Hierarchical adaptive sparse grids and Quasi Monte Carlo for option pricing under the rough Bergomi model

Christian Bayer[1], Chiheb Ben Hammouda[2], and Raul Tempone[3]

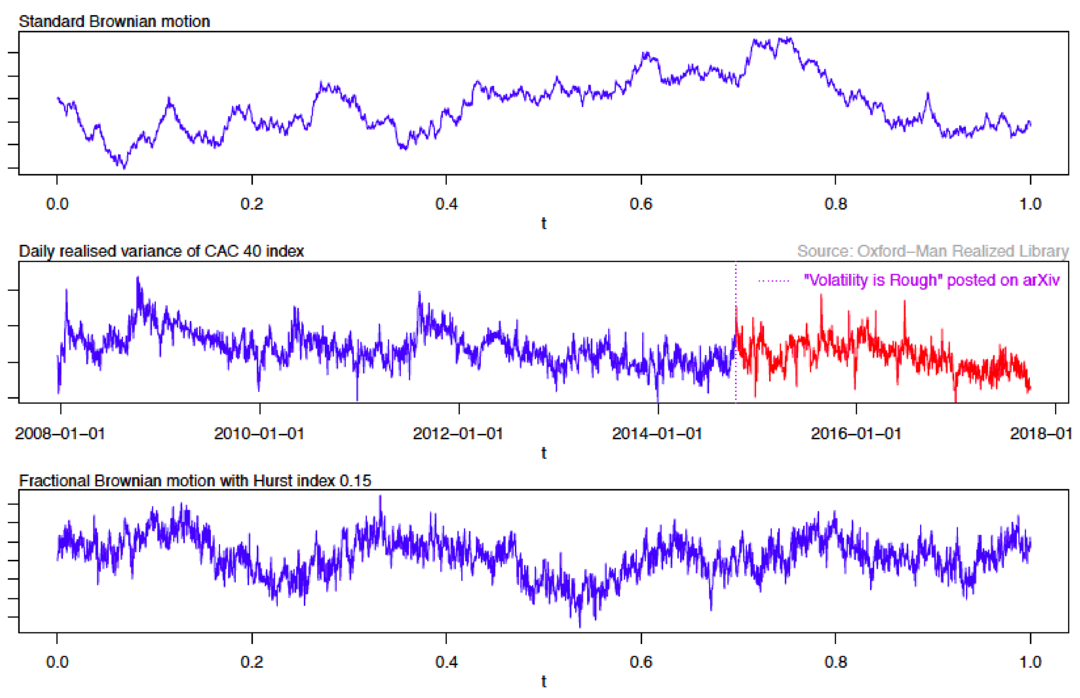
[1]Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Berlin, Germany.
[2]King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

[3]RWTH Aachen University, Germany.

Abstract

The rough Bergomi (rBergomi) model, introduced in [1], is a promising rough volatility model in quantitative finance. In the absence of analytical European option pricing methods for the model, and due to the non-Markovian nature of the fractional driver, the prevalent option is to use Monte Carlo (MC) simulation for pricing. Despite recent advances in the MC method in this context, pricing under the rBergomi model is still a time-consuming task. To overcome this issue, we design a novel, alternative, hierarchical approach, based on i) adaptive sparse grids quadrature (ASGQ), specifically using the same construction in [6], and ii) Quasi Monte Carlo (QMC). Both techniques are coupled with Brownian bridge construction and Richardson extrapolation. By uncovering the available regularity, our hierarchical methods demonstrates substantial computational gains with respect to the standard MC method, when reaching a sufficiently small error tolerance in the price estimates across different parameter constellations, even for very small values of the Hurst parameter.

