Figures 18 and 19, we indeed observe a flattening effect of λ for IV surfaces. The parameters λ and β in the gamma and power-law kernels respectively are not directly related to mean reversion, but control the decay of $g(x)$ as $x \to \infty$, and so, are related to long range dependence of tBSS processes. As we see in Figures 18 $\sim$21, the flattening effects are similar for the fOU, gamma and power-law kernels. Here, the parameters for these figures are $\alpha = -0.4$, $\eta = 0.8$, $\xi \equiv 0.3^2$, $S_0 = 1$, $\rho = -0.9$. See Figures 22$\sim$30 for the IV surfaces with longer maturities. From these figures we conclude that, at least for maturities less than 3 years, the fOU, gamma and power-law kernels produce similar IV surfaces.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alòs, E., León, J.A. and Vives, J., On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance Stoch.*, 2007, **11**, 571–589.

Barndorff-Nielsen, O.E. and Basse-O'Connor, A., Quasi Ornstein-Uhlenbeck processes. *Bernoulli*, 2011, **17**(3), 916–941.

Barndorff-Nielsen, O.E. and Schmiegel, J., Brownian semistationary processes and volatility/intermittency. *Radon Series Comp Appl Math*, 2009, **8**, 1–26.

Bayer, C., Friz, P. and Gatheral, J., Pricing under rough volatility. *Quant. Finance*, 2016, **16**(6), 887–904.

Bennedsen, M., Lunde, A. and Pakkanen, M.S., Hybrid scheme for Brownian semistationary processes. *Finance Stoch.*, 2007, **21**, 931–965.

Benth, F.E., Eyjolfsson, H. and Veraart, A.E.D., Approximating Lévy semistationary processes via Fourier methods in the context of power markets. *SIAM Financ. Math.*, 2014, **5**, 71–98.

Bingham, N.H., Goldie, C.M. and Teugels, J.L., *Regular Variation*, 1989 (Cambridge University Press: Cambridge).

Cheridito, P., Kawaguchi, H. and Maejima, M., Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.*, 2003, **8**(3), 1–14.

El Euch, O., Fukasawa, M., Gatheral, J. and Rosenbaum, M., Short-term at-the-money asymptotics under stochastic volatility models. *SIAM J. Finan. Math.*, 2019, **10**, 491–511.

Forde, M. and Zhang, H., Asymptotics for rough stochastic volatility models. *SIAM J. Financial Math.*, 2017, **8**, 114–145.

Fukasawa, M., Asymptotic analysis for stochastic volatility: Martingale expansion. *Finance and Stoch.*, 2011, **15**, 635–654.

Fukasawa, M., Short-time at-the-money skew and rough fractional volatility. *Quant. Finance*, 2017, **17**(2), 189–198.

Fukasawa, M., Volatility has to be rough. *Quant. Finance*, 2021, **21**(1), 1–8.

Fukasawa, M., Takabatake, T. and Westphal, R., Is volatility rough? 2019. Available at arXiv:1905.04852.

Garnier, J. and Solna, K., Correction to black-scholes formula due to fractional stochastic volatility. *SIAM J. Financial Math.*, 2017, **8**, 560–588.

Gatheral, J., Jaisson, T. and Rosenbaum, M., Volatility is rough. *Quant. Finance*, 2018, **18**(6), 933–949.

Horvath, B., Jacquier, A. and Tankov, P., Volatility options in rough volatility models. *SIAM J. Finan. Math.*, 2020, **11**(2), 437–469.

Jacquier, A., Martini, C. and Muguruza, A., On vix future in the rough Bergomi model. *Quant. Finance*, 2018, **18**, 45–61.

McCrickerd, R. and Pakkanen, M., Turbocharging Monte Carlo pricing for the rough Bergomi model. *Quant. Finance*, 2018, **18**, 1877–1886.

Mishura, Y., *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, 2008 (Springer: Berlin).

Romano, M. and Touzi, N., Contingent claims and market completeness in a stochastic volatility model. *Math. Finance*, 1997, **7**, 399–410.

