Comparing the results of Cholesky and Hybrid schemes for the rBergomi

March 17, 2019

1 Problem setting

1.1 The Rough Bergomi Model

We consider the rBergomi model for the price process S_t as defined in [3], normalized to $r = 0$[1], which is defined by

$$\begin{aligned} dS_t &= \sqrt{v_t} S_t dZ_t, \\ v_t &= \xi_0(t) \exp\left(\eta \widetilde{W}_t^H - \frac{1}{2}\eta^2 t^{2H}\right), \end{aligned} \tag{1.1}$$

where the Hurst parameter $0 < H < 1$ and $\eta > 0$. We refer to v_t as the variance process, and $\xi_0(t) = \mathrm{E}[v_t]$ is the forward variance curve. Here, $\widetilde{W}^H$ is a certain Riemann-Liouville fBm process[2], defined by

$$\widetilde{W}_t^H = \int_0^t K^H(t,s) dW_s^1, \quad t \geq 0, \tag{1.2}$$

where the kernel $K^H : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is

$$K^H(t,s) = \sqrt{2H}(t-s)^{H-1/2}, \quad \forall\, 0 \leq s \leq t. \tag{1.3}$$

By construction, $\widetilde{W}^H$ is a centered, locally $(H-\epsilon)$- Hölder continuous, Gaussian process with $\mathrm{Var}\left[\widetilde{W}_t^H\right] = t^{2H}$, and a dependence structure defined by

$$\mathrm{E}\left[\widetilde{W}_u^H \widetilde{W}_v^H\right] = u^{2H} G\left(\frac{v}{u}\right), \quad v > u,$$

where for $x \geq 1$ and $\gamma = \frac{1}{2} - H$

$$G(x) = 2H \int_0^1 \frac{1}{(1-s)^\gamma (x-s)^\gamma}. \tag{1.4}$$

We note that $\widetilde{W}$ is also a Brownian semi-stationary (BSS) (see Definition 1.1), which were introduced by Barndorff-Nielsen and Schmiegel [1, 2].

[1] r is the interest rate.

[2] The so-called Riemann-Liouville processes are deduced from the standard Brownian motion by applying Riemann-Liouville fractional operators, whereas the standard fBm requires a weighted fractional operator [9, 7, 8].

Definition 1.1 (Semi-stationary process)**.** X_t is called a *Brownian semi-stationary* process if

$$X_t = \int_{-\infty}^{t} K(t-s)\sigma_s dW_s \tag{1.5}$$

for some deterministic kernel function K and an adapted intermittency process σ. If the integral starts at 0 instead of ∞

$$X_t = \int_{0}^{t} K(t-s)\sigma_s dW_s \tag{1.6}$$

we call the process *truncated Brownian semi-stationary* process (TBSS).

In (1.1) and (1.2), W^1, Z denote two *correlated* standard Brownian motions with correlation $\rho \in]-1, 0]$, so that we can represent Z in terms of W^1 as

$$Z = \rho W^1 + \overline{\rho} W^{\perp} = \rho W^1 + \sqrt{1-\rho^2} W^{\perp},$$

where $(W^1, W^{\perp})$ are two independent standard Brownian motions. Therefore, the solution to (1.1), with $S(0) = S_0$, can be written as

$$\begin{aligned} S_t &= S_0 \exp\left(\int_0^t \sqrt{v(s)} dZ(s) - \frac{1}{2}\int_0^t v(s) ds\right), \quad S_0 > 0 \\ v_u &= \xi_0(u) \exp\left(\eta \widetilde{W}_u^H - \frac{\eta^2}{2} u^{2H}\right), \quad \xi_0 > 0. \end{aligned} \tag{1.7}$$

The filtration $(\mathcal{F}_t)_{t\geq 0}$ can here be taken as the one generated by the two-dimensional Brownian motion $(W^1, W^{\perp})$ under the risk neutral measure $\mathbb{Q}$, resulting in a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{Q})$. The stock price process S is clearly then a local $(\mathcal{F}_t)_{t\geq 0}$-martingale and a supermartingale. We shall henceforth use the notation $\mathrm{E}[.] = E^{\mathbb{Q}}[. \mid \mathcal{F}_0]$ unless we state otherwise.

