Adaptive sparse grids and quasi Monte Carlo for option pricing under the rough Bergomi model

Chiheb Ben Hammouda

KAUST
King Abdullah University of Science and Technology

Christian Bayer

WIAS

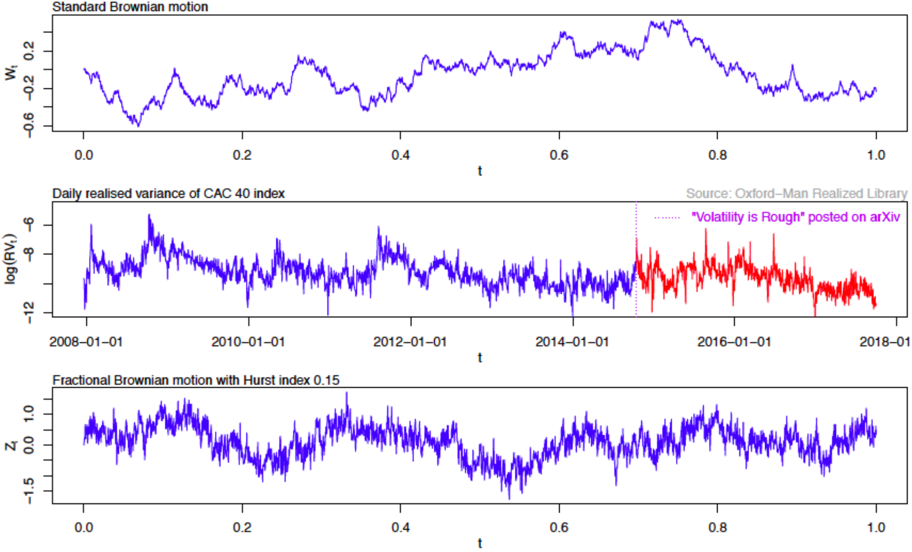
Raúl Tempone

RWTH AACHEN UNIVERSITY

3rd International Conference on Computational Finance (ICCF2019), A Coruña

8-12 July, 2019

Rough volatility [1]



[1] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. “Volatility is rough”. In: *Quantitative Finance* 18.6 (2018), pp. 933–949

The rough Bergomi model [2]

This model, under a pricing measure, is given by

$$\begin{cases} dS_t &= \sqrt{v_t} S_t dZ_t, \\ v_t &= \xi_0(t) \exp\left(\eta \widetilde{W}_t^H - \frac{1}{2}\eta^2 t^{2H}\right), \\ Z_t &:= \rho W_t^1 + \bar{\rho} W_t^\perp \equiv \rho W^1 + \sqrt{1-\rho^2} W^\perp, \end{cases} \tag{1}$$

- $(W^1, W^\perp)$: two independent standard Brownian motions
- $\widetilde{W}^H$ is Riemann-Liouville process, defined by

$$\widetilde{W}_t^H = \int_0^t K^H(t-s) dW_s^1, \quad t \geq 0,$$
$$K^H(t-s) = \sqrt{2H}(t-s)^{H-1/2}, \quad \forall\, 0 \leq s \leq t.$$

- $H \in (0, 1/2]$ controls the roughness of paths, $\rho \in [-1, 1]$ and $\eta > 0$.
- $t \mapsto \xi_0(t)$: forward variance curve, known at time 0.

[2] Christian Bayer, Peter Friz, and Jim Gatheral. "Pricing under rough volatility". In: *Quantitative Finance* 16.6 (2016), pp. 887–904

Model challenges

- **Numerically:**
  - The model is non-Markovian and non-affine ⇒ Standard numerical methods (PDEs, characteristic functions) seem inapplicable.
  - The only prevalent pricing method for mere vanilla options is Monte Carlo (MC) (Bayer, Friz, and Gatheral 2016; McCrickerd and Pakkanen 2018) ☹ still computationally expensive task.
  - Discretization methods have a poor behavior of the strong error (strong convergence rate of order $H \in [0, 1/2]$) (Neuenkirch and Shalaiko 2016) ⇒ Variance reduction methods, such as multilevel Monte Carlo (MLMC), are inefficient for very small values of H.
- **Theoretically:**
  - No proper weak error analysis done in the rough volatility context.

