Multilevel Path Branching for Digital Options

AI Haji-Ali
Joint work with Mike Giles (University of Oxford)

Heriot-Watt University

MCQMC — 21 July, 2022

The problem: Pricing a Digital option

Let X_t be a d-dimensional stochastic process satisfying the SDE for $0 < t \leq 1$

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t.$$

Let $(\mathcal{F}_t)_{0 \leq t \leq 1}$ be the natural filtration of W_t.

The problem: Pricing a Digital option

Let X_t be a d-dimensional stochastic process satisfying the SDE for $0 < t \leq 1$

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t.$$

Let $(\mathcal{F}_t)_{0 \leq t \leq 1}$ be the natural filtration of W_t.

We want to price a digital option of the form (dropping discounting)

$$\mathbb{P}[\, X_1 \in K \,] = \mathbb{E}[\, \mathbb{I}_{X_1 \in K} \,]$$

for some $K \subset \mathbb{R}^d$. Let $\{\overline{X}^{\ell}_t\}_{t=0}^1$ be an approximation of the path $\{X_t\}_{t=0}^1$ at level ℓ using $h_\ell^{-1} \equiv 2^\ell$ timesteps.

The problem: Pricing a Digital option

Let X_t be a d-dimensional stochastic process satisfying the SDE for $0 < t \leq 1$

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}W_t.$$

Let $(\mathcal{F}_t)_{0\leq t\leq 1}$ be the natural filtration of W_t.

We want to price a digital option of the form (dropping discounting)

$$\mathbb{P}[\,X_1 \in K\,] = \mathbb{E}[\,\mathbb{I}_{X_1 \in K}\,]$$

for some $K \subset \mathbb{R}^d$. Let $\{\overline{X}^{\ell}_t\}_{t=0}^1$ be an approximation of the path $\{X_t\}_{t=0}^1$ at level ℓ using $h_\ell^{-1} \equiv 2^\ell$ timesteps.

For $|\mathbb{E}[\,\mathbb{I}_{X_1\in K} - \mathbb{I}_{\overline{X}^{\ell}_1 \in K}\,]| \lesssim h_\ell^\alpha$, a Monte Carlo estimator of $\mathbb{E}[\,\mathbb{I}_{X_1\in K}\,]$ has computational complexity $\varepsilon^{-2-\alpha}$ to achieve MSE ε.

Multilevel Monte Carlo

Consider a hierarchy of corrections $\{\Delta P_\ell\}_{\ell=0}^{L}$ such that

$$\mathbb{E}[\,\Delta P_\ell\,] = \begin{cases} \mathbb{E}\left[\,\mathbb{I}_{\overline{X}_1^0 \in K}\,\right] & \ell = 0 \\ \mathbb{E}\left[\,\mathbb{I}_{\overline{X}_1^\ell \in K} - \mathbb{I}_{\overline{X}_1^{\ell-1} \in K}\,\right] & \text{otherwise.} \end{cases}$$

MLMC can be formulated as

$$\mathbb{E}\left[\,\mathbb{I}_{X_1 \in K}\,\right] = \sum_{\ell=0}^{\infty} \mathbb{E}[\,\Delta P_\ell\,] \approx \sum_{\ell=0}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Delta P_\ell^{(m)}$$

Multilevel Monte Carlo

Consider a hierarchy of corrections $\{\Delta P_\ell\}_{\ell=0}^{L}$ such that

$$\mathbb{E}[\,\Delta P_\ell\,] = \begin{cases} \mathbb{E}\left[\,\mathbb{I}_{\overline{X}_1^0 \in K}\,\right] & \ell = 0 \\ \mathbb{E}\left[\,\mathbb{I}_{\overline{X}_1^\ell \in K} - \mathbb{I}_{\overline{X}_1^{\ell-1} \in K}\,\right] & \text{otherwise.} \end{cases}$$

MLMC can be formulated as

$$\mathbb{E}\left[\,\mathbb{I}_{X_1 \in K}\,\right] = \sum_{\ell=0}^{\infty} \mathbb{E}[\,\Delta P_\ell\,] \approx \sum_{\ell=0}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Delta P_\ell^{(m)}$$

Assuming

$$\mathrm{Var}[\,\Delta P_\ell\,] \lesssim h_\ell^{\beta_{\mathrm{d}}}, \qquad |\mathbb{E}[\,\Delta P_\ell\,]| \lesssim h_\ell^{\alpha}, \qquad \mathrm{Work}(\Delta P_\ell) \lesssim h_\ell^{-1}$$

then to compute with MSE ε^2 the complexity of MLMC is $\mathcal{O}(\varepsilon^{-2-\max(1-\beta_{\mathrm{d}},0)/\alpha})$ when $\beta_{\mathrm{d}} \neq 1$ and $\mathcal{O}(\varepsilon^{-2}|\log \varepsilon|^2)$ otherwise.

