Numerical Smoothing with Multilevel Monte Carlo for Efficient Option Pricing and Density Estimation

Raúl Tempone

RWTH AACHEN UNIVERSITY

Christian Bayer

WIAS

Chiheb Ben Hammouda

KAUST
King Abdullah University of Science and Technology

MCQMC Conference 2020, University of Oxford
10 - 14 August, 2020

Outline

1. Introduction and Motivation
2. Numerical Smoothing Combined with MLMC for Efficient Option Pricing and Density Estimation
3. Numerical Experiments and Results
4. Conclusions

Options and Pricing

- **Option**: Financial security that gives the holder **the right**, but not the obligation, to buy (Call) or sell (Put) a **specified quantity** of a **specified underlying instrument** (asset) at a **specified price** (K: strike) on (European) or before (Bermudan, American) a **specified date** (T: maturity).
- **Why using options**
  - **Hedging purposes**: Effective hedge instrument against a declining stock market to limit downside losses.
  - **Speculative purposes** such as wagering on the direction of a stock.
- To **value** (price) the option is to compute the fair price of this contract.

Martingale Representation and Notation

Theorem (Fair Value of Financial Derivatives)

The fair value of a financial derivative which can be exercised at time T is given by (Harrison and Pliska 1981)

$$V(S,0) = e^{-rT} E_{\mathbb{Q}}[g(S_T)]$$

where $E_{\mathbb{Q}}$ is the expectation under the local martingale measure $\mathbb{Q}$.

- $\{\mathbf{S}_t \in \mathbb{R}^d : t \geq 0\}$ stochastic processes that represents the prices of the underlying assets at time t, defined on a continuous-time probability space $(\Omega, \mathcal{F}, \mathbb{Q})$.
- r is the risk-free interest rate
- Payoff function $g : \mathbb{R}^d \to \mathbb{R}$. E.g., European call option, $g(S_T) = \max\{S_T - K, 0\}$, where K is the strike price
- d is the number of assets

Motivation

- Issue: The payoff function g has typically low regularity
  $\Rightarrow$ ☹ Multilevel Monte Carlo (MLMC) suffers from
  - Lower strong rate of convergence $\Rightarrow$ Deteriorates the numerical complexity of the MLMC estimator.
  - High kurtosis at the deep levels $\Rightarrow$ Large cost to get reliable and robust estimates of the sample mean and sample variance.[1]
- Solution: In (Bayer, Ben Hammouda, and Tempone 2020), we develop novel methods to effectively address this issue.
  1. **Numerical smoothing**: to uncover the available regularity
     - ★ **Root finding** for determining the discontinuity location.
     - ★ **Pre-Integration (Conditional expectation)**: one dimensional integration with respect to a single well chosen variable.
  2. Approximating the resulting integral of the smoothed integrand obtained from previous step using the **MLMC estimator**.

[1] The standard deviation of the sample variance for $Y_\ell := g_\ell - g_{\ell-1}$ is $\sigma_{\mathcal{S}^2(Y_\ell)} = \frac{\text{Var}[Y_\ell]}{\sqrt{M_\ell}} \sqrt{(\kappa_\ell - 1) + \frac{2}{M_\ell - 1}}$ with κ_ℓ: the kurtosis, and M_ℓ: number of samples.

Framework in (Bayer, Ben Hammouda, and Tempone 2020)

- Approximate efficiently $\mathrm{E}[g(\mathbf{X}(T))] = \mathrm{E}[\mathrm{E}[g(\mathbf{X}(T))|B]] = \mathrm{E}[H(B)]$ using that
  - $H(B)$ is better behaved (B is a Brownian bridge).
  - $H(B)$ is approximated numerically.
- The payoff $g : \mathbb{R}^d \to \mathbb{R}$ has either jumps or kinks. Given $\phi : \mathbb{R}^d \to \mathbb{R}$
  - Hockey-stick functions: $g(\mathbf{x}) = \max(\phi(\mathbf{x}), 0)$ (put or call payoffs).
  - Indicator functions: $g(\mathbf{x}) = \mathbf{1}_{(\phi(\mathbf{x}) \geq 0)}$ (digital option, distribution functions).
  - Dirac Delta functions: $g(\mathbf{x}) = \delta_{(\phi(\mathbf{x}) = 0)}$ (density estimation, financial Greeks).
- The process $\mathbf{X}$ is approximated (via a time stepping scheme) by $\overline{\mathbf{X}}$. Typical examples:
  - One/multi-dimensional geometric Brownian motion (GBM) process.
  - Multi-dimensional stochastic volatility model: the Heston model

$$
\begin{aligned}
dX_t &= \mu X_t dt + \sqrt{v_t} X_t dW_t^X \\
dv_t &= \kappa(\theta - v_t) dt + \xi \sqrt{v_t} dW_t^v,
\end{aligned}
$$

(W_t^X, W_t^v): correlated Wiener processes with correlation ρ.