Option pricing challenges

The integration problem is challenging

- Issue 1: Time-discretization of the rough Bergomi process (large N (number of time steps)) $\Rightarrow$ S takes values in a high-dimensional space $\Rightarrow$ ☹ Curse of dimensionality when using numerical integration methods.
- Issue 2: The payoff function g is typically not smooth $\Rightarrow$ low regularity$\Rightarrow$ ☹ slow convergence of deterministic quadrature methods.

⚠ Curse of dimensionality: An exponential growth of the work (number of function evaluations) in terms of the dimension of the integration problem.

Methodology [3]

We design efficient hierarchical pricing methods based on

1. Analytic smoothing to uncover available regularity (inspired by (Romano and Touzi 1997) in the context of stochastic volatility models).
2. Approximating the option price using deterministic quadrature methods
   - **Adaptive sparse grids quadrature (ASGQ)**.
   - **Quasi Monte Carlo (QMC)**.
3. Coupling our methods with hierarchical representations
   - **Brownian bridges** as a Wiener path generation method ⇒ ↘ the effective dimension of the problem.
   - **Richardson Extrapolation** (Condition: weak error of order 1) ⇒ Faster convergence of the weak error ⇒ ↘ number of time steps (smaller dimension).

[3] Christian Bayer, Chiheb Ben Hammouda, and Raul Tempone. "Hierarchical adaptive sparse grids and quasi Monte Carlo for option pricing under the rough Bergomi model". In: *arXiv preprint arXiv:1812.08533* (2018).

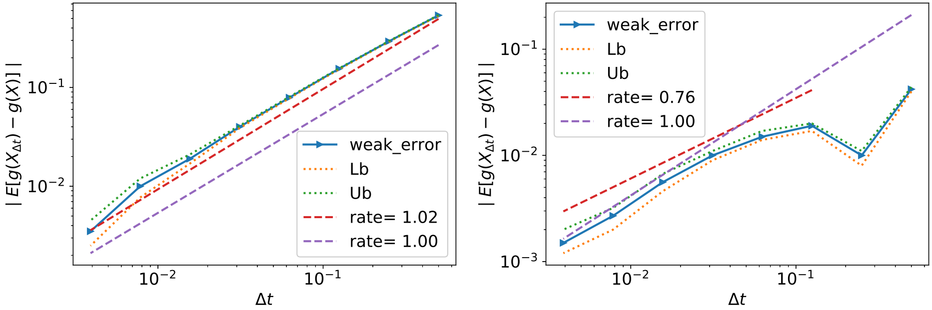
Simulation of the rough Bergomi dynamics

Goal: Simulate jointly $(W_t^1, \widetilde{W}_t^H : 0 \leq t \leq T)$, resulting in $W_{t_1}^1, \dots, W_{t_N}$ and $\widetilde{W}_{t_1}^H, \dots, \widetilde{W}_{t_N}^H$ along a given grid $t_1 < \cdots < t_N$

1. **Covariance based approach** (Bayer, Friz, and Gatheral 2016)
   - Based on Cholesky decomposition of the covariance matrix of the $(2N)$-dimensional Gaussian random vector $W_{t_1}^1, \dots, W_{t_N}^1, \widetilde{W}_{t_1}^H, \dots, \widetilde{W}_{t_N}$.
   - Exact method but slow
   - At least $\mathcal{O}\left(N^2\right)$.
2. **The hybrid scheme** (Bennedsen, Lunde, and Pakkanen 2017)
   - Based on Euler discretization but crucially improved by moment matching for the singular term in the left point rule.
   - Accurate scheme that is much faster than the Covariance based approach.
   - $\mathcal{O}(N)$ up to logarithmic factors that depend on the desired error.

On the choice of the simulation scheme

Figure 1.1: The convergence of the weak error $\mathcal{E}_B$, using MC with 6×10^6 samples, for example parameters: $H = 0.07,\ K = 1, S_0 = 1,\ T = 1,\ \rho = -0.9,$ $\eta = 1.9,\ \xi_0 = 0.0552$. The upper and lower bounds are 95% confidence intervals. a) With the hybrid scheme b) With the exact scheme.



(a) (b)

Conditional expectation for analytic smoothing

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

- $C_{\mathrm{BS}}(S_0, k, \sigma^2)$: the Black-Scholes call price, for initial spot price S_0, strike price k, and volatility σ^2.
- G maps $2N$ independent standard Gaussian random inputs to the parameters fed to Black-Scholes formula.
- ρ_N: the multivariate Gaussian density, N: number of time steps.