Rough Volatility



The rough Bergomi Model [1]

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

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

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

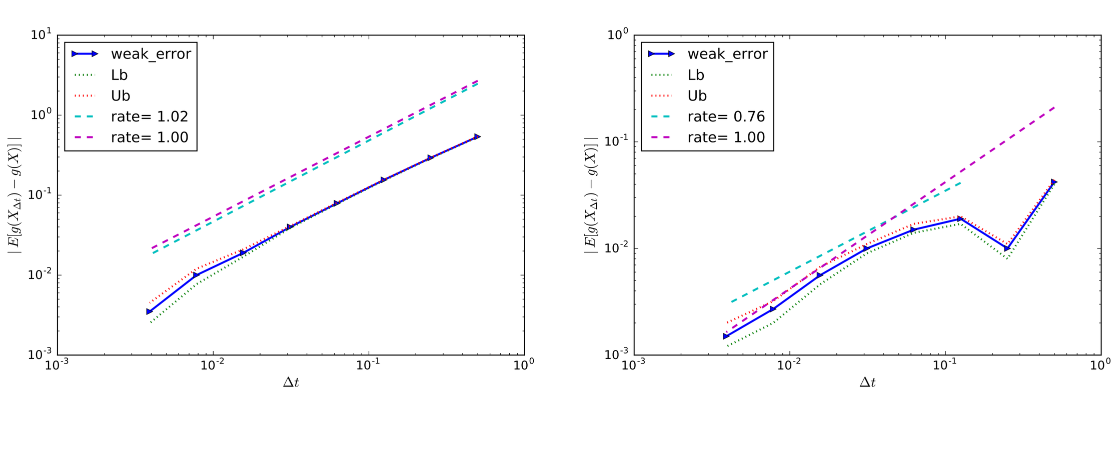
Challenges

- **Numerically:**
  - The model is non-affine and non-Markovian $\Rightarrow$ Standard numerical methods (PDEs, characteristic functions) seem inapplicable.
  - The only prevalent pricing method for mere vanilla options is Monte Carlo [1, 2, 7], still a time consuming task.
  - Discretization methods have poor behavior of the strong error, that is the convergence rate is of order of $H \in [0, 1/2]$ [8] $\Rightarrow$ Variance reduction methods, such as MLMC, are inefficient for very small values of H.
- **Theoretically:**
  - No proper weak error analysis done in the rough volatility context.

Contributions

1. We design an alternative hierarchical efficient pricing method based on:
   i) Analytic smoothing to uncover available regularity.
   ii) Approximating the option price using a deterministic quadrature method (ASGQ and QMC) coupled with Brownian bridges and Richardson Extrapolation.
2. Our hierarchical methods demonstrate substantial computational gains with respect to the standard MC method, assuming a sufficiently small relative error tolerance in the price estimates, even for small values of, H.

On the Choice of the Simulation Scheme



(a) (b)

Figure 2: *The convergence of the weak error $\mathcal{E}_B$, using MC with 6×10^6 samples, for Set 1 parameter in Table 1. The upper and lower bounds are 95% confidence intervals. a) With the hybrid scheme b) With the exact scheme.*

The Hybrid Scheme [4]

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

- 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.

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

  - N is the number of time steps
  - $\{W_j^2\}_{j=1}^N$: Artificially introduced N Gaussian random variables that are used for left-rule points in the hybrid scheme.
  - $b_k = \left(\frac{k^{H+\frac{1}{2}} - (k-1)^{H+\frac{1}{2}}}{H+\frac{1}{2}}\right)^{\frac{1}{H-\frac{1}{2}}}$.

The rough Bergomi Model: Analytic Smoothing

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

- $C_{\text{BS}}(S_0, k, \sigma^2)$ denotes 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

A quadrature estimate of $E[F]$ is

$$\mathcal{M}_{\mathcal{I}_\ell}[F] = \sum_{\boldsymbol{\beta} \in \mathcal{I}_\ell} \Delta[F_{\boldsymbol{\beta}}], \quad (\Delta: \text{mixed difference operator}) \tag{3}$$

- Product approach: $\mathcal{I}_\ell = \{\max\{\beta_1, \ldots, \beta_d\} \leq \ell;\ \boldsymbol{\beta} \in \mathbb{N}_+^d\}$
- Regular SG: $\mathcal{I}_\ell = \{|\ \boldsymbol{\beta}\ |_1 \leq \ell + d - 1;\ \boldsymbol{\beta} \in \mathbb{N}_+^d\}$
- ASGQ based on same construction as in [6]: $\mathcal{I}_\ell = \mathcal{I}^{\text{ASGQ}}$.