Contributions in the Context of MLMC Methods [2]

1. Compared to (Giles 2008b; Giles, Debrabant, and Rößler 2013), our approach can be applied when analytic smoothing is not possible.

2. Compared to the case without smoothing
   - We significantly reduce the kurtosis at the deep levels of MLMC
   - We improve the strong convergence rate ⇒ improvement of MLMC complexity from $\mathcal{O}\left(\mathrm{TOL}^{-2.5}\right)$ to $\mathcal{O}\left(\mathrm{TOL}^{-2}\log(\mathrm{TOL})^2\right)$

     ⚠ without the need to use higher order schemes such as Milstein scheme as in (Giles 2008b; Giles, Debrabant, and Rößler 2013)

3. Contrary to (Giles, Nagapetyan, and Ritter 2015), our numerical smoothing approach
   - Does not deteriorate the strong convergence behavior.
   - Easier to apply for any dynamics and QoI (no prior knowledge of the degree of smoothness of the integrand).
   - When estimating densities: our pointwise error does not increase exponentially with respect to the dimension of state vector.

[2] Christian Bayer, Chiheb Ben Hammouda, and Raúl Tempone. "Numerical smoothing and hierarchical approximations for efficient option pricing and density estimation". In: *arXiv preprint arXiv:2003.05708* (2020).

Optimal Smoothing Direction in Continuous Time (I)

- $\mathbf{X} := (X^{(1)}, \ldots, X^{(d)})$ is described by the following SDE

$$dX_t^{(i)} = a_i(\mathbf{X}_t)dt + \sum_{j=1}^{d} b_{ij}(\mathbf{X}_t)dW_t^{(j)}, \tag{1}$$

  $\{W^{(j)}\}_{j=1}^{d}$ are standard Brownian motions.
- Hierarchical representation of $\mathbf{W}$

$$\begin{aligned} W^{(j)}(t) &= \frac{t}{T} W^{(j)}(T) + B^{(j)}(t) \\ &= \frac{t}{\sqrt{T}} Z_j + B^{(j)}(t), \end{aligned}$$

  $Z_j \sim \mathcal{N}(0,1)$ (iid coarse factors); $\{B^{(j)}\}_{j=1}^{d}$: iid Brownian bridges.
- Hierarchical representation of $\mathbf{Z} := (Z_1, \ldots, Z_d)$; $\mathbf{v}$: the smoothing direction

$$\begin{aligned} \mathbf{Z} &= \underbrace{P_0 \mathbf{Z}}_{\text{One dimensional projection}} + \underbrace{P_\perp \mathbf{Z}}_{\text{Projection on the complementary}} \\ &= \underbrace{(\mathbf{Z}, \mathbf{v})}_{:= Z_v} \mathbf{v} + \mathbf{w} \end{aligned} \tag{2}$$

Optimal Smoothing Direction in Continuous Time (II)

- Using (1) and (2), observe $(H_{\mathbf{v}}(Z_v, \mathbf{w}) := g(\mathbf{X}(T)))$

$$
\begin{aligned}
\mathrm{E}[g(\mathbf{X}(T))] &= \mathrm{E}[\mathrm{E}[H_{\mathbf{v}}(Z_v, \mathbf{w}) \mid \mathbf{w}]] \\
\mathrm{Var}[g(\mathbf{X}(T))] &= \mathrm{E}[\mathrm{Var}[H_{\mathbf{v}}(Z_v, \mathbf{w}) \mid \mathbf{w}]] + \mathrm{Var}[\mathrm{E}[H_{\mathbf{v}}(Z_v, \mathbf{w}) \mid \mathbf{w}]].
\end{aligned}
$$

- The optimal smoothing direction, $\mathbf{v}$, solves

$$
\max_{\substack{\mathbf{v} \in \mathbb{R}^d \\ ||\mathbf{v}||=1}} \mathrm{E}[\mathrm{Var}[H_{\mathbf{v}}(Z_v, \mathbf{w}) \mid \mathbf{w}]] \iff \min_{\substack{\mathbf{v} \in \mathbb{R}^d \\ ||\mathbf{v}||=1}} \mathrm{Var}[\mathrm{E}[H_{\mathbf{v}}(Z_v, \mathbf{w}) \mid \mathbf{w}]]. \quad (3)
$$

- ☹ Solving the optimization problem (3) is a hard task.
- ☹ The optimal smoothing direction $\mathbf{v}$ is problem dependent.
- In (Bayer, Ben Hammouda, and Tempone 2020), we determine $\mathbf{v}$ heuristically giving the structure of problem at hand.