Sparse grids I

Notation:

- Given $F : \mathbb{R}^d \to \mathbb{R}$ and a multi-index $\boldsymbol{\beta} \in \mathbb{N}_+^d$.
- $F_{\boldsymbol{\beta}} := Q^{m(\boldsymbol{\beta})}[F]$ a quadrature operator based on a Cartesian quadrature grid ($m(\beta_n)$ points along y_n).
  ⚠ Approximating $E[F]$ with $F_{\boldsymbol{\beta}}$ is not an appropriate option due to the well-known curse of dimensionality.
- The first-order difference operators

$$\Delta_i F_{\boldsymbol{\beta}} \begin{cases} F_{\boldsymbol{\beta}} - F_{\boldsymbol{\beta} - e_i}, & \text{if } \beta_i > 1 \\ F_{\boldsymbol{\beta}} & \text{if } \beta_i = 1 \end{cases}$$

  where e_i denotes the ith d-dimensional unit vector
- The mixed (first-order tensor) difference operators

$$\Delta[F_{\boldsymbol{\beta}}] = \otimes_{i=1}^{d} \Delta_i F_{\boldsymbol{\beta}}$$

Idea: A quadrature estimate of $E[F]$ is

$$\mathcal{M}_{\mathcal{I}_\ell}[F] = \sum_{\boldsymbol{\beta} \in \mathcal{I}_\ell} \Delta[F_{\boldsymbol{\beta}}], \tag{3}$$

Sparse grids II

$$E[F] \approx \mathcal{M}_{\mathcal{I}_\ell}[F] = \sum_{\boldsymbol{\beta} \in \mathcal{I}_\ell} \Delta[F_{\boldsymbol{\beta}}],$$

- Product approach: $\mathcal{I}_\ell = \{ \mid \boldsymbol{\beta} \mid_\infty \leq \ell;\ \boldsymbol{\beta} \in \mathbb{N}_+^d \}$
- Regular sparse grids[4]: $\mathcal{I}_\ell = \{ \mid \boldsymbol{\beta} \mid_1 \leq \ell + d - 1;\ \boldsymbol{\beta} \in \mathbb{N}_+^d \}$
- Adaptive sparse grids quadrature (ASGQ): $\mathcal{I}_\ell = \mathcal{I}^{\text{ASGQ}}$ (Next slides).

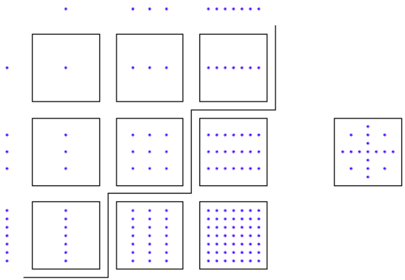


Figure 2.1: Left are product grids $\Delta_{\beta_1} \otimes \Delta_{\beta_2}$ for $1 \leq \beta_1, \beta_2 \leq 3$. Right is the corresponding SG construction.

[4] Hans-Joachim Bungartz and Michael Griebel. "Sparse grids". In: *Acta numerica* 13 (2004), pp. 147–269

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

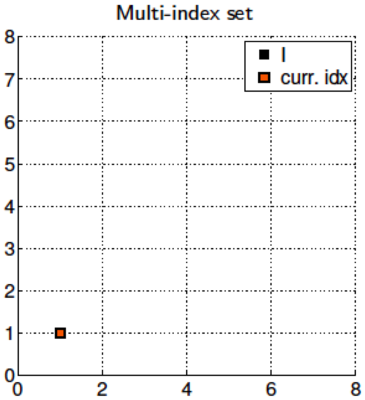


Figure 2.2: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

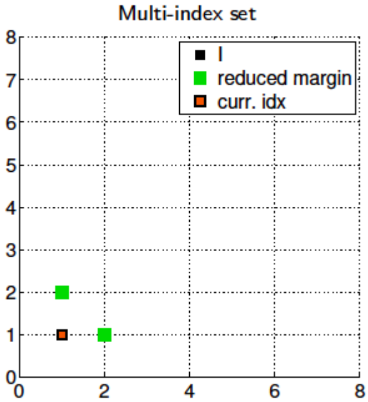


Figure 2.3: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

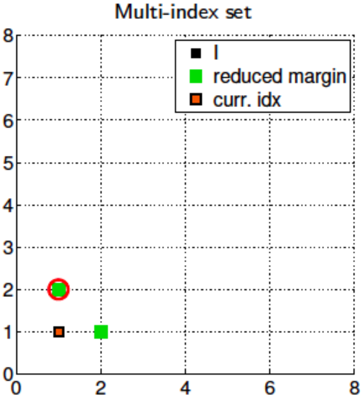


Figure 2.4: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I} \cup \{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I} \cup \{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

