Plan of action for the rBergomi project

February 24, 2019

1 Intoduction

There are mainly two point that we want to improve comparing to the current version of the rBergomi manuscript:

i) Some internal beliefs that maybe we are making wrong assumptions about the asymptotic rates of convergence for the weak error, when using the Hybrid scheme. Therefore, we suggest to test the first case of parameters with Cholesky scheme (See Section 3 for details about Cholesky scheme).

   - If we find similar results as observed with the hybrid scheme then we may add just a remark or the results of Cholesky for that case. Otherwise, we may repeat all experiments, or change the used hierarchical representation.
   - It is critical also to check if the hierarchical representation is working for the Cholesky scheme as we observed in the case of the hybrid scheme. In fact, for the hybrid scheme it was clear that $\mathbf{W}^{(1)}$ dimensions are more important than those of $\mathbf{W}^{(2)}$, reducing already the effective dimension from $2N$ to N, before even that the Brownian bridge construction creates more anisotropy for $\mathbf{W}^{(1)}$ directions. However, in the Cholesky scheme, we do not have this obvious distinction.

ii) Provide some errors bounds for the quadrature error of MISC (See Section 2 for details). This will make the method more robust and more convincing in terms of practical use. In fact, at the current stage, the errors that we provide are functions of TOL_{MISC}, that is $\mathcal{E}_Q(\text{TOL}_{\text{MISC}}, N) = f(\text{TOL}_{\text{MISC}})$ and it is not clear to us the behavior of f.

   There are two ways to achieve this:

   1. The first way relies on estimating the interpolation error and then by Monte Carlo deduce the quadrature error. We believe that the Monte Carlo error will be dominated by the statistical error and we need maybe few samples for its estimation. For this purpose, we will use the code provided by Joakim.
   2. The second way will be an alternative for the first way in case we failed to have nice error bounds. It is more expensive but more reliable. It is based on learning the error curve which will be parameterized by the different parameters involved in the pricing problem under the rough Bergomi model in addition to MISC tolerance, TOL_{MISC}, and the number of time steps, N.

2 MISC error estimate

2.0.1 MISC error estimate

In our case, we have

$$
\begin{aligned}
C_{\text{RB}}\left(T,K\right) &= \mathrm{E}\left[C_{\text{BS}}\left(S_0=\exp\left(\rho\int_0^T\sqrt{v_t}dW_t^1-\frac{1}{2}\rho^2\int_0^T v_t dt\right),\ k=K,\ \sigma^2=(1-\rho^2)\int_0^T v_t dt\right)\right]\\
&\approx \int_{\mathbb{R}^{2N}} C_{BS}\left(G(\mathbf{W}^{(1)},\mathbf{W}^{(2)})\right)\rho_N(\mathbf{W}^{(1)})\rho_N(\mathbf{W}^{(2)})d\mathbf{W}^{(1)}d\mathbf{W}^{(2)}\\
&= C_{RB}^N, \qquad (2.1)
\end{aligned}
$$

where G maps $2N$ independent standard Gaussian random inputs to the parameters fed to Black-Scholes formula, and ρ_N is the multivariate Gaussian density, given by

$$\rho_N(\mathbf{z})=\frac{1}{(2\pi)^{N/2}}e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}.$$

From (2.1), we define

$$F^N=C_{\text{BS}}(G(\mathbf{W}^{(1)},\mathbf{W}^{(2)})),$$

and introduce the set $C^0(\mathbb{R})$ of real-valued continuous functions over $\mathbb{R}$, and the subspace of polynomials of degree at most q over $\mathbb{R}$, $\mathbb{P}^q(\mathbb{R})\subset C^0(\mathbb{R})$. Next, we consider a sequence of univariate Lagrangian interpolant operators in each dimension Y_n ($1\le n\le 2N$), that is, $\{U_n^{m(\beta_n)}\}_{\beta_n\in\mathbb{N}_+}$ (we refer to the value β_n as the interpolation level). Each interpolant is built over a set of $m(\beta_n)$ collocation points, $\mathcal{H}^{m(\beta_n)}=\{y_n^1,y_n^2,\dots,y_n^{m(\beta_n)}\}\subset\mathbb{R}$, thus, the interpolant yields a polynomial approximation,

$$U^{m(\beta_n)}:C^0(\mathbb{R})\to\mathbb{P}^{m(\beta_n)-1}(\mathbb{R}),\quad U^{m(\beta_n)}[F^N](y_n)=\sum_{j=1}^{m(\beta_n)}\left(f(y_n^j)\prod_{k=1;k\neq j}^{m(\beta_n)}\frac{y_n-y_n^k}{y_n^j-y_n^k}\right).$$