Discrete Time Formulation: GBM Example

- Consider the basket option under multi-dimensional GBM model
  - The payoff function: $g(\mathbf{X}(T)) = \max\left(\sum_{j=1}^{d} c_j X^{(j)}(T) - K, 0\right)$
  - The dynamics of the stock prices: $dX_t^{(j)} = \sigma^{(j)} X_t^{(j)} dW_t^{(j)}$.
- The numerical approximation of $\{X^{(j)}(T)\}_{j=1}^{d}$, with time step Δt

$$\overline{X}^{(j)}(T) = X_0^{(j)} \prod_{n=0}^{N-1} \left[1 + \frac{\sigma^{(j)}}{\sqrt{T}} Z_1^{(j)} \Delta t + \sigma^{(j)} \Delta B_n^{(j)}\right], \quad 1 \le j \le d$$

  - $(Z_1^{(1)}, \dots, Z_N^{(1)}, Z_1^{(d)}, \dots, Z_N^{(d)})$: $N \times d$ Gaussian independent rdvs.
  - $\{\mathbf{B}^{(j)}\}_{j=1}^{d}$ are the Brownian bridges increments.
-

$$\begin{aligned} \mathrm{E}\left[g(\mathbf{X}(T))\right] &\approx \mathrm{E}\left[g\left(\overline{X}_T^{(1)}, \dots, \overline{X}_T^{(d)}\right)\right] := \mathrm{E}\left[g(\overline{\mathbf{X}}^{\Delta t}(T))\right] \\ &= \int_{\mathbb{R}^{d\times N}} G(\mathbf{z}) \rho_{N\times d}(\mathbf{z}) dz_1^{(1)} \dots dz_N^{(1)} \dots z_1^{(d)} \dots dz_N^{(d)}, \end{aligned}$$

  - $\mathbf{z} = (z_1^{(1)}, \dots, z_N^{(1)}, \dots, z_1^{(d)}, \dots, z_N^{(d)})$.
  - $\rho_{d\times N}(\mathbf{z}) = \frac{1}{(2\pi)^{d\times N/2}} e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}$.

Numerical Smoothing: Motivation

- **Idea:** Assume that the integration domain Ω can be divided into two parts Ω_1 and Ω_2.
  - In Ω_1 the integrand G is smooth and positive.
  - $G(x) = 0$ in Ω_2.
  - Along the boundary between Ω_1 and Ω_2, the integrand is non-differentiable or discontinuous.
- **Procedure**
  1. Determine Ω_2 numerically by root finding algorithm.
  2. Compute

$$\int_{\Omega} G = \int_{\Omega_1} G + \int_{\Omega_2} G$$
$$= \int_{\Omega_1} G$$

Numerical Smoothing Step

- We consider $\mathbf{Z}_1 = (Z_1^{(1)}, \dots, Z_1^{(d)})$ the most important directions.
- Design of a sub-optimal smoothing direction ($\mathcal{A}$: rotation matrix and it is payoff dependent[3] but easy to construct.)

$$\mathbf{Y} = \mathcal{A}\mathbf{Z}_1.$$

- The smoothing direction $\mathbf{v}$ (in continuous time formulation) is given by the first row of $\mathcal{A}$.
- One dimensional root finding problem to solve for y_1^*

$$K = \sum_{j=1}^{d} c_j X_0^{(j)} \prod_{n=0}^{N-1} F_n^{(j)}(y_1^*(K)), \mathbf{y}_{-1}), \tag{4}$$

$$F_n^{(j)}(y_1^*, \mathbf{y}_{-1}) = 1 + \frac{\sigma^{(j)}\Delta t}{\sqrt{T}} \left(((A)^{-1})_{j1} y_1^* + \sum_{i=2}^{d} ((A)^{-1})_{ji} y_i \right) + \sigma^{(j)} \Delta B_n^{(j)}.$$

[3] In this example, a sufficiently good choice of $\mathcal{A}$ is a rotation matrix with first row leading to $Y_1 = \sum_{i=1}^{d} Z_1^{(i)}$ up to re-scaling.