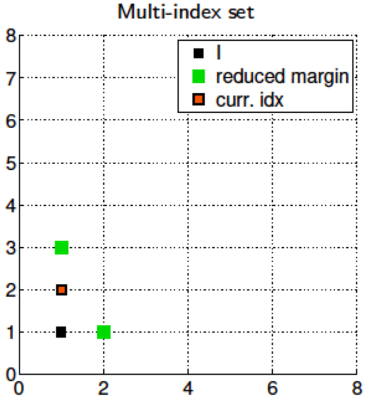


Figure 2.5: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I} \cup \{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I} \cup \{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

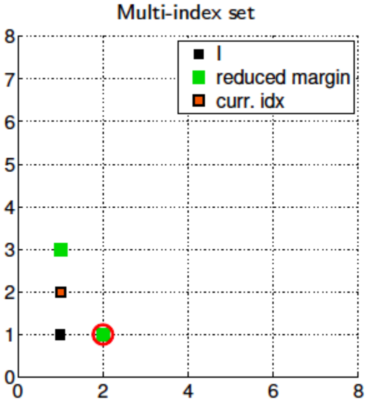


Figure 2.6: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

ASGQ in practice

- The construction of $\mathcal{I}^{\text{ASGQ}}$ is done by profit thresholding

$$\mathcal{I}^{\text{ASGQ}} = \{\boldsymbol{\beta} \in \mathbb{N}_+^d : P_{\boldsymbol{\beta}} \geq \overline{T}\}.$$

- **Profit of a hierarchical surplus** $P_{\boldsymbol{\beta}} = \frac{|\Delta E_{\boldsymbol{\beta}}|}{\Delta \mathcal{W}_{\boldsymbol{\beta}}}$.
- **Error contribution**: $\Delta E_{\boldsymbol{\beta}} = \left|\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}} - \mathcal{M}^{\mathcal{I}}\right|$.
- **Work contribution**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I}\cup\{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$.

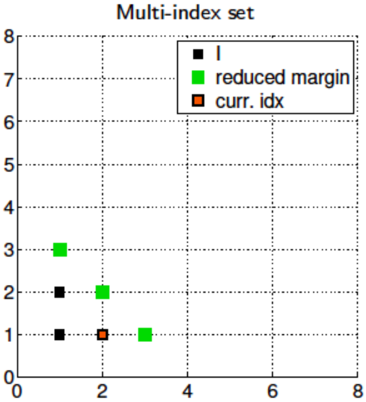


Figure 2.7: A posteriori, adaptive construction as in (Haji-Ali et al. 2016): Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one

Randomized QMC

- A (rank-1) lattice rule (Sloan 1985; Nuyens 2014) with n points

$$Q_n(f) := \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{kz \bmod n}{n}\right),$$

  where $z = (z_1, \ldots, z_d) \in \mathbb{N}^d$.
- A randomly shifted lattice rule

$$\overline{Q}_{n,q}(f) = \frac{1}{q} \sum_{i=0}^{q-1} Q_n^{(i)}(f) = \frac{1}{q} \sum_{i=0}^{q-1} \left(\frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{kz + \Delta^{(i)} \bmod n}{n}\right)\right), \qquad (4)$$

  where $\{\Delta^{(i)}\}_{i=1}^{q}$: independent random shifts, and $M^{\mathrm{QMC}} = q \times n$.
  - Unbiased approximation of the integral.
  - Practical error estimate.