Examples: Classical Method

Using $\Delta P_\ell = \mathbb{I}_{\overline{X}_1^\ell} - \mathbb{I}_{\overline{X}_1^{\ell-1}}$, note that $\mathrm{Var}[\,\Delta P_\ell\,] \lesssim h_\ell^{\beta_\mathsf{d}}$ is an implication of $\mathbb{E}\left[\left(\overline{X}_1^\ell - \overline{X}_1^{\ell-1}\right)^2\right]^{1/2} \approx \mathcal{O}(h_\ell^{\beta_\mathsf{d}})$.

- Euler-Maruyama has $\alpha = 1$ and $\beta_\mathsf{d} \approx 1/2$ and complexity is $\mathcal{O}(\varepsilon^{-5/2})$ (Compare to $\mathcal{O}(\varepsilon^{-2}|\log\varepsilon|^2)$ for a Lipschitz payoff).
- Milstein has $\alpha = 1$ and $\beta_\mathsf{d} \approx 1$ and complexity is $\mathcal{O}(\varepsilon^{-2}|\log\varepsilon|^2)$ (Compare to $\mathcal{O}(\varepsilon^{-2})$ for a Lipschitz payoff).
- Antithetic Milstein has the same rates es Euler-Maruyama (better rates possible with at least a Lipschitz payoff).

Conditional Expectation

For some $0 < \tau < 1$, let

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

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

Conditional Expectation

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 becomes relevant.

Computing ΔQ_ℓ

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

Let $g(x) = \mathbb{E}\Big[\mathbb{I}_{\overline{X}_1^\ell \in K} \,\Big|\, \overline{X}_{1-\tau}^\ell = x \Big]$, then (roughly)

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

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 complexity is $\mathcal{O}(\varepsilon^{-2})$.
- Antithetic Milstein estimator 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 (2004) and Burgos & Giles (2012) for more information on this method (for computing options and sensitivities).

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

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

  leads to worse variance.

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

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

  leads to worse work.

Solution: More splitting

For $\tau' > \tau$

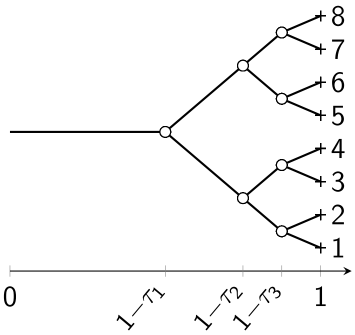
$$\begin{aligned}\Delta Q'_\ell &:= \mathbb{E}[\, \Delta Q_\ell \,|\, \mathcal{F}_{1-\tau'} \,] \\ &= \mathbb{E}[\, \mathbb{E}[\, \Delta P_\ell \,|\, \mathcal{F}_{1-\tau} \,] \,|\, \mathcal{F}_{1-\tau'} \,]\end{aligned}$$

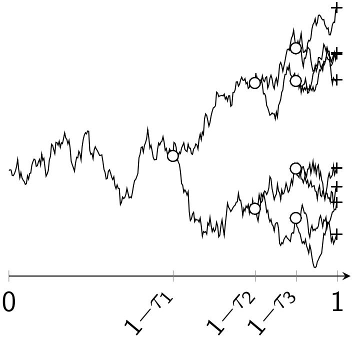
$$\text{Again} \qquad \mathbb{E}[\, \Delta Q'_\ell \,] = \mathbb{E}[\, \Delta P \,]$$

Now we have finer control over τ, τ' and the number of samples we can use to compute the two expectations.

Path Branching

- Let $1 - \tau_{\ell'} = 1 - 2^{-\ell'}$ for $\ell' \in \{1, \ldots, \ell\}$.
- For every ℓ', starting from $X_{1-\tau_{\ell'}}$ at time $1 - \tau_{\ell'}$, create two sample paths $\{X_t\}_{1-\tau_{\ell'} \le t \le 1-\tau_{\ell'+1}}$ which depend on two independent samples of the Brownian motion $\{W_t\}_{1-\tau_{\ell'} \le t \le 1-\tau_{\ell'+1}}$.
- Evaluate the payoff difference $\Delta P_\ell^{(i)}$ for every $X_1^{(i)}$ for $i \in \{1, \ldots, 2^\ell\}$
- Define the Monte Carlo average as $\Delta \mathcal{P}_\ell := 2^{-\ell} \sum_{i=1}^{2^\ell} \Delta P_\ell^{(i)}$





Main Assumptions & Bounds

Another way to see this: We have 2^ℓ extra samples. Cost (identical would be too correlated)? Correlation (independent would be too costly)?