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**: $\Delta \mathcal{W}_{\boldsymbol{\beta}} = \text{Work}[\mathcal{M}^{\mathcal{I} \cup \{\boldsymbol{\beta}\}}] - \text{Work}[\mathcal{M}^{\mathcal{I}}]$

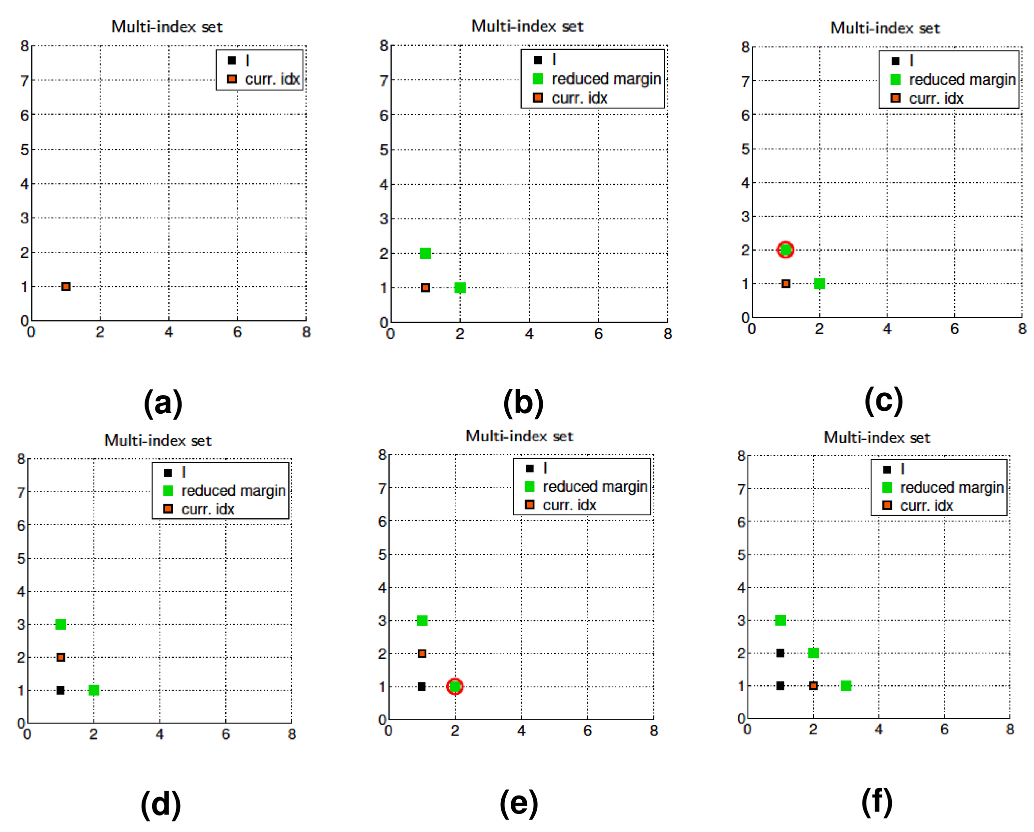


Figure 3: *Construction of the index set for ASGQ method. A posteriori, adaptive construction: Given an index set $\mathcal{I}_k$, compute the profits of the neighbor indices and select the most profitable one.*

Error Comparison

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

- When using ASGQ estimator, Q_N

$$\mathcal{E}_{\text{tot}} \leq \left|C_{\text{RB}} - C_{\text{RB}}^N\right| + \left|C_{\text{RB}}^N - Q_N\right| \leq \mathcal{E}_B(N) + \mathcal{E}_Q(\text{TOL}_{\text{ASGQ}}, N), \tag{4}$$