Wiener path generation methods

$\{t_i\}_{i=0}^N$: Grid of time steps, $\{B_{t_i}\}_{i=0}^N$: Brownian motion increments

- **Random Walk**
  - Proceeds incrementally, given B_{t_i},

$$B_{t_{i+1}} = B_{t_i} + \sqrt{\Delta t} z_i, \; z_i \sim \mathcal{N}(0,1).$$

  - All components of $\mathbf{z} = (z_1, \ldots, z_N)$ have the same scale of importance: isotropic.
- **Hierarchical Brownian Bridge**
  - Given a past value B_{t_i} and a future value B_{t_k}, the value B_{t_j} (with $t_i < t_j < t_k$) can be generated according to ($\rho = \frac{j-i}{k-i}$)

$$B_{t_j} = (1-\rho) B_{t_i} + \rho B_{t_k} + \sqrt{\rho(1-\rho)(k-i)\Delta t} z_j, \; z_j \sim \mathcal{N}(0,1).$$

  - The most important values (determine the large scale structure of Brownian motion) are the first components of $\mathbf{z} = (z_1, \ldots, z_N)$.
  - $\searrow$ the effective dimension (# important dimensions) by $\nearrow$ anisotropy between different directions $\Rightarrow$ Faster ASGQ and QMC convergence.

Richardson Extrapolation (Talay and Tubaro 1990)

Motivation

- $(X_t)_{0\leq t\leq T}$ a certain stochastic process, $(\widehat{X}^h_{t_i})_{0\leq t_i\leq T}$ its approximation using a suitable scheme with a time step h.
- For sufficiently small h, and a suitable smooth function f, assume

$$\mathrm{E}\left[f(\widehat{X}^h_T)\right] = \mathrm{E}\left[f(X_T)\right] + ch + \mathcal{O}\left(h^2\right).$$

$$\Rightarrow 2\mathrm{E}\left[f(\widehat{X}^{2h}_T)\right] - \mathrm{E}\left[f(\widehat{X}^h_T)\right] = \mathrm{E}\left[f(X_T)\right] + \mathcal{O}\left(h^2\right).$$

General Formulation

$\{h_J = h_0 2^{-J}\}_{J\geq 0}$: grid sizes, K_{R}: level of Richardson extrapolation, $I(J,K_{\mathrm{R}})$: approximation of $\mathrm{E}\left[f(X_T)\right]$ by terms up to level K_{R}

$$I(J,K_{\mathrm{R}}) = \frac{2^{K_{\mathrm{R}}} I(J,K_{\mathrm{R}}-1) - I(J-1,K_{\mathrm{R}}-1)}{2^{K_{\mathrm{R}}}-1}, \quad J = 1,2,\ldots, K_{\mathrm{R}} = 1,2,\ldots \tag{5}$$

Advantage

Applying level K_{R} of Richardson extrapolation dramatically reduces the bias
$\Rightarrow\searrow$ the number of time steps N needed to achieve a certain error tolerance
$\Rightarrow\searrow$ the total dimension of the integration problem.

Numerical experiments

Table 1: Reference solution (using MC with 500 time steps and number of samples, $M = 8 \times 10^6$) of call option price under the rough Bergomi model, for different parameter constellations. The numbers between parentheses correspond to the statistical errors estimates.

Parameters	**Reference solution**
$H = 0.07, K = 1, S_0 = 1, T = 1, \rho = -0.9, \eta = 1.9, \xi_0 = 0.0552$	0.0791 (5.6e−05)
$H = 0.02, K = 1, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.1246 (9.0e−05)
$H = 0.02, K = 0.8, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.2412 (5.4e−05)
$H = 0.02, K = 1.2, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.0570 (8.0e−05)

- Set 1 is the closest to the empirical findings (Gatheral, Jaisson, and Rosenbaum 2018), suggesting that $H \approx 0.1$. The choice $\nu = 1.9$ and $\rho = -0.9$ is justified by (Bayer, Friz, and Gatheral 2016).
- For the remaining three sets, we test the potential of our method for a very rough case, where variance reduction methods are inefficient.

Error comparison

$\mathcal{E}_{\text{tot}}$: the total error of approximating the expectation in (2).

- When using ASGQ estimator, Q_N

$$\mathcal{E}_{\text{tot}} \leq \left|C_{\text{RB}} - C_{\text{RB}}^N\right| + \left|C_{\text{RB}}^N - Q_N\right| \leq \mathcal{E}_B(N) + \mathcal{E}_Q(\text{TOL}_{\text{ASGQ}}, N),$$

  where $\mathcal{E}_Q$ is the quadrature error, $\mathcal{E}_B$ is the bias, TOL_{ASGQ} is a user selected tolerance for ASGQ method.

