Hierarchical Methods for Risk Assessment

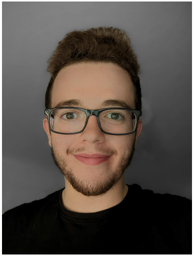
Abdul-Lateef Haji-Ali

Heriot-Watt University & Alexander von Humboldt Foundation

SNIPS, Växjö, Sweden — August 29, 2025

Joint work with



M. Giles [1,2,3] (Oxford)

A. Teckentrup[4] (UoE)

J. Spence [2,4] (UoE)

I. Powell [5] (HWU)

---

[1] Giles and H-A (2019) "Multilevel Nested Simulation for Efficient Risk Estimation", *SIAM/ASA Journal on Uncertainty Quantification*

[2] Giles, H-A, and Spence (2023) *Efficient Risk Estimation for the Credit Valuation Adjustment*,

[3] Giles and H-A (2024) "Multilevel Path Branching for Digital Options", *Annals of Applied Probability*

[4] H-A, Spence, and Teckentrup (2022) "Adaptive Multilevel Monte Carlo for Probabilities", *SIAM Journal on Numerical Analysis*

[5] Croci, H-A, and Powell (2025+) *An Adaptive Sampling Scheme for Level-set Approximation*,

The problem: Risk assessment

$$\mathbb{E}[\, f(X)\, \mathbb{I}_{X\in\Omega}\,]$$

where X is a d-dimensional random variable and $\Omega \subset \mathbb{R}^d$.

The problem: Risk assessment

$$\mathbb{E}[\, f(X)\, \mathbb{I}_{X\in\Omega}\,]$$

where X is a d-dimensional random variable and $\Omega \subset \mathbb{R}^d$.

Three reasons this problem can be challenging:

1. Dimensionality of X and Ω,
2. the complexity of sampling X or computing f,
3. and rarity of event – use (sequential) importance sampling.

The problem: Risk assessment

$$\mathbb{E}[\, f(X)\, \mathbb{I}_{X\in\Omega}\,]$$

where X is a d-dimensional random variable and $\Omega \subset \mathbb{R}^d$.

Three reasons this problem can be challenging:

1. Dimensionality of X and Ω,
2. the complexity of sampling X or computing f,
3. and rarity of event – use (sequential) importance sampling.

Three reasons why this problem can be easy:

1. Structure in approximations of X (and f).
2. Regularity of Ω.
3. Regularity of the density of X.

For the rest of this talk, take $f(X) = 1$ for simplicity.

The examples: Computing probabilities

- Financial risk assessment $X \coloneqq \mathbb{E}[\, Y \,|\, R\,] - \mathrm{MaxLoss}$ (prob. of loss, VaR, CVaR)

$$\mathbb{P}[\,\mathbb{E}[\, Y \,|\, R\,] > \mathrm{MaxLoss}\,]$$

- Digital options $X \coloneqq S(T)$ where S is an asset price satisfying an SDE

$$\mathbb{P}[\, S(T) \in \Omega\,]$$

- Nuclear leakage: $X = u(Y)$ depends on the solution of an advection-dispersion-decay PDE with random porosity Y

$$\mathbb{P}[\, u(Y) \in \Omega\,]$$

The examples: Computing probabilities

- Financial risk assessment $X := \mathbb{E}[\, Y \mid R\,] - \mathrm{MaxLoss}$ (prob. of loss, VaR, CVaR)

$$\mathbb{P}[\,\mathbb{E}[\, Y \mid R\,] > \mathrm{MaxLoss}\,] \approx \mathbb{P}\left[\frac{1}{N}\sum_{i=1}^{N} Y^{(i)}(R) > \mathrm{MaxLoss}\right]$$

- Digital options $X := S(T)$ where S is an asset price satisfying an SDE

$$\mathbb{P}[\, S(T) \in \Omega\,] \approx \mathbb{P}[\, S_h(T) \in \Omega\,]$$

  where S_h is an Euler-Maruyama or Milstein approximations with step size h.

- Nuclear leakage: $X = u(Y)$ depends on the solution of an advection-dispersion-decay PDE with random porosity Y

$$\mathbb{P}[\, u(Y) \in \Omega\,] \approx \mathbb{P}[\, u_h(Y) \in \Omega\,]$$

  where g_h is a Finite Element approximation with grid size/time-step h.

Multilevel Monte Carlo

Focus on computing $\mathbb{E}[\, g(X)\,]$ for some function $g : \mathbb{R}^d \to \mathbb{R}$,

Assumptions

Assume we can approximate $X \approx \overline{X}_\ell$ with $\ell \in \mathbb{N}$

- Work of $\overline{X}_\ell$ is $\propto 2^{\gamma\ell}$.
- Bias: $|\mathbb{E}[\, g(\overline{X}_\ell) - g(X)\,]| \propto 2^{-\alpha\ell}$.
- Variance: $\mathbb{E}[\, \|\overline{X}_\ell - X\|^2\,] \propto 2^{-\beta\ell}$.

Multilevel Monte Carlo (MLMC)

$$\frac{1}{M_0}\sum_{m=1}^{M_0} g(\overline{X}_0^{(0,m)}) + \sum_{\ell=1}^{L} \frac{1}{M_\ell}\sum_{m=1}^{M_\ell} g(\overline{X}_\ell^{(\ell,m)}) - g(\overline{X}_{\ell-1}^{(\ell,m)})$$

Theorem

For Lipschitz g, the overall cost of MLMC for computing $\mathbb{E}[g(X)]$ to accuracy ε using optimal $L, \{M_\ell\}_{\ell=0}^{L}$ is (up to logarithmic terms) is

$$\mathcal{O}\left(\varepsilon^{-2-\max\left(\frac{\gamma-\beta}{\alpha},0\right)}\right).$$

Classical example: $g(x) = \max(x-K, 0)$, Euler-Maruyama approximation, MLMC complexity is $\mathcal{O}\left(\varepsilon^{-2}(\log\varepsilon)^2\right)$ vs $\mathcal{O}\left(\varepsilon^{-3}\right)$ for Monte Carlo.

