Comparing the results of Cholesky and Hybrid schemes for the rBergomi

March 9, 2019

1 Problem setting

1.1 The rBergomi model

We consider the rBergomi model for the price process S_t, 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 $\xi_0(t) = \mathrm{E}\left[v_t\right]$ 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

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

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 $\gamma = \frac{1}{2} - H$

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

[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.

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

$$\begin{aligned} S_t &= S_0 \exp\left(\int_0^t \sqrt{v(s)} dZ(s) - \frac{1}{2} \int_0^t v(s) ds \right), \quad S_0 > 0 \\ v_u &= \xi_0(u) \exp\left(\eta \widetilde{W}_u^H - \frac{\eta^2}{2} u^{2H} \right), \quad \xi_0 > 0. \end{aligned} \tag{1.3}$$

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.

1.2 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 $W_{t_1}^1, \dots, W_{t_N}^1$ and $\widetilde{W}_{t_1}^H, \dots, \widetilde{W}_{t_N}^H$ along a given time grid $t_1 < \dots < t_N$. In the literature, there are essentially two suggested ways to achieve this:

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

ii) **The hybrid scheme**: 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 $W_{t_1}^1, \dots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \dots, \widetilde{W}_{t_N}$, 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 $W_{t_1}^1, \dots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \dots, \widetilde{W}_{t_N}$.

In this work, we adopt approach (ii) for the simulation of the rBergomi dynamics. The hybrid scheme discretizes the $\widetilde{W}^H$ 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 (see (1.4)). We utilize the hybrid scheme with $\kappa = 1$[3], which is based on the following

[3] There are different variants of the hybrid scheme depending on the value of κ.

approximation

$$\widetilde{W}^H_{\frac{i}{N}} \approx \overline{W}^H_{\frac{i}{N}} = \sqrt{2H}\left(W^2_i + \sum_{k=2}^{i}\left(\frac{b_k}{N}\right)^{H-\frac{1}{2}}\left(W^1_{\frac{i-(k-1)}{N}} - W^1_{\frac{i-k}{N}}\right)\right), \tag{1.4}$$

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 (1.4) requires the most computational effort in the simulation. Given that (1.4) can be seen as discrete convolution, 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 $\overline{W}^H_0, \overline{W}^H_1, \dots, \overline{W}^H_{\frac{[Nt]}{N}}$ 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 $(\mathbf{W}^{(1)}, \mathbf{W}^{(2)})$.

2 Details of Cholesky scheme coupled with hierarchical reresentation

Let us denote by the matrix A, the computable covariance matrix of $\widetilde{W}^H_{t_1}, \dots, \widetilde{W}_{t_N}, W^1_{t_1}, \dots, W^1_{t_N}$. We can use Cholesky decomposition of A to produce exact samples of $W^1_{t_1}, \dots, W^1_{t_N}, \widetilde{W}^H_{t_1}, \dots, \widetilde{W}_{t_N}$.

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, $\mathbf{X} = (X_1, \dots, X_N, X_{N+1}, \dots, X_{2N})'$, we have

$$\mathbf{W}^{(1)} = L_1 \mathbf{X}_{1:N}, \quad \widetilde{\mathbf{W}} = \left(\begin{array}{c|c} L_2 & L_3 \end{array}\right)\mathbf{X}. \tag{2.1}$$

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

$$\mathbf{X} = \left(\begin{array}{c|c} L_1^{-1}G & 0 \\ 0 & I_N \end{array}\right)\mathbf{Z}. \tag{2.2}$$

We need to make sure that $\mathbf{X}$ has Gaussian distribution as an outcome of the construction (2.2). Consequently, we need to compute carefully L_1^{-1}. Actually, I observed that $L_1 = I_{N\times N}$. Therefore, $\mathbf{X}$ has Gaussian distribution as an outcome of the construction (2.2).

3 Comparing Cholesky and Hybrid schemes results

In this section, we compare the weak rates obtained for set 1 in Table 3.1. We compare three different cases: i) rBergomi simulated using Hybrid scheme with hierarchical construction (see Figure 3.1), ii) rBergomi simulated using Cholesky scheme without hierarchical construction (see Figure 3.3a), and iii) rBergomi simulated using Cholesky scheme with hierarchical construction as in Section 2 (see Figure 3.3b). From those plots, it seems to me that we have a better behavior of the hybrid scheme, in terms of weak error, 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. However, 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.