The $2N$-variate Lagrangian interpolant can then be built by a tensorization of univariate interpolants: denote by $C^0(\mathbb{R}^{2N})$ the space of real-valued $2N$-variate continuous functions over $\mathbb{R}^{2N}$ and by $\mathbb{P}^{\mathbf{q}}(\mathbb{R}^{2N})=\otimes_{n=1}^{2N}\mathbb{P}^{\mathbf{q}_n}(\mathbb{R})$ the subspace of polynomials of degree at most q_n over $\mathbb{R}$, with $\mathbf{q}=(q_1,\dots,q_{2N})\in\mathbb{N}^{2N}$, and consider a multi-index $\boldsymbol{\beta}\in\mathbb{N}_+^{2N}$ assigning the interpolation level in each direction, y_n, then the multivariate interpolant can then be written as

$$U^{m(\boldsymbol{\beta})}:C^0(\mathbb{R}^{2N})\to\mathbb{P}^{m(\boldsymbol{\beta})-1}(\mathbb{R}^{2N}),\quad U^{m(\boldsymbol{\beta})}[F^N](\mathbf{y})=\bigotimes_{n=1}^{2N}U^{m(\beta_n)}[F^N](\mathbf{y}),$$

Given this construction, we can define the MISC interpolant for approximating F^N, using a set of multi indices $\mathcal{I}\in\mathbb{N}^{2N}$ as

$$I^{\mathcal{I}}[F^N]=\sum_{\boldsymbol{\beta}\in\mathcal{I}}\Delta U_N^{\boldsymbol{\beta}}, \qquad (2.2)$$

where

$$\Delta_i U_N^{\boldsymbol{\beta}} = \begin{cases} U_N^{\boldsymbol{\beta}} - U_N^{\boldsymbol{\beta}'}, \text{ with } \boldsymbol{\beta}' = \boldsymbol{\beta} - e_i, \text{ if } \boldsymbol{\beta}_i > 0 \\ U_N^{\boldsymbol{\beta}}, \quad \text{otherwise,} \end{cases}$$

where e_i denotes the ith $2N$-dimensional unit vector. Then, $\Delta U_N^{\boldsymbol{\beta}}$ is defined as

$$\Delta U_N^{\boldsymbol{\beta}} = \left(\prod_{i=1}^{2N} \Delta_i \right) U_N^{\boldsymbol{\beta}}.$$

We define the interpolation error induced by MISC as

$$e_N = F^N - I^{\mathcal{I}}[F^N]. \tag{2.3}$$

One can have a bound on the interpolation error of MISC, e_N, by tensorizing one dimensional error estimates, and then simply integrate that bound to get the MISC quadrature error, $\mathcal{E}_Q(\mathrm{TOL}_{\mathrm{MISC}}, N)$. However, we think that this will not lead to a sharp error estimate for MISC. Another strategy for estimating the MISC quadrature error, is to estimate $\mathrm{E}\left[e_N\right]$ using MC by sampling directly e_N.

If we define $Y = F^N + (Q_N^{\mathcal{I}} - I^{\mathcal{I}}[F^N])$ (where $Q_N^{\mathcal{I}}$ is the MISC quadrature estimator, then we have

$$\begin{aligned} \mathrm{E}\left[Y\right] &= \mathrm{E}\left[F^N\right] \\ Var[Y] &= Var[e_N] < Var[\mathcal{A}_{\mathrm{MC}}], \end{aligned} \tag{2.4}$$

where $\mathcal{A}_{\mathrm{MC}}$ is the MC estimator for $\mathrm{E}\left[F^N\right]$.

(2.4) shows that MISC can be seen as a control variate for MC estimator and consequently as a powerful variance reduction tool.

TO-DO 1: Estimate numerically $\mathrm{E}\left[e_N\right]$ using MC by sampling directly e_N.

TO-DO 2: Show numerically (2.4), that is ISC can be seen as a control variate for MC estimator.

3 Details of Cholesky scheme coupled with hierarchical reresentation

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

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

$$\mathbf{W}^{(1)} = L_1 \mathbf{X}_{1:N}, \quad \widetilde{\mathbf{W}} = \left(\begin{array}{c|c} L_2 & L_3 \end{array} \right) \mathbf{X}. \tag{3.1}$$

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

$$\mathbf{X} = \left(\begin{array}{c|c} L_1^{-1}G & 0 \\ 0 & I_N \end{array} \right) \mathbf{Z}. \tag{3.2}$$

We need to make sure that $\mathbf{X}$ has Gaussian distribution as an outcome of the construction (3.2). 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 (3.2).

TO-DO 1: Implement the appropriate Cholesky scheme, taking into account the above construction, and check if the hierarchical construction is giving good results.