Discontinuous g: Key assumptions

Assumptions

When $g(x) = \mathbb{I}_{x \in \Omega}$, for some random variable $\sigma_\ell > 0$ and all $\ell \in \mathbb{N}$, assume that

1. There is $q > 2$ such that

$$\left(\mathbb{E}\left[\left(\frac{\|\overline{X}_\ell - X\|}{\sigma_\ell} \right)^q \right] \right)^{1/q} \lesssim 2^{-\beta\ell/2}.$$

2. There is $\delta_0 > 0$ such that for $\delta \le \delta_0$ we have

$$\mathbb{P}\left[\frac{\mathrm{dist}_{\partial\Omega}(\overline{X}_\ell)}{\sigma_\ell} \le \delta \right] \lesssim \delta.$$

MLMC analysis

Lemma

$$\mathrm{Var}[\,\mathbb{I}_{\overline{X}_\ell\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1}\in\Omega}\,] \lesssim 2^{-(\lambda_{q,1}\beta/2)\ell}$$

Proof. $\lambda_{q,c} := \frac{q}{q+c} \uparrow 1$ and $|X - \overline{X}_\ell| \approx \mathcal{O}(2^{-\ell\beta/2})$ □

Theorem

For discontinuous g, the overall cost of MLMC for computing $\mathbb{E}[\,g(X)\,]$ to accuracy ε using optimal $L, \{M_\ell\}_{\ell=0}^{L}$ is (up to logarithmic terms) is

$$\mathcal{O}\left(\varepsilon^{-2-\max\left(\frac{\gamma-\lambda_{q,1}\beta/2}{\alpha},0\right)}\right).$$

Mission: Find better estimators with improved variance convergence. Hopefully easy to apply to a large class of problems, dimensions and approximations, e.g., Euler-Maruyama and Milstein.

Previous research (non-exhaustive)

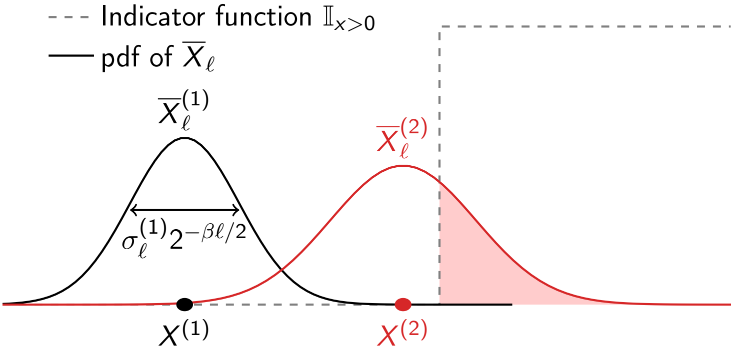
1. **Path splitting:** for SDEs (Glasserman 2003; Burgos and Giles 2012).
2. **Importance Sampling on the difference:** (Xia and Giles 2012).
3. **Explicit Smoothing:** $\mathbb{I}_{x>0} \approx \Phi_L(x)$ (Giles, Nagapetyan, and Ritter 2015) or $\mathbb{I}_{x>0} \approx \Phi_\ell(x)$ (Xu, He, and Wang 2024).
4. **Numerical smoothing:** with per-sample root finding (Achtsis, Cools, and Nuyens 2013; Bayer, Siebenmorgen, and Tempone 2018; Bayer, Ben Hammouda, and Tempone 2024).
5. **Integration by parts using Malliavin calculus:** For SDEs, requires evaluation of derivative (Altmayer and Neuenkirch 2015).
6. **Integration then differentiation:** (Krumscheid and Nobile 2018).
7. **Higher-order approx.:** Quasi-Monte Carlo (Xu, He, and Wang 2024).
8. **Adaptivity in Monte Carlo**,
   - For level-set estimation (with limited analysis), (Min and Gibou 2007).
   - For nested expectations (Broadie, Du, and Moallemi 2011).
   - PDEs with a.s. bounds (Elfverson, Hellman, and Målqvist 2016).

Adaptive Multilevel Monte Carlo: Algorithm

Refine samples of $\overline{X}_\ell$ to $\overline{X}_{\ell+\eta_\ell}$, where $0 \le \eta_\ell \le \lceil \theta\ell \rceil$ is the smallest integer for which

$$\delta_{\ell+\eta_\ell} := \frac{\mathrm{dist}_{\partial\Omega}(\overline{X}_\ell)}{\sigma_\ell} \ge a_{\ell+\eta_\ell}$$

for some $0 \le \theta \le 1$.

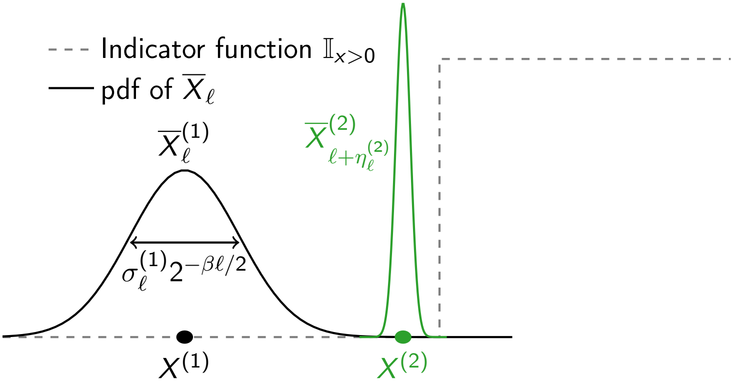


Adaptive Multilevel Monte Carlo: Algorithm

Refine samples of $\overline{X}_\ell$ to $\overline{X}_{\ell+\eta_\ell}$, where $0 \le \eta_\ell \le \lceil \theta\ell \rceil$ is the smallest integer for which

$$\delta_{\ell+\eta_\ell} := \frac{\mathrm{dist}_{\partial\Omega}(\overline{X}_\ell)}{\sigma_\ell} \ge a_{\ell+\eta_\ell}$$

for some $0 \le \theta \le 1$.