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, $Q_N^{\text{MC (QMC)}}$

$$\mathcal{E}_{\text{tot}} \leq \left|C_{\text{RB}} - C_{\text{RB}}^N\right| + \left|C_{\text{RB}}^N - Q_N^{\text{MC (QMC)}}\right| \leq \mathcal{E}_B(N) + \mathcal{E}_S(M, N), \tag{5}$$

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

- The number of samples, M^{QMC} and M^{MC}, are chosen so that the statistical errors of QMC, $\mathcal{E}_{S,\text{QMC}}(M^{\text{QMC}})$, and MC, $\mathcal{E}_{S,\text{MC}}(M^{\text{MC}})$, satisfy

$$\mathcal{E}_{S,\text{QMC}}(M^{\text{QMC}}) = \mathcal{E}_{S,\text{MC}}(M^{\text{MC}}) = \mathcal{E}_B(N) = \frac{\mathcal{E}_{\text{tot}}}{2}, \tag{6}$$

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
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 $(5.6e{-}05)$
Set 2: $H = 0.02, K = 1, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.1246 $(9.0e{-}05)$
Set 3: $H = 0.02, K = 0.8, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.2412 $(5.4e{-}05)$
Set 4: $H = 0.02, K = 1.2, S_0 = 1, T = 1, \rho = -0.7, \eta = 0.4, \xi_0 = 0.1$	0.0570 $(8.0e{-}05)$

- The first set is the closest to the empirical findings [5, 3], suggesting that $H \approx 0.1$. The choice of values $\nu = 1.9$ and $\rho = -0.9$ is justified by [1].
- For the remaining three sets, we wanted to test the potential of our method for a very rough case, where variance reduction methods are inefficient.

Relative Errors and Computational Gains of the Different Methods.

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.*

Parameter set	Total relative error	CPU time ratio (MC/ASGQ)	CPU time ratio (MC/QMC)
Set 1	1%	15	10
Set 2	0.2%	21.5	73.2
Set 3	0.4%	26.7	21.3
Set 4	2%	5	10