Numerical Smoothing and Pre-Integration Steps

$$\mathrm{E}\left[g(\mathbf{X}(T))\right] \approx \mathrm{E}\left[g(\overline{\mathbf{X}}^{\Delta t}(T))\right] = \int_{\mathbb{R}^{d\times N}} G(\mathbf{z})\rho_{d\times N}(\mathbf{z}) dz_1^{(1)} \dots dz_N^{(1)} \dots dz_1^{(d)} \dots dz_N^{(d)}$$
$$= \int_{\mathbb{R}^{dN-1}} I(\mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{d-1}(\mathbf{y}_{-1}) d\mathbf{y}_{-1} \rho_{dN-d}(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) d\mathbf{z}_{-1}^{(1)} \dots d\mathbf{z}_{-1}^{(d)}$$
$$= \mathrm{E}\left[I(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)})\right] \approx \mathrm{E}\left[\bar{I}(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)})\right],$$

$$I(\mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) = \int_{\mathbb{R}} G(y_1, \mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{y_1}(y_1) dy_1$$
$$= \int_{-\infty}^{y_1^*} G(y_1, \mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{y_1}(y_1) dy_1 + \int_{y_1^*}^{+\infty} G(y_1, \mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{y_1}(y_1) dy_1$$
$$\approx \bar{I}(\mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) := \sum_{k=0}^{N_q} \eta_k G\left(\zeta_k\left(\bar{y}_1^*\right), \mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}\right), \tag{5}$$

Notation

- G maps $N \times d$ Gaussian random inputs to $g(\overline{\mathbf{X}}^{\Delta t}(T))$;
- $\mathbf{Y} = \mathcal{A}\mathbf{Z}_1$ with $\mathbf{Z}_1 = (Z_1^{(1)}, \dots, Z_1^{(d)})$ the most important directions; $\mathcal{A}$: a problem dependent rotation matrix;
- y_1^*: the exact discontinuity location; $\bar{y}_1^*$: the approximated discontinuity location via root finding;
- N_q: the number of Laguerre quadrature points $\zeta_k \in \mathbb{R}$, and corresponding weights η_k;
- $\rho_{d\times N}(\mathbf{z}) = \frac{1}{(2\pi)^{d\times N/2}} e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}$.

Numerical Smoothing and Pre-Integration Steps

$$
\begin{aligned}
\mathrm{E}\left[g(\mathbf{X}(T))\right] \approx \mathrm{E}\left[g(\overline{\mathbf{X}}^{\Delta t}(T))\right] &= \int_{\mathbb{R}^{d\times N}} G(\mathbf{z})\rho_{d\times N}(\mathbf{z}) dz_1^{(1)}\dots dz_N^{(1)}\dots dz_1^{(d)}\dots dz_N^{(d)} \\
&= \int_{\mathbb{R}^{dN-1}} I(\mathbf{y}_{-1}, \mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) \rho_{d-1}(\mathbf{y}_{-1}) d\mathbf{y}_{-1} \rho_{dN-d}(\mathbf{z}_{-1}^{(1)}, \dots, \mathbf{z}_{-1}^{(d)}) d\mathbf{z}_{-1}^{(1)} \dots d\mathbf{z}_{-1}^{(d)} \\
&= \mathrm{E}\left[I(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)})\right] \approx \mathrm{E}\left[\bar{I}(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)})\right], \qquad (6)
\end{aligned}
$$

In (Bayer, Ben Hammouda, and Tempone 2020) we show

- I and $\bar{I}$ are highly smooth functions.
- Our methodology on how to choose the transformation $\mathcal{A}$ and how to approximate $\bar{I}$.
- How MLMC efficiently approximate the expectation in (6).

Notation

- G maps $N \times d$ Gaussian random inputs to $g(\overline{\mathbf{X}}^{\Delta t}(T))$;
- $\mathbf{Y} = \mathcal{A}\mathbf{Z}_1$ with $\mathbf{Z}_1 = (Z_1^{(1)}, \dots, Z_1^{(d)})$ the most important directions; $\mathcal{A}$: a problem dependent rotation matrix;
- y_1^*: the exact discontinuity location; $\bar{y}_1^*$: the approximated discontinuity location via root finding;
- N_q: the number of Laguerre quadrature points $\zeta_k \in \mathbb{R}$, and corresponding weights η_k;
- $\rho_{d\times N}(\mathbf{z}) = \frac{1}{(2\pi)^{d\times N/2}} e^{-\frac{1}{2}\mathbf{z}^T\mathbf{z}}$.