Adaptive Multilevel Monte Carlo: Analysis

Theorem (Giles and H-A 2019; H-A, Spence, and Teckentrup 2022)

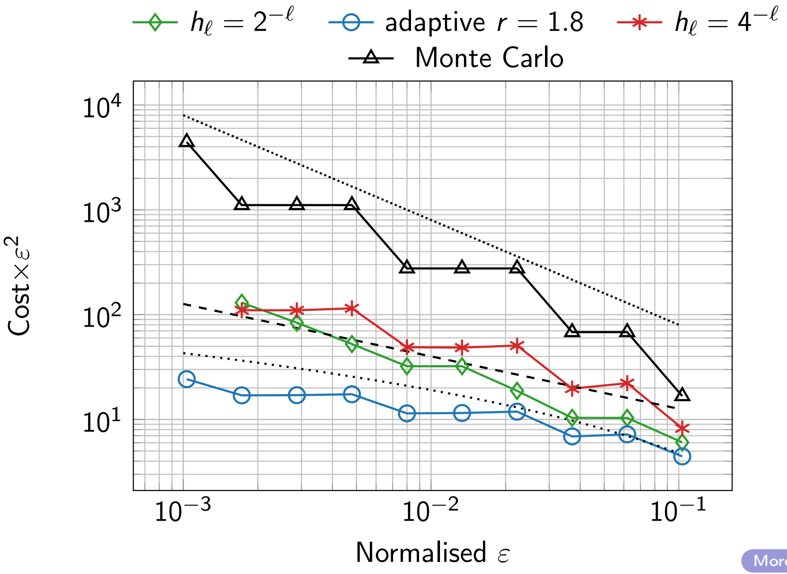
There is $\{a_{\ell+k}\}_{k=0,\dots,\lceil\theta\ell\rceil}$, such that

- The expected work of sampling $\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega}$ is: $W_\ell \propto 2^{\gamma\ell}$.
- The variance is: $\mathrm{Var}[\,\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\,] \propto 2^{-\frac{1+\theta}{2}\lambda_{q,1}\beta\ell}$,

$$\text{where} \qquad \theta = \begin{cases} \left(\frac{2\gamma}{\beta\lambda_{q,1}} - 1\right)^{-1} & \beta < \gamma/\lambda_{q,1} \\ 1 & \beta > \gamma/\lambda_{q,1} \end{cases}.$$

- **Example**: Euler-Maruyama approximation, Adaptive MLMC complexity is $\mathcal{O}\left(\varepsilon^{-2}(\log\varepsilon)^2\right)$ (same as for Lipschitz g) vs $\mathcal{O}\left(\varepsilon^{-5/2}\right)$ for classical MLMC.

Digital option on GBM with E-M: Adaptivity



More

The Good and the Bad of sample adaptivity

The Good:

- Achieves MLMC complexity as if applied to Lipschitz functions.
- Applicable to (almost) any model and approximation method.

The Bad:

- Adaptivity mitigates the discontinuity but does not remove it.
- The main issue: Unlike $\overline{X}_\ell$, higher moments of differences of $\mathbb{I}_{\overline{X}_\ell \in \Omega}$ will have relatively worse convergence.

$$\mathbb{E}\left[\left|\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\right|^q\right]^{1/q} = \mathbb{E}\left[\left|\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\right|^2\right]^{1/q}$$

- For example: the Kurtosis of the difference is

$$\frac{\mathbb{E}\left[\left|\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\right|^4\right]}{\mathbb{E}\left[\left|\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\right|^2\right]^2} = \mathbb{E}\left[\left|\mathbb{I}_{\overline{X}_{\ell+\eta_\ell}\in\Omega} - \mathbb{I}_{\overline{X}_{\ell-1+\eta_{\ell-1}}\in\Omega}\right|^2\right]^{-1}$$

The Good and the Bad of sample adaptivity (cont.)

- While the second moment is decreasing at a good rate (for MLMC), the Kurtosis is increasing at an equally bad rate (for MLMC).
- Kurtosis issue can/must be resolved through algorithmic means.
- Nesting estimators of this kind is limited.
  - Nesting MLMC estimators of Lipschitz functions is surprisingly possible. Assuming good control of higher moments, or using methods that don't require such high moments, i.e., biased MLMC instead of unbiased MLMC.
  - We applied these ideas to computing Credit Valuation Adjustment involving triply nested expectations (Giles, H-A, and Spence 2023).
- Antithetic estimators are not applicable – quantity of interest is not sufficiently smooth.

Other risk measures: VaR

- VaR is defined for a given confidence level $\eta \in (0,1)$, as

$$\mathrm{VaR}_\eta(X) = \inf\{\xi \in \mathbb{R} \,:\, \mathbb{P}[X < \xi] \geq \eta\}.$$

  This can be estimated by root-finding algorithm, with the acceptable error ε of estimating the probability being steadily reduced during the iteration.

Other risk measures: VaR

- VaR is defined for a given confidence level $\eta \in (0, 1)$, as
$$\mathrm{VaR}_\eta(X) = \inf\{\xi \in \mathbb{R} \,:\, \mathbb{P}[X < \xi] \geq \eta\}.$$
This can be estimated by root-finding algorithm, with the acceptable error ε of estimating the probability being steadily reduced during the iteration.
- A different approach is to construct an MLMC version of the stochastic approximation algorithm (Bardou, Frikha, and Pagès 2009)
$$\xi^{(\mathrm{MLMC})} = \xi_\ell^{(N_0)} + \sum_{\ell=1}^{L} \xi_\ell^{(N_\ell)} - \xi_{\ell-1}^{(N_\ell)}$$
$$\xi_\ell^{(n+1)} = \xi_\ell^{(n)} - \gamma_{n+1}\left(1 - \frac{1}{1-\eta}\mathbb{I}_{\overline{X}_\ell^{(n+1)} > \xi^{(n)}}\right)$$
and directly apply adaptive sampling there (Crépey, Frikha, Louzi, and Spence 2024).