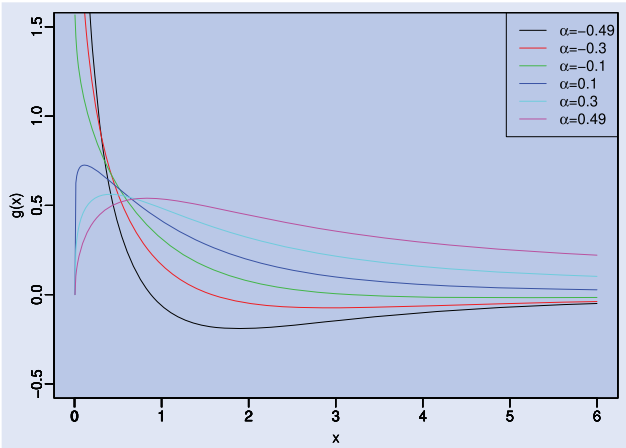

Figure A1. $\hat{g}$ ($\lambda = 1$).

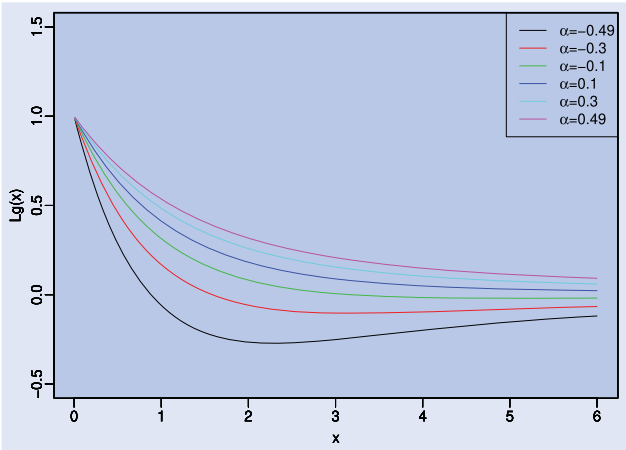

Figure A2. $\hat{L}_g$ ($\lambda = 1$).

Appendices

Appendix 1. Fractional Ornstein-Uhlenbeck process

Here we review the fOU process and show it is a BSS process satisfying the assumptions (A1) $\sim$ (A3) in Section 2.2.

A.1. A construction of fOU processes

PROPOSITION A.2 *Let $(B_t^H)_{t\in\mathbb{R}}$ be a fractional Brownian motion with the Hurst parameter $H \in (0,1)$ and $\xi \in L^0(\Omega)$. Let $-\infty \le a < \infty$ and $\lambda, \sigma > 0$. Then for almost all $\omega \in \Omega$, we have the following:*

(1) *For all $t > a$,*

$$\int_a^t e^{\lambda u}\, dB_u^H(\omega)$$

exists as a Riemann-Stieltjes integral and is equal to

$$\int_a^t e^{\lambda u}\, dB_u^H(\omega) = e^{\lambda t} B_t^H(\omega) - e^{\lambda a} B_a^H(\omega) - \lambda \int_a^t B_u^H(\omega) e^{\lambda u}\, du.$$

(2) *The function*

$$\int_a^t e^{\lambda u}\, dB_u^H(\omega), \quad t > a$$

is continuous in t.

(3) *The unique continuous function $y(t)$ that solves the equation,*

$$y(t) = \xi(\omega) - \lambda \int_0^t y(s)\, ds + \sigma B_t^H(\omega), \quad t \ge 0 \tag{A1}$$

is given by

$$y(t) = e^{-\lambda t}\left\{\xi(\omega) + \sigma \int_0^t e^{\lambda u}\, dB_u^H(\omega)\right\}, \quad t \ge 0.$$

In particular, when $\xi(\omega) = \sigma \int_{-\infty}^0 e^{\lambda u} dB_u^H(\omega)$, the unique continuous solution of the equation (A1) is given by

$$y(t) = \sigma \int_{-\infty}^t e^{-\lambda(t-u)}\, dB_u^H(\omega), \quad t \ge 0.$$

See Cheridito *et al.* (2003) for the proof.