Remark 1.2. The rBergomi model is non-Markovian in the instantaneous variance v_t, that is $E^{\mathbb{Q}}[v_u \mid \mathcal{F}_t] \neq= E^{\mathbb{Q}}[v_u \mid v_t]$. However, it is Markovian in the state vector by definition, that is $E^{\mathbb{Q}}[v_u \mid \mathcal{F}_t] = \xi_t(u)$.

2 Simulation of the rBergomi Model

One of the numerical challenges encountered in the simulation of the rBergomi dynamics is the computation of $\int_0^T \sqrt{v_t} dW_t^1$ and $V = \int_0^T v_t dt$ in (**??**), mainly because of the singularity of the Volterra kernel $K^H(s,t)$ at the diagonal $s = t$. In fact, one needs to jointly simulate two Gaussian processes $(W_t^1, \widetilde{W}_t^H : 0 \leq t \leq T)$, resulting in $W_{t_1}^1, \ldots, W_{t_N}^1$ and $\widetilde{W}_{t_1}^H, \ldots, \widetilde{W}_{t_N}^H$ along a given time grid $t_1 < \cdots < t_N$. In the literature, there are essentially two suggested ways to achieve this:

i) **Covariance based approach (exact simulation)** [3, 5]: Given that $W_{t_1}^1, \ldots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \ldots, \widetilde{W}_{t_N}$ together form a $(2N)$-dimensional Gaussian random vector with computable covariance matrix, A, one can use Cholesky decomposition of the covariance matrix to produce exact samples of $W_{t_1}^1, \ldots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \ldots, \widetilde{W}_{t_N}$ from $2N$-dimensional Gaussian random vector as input. This method is exact but slow. The simulation requires $\mathcal{O}\left(N^2\right)$ flops. Note that the offline cost is $\mathcal{O}\left(N^3\right)$ flops.

ii) **The hybrid scheme of [6]**: This scheme uses a different approach, which is essentially based on Euler discretization but crucially improved by moment matching for the singular term in the left point rule. It is also inexact in the sense that samples produced here do not exactly have the distribution of $W^1_{t_1}, \dots, W^1_{t_N}, \widetilde{W}^H_{t_1}, \dots, \widetilde{W}_{t_N}$, however they are much more accurate than samples produced from simple Euler discretization, but much faster than method (i). As in method (i), in this case, we need a $2N$-dimensional Gaussian random input vector to produce one sample of $W^1_{t_1}, \dots, W^1_{t_N}, \widetilde{W}^H_{t_1}, \dots, \widetilde{W}_{t_N}$.

In the following, we give more details for each scheme.

2.1 The Hybrid Scheme

The hybrid scheme was developed in [6] for the simulation of BSS (1.5) and TBSS (1.6) processes with kernel functions, that are similar to a power function near zero, i.e. $K(x)$ is similar to x^α for some $\alpha \in (-\frac{1}{2}, \frac{1}{2})$ when $x > 0$ is near zero.

In this section, we start by introducing the general idea of the hybrid scheme. Then, we illustrate how it can be used in simulating rBergomi dynamics.

2.1.1 General Idea of the Hybrid Scheme

Before describing the general hybrid scheme, there are two assumptions (conditions) that should be fulfilled by the kernel function K in order to apply the scheme. In fact, we assume that the kernel function $K : (0, \infty) \to [0, \infty)$ satisfy the following

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

$$K(x) = x^\alpha L_K(x), \quad x \in (0, 1],$$

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

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

2. The function K is continuously differentiable on $(0, \infty)$, with derivative K' that is ultimately monotonic and also satisfies $\int_1^\infty K'(x)^2 dx < \infty$.

The hybrid scheme starts by discretizing a TBSS process X_t giving by (1.6) on the grid $G^n_t = \{t, t - \frac{1}{n}, t - \frac{2}{n}, \dots\}$ for $n \in \mathbb{N}$, so that we get

$$X_t \approx \sum_{k=1}^{N_n} \sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} K(t-s) dW_s, \tag{2.1}$$

where $N_n \in \mathbb{N}$ is the truncation parameter and denotes the number of discretization steps, and we assume that σ doesnt vary too much (valid in the rBergomi model since σ is constant), therefore constant in each discretization cell.