Numerical Complexity of the Different Methods with the Different Configurations

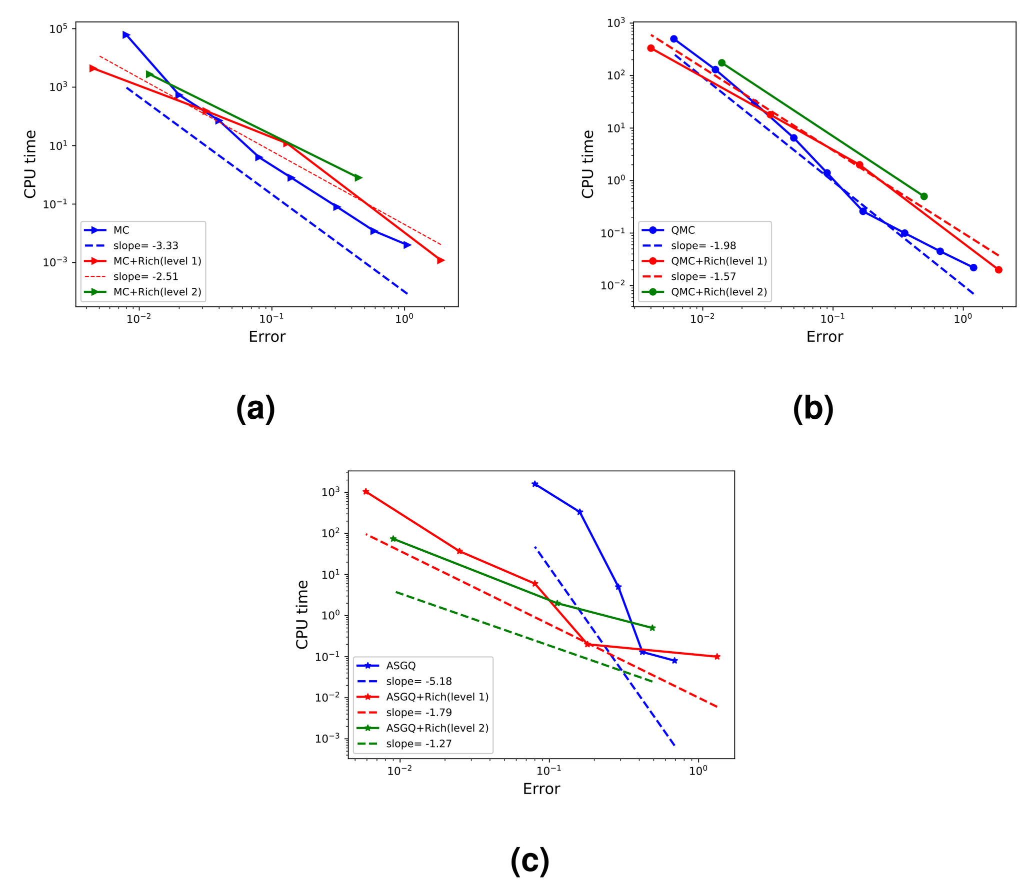


Figure 4: *Comparing the numerical complexity of the different methods with the different configurations in terms of the level of Richardson extrapolation, for the case of parameter set 1 in Table 1. a) MC methods. b) QMC methods. d) ASGQ methods.*

Comparing the Numerical Complexity of the Best Configurations

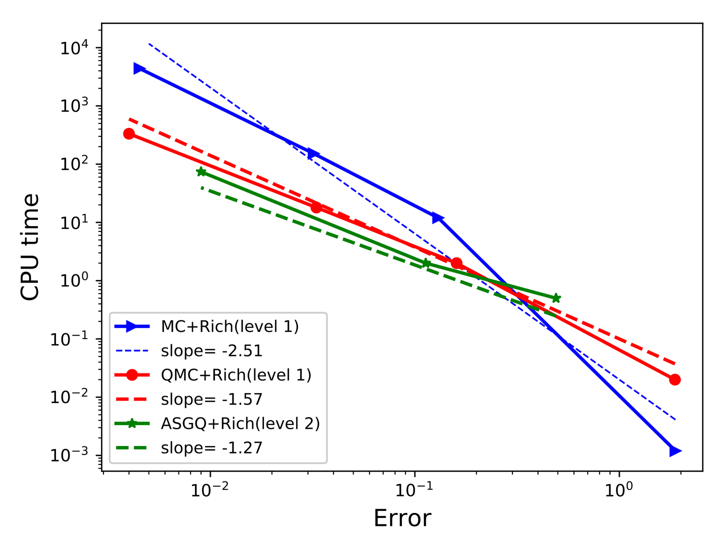


Figure 5: *Computational work comparison for the different methods with the best configurations concluded from Figure 4, for the case of parameter set 1 in Table 1.*

Acknowledgements

C. Bayer gratefully acknowledges support from the German Research Foundation (DFG, grant BA5484/1). This work was supported by the KAUST Office of Sponsored Research (OSR) under Award No. URF/1/2584-01-01 and the Alexander von Humboldt Foundation. C. Ben Hammouda and R. Tempone are members of the KAUST SRI Center for Uncertainty Quantification.

References

[1] C. Bayer, P. Friz, and J. Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016.

[2] C. Bayer, P. K. Friz, P. Gassiat, J. Martin, and B. Stemper. A regularity structure for rough volatility. *arXiv preprint arXiv:1710.07481*, 2017.

[3] M. Bennedsen, A. Lunde, and M. S. Pakkanen. Decoupling the short-and long-term behavior of stochastic volatility. *arXiv preprint arXiv:1610.00332*, 2016.

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

[5] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[6] A.-L. Haji-Ali, F. Nobile, L. Tamellini, and R. Tempone. Multi-index stochastic collocation for random pdes. *Computer Methods in Applied Mechanics and Engineering*, 306:95–122, 2016.

[7] R. McCrickerd and M. S. Pakkanen. Turbocharging Monte Carlo pricing for the rough Bergomi model. *Quantitative Finance*, pages 1–10, 2018.

[8] A. Neuenkirch and T. Shalaiko. The order barrier for strong approximation of rough volatility models. *arXiv preprint arXiv:1606.03854*, 2016.