Other risk measures: CVaR

Given an estimate $\widetilde{\mathrm{VaR}}_\eta$, CVaR is then (Rockafellar and Uryasev 2002)

$$\begin{aligned}
\mathbb{E}[\,X \,|\, X > \mathrm{VaR}_\eta\,] &= \mathrm{VaR}_\eta + (1-\eta)^{-1}\mathbb{E}[\,\max(0, X - \mathrm{VaR}_\eta)\,] \\
&= \inf_{\xi\in\mathbb{R}} \left\{\xi + (1-\eta)^{-1}\mathbb{E}[\,\max(0, X-\xi)\,]\right\} \\
&= \widetilde{\mathrm{VaR}}_\eta + (1-\eta)^{-1}\mathbb{E}\left[\,\max\left(0, X - \widetilde{\mathrm{VaR}}_\eta\right)\,\right] \\
&\qquad + \mathcal{O}\left(\left(\widetilde{\mathrm{VaR}}_\eta - \mathrm{VaR}_\eta\right)^2\right)
\end{aligned}$$

For ε RMS error, first estimate $\widetilde{\mathrm{VaR}}_\eta$ to accuracy $\mathcal{O}(\varepsilon^{1/2})$ at cost $o(\varepsilon^{-2})$.

Then estimate $\mathbb{E}[\,\max(0, X - \widetilde{\mathrm{VaR}}_\eta)\,]$ to accuracy ε using MLMC + uniform sampling (for a Lipschitz function) – complexity is not affected by the discontinuity.

Other risk measures: Level-set of a function

Let $D \subset\subset \mathbb{R}^d$ be a d-dimensional domain with compact closure and a sufficiently smooth boundary. We are interested in approximating the zero level set of a function f,

$$\mathcal{L}_0 := \{\mathrm{y} \in \overline{D} \,:\, \mathbb{E}[\,X(y)\,] = 0\}$$

for some random function, $X : D \to \mathbb{R}$, which can be evaluated pointwise.

Other risk measures: Level-set of a function

Let $D \subset\subset \mathbb{R}^d$ be a d-dimensional domain with compact closure and a sufficiently smooth boundary. We are interested in approximating the zero level set of a function f,

$$\mathcal{L}_0 := \{\mathrm{y} \in \overline{D} \, : \, \mathbb{E}[\, X(y) \,] = 0\}$$

for some random function, $X : D \to \mathbb{R}$, which can be evaluated pointwise.

For any $y \in \overline{D}$, we can use iid samples $\{X^{(i)}(y)\}_{i=1}^{M_\ell}$,

$$\widetilde{E}_{M_\ell}[X(\mathrm{y})] = \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} X^{(i)}(y),$$

to get a corresponding approximation of $\mathcal{L}_0$.

Adaptive Level-set Computation

Similar to (Min and Gibou 2007), our method is cell-based: for each cell, $\square$, in a grid starting from a uniform refinement of $2^{d\ell_0}$ cells, we

- Select N points in $\square$, say $\mathrm{y}_1^\square, \ldots, \mathrm{y}_N^\square$, deterministically,
- evaluate the approximations $\widetilde{E}_{M_\ell} X\left(\mathrm{y}_1^\square\right), \ldots, \widetilde{E}_{M_\ell} X\left(\mathrm{y}_N^\square\right)$.
- Obtain an approximate function $\hat{X}_\ell^\square$ via a known approximation (or interpolation) scheme on the N samples in $\square$.
- Compute

$$\hat{\delta}_\ell^\square = \frac{\inf_{y \in \square} \left| \hat{X}_\ell^\square(y) \right|}{\mathrm{Error}_\ell}$$

- If $\ell < L$ and $\hat{\delta}_\ell^\square \leq a_\ell$, refine the cell $\square$ into 2^d cells

At the end of the algorithm, return the union of zero level-sets of $\{\hat{X}_\ell^\square\}_\square$.

Adaptive Level-set Computation: Complexity analysis (Croci, H-A, and Powell 2025+)

Assume the approximation scheme converges with rate α and that there exist some $\delta_0, \rho_0 > 0$ such that for μ is the d-dimensional Lebesgue measure and all $0 < b < \delta_0$ we have

$$\mu(\{x \in \overline{D} : |f(x)| \leq b\}) \leq \rho_0 b.$$

Then, there is a choice of $\ell_0 \approx 0, L \approx \mathcal{O}(\log(\varepsilon^{-1/(\alpha\lambda_{q,1})}))$ and $\{a_\ell\}_{\ell=\ell_0}^{L}$ such that the adaptive/non-adaptive algorithms have computational complexities

$$\mathcal{O}\left(\varepsilon^{-\left(2+\frac{d-1}{\alpha}\right)/\lambda_{q,1}}\right) \qquad \text{vs.} \qquad \mathcal{O}\left(\varepsilon^{-\left(2+\frac{d}{\alpha}\right)/\lambda_{q,1}}\right)$$

Dynamical Conditional Expectation: SDEs

- Utilizing the smoothness of the density requires methods (and analysis) that is specific to models.
- Focus on SDEs: Let X be the solution to an SDE at time 1 and denotes its ℓ-level approximation by $\overline{X}_\ell$. Denote the filtration at time t by $\mathcal{F}_t$.
- Define $\Delta P_\ell := \mathbb{I}_{\overline{X}_\ell \in \Omega} - \mathbb{I}_{\overline{X}_{\ell-1} \in \Omega}$.
- For some $0 < \tau < 1$, let

$$\Delta Q_\ell := \mathbb{E}[\, \Delta P_\ell \,|\, \mathcal{F}_{1-\tau} \,].$$

$$\text{Note} \qquad \mathbb{E}[\, \Delta Q_\ell \,] = \mathbb{E}[\, \Delta P_\ell \,].$$

Dynamical Conditional Expectation: SDEs

- Utilizing the smoothness of the density requires methods (and analysis) that is specific to models.
- Focus on SDEs: Let X be the solution to an SDE at time 1 and denotes its ℓ-level approximation by $\overline{X}_\ell$. Denote the filtration at time t by $\mathcal{F}_t$.
- Define $\Delta P_\ell := \mathbb{I}_{\overline{X}_\ell \in \Omega} - \mathbb{I}_{\overline{X}_{\ell-1} \in \Omega}$.
- For some $0 < \tau < 1$, let

$$\Delta Q_\ell := \mathbb{E}[\,\Delta P_\ell \,|\, \mathcal{F}_{1-\tau}\,].$$

$$\text{Note} \qquad \mathbb{E}[\,\Delta Q_\ell\,] = \mathbb{E}[\,\Delta P_\ell\,].$$

We can consider the MLMC estimator based on ΔQ_ℓ instead of ΔP_ℓ. The work and (hopefully improved) variance convergence of ΔQ_ℓ become relevant.