The idea of the hybrid scheme is to split the kernel function into two parts 2.2, depending on a small parameter κ[3]

$$K(t-s) \approx \begin{cases} (t-s)^{\alpha} L_K\left(\frac{k}{n}\right), & t-s \in \left[\frac{k-1}{n}, \frac{k}{n}\right] \setminus \{0\} \quad , \quad \text{if } k \leq \kappa \\ K\left(\frac{b_k}{n}\right), & t-s \in \left[\frac{k-1}{n}, \frac{k}{n}\right] \quad , \quad \text{if } k > \kappa, \end{cases} \tag{2.2}$$

and then discretizing the X_t process into Wiener integrals of power functions and a Riemann sum, appearing from approximating the kernel by power functions near the origin and step functions elsewhere, resulting in 2.3

$$X_t \approx \sum_{k=1}^{\kappa} L_K\left(\frac{k}{n}\right) \sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} (t-s)^{\alpha} dW_s + \sum_{k=\kappa+1}^{N_n} L_K\left(\frac{k}{n}\right) K\left(\frac{b_k}{n}\right) \sigma_{t-\frac{k}{n}} \int_{t-\frac{k}{n}}^{t-\frac{k}{n}+\frac{1}{n}} dW_s. \tag{2.3}$$

Remark 2.1. $(b_k)_{k=\kappa+1}^{\infty}$ is a sequence of real numbers that only needs to satisfy $b_k \in [k-1, k] \setminus \{0\}$. But there is an optimal sequence $(b_k^*)_{k=\kappa+1}^{\infty}$, that minimizes the asymptotic MSE (see [6] for more details) and it is given by

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

2.1.2 The hybrid Scheme in the rBergomi Context

In the context of the rBergomi model, we have $\sigma = 1$, $\alpha = H - \frac{1}{2}$ and we can use the hybrid scheme to simulate $\widetilde{W}_t^H$, which is a TBSS process, with the kernel function K^H given by (1.3). In this context, we can show that by choosing $L_{K^H}(x) = 1$, we have K^H satisfy the conditions required to apply the hybrid scheme (see [6]) and therefore we end up with the following scheme on equidistant grid $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{nT}{n}\}$

$$\widetilde{W}_{\frac{i}{n}}^H \approx \overline{W}_{\frac{i}{n}}^H = \sqrt{2H} \left(\sum_{k=1}^{\min(i,\kappa)} \int_{\frac{i}{n}-\frac{k}{n}}^{\frac{i}{n}-\frac{k}{n}+\frac{1}{n}} \left(\frac{i}{n} - s\right)^{\alpha} dW_s + \sum_{k=\kappa+1}^{i} \left(\frac{b_k}{n}\right)^{\alpha} \int_{\frac{i}{n}-\frac{k}{n}}^{\frac{i}{n}-\frac{k}{n}+\frac{1}{n}} dW_s \right), \tag{2.4}$$

which results for $\kappa = 1$ in (2.5).

$$\widetilde{W}_{\frac{i}{N}}^H \approx \overline{W}_{\frac{i}{N}}^H = \sqrt{2H} \left(W_i^2 + \sum_{k=2}^{i} \left(\frac{b_k}{N}\right)^{H-\frac{1}{2}} \left(W^1_{\frac{i-(k-1)}{N}} - W^1_{\frac{i-k}{N}} \right) \right), \tag{2.5}$$

where N is the number of time steps and

$$b_k = \left(\frac{k^{H+\frac{1}{2}} - (k-1)^{H+\frac{1}{2}}}{H+\frac{1}{2}} \right)^{\frac{1}{H-\frac{1}{2}}}.$$

The sum in (2.5) requires the most computational effort in the simulation. Given that (2.5) can be seen as discrete convolution (see [6]), we employ the fast Fourier transform to evaluate it, which results in $\mathcal{O}(N \log N)$ floating point operations.

We note that the variates $\overline{W}_0^H, \overline{W}_1^H, \ldots, \overline{W}_{\frac{[Nt]}{N}}^H$ are generated by sampling $[Nt]$ i.i.d draws from a $(\kappa+1)$-dimensional Gaussian distribution and computing a discrete convolution. We denote these pairs of Gaussian random variables from now on by $(\mathbf{W}^{(1)}, \mathbf{W}^{(2)})$.

[3] There are different variants of the hybrid scheme depending on the value of κ (see [6] for more details).