Density Estimation

- **Goal:** Approximate the density ρ_X at u, for a stochastic process X

$$\rho_X(u) = \mathrm{E}\left[\delta(X-u)\right], \quad \delta \text{ is the Dirac delta function.}$$

⚠ Without any smoothing techniques (regularization, kernel density,...) MC and MLMC fail due to the infinite variance caused by the singularity of the distribution δ.

- **Strategy:** in (Bayer, Ben Hammouda, and Tempone 2020)
  1. Exact conditioning with respect to the Brownian bridge

$$\rho_X(u) = \frac{1}{\sqrt{2\pi}}\mathrm{E}\left[\exp\left(-\left(y_1^*(u)\right)^2/2\right)\left|\frac{dy_1^*}{dx}(u)\right|\right]$$
$$\approx \frac{1}{\sqrt{2\pi}}\mathrm{E}\left[\exp\left(-\left(\bar{y}_1^*(u)\right)^2/2\right)\left|\frac{d\bar{y}_1^*}{dx}(u)\right|\right], \tag{7}$$

     y_1^*: the exact discontinuity (see (4)); $\bar{y}_1^*$: the approximated discontinuity.
  2. We use MLMC method to efficiently approximate (7).

⚠ Kernel density techniques or parametric regularization as in (Giles, Nagapetyan, and Ritter 2015) ⇒ a pointwise error that increases exponentially with respect to the dimension of the state vector X.

Why not Kernel Density Techniques in Multiple Dimensions?

- Similar to approaches based on parametric regularization as in (Giles, Nagapetyan, and Ritter 2015).
- This class of approaches has a pointwise error that increases exponentially with respect to the dimension of the state vector $\mathbf{X}$.
- For a d-dimensional problem, a kernel density estimator with a bandwidth matrix, $\mathcal{H} = \text{diag}(h, \dots, h)$

$$\mathbf{MSE} \approx c_1 M^{-1} h^{-d} + c_2 h^4. \tag{8}$$

  M is the number of samples, and c_1 and c_2 are constants.
- **Our approach in high dimension**: For $\mathbf{u} \in \mathbb{R}^d$

$$\begin{aligned} \rho_{\mathbf{X}}(\mathbf{u}) = \mathrm{E}\left[\delta(\mathbf{X} - \mathbf{u})\right] &= \mathrm{E}\left[\rho_d\left(\mathbf{y}^*(\mathbf{u})\right)\left|\det\left(\mathbf{J}\left(\mathbf{u}\right)\right)\right|\right] \\ &\approx \mathrm{E}\left[\rho_d\left(\overline{\mathbf{y}}^*(\mathbf{u})\right)\left|\det\left(\overline{\mathbf{J}}\left(\mathbf{u}\right)\right)\right|\right], \end{aligned} \tag{9}$$

  $\mathbf{J}$ is the Jacobian matrix, with $\mathbf{J}_{ij} = \frac{\partial y_i^*}{\partial x_j}$; $\rho_d(.)$ is the multivariate Gaussian density.
- Thanks to the exact conditional expectation with respect to the Brownian bridge $\Rightarrow$ the error of our approach only restricted to the error for finding an approximated location of the discontinuity
  $\Rightarrow$ the error in our approach is insensitive to the dimension of the problem.

Multilevel Monte Carlo (MLMC) (Giles 2008a)

- **Aim**: Estimate efficiently $\mathrm{E}\left[\bar{I}(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)})\right]$ in (6).
- **Setting**
  - A hierarchy of nested meshes of $[0,T]$, indexed by $\{\ell\}_{\ell=0}^{L}$.
  - $\Delta t_\ell = K^{-\ell}\Delta t_0$: The time steps size for levels $\ell \geq 1$; $K>1$, $K \in \mathbb{N}$.
  - $\bar{I}_\ell$: the level ℓ approximation of $\bar{I}$, computed with: step size of Δt_ℓ; $N_{q,\ell}$ Laguerre quadrature points; $\mathrm{TOL}_{\mathrm{Newton},\ell}$ as the tolerance of the Newton method at level ℓ
- **MLMC idea**

$$\mathrm{E}\left[\bar{I}_L\right] = \mathrm{E}\left[\bar{I}_0\right] + \sum_{\ell=1}^{L}\mathrm{E}\left[\bar{I}_\ell - \bar{I}_{\ell-1}\right]$$