Computing ΔQ_ℓ

In 1D, taking $\tau \equiv h_\ell$, the step-size, and using Euler-Maruyama for the last step we know that the conditional distribution of $\overline{X}_\ell(1)$ given $\mathcal{F}_{1-h_\ell}$ is Gaussian and we can compute ΔQ_ℓ exactly.

Making time dependence explicit, let
$g(x) = \mathbb{E}\Big[\mathbb{I}_{\overline{X}_\ell(1)\in\Omega} \,\Big|\, \overline{X}_\ell(1-h_\ell) = x \Big]$, then (roughly)

$$\begin{aligned}
\mathbb{E}[\,\Delta Q_\ell^2\,] &\approx \mathbb{E}\Big[\big(g(\overline{X}_\ell(1-h_\ell)) - g(\overline{X}_{\ell-1}(1-h_\ell))\big)^2 \Big] \\
&\lesssim \mathbb{E}\Big[(g'(\overline{X}_\ell(1-h_\ell)))^2 \, \big|\overline{X}_\ell(1-h_\ell) - \overline{X}_{\ell-1}(1-h_\ell)\big|^2 \Big] + \ldots \\
&\lesssim \mathcal{O}\Big(h_\ell^{1/2} \, (h_\ell^{-1/2})^2 \, h_\ell^{\beta} \Big) = \mathcal{O}(h_\ell^{-1/2+\beta})
\end{aligned}$$

Examples: Conditional Expectations

- Euler-Maruyama has $\beta = 1$, hence $\mathrm{Var}[\,\Delta Q_\ell\,] \approx \mathcal{O}(h_\ell^{1/2})$. Using the Conditional expectation does not offer an advantage over the classical method.
- Milstein has $\beta = 2$, hence $\mathrm{Var}[\,\Delta Q_\ell\,] \approx h_\ell^{3/2}$ and the MLMC complexity is $\mathcal{O}(\varepsilon^{-2})$.
- Antithetic estimator with truncated Milstein has similar complexity to Euler-Maruyama. We do have $\beta = 2$ but would involve the second derivative $\mathbb{E}[\,(g'')^2\,] \propto h_\ell^{-3/2}$.

Path splitting to estimate ΔQ_ℓ

More generally, for any method and any τ, we can use path splitting (Monte Carlo) with sufficient number of samples, leading to increased work. See, e.g., (Glasserman 2003; Burgos and Giles 2012) (applied to computing sensitivities).

- When $\tau \to 0$, i.e., splitting late,

$$\mathrm{Var}[\,\Delta Q_\ell\,] \le \mathbb{E}\Big[\,(\mathbb{E}[\,\Delta P_\ell \,|\, \mathcal{F}_{1-\tau}\,])^2\,\Big] = \mathbb{E}\Big[\,(\Delta P_\ell)^2\,\Big] \approx \mathcal{O}(h_\ell^{\beta/2})$$

  leads to worse variance.

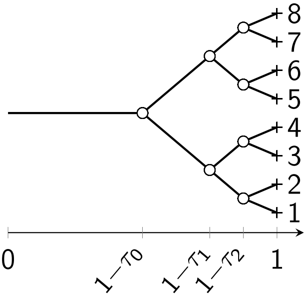
- When $\tau \to 1$, i.e., splitting early,

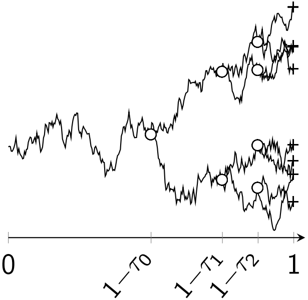
$$\mathrm{Var}[\,\Delta Q_\ell\,] \le \mathbb{E}\Big[\,(\mathbb{E}[\,\Delta P_\ell \,|\, \mathcal{F}_{1-\tau}\,])^2\,\Big] = (\mathbb{E}[\,\Delta P_\ell\,])^2 \approx \mathcal{O}(h_\ell^{\beta})$$

  leads to worse work.

Path Branching

- Let $1 - \tau_{\ell'} = 1 - 2^{-\ell'}$ for $\ell' \in \{1, \ldots, \ell\}$.
- For every ℓ', given $\mathcal{F}_{1-\tau_{\ell'}}$ at time $1 - \tau_{\ell'}$, create two sample paths from time $1 - \tau_{\ell'}$ to time $1 - \tau_{\ell'+1}$ which depend on two independent samples of the Brownian motion $\{W_t\}_{1-\tau_{\ell'} \leq t \leq 1-\tau_{\ell'+1}}$.
- Evaluate the payoff difference $\Delta P_\ell^{(i)}$ for $i \in \{1, \ldots, 2^\ell\}$, each corresponding to a sample of X.
- Define the Monte Carlo average as $\Delta \mathcal{P}_\ell := 2^{-\ell} \sum_{i=1}^{2^\ell} \Delta P_\ell^{(i)}$





Work/Variance bounds

Theorem

Assume that for some $\delta_0 > 0$ *and all* $0 < \delta \leq \delta_0$ *and* $0 < \tau \leq 1$*, we have*

$$\mathbb{E}\Big[(\mathbb{P}[\, dist_{\partial\Omega}(X) \leq \delta \,|\, \mathcal{F}_{1-\tau} \,])^2 \Big] \lesssim \frac{\delta^2}{\tau^{1/2}}.$$

Then

$$\mathbb{E}[\, \Delta\mathcal{P}_\ell \,] = \mathbb{E}[\, \Delta P_\ell \,]$$
$$\text{Work}(\Delta\mathcal{P}_\ell) \lesssim \ \ell\, \text{Work}(\Delta P_\ell)$$
$$\text{Var}[\, \Delta\mathcal{P}_\ell \,] \lesssim \ h_\ell \text{Var}[\, \Delta P_\ell \,] + h_\ell^{\lambda_{q,2\beta}}$$