2.2 The Exact Scheme Coupled with Hierarchical Representation

The exact scheme is based on simulating exactly the joint covariance of $(\widetilde{W}^H, W^1)$, which is given by

$$\begin{cases} \mathrm{Var}(\widetilde{W}_t) = t^{2H} & , \quad t \geq 0 \\ \mathrm{Cov}(\widetilde{W}_t, \widetilde{W}_s) = t^{2H} G\left(\frac{s}{t}\right) & , \quad s > t \geq 0 \\ \mathrm{Cov}(\widetilde{W}_t, W_s^1) = \rho D_H \left(t^{H+\frac{1}{2}} - \left(t - \min(t,s)^{H+\frac{1}{2}}\right)\right) & , \quad s,t \geq 0 \\ \mathrm{Cov}(W_t^1, W_s^1) = \min(t,s) & , \quad s,t \geq 0, \end{cases} \tag{2.6}$$

with

$$D_H = \frac{\sqrt{2H}}{H + \frac{1}{2}}$$

and G is given by (1.4) for $x \geq 1$ and $\gamma = \frac{1}{2} - H$.

Different ways can be used to compute (1.4). For instance, in the literature we find that in [5]

$$G_2(x) = 2H \left(\frac{1}{1-\gamma} x^{-\gamma} + \frac{\gamma}{1-\gamma} x^{-(1+\gamma)} \frac{1}{2-\gamma} {}_2F_1\left(1, 1+\gamma, 3-\gamma, x^{-1}\right) \right), \tag{2.7}$$

where

$${}_2F_1(a,b,c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a} dt.$$

Remark 2.2. Because the covariance matrix does not have any 0 values, if $\rho \neq 0$, the computational effort for the exact simulation is quite high and the algorithm is very slow.

Let us denote by the matrix A, the computable covariance matrix of $\widetilde{W}_{t_1}^H, \ldots, \widetilde{W}_{t_N}, W_{t_1}^1, \ldots, W_{t_N}^1$ using (2.6). We can use Cholesky decomposition of A to produce exact samples of $W_{t_1}^1, \ldots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \ldots, \widetilde{W}_{t_N}$.

In fact let us denote by L the triangular matrix resulting from Cholesky decomposition such that

$$L = \left(\begin{array}{c|c} L_1 & 0 \\ L_2 & L_3 \end{array} \right),$$

where L_1, L_2, L_3 are $N \times N$ matrices, such that L_1 and L_3 are triangular.

Then, given a $2N \times 1$-dimensional Gaussian random input vector, $\mathbf{X} = (X_1, \ldots, X_N, X_{N+1}, \ldots, X_{2N})'$, we have

$$\mathbf{W}^{(1)} = L_1 \mathbf{X}_{1:N}, \quad \widetilde{\mathbf{W}} = \left(\begin{array}{c|c} L_2 & L_3 \end{array} \right) \mathbf{X}. \tag{2.8}$$

On the other hand, let us assume that we can construct $\mathbf{W}^{(1)}$ hierarchically through Brownian bridge construction defined by the linear mapping given by the matrix G, then given a N-dimensional Gaussian random input vector, $\mathbf{Z}'$, we can write

$$\mathbf{W}^{(1)} = G\mathbf{Z}',$$

and consequently

$$\mathbf{X}_{1:N} = L_1^{-1} G \mathbf{Z}'.$$

Therefore, given a $2N$-dimensional Gaussian random input vector, $\mathbf{Z} = (\mathbf{Z}', \mathbf{Z}'')$, we define our hierarchical representation by

$$\mathbf{X} = \left(\begin{array}{c|c} L_1^{-1} G & 0 \\ 0 & I_N \end{array} \right) \mathbf{Z}. \tag{2.9}$$

As described in Section 4 in [3], the exact scheme is given by the following. For N number of time steps and M the number of simulations,

1. Construct the joint covariance matrix for the Volterra process $\tilde{W}$ and the Brownian motion W^1 and compute its Cholesky decomposition.
2. For each time, generate iid normal random vectors and multiply them by the lower-triangular matrix obtained by the Cholesky decomposition to get a $M \times 2N$ matrix of paths of $\tilde{W}$ and W^1 with the correct joint marginals.
3. With these paths held in memory, we may evaluate the expectation under of any payoff of interest.