Let $\lambda, \sigma > 0$ and initial condition $\xi \in L^0(\Omega)$. Since the Langevin equation,

$$X_t = \xi - \lambda \int_0^t X_s\, \mathrm{d}s + N_t, \quad t \ge 0 \tag{A2}$$

only involves an integral with respect to t, it can be solved path-wise for much more general noise process $(N_t)_{t\le 0}$ than Brownian motion.

Now, we consider that $N_t = \sigma B_t^H(\omega)$. It follows from Proposition A.1 that

$$X_t^{H,\xi} := e^{-\lambda t}\left(\xi + \sigma \int_0^t e^{\lambda u}\, \mathrm{d}B_u^H\right), \quad t \ge 0 \tag{A3}$$

is the unique almost surely continuous process that solves the equation (A2). In particular, the almost surely continuous process

$$X_t^H := \sigma \int_{-\infty}^t e^{-\lambda(t-u)}\,\mathrm{d}B_u^H, \quad t \in \mathbb{R} \tag{A4}$$

solves the equation (A2) with initial condition $\xi = \sigma \int_{-\infty}^0 e^{\lambda u}\,\mathrm{d}B_u^H(\omega)$.

$(X_t^H)_{t\in\mathbb{R}}$ is Gaussian process and it follows from the stationarity of the increments of fBm that it is stationary. Furthermore, as in the Brownian motion case, for every $\xi \in L^0(\Omega)$,

$$X_t^H - X_t^{H,\xi} = e^{-\lambda t}(Y_0^H - \xi) \to 0, \quad t \to \infty \quad \text{a.s.}$$

which implies that every stationary solution of (A2) has the same distribution as $(X_t^H)_{t\geq 0}$ Cheridito *et al.* (2003).

Definition A.3 (fOU process) *We call* $(X_t^{H,\xi})_{t\geq 0}$ *a fractional Ornstein-Uhlenbeck process with initial condition* ξ *and* $(X_t^H)_{t\in\mathbb{R}}$ *a stationary fractional Ornstein-Uhlenbeck process.*

It follows from Proposition A.1 that

$$\begin{aligned} X_t^H &= \sigma \lim_{a\to-\infty} \int_a^t e^{-\lambda(t-u)}\,\mathrm{d}B_u^H, \quad t \in \mathbb{R} \\ &= \sigma e^{-\lambda t} \lim_{a\to-\infty} \left\{ e^{\lambda t} B_t^H(\omega) - e^{\lambda a} B_a^H(\omega) - \lambda \int_a^t B_u^H(\omega) e^{\lambda u}\,\mathrm{d}u \right\} \end{aligned}$$

and

$$\lim_{a\to-\infty} |e^{\lambda a} B_a^H(\omega)| \leq \lim_{a\to-\infty} \frac{|B_a^H(\omega)|}{|a|^\gamma} = 0, \quad \text{for all } \gamma > H.$$

A.2. Wiener integral representation

A representation of the fOU process as a BSS process has been given in Barndorff-Nielsen and Basse-O'Connor (2011), Garnier and Solna (2017). Here we derive it in a self-contained manner and show that the kernel function satisfies all the assumptions for the Hybrid and 3R Hybrid schemes to be applicable. Using the Mandelbrot-van Ness representation of the fractional Brownian motion (Mishura 2008, Theorem 1.3.1), the stochastic Fubini theorem implies that