For the antithetic estimator with truncated-Milstien, we also have (with higher regularity conditions that we assume/show)

$$\text{Var}[\, \Delta\mathcal{P}_\ell \,] \lesssim \ h_\ell \text{Var}[\, \Delta P_\ell \,] + h_\ell^{2\lambda_{q,5}}$$

Examples: Path Branching

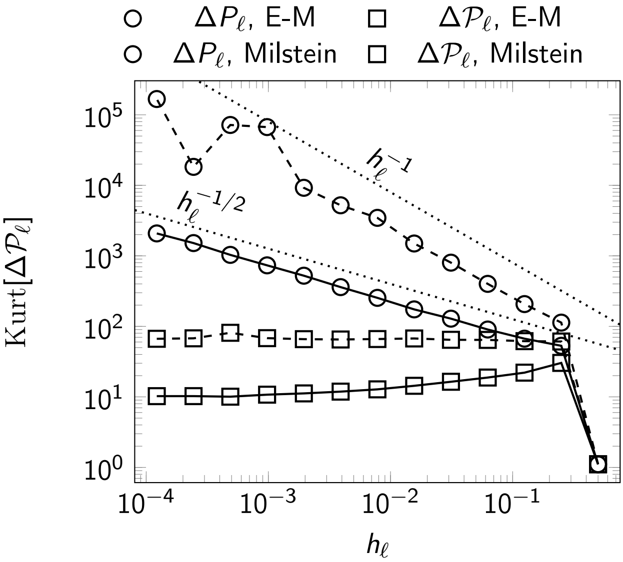
- Euler-Maruyama has $\beta = 1$ hence $\mathrm{Var}[\,\Delta\mathcal{P}_\ell\,] \approx \mathcal{O}(h_\ell)$.
  The complexity is $\mathcal{O}(\varepsilon^{-2}|\log\varepsilon|^3)$ (Compare to $\mathcal{O}(\varepsilon^{-2}|\log\varepsilon|^2)$ for a Lipschitz payoff).

- Milstein has $\beta = 2$ hence $\mathrm{Var}[\,\Delta\mathcal{P}_\ell\,] \approx \mathcal{O}(h_\ell^2)$ and complexity is $\mathcal{O}(\varepsilon^{-2})$ (Same as for a Lipschitz payoff).

- Antithetic estimator with truncated Milstein estimator has better rates than Euler-Maruyama!
  $\mathrm{Var}[\,\Delta\mathcal{P}_\ell\,] \approx \mathcal{O}(h_\ell^{3/2})$ hence complexity is $\mathcal{O}(\varepsilon^{-2})$ (Same as for a Lipschitz payoff).

We also show improved bounds on the Kurotisis.

Digital option on GBM: Branching



More

Conclusions

- Adaptive methods are versatile and easy to use, but still suffer when higher moments are needed.
- Branching is applicable to a large class of SDEs, is similar to numerical smoothing, but is a more natural way to exploit smoothness in a Monte Carlo setting without the need for root finding.
- Assumptions of the form

$$\mathbb{E}\Big[(\mathbb{P}[\,\mathrm{dist}_{\partial\Omega}(X) \le \delta \,|\, \mathcal{F}_{1-\tau}\,])^2 \Big] \lesssim \frac{\delta^2}{\tau^{1/2}},$$

can be related to smoothness conditions on the set Ω and the density of X (or the coefficients of the SDE).

Current/Future work

- Computing sensitivities: Using bumping, the variance increases as the bump distance decreases. Branching can help.
- Pricing other options (Barrier); not clear extension, combine with adaptive splitting?
- Applications to other applications: Particle systems (probably requires Multi-index Monte Carlo).
- Approximate CDFs – adaptivity will probably fare worse than branching. More analysis required.
- Parabolic SPDEs: Branching extends naturally, but analysis could be more challenging.

References I

- Achtsis, N., R. Cools, and D. Nuyens (2013). "Conditional sampling for barrier option pricing under the LT method". In: *SIAM Journal on Financial Mathematics* 4.1, pp. 327–352.
- Altmayer, M. and A. Neuenkirch (2015). "Multilevel Monte Carlo quadrature of discontinuous payoffs in the generalized Heston model using Malliavin integration by parts". In: *SIAM Journal on Financial Mathematics* 6.1, pp. 22–52.
- Bardou, O., N. Frikha, and G. Pagès (Jan. 2009). "Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling". In: *Monte Carlo Methods and Applications* 15.3, pp. 173–210. ISSN: 0929-9629, 1569-3961. DOI: 10.1515/mcma.2009.011.
- Bayer, C., C. Ben Hammouda, and R. Tempone (2024). "Multilevel Monte Carlo with Numerical Smoothing for Robust and Efficient Computation of Probabilities and Densities". In: *SIAM Journal on Scientific Computing* 46.3, A1514–A1548. DOI: 10.1137/22M1495718.
- Bayer, C., M. Siebenmorgen, and R. Tempone (2018). "Smoothing the payoff for efficient computation of basket option prices". In: *Quantitative Finance* 18.3, pp. 491–505. DOI: 10.1080/14697688.2017.1308003.
- Broadie, M., Y. Du, and C. C. Moallemi (June 2011). "Efficient Risk Estimation via Nested Sequential Simulation". In: *Management Science* 57.6, pp. 1172–1194. ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.1110.1330.

References II

Burgos, S. and M. B. Giles (2012). "Computing Greeks Using Multilevel Path Simulation". In: *Monte Carlo and Quasi-Monte Carlo Methods 2010*. Springer Berlin Heidelberg, pp. 281–296. ISBN: 9783642274398, 9783642274404. DOI: 10.1007/978-3-642-27440-4_13.

Crépey, S., N. Frikha, A. Louzi, and J. Spence (2024). "Adaptive multilevel stochastic approximation of the Value-at-Risk". In: *arXiv preprint arXiv:2408.06531*.

Croci, M., A.-L. H-A, and I. C. J. Powell (2025+). *An Adaptive Sampling Scheme for Level-set Approximation*.