3 Weak error analysis

In this work, we are interested in approximating $\mathrm{E}\left[g(X_T\right]$, where g is some smooth function and X is the asset price under rBergomi dynamics such that $X_t = X_t(W_t^{(1)}, \widetilde{W}_t)$ where $W^{(1)}$ is standard Brownian motion and $\widetilde{W}_t$ is the fractional Brownian motion as given by (1.2). Then we can express the Hybrid and Cholesky scheme as the following

$$\mathrm{E}\left[g\left(X_T\left(W_t^{(1)}, \widetilde{W}_t\right)\right)\right] \approx \mathrm{E}\left[g\left(\overline{X}_N\left(\overline{W}_1^{(1)}, \ldots, \overline{W}_N^{(1)}, \overline{\overline{W}}_1, \ldots, \overline{\overline{W}}_N\right)\right)\right] : \quad \textbf{(Hybrid scheme)} \tag{3.1}$$

$$\mathrm{E}\left[g\left(X_T\left(W_t^{(1)}, \widetilde{W}_t\right)\right)\right] \approx \mathrm{E}\left[g\left(\overline{X}_N\left(W_1^{(1)}, \ldots, W_N^{(1)}, \widetilde{W}_1, \ldots, \widetilde{W}_N\right)\right)\right] : \quad \textbf{(Cholesky scheme)}, \tag{3.2}$$

To simplify notation, let $\overline{\mathbf{W}^1} = (\overline{W}_1^{(1)}, \ldots, \overline{W}_N^{(1)})$, $\overline{\overline{\mathbf{W}}} = (\overline{\overline{W}}_1, \ldots, \overline{\overline{W}}_N)$, $\mathbf{W}^1 = (W_1^{(1)}, \ldots, W_N^{(1)})$ and $\widetilde{\mathbf{W}} = (\widetilde{W}_1, \ldots, \widetilde{W}_N)$ .If we denote by ε_B^{Hyb} and ε_B^{Chol} the weak errors produced by the hybrid and Cholesky scheme respectively, then we can write

$$\begin{aligned} \varepsilon_B^{Hyb} &= \left|\mathrm{E}\left[g\left(X_T\left(W_t^{(1)}, \widetilde{W}_t\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}^1}, \overline{\overline{\mathbf{W}}}\right)\right)\right]\right| \\ &\leq \left|\mathrm{E}\left[g\left(X_T\left(W_t^{(1)}, \widetilde{W}_t\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1, \widetilde{\mathbf{W}}\right)\right)\right]\right| + \left|\mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}^1}, \overline{\overline{\mathbf{W}}}\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1, \widetilde{\mathbf{W}}\right)\right)\right]\right| \\ &\leq \varepsilon_B^{Chol} + \left|\mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}^1}, \overline{\overline{\mathbf{W}}}\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1, \widetilde{\mathbf{W}}\right)\right)\right]\right| \end{aligned} \tag{3.3}$$

From the construction of the Cholesky scheme (See Section 2.2), we expect that the weak error is purely the discretization error, that is

$$\varepsilon_B^{Chol} = \mathcal{O}\left(\Delta t\right),$$

which was confirmed by our numerical experiments (See Figure 4.2b for case Set 1 in Table 4.1).

Therefore, it remains to understand how the second term in 3.1 behaves with respect to Δt and H.From our numerical experiments it seems that that term is also of order Δt and its rate of convergence is independent of H (see Figures 4.1 and 4.4).

4 Comparing Cholesky and Hybrid schemes results

4.1 Weak error

In this section, we compare the weak rates obtained for set 1 in Table 4.1. We compare three different cases: i) rBergomi simulated using Hybrid scheme with hierarchical construction (see Figure 4.1), ii) rBergomi simulated using Cholesky scheme without hierarchical construction (see Figure 4.2a), and iii) rBergomi simulated using Cholesky scheme with hierarchical construction as in Section **??** (see Figure 4.2b). From those plots, it seems to me that we have a better behavior of the hybrid scheme, in terms of weak error rate, than the Cholesky scheme (for both cases (with/without hierarchical representation)), at least in the pre-asymptotic regime, which justifies our use of Richardson extrapolation with the hybrid scheme. On the other hand, it seems that the weak error when using Cholesky scheme is smaller in magnitude compared to the one using the hybrid scheme for a fixed number of time steps N. Therefore, with the observed weak convergence behavior using Cholesky scheme, it is not clear to me if it makes sense to use Richardson extrapolation since the rates are too low. We note that we observed similar behavior for sets with $H = 0.02$ (see Figure 4.3 for weak error observed for set 2 using Cholesky scheme).