$$\begin{aligned} X_t^H &= \sigma B_t^H(\omega) - \sigma\lambda \int_{-\infty}^t e^{-\lambda(t-u)} B_u^H(\omega)\,\mathrm{d}u \\ &= \sigma C_H \left\{ \int_{\mathbb{R}} (t-x)_+^\alpha - (-x)_+^\alpha \mathrm{d}W_x - \lambda \int_{-\infty}^t e^{-\lambda(t-u)} \right. \\ &\quad \times \left. \left(\int_{\mathbb{R}} (u-x)_+^\alpha - (-x)_+^\alpha \mathrm{d}W_x \right) \mathrm{d}u \right\} \\ &= \sigma C_H \int_{\mathbb{R}} \left\{ (t-x)_+^\alpha - (-x)_+^\alpha \right. \\ &\quad \left. - \lambda \int_{-\infty}^t e^{-\lambda(t-u)} \left((u-x)_+^\alpha - (-x)_+^\alpha \right) \mathrm{d}u \right\} \mathrm{d}W_x \\ &= \sigma C_H \int_{-\infty}^t \left\{ (t-x)^\alpha - \lambda \int_x^t e^{-\lambda(t-u)} (u-x)^\alpha\,\mathrm{d}u \right\} \mathrm{d}W_x \\ &= \sigma C_H \int_{-\infty}^t \left\{ (t-x)^\alpha - \lambda e^{-\lambda(t-x)} \int_0^{t-x} e^{\lambda y} y^\alpha\,\mathrm{d}y \right\} \mathrm{d}W_x \\ &\because \text{replace } (u-x) \text{ for } y, \end{aligned}$$

where

$$\begin{aligned} C_H &= \left(\int_{\mathbb{R}_+} ((1+s)^\alpha - s^\alpha)^2\,\mathrm{d}s + \frac{1}{2H} \right)^{-\frac{1}{2}} \\ &= \frac{(2H\sin(\pi H)\Gamma(2H))^{1/2}}{\Gamma(H+1/2)}. \end{aligned}$$

Consequently, the stationary fOU process X_t^H be expressed as (A5):

$$\begin{aligned} X_t^H &= \sigma C_H \int_{-\infty}^t \hat{g}(t-s)\,\mathrm{d}W_s \quad \text{where} \\ \hat{g}(x) &= x^\alpha - \lambda e^{-\lambda x} \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s. \end{aligned} \tag{A5}$$

Proposition A.4 *The fOU kernel* $g = \hat{g}$ *satisfies the assumptions* (A1) ∼ (A3).

Proof Let $\alpha = H - \frac{1}{2} \in (-\frac{1}{2}, \frac{1}{2})$. First, we check that $\hat{g}$ satisfies (A1).

$$\begin{aligned} \frac{\hat{g}(tx)}{\hat{g}(x)} &= \frac{(tx)^\alpha - \lambda e^{-\lambda tx} \int_0^{tx} s^\alpha e^{\lambda s}\,\mathrm{d}s}{x^\alpha - \lambda e^{-\lambda x} \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s} \\ &= \frac{t^\alpha - \lambda e^{-\lambda tx} \int_0^{tx} \left(\frac{s}{x}\right)^\alpha e^{\lambda s}\,\mathrm{d}s}{1 - \lambda e^{-\lambda x} \int_0^x \left(\frac{s}{x}\right)^\alpha e^{\lambda s}\,\mathrm{d}s}, \quad x \in (0,1], \quad t \in \mathbb{R} \end{aligned}$$

Changing variables, we have

$$\int_0^{tx} \left(\frac{s}{x}\right)^\alpha e^{\lambda s}\,\mathrm{d}s = \int_0^t u^\alpha e^{\lambda x u} x\,\mathrm{d}u \to 0$$