Elfverson, D., F. Hellman, and A. Målqvist (Jan. 2016). "A Multilevel Monte Carlo Method for Computing Failure Probabilities". In: *SIAM/ASA Journal on Uncertainty Quantification* 4.1, pp. 312–330. ISSN: 2166-2525. DOI: 10.1137/140984294.

Giles, M. B. and A.-L. H-A (Jan. 2019). "Multilevel Nested Simulation for Efficient Risk Estimation". In: *SIAM/ASA Journal on Uncertainty Quantification* 7.2, pp. 497–525. ISSN: 2166-2525. DOI: 10.1137/18m1173186.

— (2024). "Multilevel Path Branching for Digital Options". In: *Annals of Applied Probability* 34.5, pp. 4836–4862. ISSN: 1050-5164. DOI: 10.1214/24-AAP2083.

Giles, M. B., A.-L. H-A, and J. Spence (2023). *Efficient Risk Estimation for the Credit Valuation Adjustment*. DOI: 10.48550/arxiv.2301.05886. arXiv: 2301.05886 [q-fin.CP].

References III

- Giles, M. B., T. Nagapetyan, and K. Ritter (Jan. 2015). "Multilevel Monte Carlo Approximation of Distribution Functions and Densities". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1, pp. 267–295. ISSN: 2166-2525. DOI: 10.1137/140960086.
- Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Vol. 53. Springer New York. ISBN: 9781441918222, 9780387216171. DOI: 10.1007/978-0-387-21617-1.
- H-A, A.-L., J. Spence, and A. L. Teckentrup (Aug. 2022). "Adaptive Multilevel Monte Carlo for Probabilities". In: *SIAM Journal on Numerical Analysis* 60.4, pp. 2125–2149. ISSN: 0036-1429, 1095-7170. DOI: 10.1137/21m1447064.
- Krumscheid, S. and F. Nobile (2018). "Multilevel Monte Carlo approximation of functions". In: *SIAM/ASA Journal on Uncertainty Quantification* 6.3, pp. 1256–1293.
- Min, C. and F. Gibou (2007). "A second order accurate level set method on non-graded adaptive Cartesian grids". In: *Journal of Computational Physics* 225.1, pp. 300–321.
- Rockafellar, R. T. and S. Uryasev (July 2002). "Conditional Value-at-Risk for general loss distributions". In: *Journal of Banking & Finance* 26.7, pp. 1443–1471. ISSN: 0378-4266. DOI: 10.1016/s0378-4266(02)00271-6.
- Xia, Y. and M. B. Giles (2012). "Multilevel path simulation for jump-diffusion SDEs". In: *Monte Carlo and Quasi-Monte Carlo Methods 2010*. Springer, pp. 695–708.

References IV

Xu, Z., Z. He, and X. Wang (2024). "Efficient risk estimation via nested multilevel quasi-Monte Carlo simulation". In: *Journal of Computational and Applied Mathematics* 443, p. 115745.

Numerical Tests: Digital Options

For constant $\mu, \sigma, S(0)$ consider the asset

$$\mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}W(t).$$

Compute

$$\mathbb{E}[\,\mathbb{I}_{X>0}\,] \coloneqq \mathbb{E}[\,\mathbb{I}_{S(T)-K>0}\,]$$

for some strike price $K > 0$. We use Euler-Maruyama with a step size $h_\ell = 2^{-\ell}$ to approximate $S_{h_\ell}(\cdot) \approx S(\cdot)$ and set

$$X \coloneqq S_{h_\ell}(T) - K.$$

The assumptions are satisfied using constant $\sigma_\ell \equiv 1$ for $\alpha = \beta = \gamma = 1$ and any $q < \infty$ giving complexity $\mathcal{O}\big(\varepsilon^{-2.5-\nu}\big)$ for standard Multilevel Monte Carlo and $\mathcal{O}(\varepsilon^{-2-\nu})$ for any $\nu > 0$ using adaptive Multilevel Monte Carlo.

Digital option on GBM

Consider the assets

$$\mathrm{d}S^{(i)}(t) = \mu^i S^{(i)}(t) + \sigma^i S^{(i)}(t)\mathrm{d}W^{(i)}(t)$$

where

$$W^i(t) = \rho W^{(i)}_{\text{com}}(t) + \sqrt{1-\rho^2} W^{(i)}_{\text{ind}}(t)$$

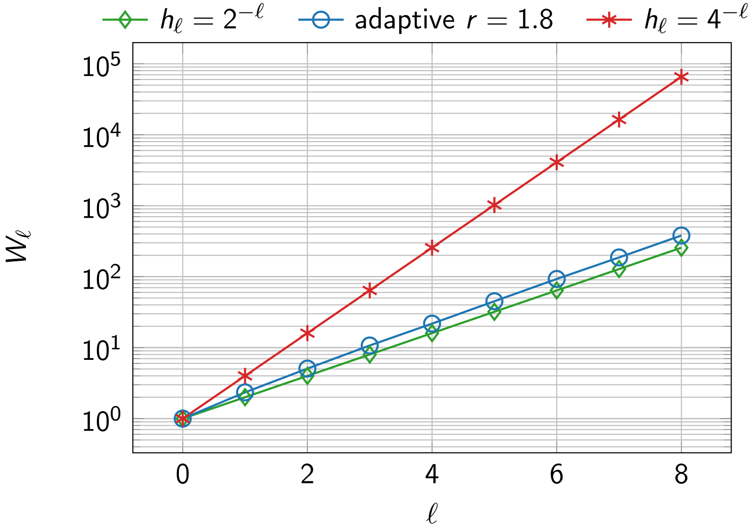
for $1 \leq i \leq 10$. Consider the digital option with payoff