Parameters	Reference solution
Set 1: $H = 0.07, K = 1, S_0 = 1, T = 1, \rho = -0.9, \eta = 1.9, \xi_0 = 0.235^2$	0.0791 $(7.9e{-}05)$
Set 2: $H = 0.02, K = 1, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.1248 $(1.3e{-}04)$
Set 3: $H = 0.02, K = 0.8, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.2407 $(5.6e{-}04)$
Set 4: $H = 0.02, K = 1.2, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.0568 $(2.5e{-}04)$
Set 5: $H = 0.43, K = 1, S_0 = 1, T = 1, \rho = -0.9, \eta = 1.9, \xi_0 = 0.235^2$	0.0712 $(7.9e{-}05)$

Table 4.1: Reference solution, which is the approximation of the call option price under the rBergomi model, using MC with 500 time steps and number of samples, $M = 10^6$, for different parameter constellations. The numbers between parentheses correspond to the statistical errors estimates.

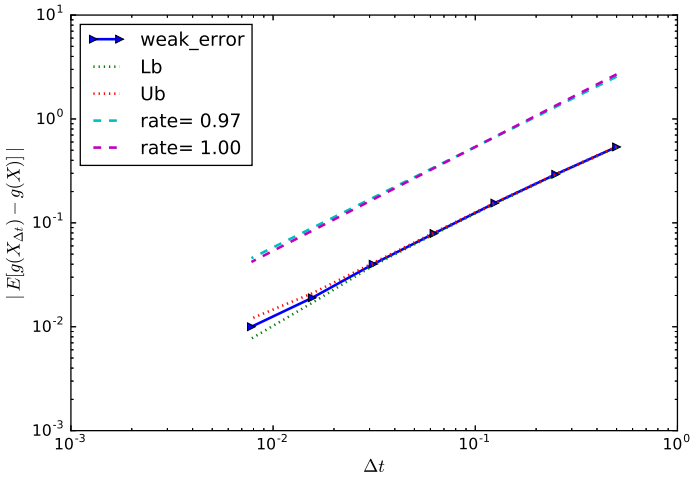
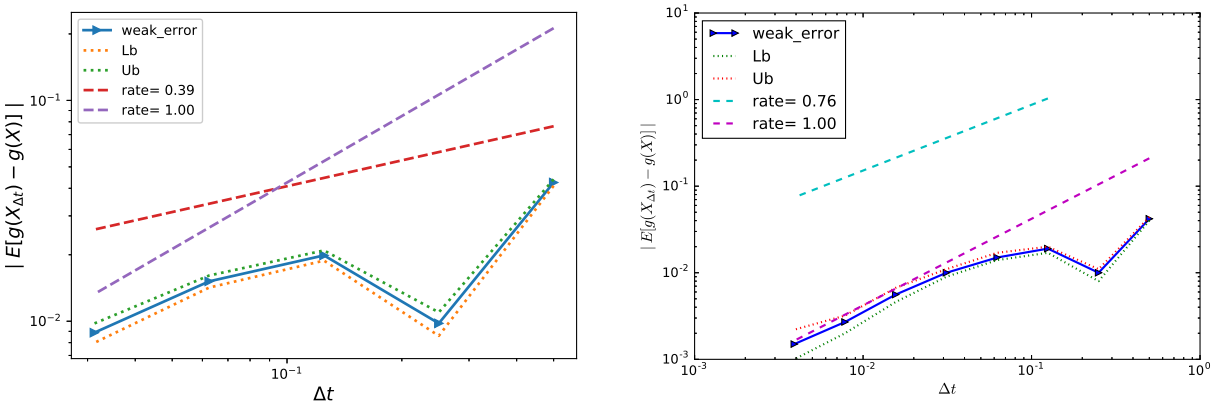


Figure 4.1: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 10^6$) with hierarchical hybrid scheme, for set 1 parameter in Table 4.1. We refer to C_{RB} as $\text{E}\left[g(X)\right]$, and to C^N_{RB} as $\text{E}\left[g(X_{\Delta t})\right]$. The upper and lower bounds are 95% confidence intervals.



(a) (b)

Figure 4.2: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 6 \times 10^6$) with Cholesky scheme, for set 1 parameter in Table 4.1. We refer to C_{RB} as $\text{E}\left[g(X)\right]$, and to C^N_{RB} as $\text{E}\left[g(X_{\Delta t})\right]$. The upper and lower bounds are 95% confidence intervals. a) With hierarchical representation. b) Without hierarchical representation.