- When using randomized QMC or MC estimator, $Q_N^{\text{MC (QMC)}}$

$$\mathcal{E}_{\text{tot}} \leq \left|C_{\text{RB}} - C_{\text{RB}}^N\right| + \left|C_{\text{RB}}^N - Q_N^{\text{MC (QMC)}}\right| \leq \mathcal{E}_B(N) + \mathcal{E}_S(M, N),$$

  where $\mathcal{E}_S$ is the statistical error, M is the number of samples used for MC or randomized QMC method.

- M^{QMC} and M^{MC}, are chosen so that $\mathcal{E}_{S,\text{QMC}}(M^{\text{QMC}})$ and $\mathcal{E}_{S,\text{MC}}(M^{\text{MC}})$ satisfy

$$\mathcal{E}_{S,\text{QMC}}(M^{\text{QMC}}) = \mathcal{E}_{S,\text{MC}}(M^{\text{MC}}) = \mathcal{E}_B(N) = \frac{\mathcal{E}_{\text{tot}}}{2}.$$

Relative errors and computational gains

Table 2: In this table, we highlight the computational gains achieved by ASGQ and QMC over MC method to meet a certain error tolerance. We note that the ratios are computed for the best configuration with Richardson extrapolation for each method. The ratios (ASGQ/MC) and (QMC/MC) are referred to CPU time ratios.

Parameters	**Relative error**	**(ASGQ/MC)**	**(QMC/MC)**
Set 1	1%	7%	10%
Set 2	0.2%	5%	1%
Set 3	0.4%	4%	5%
Set 4	2%	20%	10%

Computational work of the MC method with different configurations

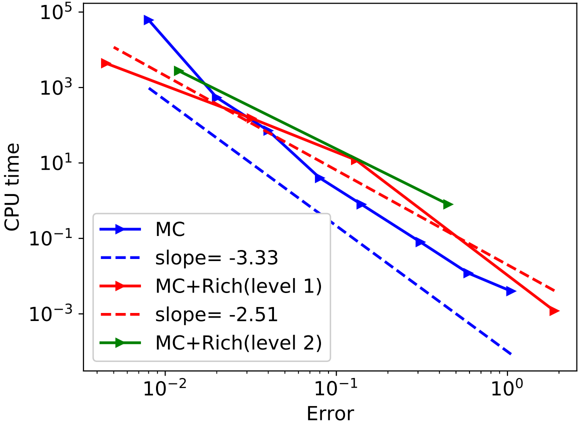


Figure 3.1: Computational work of the MC method with the different configurations in terms of Richardson extrapolation 's level. Case of parameter set 1 in Table 1.

Computational work of the QMC method with different configurations

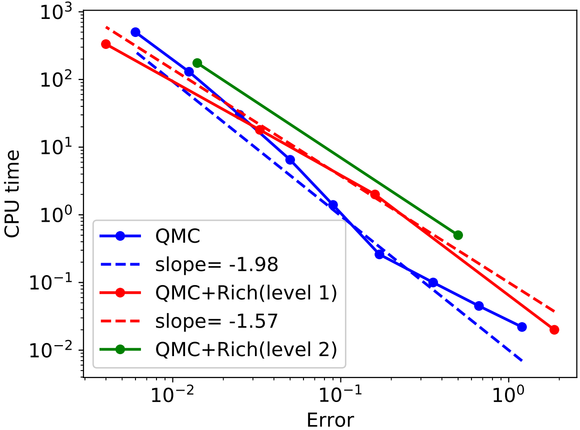


Figure 3.2: Computational work of the QMC method with the different configurations in terms of Richardson extrapolation 's level. Case of parameter set 1 in Table 1.

Computational work of the ASGQ method with different configurations

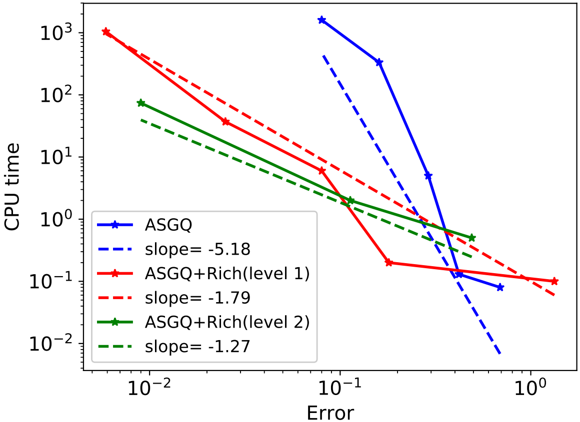


Figure 3.3: Computational work of the ASGQ method with the different configurations in terms of Richardson extrapolation 's level. Case of parameter set 1 in Table 1.