as $x \to 0$, which means

$$\lim_{x\to 0} \frac{\hat{g}(tx)}{\hat{g}(x)} = t^\alpha$$

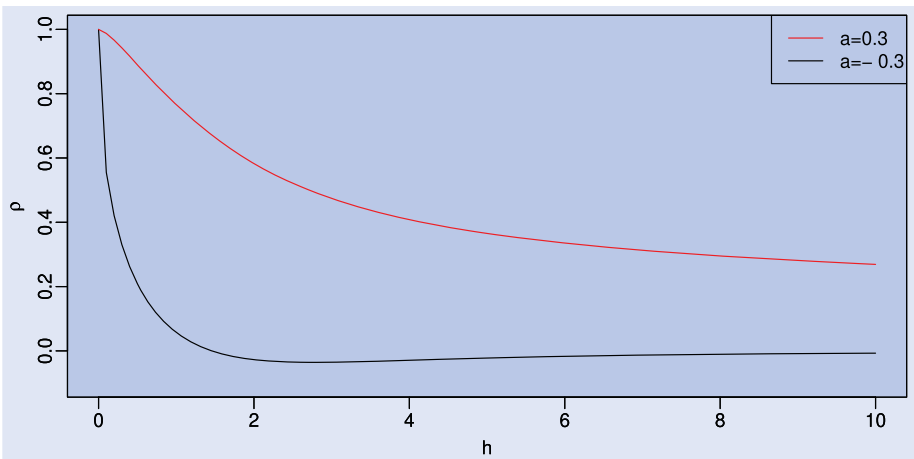

Figure A3. The auto-correlation coefficient of fOU ($\sigma C_H = 1,\ \lambda = 1$).

so $\hat{g}(x)$ is a regular varying function and

$$\hat{L}_g(x) := 1 - \lambda e^{-\lambda x} \int_0^x \left(\frac{s}{x}\right)^\alpha e^{\lambda s}\,\mathrm{d}s$$

is a slowly varying function satisfying $\lim_{x\to 0} \hat{L}_g(x) = 1$. In addition,

$$\begin{aligned}\hat{L}_g'(x) &= \left(1 - \lambda e^{-\lambda x} x^{-\alpha} \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s\right)' \\ &= -\left(-\lambda^2 e^{-\lambda x} x^{-\alpha} - \alpha\lambda e^{-\lambda x} x^{-\alpha-1}\right) \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s \\ &\quad - \lambda e^{-\lambda x} x^{-\alpha} \left(x^\alpha e^{\lambda x}\right)\end{aligned}$$

so the derivative $\hat{L}_g'$ satisfies

$$\begin{aligned}\left|\hat{L}_g'(x)\right| &\le \lambda e^{-\lambda x} x^{-\alpha} |\lambda + \alpha x^{-1}| \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s + \lambda \\ &\le \lambda x^{-\alpha} |\lambda + \alpha x^{-1}| \frac{1}{\alpha+1} x^{\alpha+1} + \lambda \\ &\le 2\lambda x \left|\lambda + \frac{1}{2} x^{-1}\right| + \lambda \quad \because \alpha \in \left(-\frac{1}{2}, \frac{1}{2}\right) \\ &= 2\lambda(1 + \lambda x).\end{aligned}$$

Second, for (A2), it is clear that $\hat{g}$ is continuously differentiable on $(0, \infty)$ and its derivative is

$$\begin{aligned}\hat{g}'(x) &= \alpha x^{\alpha-1} + \lambda^2 e^{-\lambda x} \int_0^x s^\alpha e^{\lambda s}\,\mathrm{d}s - \lambda x^\alpha \\ &= \alpha x^{\alpha-1} - \lambda \hat{g}(x). \qquad \text{(A6)}\end{aligned}$$

By changing variable $\frac{s}{x} = 1 - u$;

$$\begin{aligned}\hat{g}(x) &= x^\alpha \left\{1 - \lambda \int_0^x \left(\frac{s}{x}\right)^\alpha e^{\lambda(s-x)}\,\mathrm{d}s\right\} \\ &= x^\alpha \left\{1 - \lambda \int_0^1 (1-u)^\alpha e^{-\lambda x u} x\,\mathrm{d}u\right\} \\ &= x^\alpha \left\{\lambda x \int_0^\infty e^{-\lambda x u}\,\mathrm{d}u - \lambda x \int_0^1 (1-u)^\alpha e^{-\lambda x u}\,\mathrm{d}u\right\} \\ &= \underbrace{\lambda x^{\alpha+1} \int_0^1 (1 - (1-u)^\alpha) e^{-\lambda x u}\,\mathrm{d}u}_{(*)} + \underbrace{\lambda x^{\alpha+1} \int_1^\infty e^{-\lambda x u}\,\mathrm{d}u}_{(**)}.\end{aligned}$$