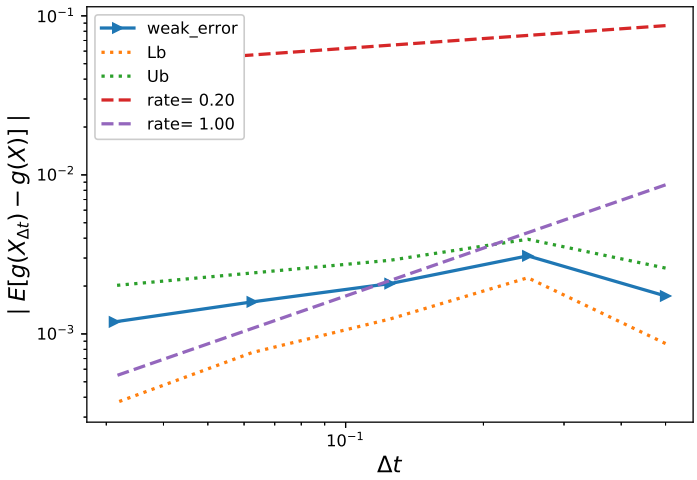
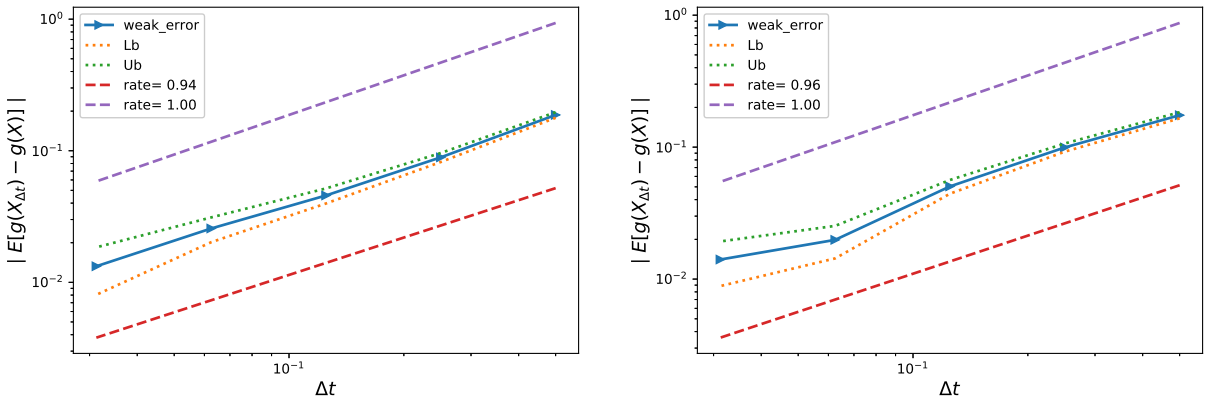


Figure 4.3: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 6 \times 10^6$) with hierarchical Cholesky scheme, for set 2 parameter in Table 4.1. We refer to C_{RB} as $\mathrm{E}\,[g(X)]$, and to C^N_{RB} as $\mathrm{E}\,[g(X_{\Delta t})]$. The upper and lower bounds are 95% confidence intervals.

To investigate more the behavior observed for the Cholesky scheme, we test the case of set 5 in table 4.1 which is close to the Gaussian case for $H = 1/2$ (see Figure 4.4). We observed a weak convergence rate of order almost 1. This observation confirms first that maybe the hybrid scheme is more robust, in terms of weak error, than Cholesky for the simulation of the rough Bergomi dynamics. Furthermore, we believe that the weak error in the Cholesky scheme depends on H, and that the common error in both the Cholesky and Hybrid scheme is dominated by the second kind of weak error involved in the hybrid scheme with is of order Δt that is why we observed more robust rate for the hybrid scheme. We try in Section 3 to provide an analysis for the weak rate.



(a) (b)

Figure 4.4: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 10^5$) with Cholesky scheme, for set 5 parameter in Table 4.1. We refer to C_{RB} as $\mathrm{E}\left[g(X)\right]$, and to C^N_{RB} as $\mathrm{E}\left[g(X_{\Delta t})\right]$. The upper and lower bounds are 95% confidence intervals. a) With hierarchical representation. b) Without hierarchical representation.

Remark 4.1. Our observations are in harmony with results observed in [4], where it was observed that the weak error for pricing European option under the rBergomi, simulated using Cholesky scheme and for a particular choice of test function is of order $2H$ across the full range of $0 < H < \frac{1}{2}$ (see Figure 3 in [4]). On the other hand, I suspect that the results reported in the Master thesis provided by Christian are reported on opposite way, that is the results reported for the hybrid scheme correspond to the Cholesky scheme (to be checked).