Computational work of the different methods with their best configurations

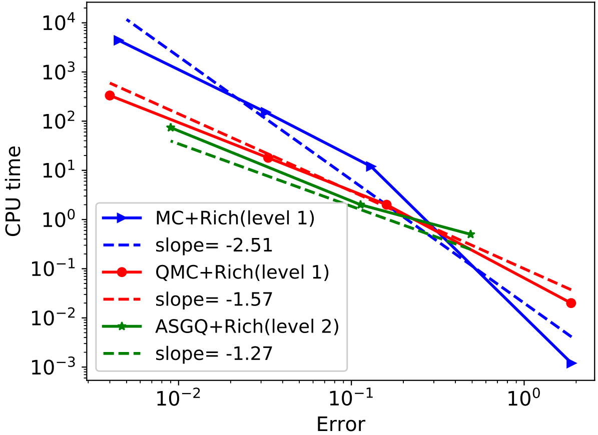


Figure 3.4: Computational work comparison of the different methods with the best configurations, for the case of parameter set 1 in Table 1.

Conclusions

- Proposed novel fast option pricers, for options whose underlyings follow the rough Bergomi model, based on
  - Conditional expectations for numerical smoothing.
  - Hierarchical deterministic quadrature methods.
- Given a sufficiently small relative error tolerance, our proposed methods demonstrate substantial computational gains over the standard MC method, for different parameter constellations.
  $\Rightarrow$ Huge cost reduction when calibrating under the rough Bergomi model.
- Accelerating our novel methods can be achieved by using better QMC or ASGQ methods.
- More details can be found in
  Christian Bayer, Chiheb Ben Hammouda, and Raul Tempone. "Hierarchical adaptive sparse grids and quasi Monte Carlo for option pricing under the rough Bergomi model". In: *arXiv preprint arXiv:1812.08533* (2018).

References I

Christian Bayer, Peter Friz, and Jim Gatheral. “Pricing under rough volatility”. In: *Quantitative Finance* 16.6 (2016), pp. 887–904.

Christian Bayer, Chiheb Ben Hammouda, and Raul Tempone. “Hierarchical adaptive sparse grids and quasi Monte Carlo for option pricing under the rough Bergomi model”. In: *arXiv preprint arXiv:1812.08533* (2018).

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. “Hybrid scheme for Brownian semistationary processes”. In: *Finance and Stochastics* 21.4 (2017), pp. 931–965.

Hans-Joachim Bungartz and Michael Griebel. “Sparse grids”. In: *Acta numerica* 13 (2004), pp. 147–269.

References II

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. "Volatility is rough". In: *Quantitative Finance* 18.6 (2018), pp. 933–949.

Abdul-Lateef Haji-Ali et al. "Multi-index stochastic collocation for random PDEs". In: *Computer Methods in Applied Mechanics and Engineering* (2016).

Ryan McCrickerd and Mikko S Pakkanen. "Turbocharging Monte Carlo pricing for the rough Bergomi model". In: *Quantitative Finance* (2018), pp. 1–10.

Andreas Neuenkirch and Taras Shalaiko. "The Order Barrier for Strong Approximation of Rough Volatility Models". In: *arXiv preprint arXiv:1606.03854* (2016).

Dirk Nuyens. *The construction of good lattice rules and polynomial lattice rules.* 2014.

References III

Marc Romano and Nizar Touzi. "Contingent claims and market completeness in a stochastic volatility model". In: *Mathematical Finance* 7.4 (1997), pp. 399–412.

Ian H Sloan. "Lattice methods for multiple integration". In: *Journal of Computational and Applied Mathematics* 12 (1985), pp. 131–143.

Denis Talay and Luciano Tubaro. "Expansion of the global error for numerical schemes solving stochastic differential equations". In: *Stochastic analysis and applications* 8.4 (1990), pp. 483–509.

Thank you for your attention