Assumption

Assume that there exists $\beta_{\mathsf{d}}, \beta_{\mathsf{c}}, p > 0$ such that for all $\tau > h_\ell$

$$\mathbb{E}[\,(\Delta P_\ell)^2\,] \lesssim h_\ell^{\beta_{\mathsf{d}}}$$

and $$\mathbb{E}\Big[\,(\mathbb{E}[\,\Delta P_\ell \,|\, \mathcal{F}_{1-\tau}\,])^2\,\Big] \lesssim \frac{h_\ell^{\beta_{\mathsf{c}}}}{\tau^{1/2}}$$

Theorem (Work/Variance bounds)

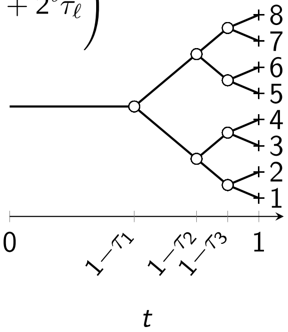
$$\mathbb{E}[\,\Delta \mathcal{P}_\ell\,] = \mathbb{E}[\,\Delta P_\ell\,]$$
$$\text{Work}(\Delta \mathcal{P}_\ell) \lesssim \ell\, h_\ell^{-1}$$
$$\text{Var}[\,\Delta \mathcal{P}_\ell\,] \lesssim h_\ell^{\beta_{\mathsf{d}}+1} + h_\ell^{\beta_{\mathsf{c}}}$$

Proof

Recall $\tau_{\ell'} = 2^{-\ell'}$

$$\mathrm{Work}(\Delta\mathcal{P}_\ell) \le h_\ell^{-1}\left((1-\tau_1) + \sum_{\ell'=1}^{\ell-1} 2^{\ell'}(\tau_{\ell'} - \tau_{\ell'+1}) + 2^\ell \tau_\ell\right)$$

$$\lesssim \ell\, h_\ell^{-1}$$



$$\mathrm{Var}[\,\Delta\mathcal{P}_\ell\,] \le \mathbb{E}\left[\left(\frac{1}{2^\ell}\sum_{i=1}^{2^\ell} \Delta P_\ell^{(i)}\right)^2\right]$$

$$\le \frac{1}{2^\ell}\mathbb{E}[\,\Delta P_\ell^2\,] + \frac{1}{2^{2\ell}}\sum_{i=1}^{2^\ell}\sum_{j=1,i\neq j}^{2^\ell} \mathbb{E}[\,\Delta P_\ell^{(i)} \Delta P_\ell^{(j)}\,]$$

$$\le \frac{1}{2^\ell}\mathbb{E}[\,\Delta P_\ell^2\,] + \frac{1}{2^{2\ell}}\sum_{i=1}^{2^\ell}\sum_{j=1,i\neq j}^{2^\ell} \mathbb{E}[\,(\mathbb{E}[\,\Delta P_\ell \,|\, \mathcal{F}_{1-\tau^{(i,j)}}\,])^2\,]$$

Examples: Path Branching

- Euler-Maruyama has $\beta_{\mathsf{d}} \approx 1/2$ and $\beta_{\mathsf{c}} \approx 1$ 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 $\beta_{\mathsf{d}} \approx 1$ and $\beta_{\mathsf{c}} \approx 2$ 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 Milstein estimator has better rates than Euler-Maruyama! Different analysis shows $\mathrm{Var}[\,\Delta\mathcal{P}_\ell\,] \approx \mathcal{O}(h_\ell^{3/2})$ hence complexity is $\mathcal{O}(\varepsilon^{-2})$ (Same as for a Lipschitz payoff).

Simplified Assumptions on SDE solution/Approximation

Theorem (Based on SDE solution and approximation)

Assume that for some $\delta_0 > 0$ *and all* $0 < \delta \leq \delta_0$ *and* $0 < \tau \leq 1$, *and letting* $d_{\partial K}(x) = \min_{y \in \partial K} \|x - y\|$, *there is a constant* C *independent of* δ, τ *and* $\mathcal{F}_{1-\tau}$ *such that*

$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(X_1) \leq \delta \,|\, \mathcal{F}_{1-\tau} \,]\right)^2 \Big] \leq C \, \frac{\delta^2}{\tau^{1/2}}.$$

Assume additionally that there is $q > 2$ *and* $\beta > 0$ *such that*

$$\mathbb{E}\Big[\left(X_1 - \overline{X}_1^{\ell} \right)^q \Big]^{1/q} \lesssim h_\ell^{\beta/2}$$

Then $\quad \beta_{\mathrm{d}} = \dfrac{\beta}{2} \times \left(1 - \dfrac{1}{q+1}\right) \quad$ and $\quad \beta_{\mathrm{c}} = \beta \times \left(1 - \dfrac{2}{q+2}\right)$