4.2 Comparing the different errors and computational time for MC and MISC for the Cholesky scheme

Due to the behavior of the weak error for small values of H, we can not apply Richardson extrapolation. However, since the weak error when using Cholesky scheme seems to be one order smaller than the one obtained when using the hybrid scheme, we try to compare the numerical complexity of MC and MISC for the case of set 1 parameter in Table 4.1, without using Richardson extrapolation. From Tables 4.2 and 4.3, we see that to get a total relative error below 1%, we need more than 32 time steps for both MC and MISC, implying (given our previous results with the hybrid scheme) that using the hybrid scheme coupled with Richardson extrapolation gives better results for MC and MISC, than those obtained with Cholesky scheme.

Method	Steps				
	2	4	8	16	32
MISC ($\text{TOL}_{\text{MISC}} = 10^{-1}$)	**0.122** (0.042,0.080)	**0.093** (0.010,0.083)	**0.193** (0.020,0.173)	**0.150** (0.015,0.135)	–
MISC ($\text{TOL}_{\text{MISC}} = 10^{-2}$)	**0.122** (0.042,0.080)	**0.016** (0.010,0.006)	**0.028** (0.020,0.008)	**0.029** (0.015,0.014)	–
MISC ($\text{TOL}_{\text{MISC}} = 5.10^{-3}$)	**0.057** (0.042,0.015)	**0.013** (0.010,0.003)	–	–	–
MC	**0.084** (0.042,0.042)	**0.020** (0.010,0.010)	**0.039** (0.020,0.019)	**0.030** (0.015,0.015)	**0.017** (0.009,0.008)
M(# MC samples)	10^3	3×10^4	4×10^3	6×10^3	2×10^4

Table 4.2: Total relative error of MISC, without Richardson extrapolation, with different tolerances, and MC to compute the call option prices for different numbers of time steps, where the rBergomi dynamics are simulated using Cholesky scheme. The values between parentheses correspond to the different errors contributing to the total relative error: for MISC we report the bias and quadrature errors and for MC we report the bias and the statistical errors estimates.

Method	Steps				
	2	4	8	16	32
MISC ($\text{TOL}_{\text{MISC}} = 10^{-1}$)	0.1	0.08	8	670	–
MISC ($\text{TOL}_{\text{MISC}} = 10^{-2}$)	0.1	2.5	15	2040	–
MISC ($\text{TOL}_{\text{MISC}} = 5.10^{-3}$)	0.3	10	–	–	–
MC method	0.2	5.6	0.8	1.6	22

Table 4.3: Comparison of the computational time (in seconds) of MC and MISC, to compute the call option price of the rBergomi model, simulated using Cholesky scheme, for different numbers of time steps. The average MC CPU time is computed over 100 runs.

References Cited

[1] Ole E Barndorff-Nielsen and Jürgen Schmiegel. Ambit processes; with applications to turbulence and tumour growth. In *Stochastic analysis and applications*, pages 93–124. Springer, 2007.

[2] Ole E Barndorff-Nielsen and Jürgen Schmiegel. Brownian semistationary processes and volatility/intermittency. *Advanced financial modelling*, 8:1–26, 2009.

[3] Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[4] Christian Bayer, Peter K Friz, Paul Gassiat, Joerg Martin, and Benjamin Stemper. A regularity structure for rough volatility. *arXiv preprint arXiv:1710.07481*, 2017.

[5] Christian Bayer, Peter K Friz, Archil Gulisashvili, Blanka Horvath, and Benjamin Stemper. Short-time near-the-money skew in rough fractional volatility models. *Quantitative Finance*, pages 1–20, 2018.

[6] Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, 21(4):931–965, 2017.

[7] Domenico Marinucci and Peter M Robinson. Alternative forms of fractional Brownian motion. *Journal of statistical planning and inference*, 80(1-2):111–122, 1999.

[8] Jean Picard. Representation formulae for the fractional Brownian motion. In *Séminaire de Probabilités XLIII*, pages 3–70. Springer, 2011.

[9] VM Sithi and SC Lim. On the spectra of riemann-liouville fractional brownian motion. *Journal of Physics A: Mathematical and General*, 28(11):2995, 1995.