$$\mathrm{Var}\left[\bar{I}_0\right] \gg \mathrm{Var}\left[\bar{I}_\ell - \bar{I}_{\ell-1}\right] \searrow \text{ as } \ell \nearrow$$

$$M_0 \gg M_\ell \searrow \text{ as } \ell \nearrow$$

- **MLMC estimator**

$$\widehat{Q} := \sum_{\ell=0}^{L}\widehat{Q}_\ell. \tag{10}$$

with

$$\widehat{Q}_0 := \frac{1}{M_0}\sum_{m_0=1}^{M_0}\bar{I}_{0,[m_0]}; \quad \widehat{Q}_\ell := \frac{1}{M_\ell}\sum_{m_\ell=1}^{M_\ell}\left(\bar{I}_{\ell,[m_\ell]} - \bar{I}_{\ell-1,[m_\ell]}\right),\ 1 \leq \ell \leq L$$

MLMC Complexity (Cliffe et al. 2011)

$$\mathcal{O}\left(\mathrm{TOL}^{-2-\max\left(0,\frac{\gamma-\beta}{\alpha}\right)}\log\left(\mathrm{TOL}\right)^{2\times\mathbf{1}_{\{\beta=\gamma\}}}\right) \tag{11}$$

i) Weak rate: $\left|\mathrm{E}\left[g\left(\mathbf{Z}_\ell(T)\right) - g\left(\mathbf{X}(T)\right)\right]\right| \leq c_1 2^{-\alpha\ell}$

ii) Strong rate: $\mathrm{Var}\left[g\left(\mathbf{Z}_\ell(T)\right) - g\left(\mathbf{Z}_{\ell-1}(T)\right)\right] \leq c_2 2^{-\beta\ell}$

iii) Work rate: $W_\ell \leq c_3 2^{\gamma\ell}$ (W_ℓ: expected cost)

Error and Work Discussion for MLMC (I)

- $\hat{Q}$: the MLMC estimator, as defined in (10).
-

$$\begin{aligned} \mathrm{E}\left[g(X(T)\right] - \hat{Q} &= \underbrace{\mathrm{E}\left[g(X(T))\right] - \mathrm{E}\left[g(\overline{\mathbf{X}}^{\Delta t_L}(T))\right]}_{\text{Error I: bias or weak error}} \\ &+ \underbrace{\mathrm{E}\left[I_L\left(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)}\right)\right] - \mathrm{E}\left[\bar{I}_L\left(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)}\right)\right]}_{\text{Error II: numerical smoothing error}} \\ &+ \underbrace{\mathrm{E}\left[\bar{I}_L\left(\mathbf{Y}_{-1}, \mathbf{Z}_{-1}^{(1)}, \dots, \mathbf{Z}_{-1}^{(d)}\right)\right] - \hat{Q}}_{\text{Error III: MLMC statistical error}}. \end{aligned} \tag{12}$$

- Schemes based on forward Euler to simulate asset dynamics

$$\text{Error I} = \mathcal{O}\left(\Delta t_L\right).$$

-

$$\text{Error III} = \sqrt{\sum_{\ell=L_0}^{L} M_\ell^{-1} V_\ell} = \mathcal{O}\left(\sqrt{\sum_{\ell=L_0}^{L} \sqrt{N_{q,\ell} \log\left(\mathrm{TOL}_{\mathrm{Newton},\ell}^{-1}\right)}}\right).$$

⚠ **Notation**: $V_\ell := \mathrm{Var}\left[\bar{I}_\ell - \bar{I}_{\ell-1}\right]$; M_ℓ: number of samples at level ℓ.

Error and Work Discussion for MLMC (II)

$$
\begin{aligned}
\text{Error II} &:= \mathrm{E}\left[I\left(\mathbf{Y}_{-1},\mathbf{Z}_{-1}^{(1)},\dots,\mathbf{Z}_{-1}^{(d)}\right)\right]-\mathrm{E}\left[\bar{I}_L\left(\mathbf{Y}_{-1},\mathbf{Z}_{-1}^{(1)},\dots,\mathbf{Z}_{-1}^{(d)}\right)\right]\\
&\leq \sup_{\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)}}\left|I\left(\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)}\right)-\bar{I}_L\left(\mathbf{y}_{-1},\mathbf{z}_{-1}^{(1)},\dots,\mathbf{z}_{-1}^{(d)}\right)\right|\\
&= \mathcal{O}\left({N_{q,L}}^{-s}\right)+\mathcal{O}\left(\left|y_1^*-\bar{y}_1^*\right|^{\kappa+1}\right)\\
&= \mathcal{O}\left(N_{q,L}^{-s}\right)+\mathcal{O}\left(\mathrm{TOL}_{\mathrm{Newton},L}^{\kappa+1}\right)
\end{aligned}
\tag{13}
$$