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 3.1: 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.

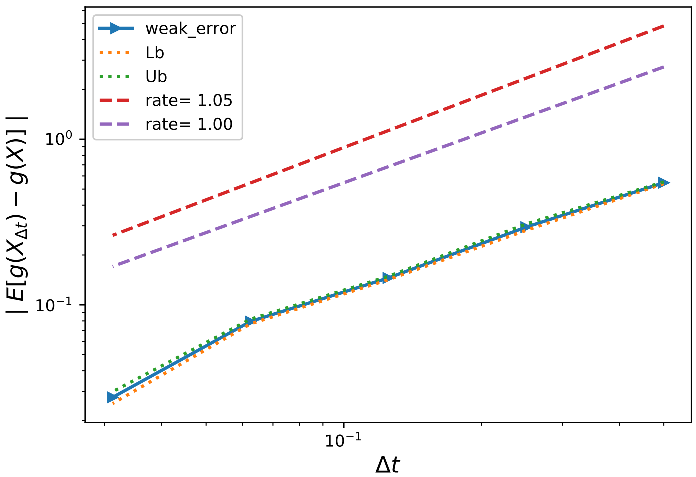
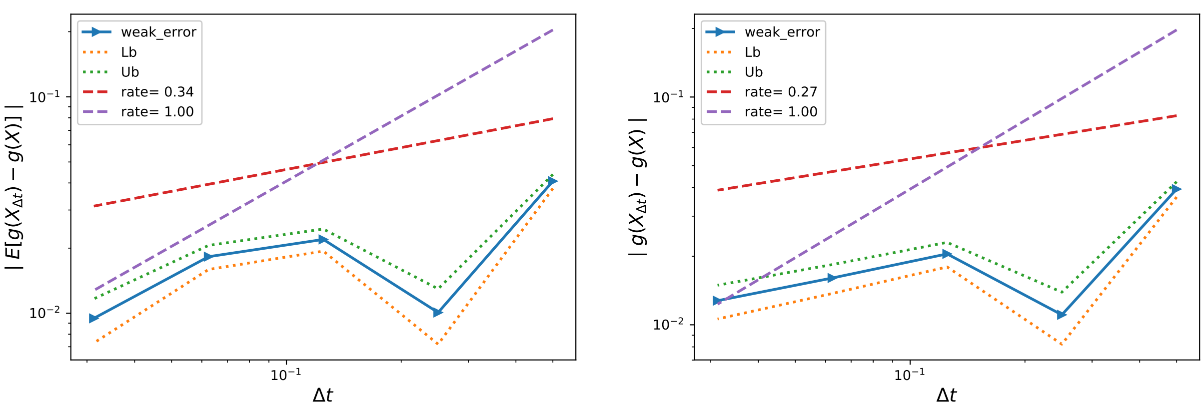


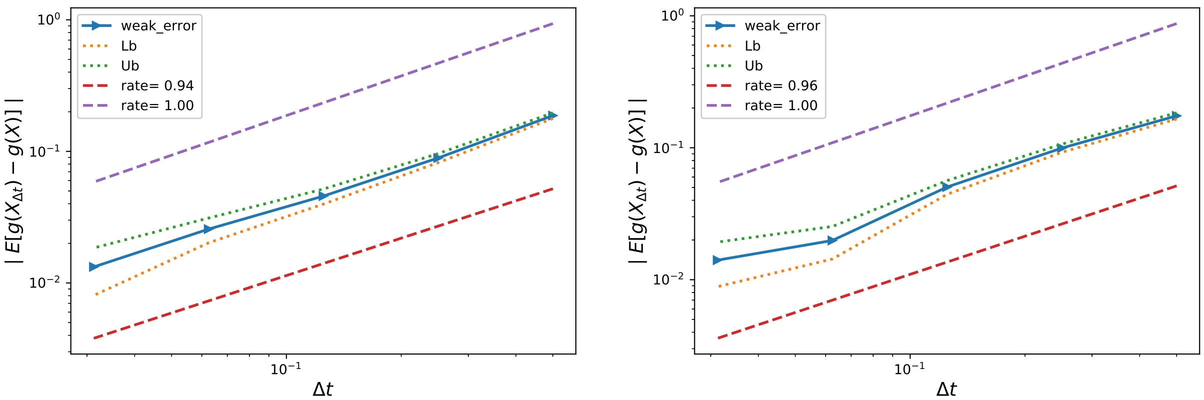
Figure 3.1: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 10^6$) with hierarchical hybrid scheme, for set 1 parameter in Table 3.1. We refer to C_{RB} as $\mathrm{E}\left[g(X)\right]$, and to C_{RB}^N as $\mathrm{E}\left[g(X_{\Delta t})\right]$. The upper and lower bounds are 95% confidence intervals.