For $(**)$, we have

$$\lambda x^{\alpha+1} \int_1^\infty e^{-\lambda x u}\,\mathrm{d}u = x^\alpha e^{-\lambda x} = o(x^{\alpha-n})$$

as $x \to \infty$ for any $n \in \mathbb{N}$. For $(*)$, for any $n \in \mathbb{N}$,

$$\begin{aligned}&\left|\lambda x^{\alpha+1} \int_{1/2}^1 (1 - (1-u)^\alpha) e^{-\lambda x u}\,\mathrm{d}u\right| \\ &\quad \le \lambda x^{\alpha+1} e^{-\lambda x/2} \int_{1/2}^1 |1 - (1-u)^\alpha|\,\mathrm{d}u = o(x^{\alpha-n})\end{aligned}$$

and

$$\begin{aligned}&x \int_0^{1/2} (1 - (1-u)^\alpha) e^{-\lambda x u}\,\mathrm{d}u \\ &\quad = \int_0^{x/2} \left(1 - \left(1 - \frac{y}{x}\right)^\alpha\right) e^{-\lambda y}\,\mathrm{d}y \\ &\quad = -\sum_{k=1}^{n-1} (-1)^k \binom{\alpha}{k} \int_0^{x/2} \left(\frac{y}{x}\right)^k e^{-\lambda y}\,\mathrm{d}y \\ &\qquad - (-1)^n \binom{\alpha}{n} n \left(\frac{y}{x}\right)^n \int_0^{x/2} \\ &\qquad \times \int_0^1 \left(1 - \frac{y}{x} s\right)^{\alpha-n} (1-s)^{n-1}\,\mathrm{d}s e^{-\lambda y}\,\mathrm{d}y \\ &\quad = \sum_{k=1}^n (-1)^{k+1} \binom{\alpha}{k} \frac{k!}{\lambda^{k+1} x^k} + o(x^{-n}).\end{aligned}$$

as $x \to \infty$ by the dominated convergence theorem:

$$1_{[0,x/2]}(y) \left|1 - \frac{y}{x} s\right|^{\alpha-n} \le \left|1 - \frac{s}{2}\right|^{\alpha-n}.$$

Here, we have used the Taylor formula

$$(1+v)^\alpha = \sum_{k=0}^{n-1} \binom{\alpha}{k} v^k + \binom{\alpha}{n} n v^n \int_0^1 (1+vs)^{\alpha-n} (1-s)^{n-1}\,\mathrm{d}s.$$

Consequently, as $x \to \infty$,

$$\begin{aligned}\hat{g}(x) &= \frac{\alpha}{\lambda} x^{\alpha-1} - \frac{\alpha(\alpha-1)}{\lambda^2} x^{\alpha-2} \\ &\quad + \frac{\alpha(\alpha-1)(\alpha-2)}{\lambda^3} x^{\alpha-3} + o(x^{\alpha-3})\end{aligned}$$

and so, by (A6),

$$\begin{aligned}\hat{g}'(x) &= \alpha x^{\alpha-1} - \lambda \hat{g}(x) = \frac{\alpha(\alpha-1)}{\lambda} x^{\alpha-2} \\ &\quad - \frac{\alpha(\alpha-1)(\alpha-2)}{\lambda^2} x^{\alpha-3} + o(x^{\alpha-3}).\end{aligned}$$

This implies that $\hat{g}'$ is square-integrable on $(1, \infty)$:

$$\int_1^\infty \hat{g}'(x)^2\,\mathrm{d}x < \infty.$$

Also,

$$\begin{aligned}\hat{g}''(x) &= \alpha(\alpha-1) x^{\alpha-2} - \lambda \hat{g}'(x) \\ &= \frac{\alpha(\alpha-1)(\alpha-2)}{\lambda} x^{\alpha-3} + o(x^{\alpha-3}),\end{aligned}$$

which implies that $\hat{g}''(x) \neq 0$ and continuous for enough large x. Therefore $\hat{g}'(x)$ is ultimately monotonic.

Finally, for (A3), we have already seen

$$\hat{g}(x) = \mathcal{O}(x^{\alpha-1}), \quad x \to \infty.$$

■

A.3. Dependence structure

First we give some graphs of the fOU kernel that will help the readers to understand the nature of our approximation.