$$\mathbb{I}_{\left(\frac{1}{10}\sum_{i=1}^{10} S^{(i)}(t)\right) > K}.$$

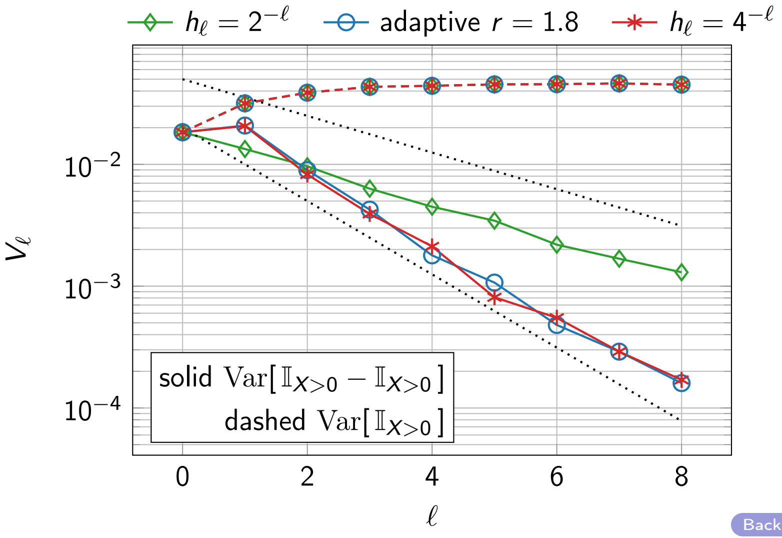
Thus, compute

$$\mathbb{E}\left[\mathbb{I}_{\left(\frac{1}{10}\sum_{i=1}^{10} S^{(i)}(t)\right) > K}\right].$$

Digital option on GBM with E-M: Adaptivity

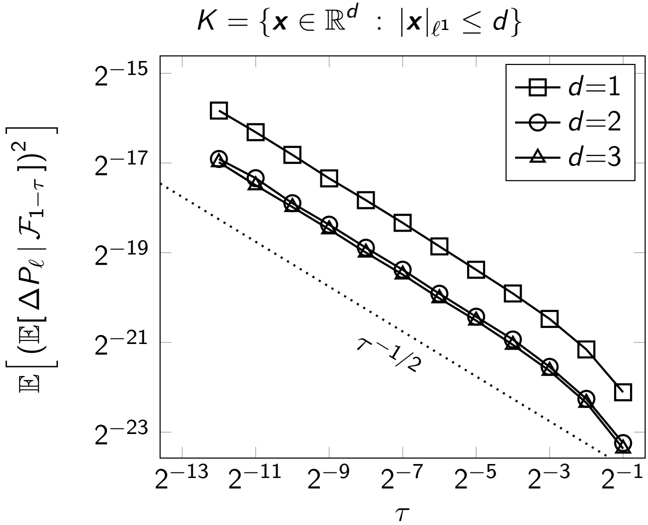


Digital option on GBM with E-M: Adaptivity

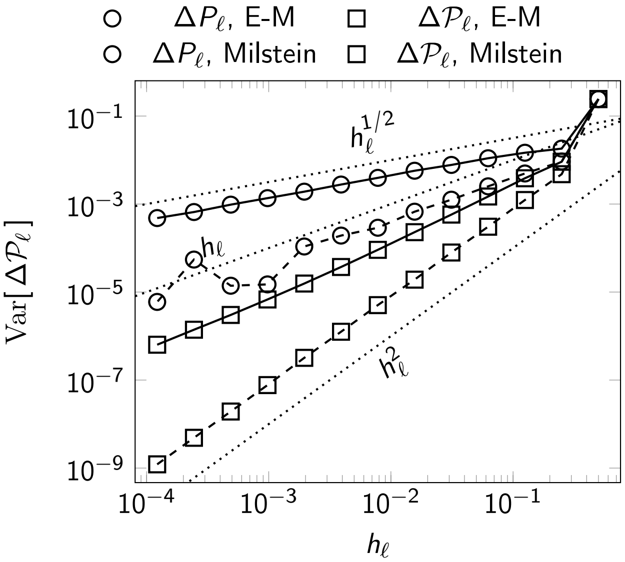


Back

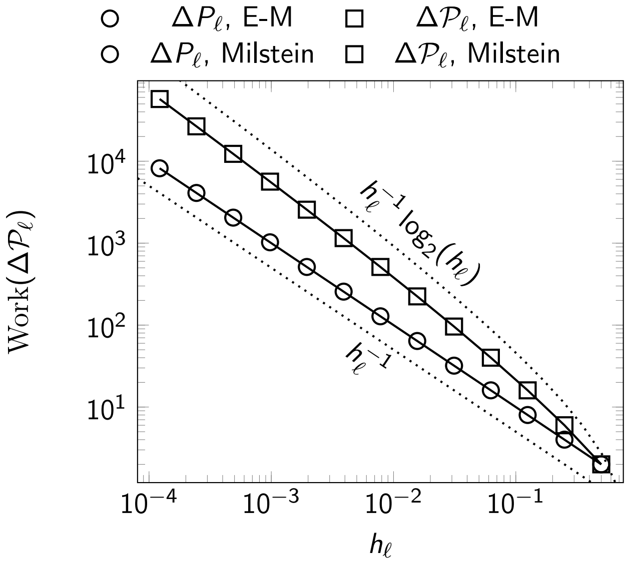
Digital option on GBM: Branching



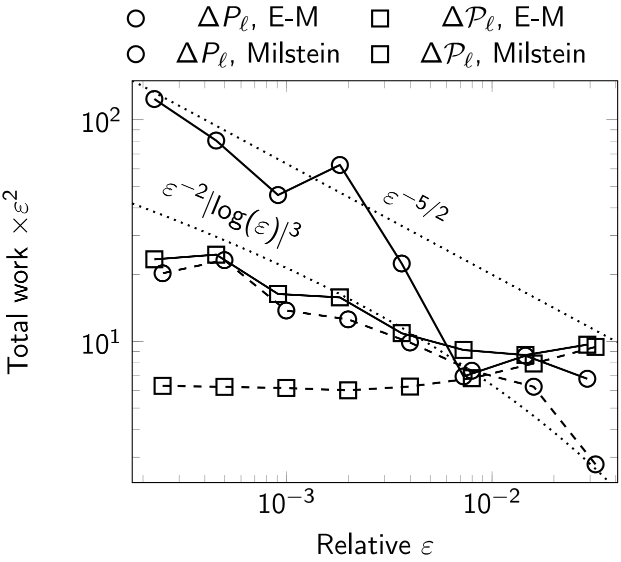
Digital option on GBM: Branching



Digital option on GBM: Branching



Digital option on GBM: Branching



Back