(a) (b)

Figure 3.2: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 10^6$) with Cholesky scheme, for set 1 parameter in Table 3.1. We refer to C_{RB} as $\mathrm{E}\left[g(X)\right]$, and to C_{RB}^N 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.

To investigate more the behavior observed for the Cholesky scheme, we test the case of set 5 in table 3.1 which is close to the Gaussian case for $H = 1/2$ (see Figure 3.3). Surprisingly, we observed a weak convergence rate of order almost 1. This observations 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 in the hybrid scheme with is of order Δ that is why we observed more robust rate for the hybrid scheme. We try in Section 4 to provide an analysis for the weak rate.



(a) (b)

Figure 3.3: The convergence of the weak error $\mathcal{E}_B(N)$, using MC ($M = 10^5$) with Cholesky scheme, for set 5 parameter in Table 3.1. We refer to C_{RB} as $\mathrm{E}\left[g(X)\right]$, and to C_{RB}^N 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 3.1. Our observations are in harmony with results observed in [1], 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 [1]). 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 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

(4.1)
$$\mathrm{E}\left[g\left(X_T\left(W_t^{(1)}, \widetilde{W}_t\right)\right)\right] \approx \mathrm{E}\left[g\left(\overline{X}_N\left(\overline{W}_1^{(1)}, \dots, \overline{W}_N^{(1)}, \overline{\overline{W}}_1, \dots, \overline{\overline{W}}_N\right)\right)\right]: \quad \textbf{(Hybrid scheme)}$$

(4.2)

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

To simplify notation, let $\overline{\mathbf{W}}^1 = (\overline{W}_1^{(1)},\dots,\overline{W}_N^{(1)})$, $\overline{\overline{\mathbf{W}}} = (\overline{\overline{W}}_1,\dots,\overline{\overline{W}}_N)$, $\mathbf{W}^1 = (W_1^{(1)},\dots,W_N^{(1)})$ and $\widetilde{\mathbf{W}} = (\widetilde{W}_1,\dots,\widetilde{W}_N)$ .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

$$\begin{aligned}
\varepsilon_B^{Hyb} &= \left|\mathrm{E}\left[g\left(X_T\left(W_t^{(1)},\widetilde{W}_t\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}}^1,\overline{\overline{\mathbf{W}}}\right)\right)\right]\right| \\
&\leq \left|\mathrm{E}\left[g\left(X_T\left(W_t^{(1)},\widetilde{W}_t\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1,\widetilde{\mathbf{W}}\right)\right)\right]\right| + \left|\mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}}^1,\overline{\overline{\mathbf{W}}}\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1,\widetilde{\mathbf{W}}\right)\right)\right]\right| \\
&\leq \varepsilon_B^{Chol} + \left|\mathrm{E}\left[g\left(\overline{X}_N\left(\overline{\mathbf{W}}^1,\overline{\overline{\mathbf{W}}}\right)\right)\right] - \mathrm{E}\left[g\left(\overline{X}_N\left(\mathbf{W}^1,\widetilde{\mathbf{W}}\right)\right)\right]\right| \qquad (4.3)
\end{aligned}$$

To-DO: We need to analyze both terms involved in the right hand-side of (4.3).

References Cited

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