Second we give the graph of the auto-correlation coefficient defined by

$$\rho(h) = \frac{\mathbb{E}[X_t X_{t+h}]}{\sqrt{\mathbb{V}[X_t] V[X_{t+h}]}}$$

for $h > 0$ in order to illustrate the dependence structure of the fOU process. A fOU process X_t has the property of short-range dependence for $\alpha \in (-\frac{1}{2}, 0)$, and the property of long-range dependence for $\alpha \in (0, \frac{1}{2})$; see Cheridito *et al.* (2003) for the detail.Appendix 2.

Table A1. Computational costs.

	Hybrid scheme		3R Hybrid scheme	
Number of steps	Algorithm	Time(sec)	Algorithm	Time(sec)
1st step	Calculating $\left\{L_g\left(\frac{k}{n}\right)\right\}_{k=1,\ldots,\kappa}$ and $\left\{g\left(\frac{c_k}{n}\right)\right\}_{k=\kappa+1,\ldots,N_n}$	3.74	Calculating $\left\{L_g\left(\frac{k}{n}\right)\right\}_{k=1,\ldots,\kappa'}$ and $\left\{g\left(\frac{c_k}{n}\right)\right\}_{k=\kappa'+1,\ldots,N_n}$	3.74
2nd step	Calculating Σ	0.03	Calculating Σ	0.03
3rd step			Calculating $\{a_k\}_{k=\kappa+1,\ldots,\kappa'}, \{b_k\}_{k=\kappa+1,\ldots,\kappa'}$	0.04
4th step	Generating $(N_n + \lfloor nT \rfloor)$ pieces of random numbers following $N(0, \Sigma)$*1	134.99*	Generating $(N_n + \lfloor nT \rfloor)$ pieces of random numbers following $N(0, \Sigma)$*1	134.99*
5th step	Convolution and addition*2	87.86*	Convolution and addition*3	95.78*
Total time	226.62(sec)		234.58(sec)	
PC spec	Processor : Intel(R)Core(TM)i5-8250U CPU @ 1.60GHz 1.80GHz RAM : 8.00GB			

Note: (∗) The underlined parts depend on m.
(∗1) The number of random numbers is $(\kappa + 1) \times (N_n + \lfloor nT \rfloor)$ because the generated random number vectors are $\kappa + 1$ dimension.
(∗2) Refer to (8).
(∗3) Refer to (14).

Appendix 2. Computational costs: an experiment

Here we report a numerical experiment to examine computational costs of the Hybrid and the 3R Hybrid schemes. The grid is set to be $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{\lfloor nT \rfloor}{n}\}$ and, the parameters are set as

$$n = 8,192, \quad \alpha = -0.4, \quad \kappa = 2,$$
$$T = 1, \quad \gamma = 0.2, \quad \sigma = 1, \quad \kappa' = 10,$$

and g is a fOU kernel. We simulated $m = 10,000$ sample paths by each of the schemes, for which the computation time is reported in Table A1. The programming language was R and and the computation times ('user time') for modules were measured by the R function system.time.

The first sum of (8) and that of (15), the second sum of (8) and the third sum of (15) are respectively equal in computation time. Therefore, the difference in computation time between the Hybrid and the 3R Hybrid schemes is only due to the second sum of (15). In this experiment with κ' of modest size, this second sum, as well as the first one, was computed without the FFT. Only the third term of (15), as well as the second term of (8), was by the FFT. In the FFT convolution, for a faster computation, we embeded vectors into larger ones whose length was of the form 2^p by the zero padding. According to Table A1, the additional computation cost for the 3R Hybrid scheme is marginal and consequently, the Hybrid scheme with $\kappa = 2$ and the 3R Hybrid scheme with $\kappa = 2$, $\kappa' = 10$ require almost the same computational costs. Recall that the 3R Hybrid scheme with $\kappa' = 10$ enjoys almost the same accuracy as the Hybrid scheme with $\kappa = 10$ that requires not only more computation time due to the increased random number generations, but also a special care for the Cholesky decomposition of Σ that is almost degenerage.