- y_1^*: the exact location of the non smoothness.
- $\bar{y}_1^*$: the approximated location of the non smoothness obtained by Newton iteration $\Rightarrow \left|y_1^*-\bar{y}_1^*\right| = \mathrm{TOL}_{\mathrm{Newton},L}$
- $\kappa \geq 0$ ($\kappa = 0$: heavy-side payoff (digital option), and $\kappa = 1$: call or put payoffs).
- $N_{q,L}$ is the number of points used by the Laguerre quadrature, at level L, for the one dimensional pre-integration step.
- $s > 0$: Derivatives of G with respect to y_1 are bounded up to order s.

Error and Work Discussion for MLMC (III)

An optimal performance of MLMC is given by

$$\begin{cases} \min\limits_{(L, L_0, \{M_\ell\}_{\ell=0}^{L}, N_q, \mathrm{TOL}_{\mathrm{Newton}})} \mathrm{Work}_{\mathrm{MLMC}} \propto \sum_{\ell=L_0}^{L} M_\ell \left(N_{q,\ell} \Delta t_\ell^{-1} \right) \\ s.t.\ \mathcal{E}_{\mathrm{total,MLMC}} = \mathrm{TOL}. \end{cases} \tag{14}$$

$$\begin{aligned} \mathcal{E}_{\mathrm{total,MLMC}} &:= \mathrm{E}\left[g(X(T)) \right] - \hat{Q} \\ &= \mathcal{O}\left(\Delta t_L \right) + \mathcal{O}\left(\sqrt{\sum_{\ell=L_0}^{L} \sqrt{N_{q,\ell} \log\left(\mathrm{TOL}_{\mathrm{Newton},\ell}{}^{-1} \right)}} \right) + \mathcal{O}\left(N_{q,L}{}^{-s} \right) \\ &+ \mathcal{O}\left(\mathrm{TOL}_{\mathrm{Newton},L}{}^{\kappa+1} \right). \end{aligned}$$

Numerical Results for MLMC

Method	κ_L	α	β	γ	Numerical Complexity
Without smoothing + digital under GBM	709	1	1/2	1	$\mathcal{O}\left(\text{TOL}^{-2.5}\right)$
With numerical smoothing + digital under GBM	3	1	1	1	$\mathcal{O}\left(\text{TOL}^{-2}\left(\log(\text{TOL})\right)^2\right)$
Without smoothing + digital under Heston	245	1	1/2	1	$\mathcal{O}\left(\text{TOL}^{-2.5}\right)$
With numerical smoothing + digital under Heston	7	1	1	1	$\mathcal{O}\left(\text{TOL}^{-2}\log(\text{TOL})^2\right)$
With numerical smoothing + GBM density	5	1	1	1	$\mathcal{O}\left(\text{TOL}^{-2}\left(\log(\text{TOL})\right)^2\right)$
With numerical smoothing + Heston density	8	1	1	1	$\mathcal{O}\left(\text{TOL}^{-2}\left(\log(\text{TOL})\right)^2\right)$

Table 1: Summary of the MLMC numerical results observed different examples. κ_L is the kurtosis at the deepest levels of MLMC, (α, β, γ) are weak, strong and work rates respectively. TOL is the user-selected MLMC tolerance.

Digital Option under the Heston Model: Without Smoothing

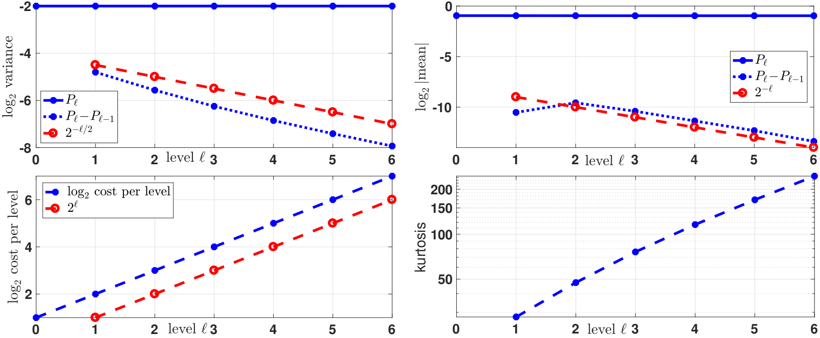

Figure 3.1: Digital option under Heston: Convergence plots of MLMC without smoothing, combined with the fixed truncation scheme.

Digital Option under the Heston Model With Numerical Smoothing

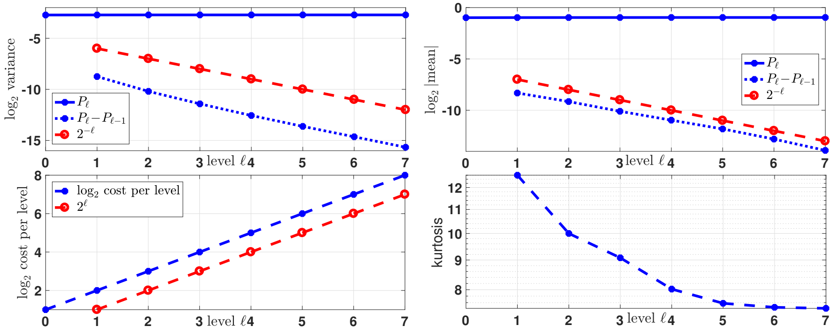

Figure 3.2: Digital option under Heston: Convergence plots for MLMC with numerical smoothing, combined with the Heston OU based scheme.

Digital Option under the Heston Model: Numerical Complexity Comparison

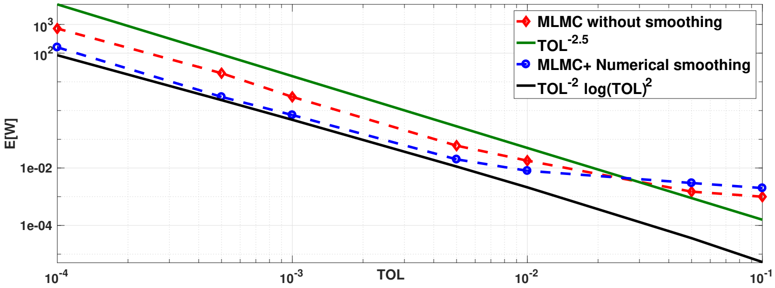

Figure 3.3: Digital option under Heston: Comparison of the numerical complexity of i) standard MLMC (based on fixed truncation scheme), and ii) MLMC with numerical smoothing (based on Heston OU based scheme).

Conclusions

1. In (Bayer, Ben Hammouda, and Tempone 2020), we propose a numerical smoothing approach that can be combined with MLMC estimator for efficient option pricing and density estimation.
2. Our approach
   - is relevant for cases where one can not apply analytic smoothing.
   - improves the strong convergence rate without the need for using higher order schemes as in (Giles 2008b; Giles, Debrabant, and Rößler 2013).
   - can be easily extended to any model dynamics and payoff structure.
   - can be easily extended to computing financial Greeks.
3. In (Bayer, Ben Hammouda, and Tempone 2020), we show the advantages of combining numerical smoothing with deterministic quadrature methods.
4. More details can be found in
   Christian Bayer, Chiheb Ben Hammouda, and Raúl Tempone. "Numerical smoothing and hierarchical approximations for efficient option pricing and density estimation". In: *arXiv preprint arXiv:2003.05708* (2020)

References I

Christian Bayer, Chiheb Ben Hammouda, and Raúl Tempone. “Numerical smoothing and hierarchical approximations for efficient option pricing and density estimation”. In: *arXiv preprint arXiv:2003.05708* (2020).

K Andrew Cliffe et al. “Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients”. In: *Computing and Visualization in Science* 14.1 (2011), p. 3.

Michael Giles. “Multi-level Monte Carlo path simulation”. In: *Operations Research* 53.3 (2008), pp. 607–617.

Michael B Giles. “Improved multilevel Monte Carlo convergence using the Milstein scheme”. In: *Monte Carlo and Quasi-Monte Carlo Methods 2006*. Springer, 2008, pp. 343–358.

References II

Michael B Giles, Kristian Debrabant, and Andreas Rößler. "Numerical analysis of multilevel Monte Carlo path simulation using the Milstein discretisation". In: *arXiv preprint arXiv:1302.4676* (2013).

Michael B Giles, Tigran Nagapetyan, and Klaus Ritter. "Multilevel Monte Carlo approximation of distribution functions and densities". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 267–295.

J Michael Harrison and Stanley R Pliska. "Martingales and stochastic integrals in the theory of continuous trading". In: *Stochastic processes and their applications* 11.3 (1981), pp. 215–260.

Thank you for your attention