MLMC Complexity

When q is arbitrary,

$$\beta_{\mathsf{d}} \approx \frac{\beta}{2} \qquad \text{and} \qquad \beta_{\mathsf{c}} \approx \beta$$

and for $\beta \leq 2$

$$\mathrm{Var}[\,\Delta \mathcal{P}_\ell\,] \approx \mathcal{O}(h_\ell^{\beta})$$
$$\mathrm{Work}(\Delta \mathcal{P}_\ell) = \mathcal{O}(\ell h_\ell^{-1})$$

- Using Euler-Maryama: $\beta = 1$ and the MLMC computational complexity is approximately $o\big(\varepsilon^{-2+\nu}\big)$ for any $\nu > 0$ and for MSE ε.
- Using Milstein: $\beta = 2$ and the complexity is $\mathcal{O}(\varepsilon^{-2})$.

SDEs with Gaussian Transition Kernels

Lemma

Assume that a and σ are bounded and uniformly Hölder continuous and σ is uniformly elliptic and when ∂K is "nice" then there is $C > 0$ such that

$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(X_1) \leq \delta \,|\, \mathcal{F}_{1-\tau} \,]\right)^2 \Big] \leq C \, \frac{\delta^2}{\tau^{1/2}}$$

SDEs with Gaussian Transition Kernels

Lemma

Assume that a and σ are bounded and uniformly Hölder continuous and σ is uniformly elliptic and when ∂K is "nice" then there is $C > 0$ such that

$$\mathbb{E}\Big[\big(\mathbb{P}[\, d_{\partial K}(X_1) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\big)^2 \Big] \le C\, \frac{\delta^2}{\tau^{1/2}}$$

and

$$\mathbb{E}\Big[\big(\mathbb{P}[\, d_{\partial K}(\exp(X_1)) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\big)^2 \Big] \le C\, \frac{\delta^2}{\tau^{1/2}}.$$

SDEs with Gaussian Transition Kernels

Lemma

Assume that a and σ are bounded and uniformly Hölder continuous and σ is uniformly elliptic and when ∂K is "nice" then there is $C > 0$ such that

$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(X_1) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\right)^2 \Big] \le C \, \frac{\delta^2}{\tau^{1/2}}$$

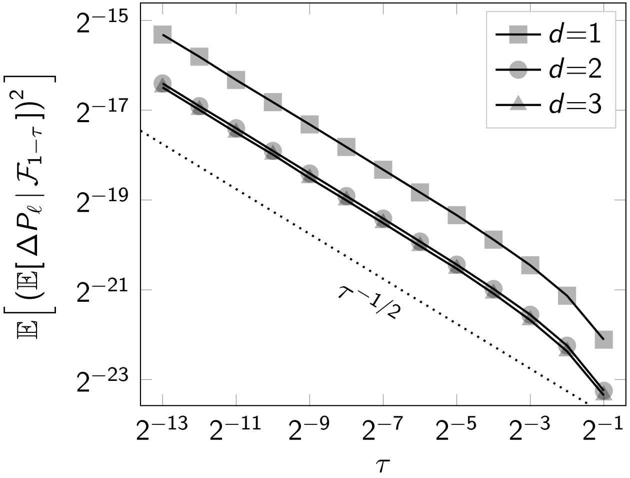
and

$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(\exp(X_1)) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\right)^2 \Big] \le C \, \frac{\delta^2}{\tau^{1/2}}.$$

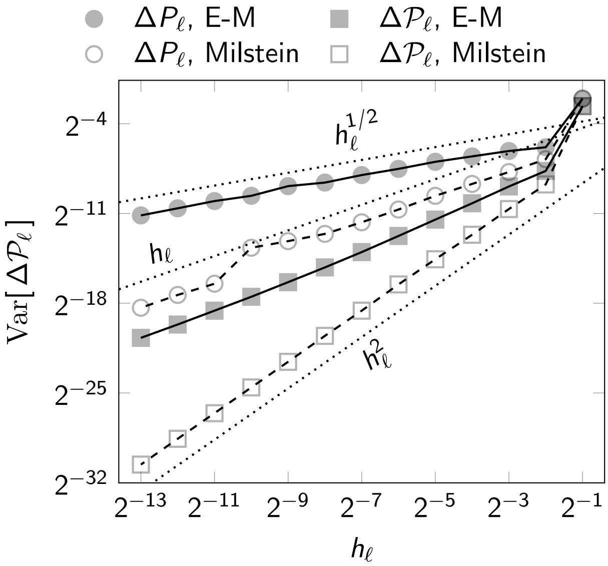
Proof. Based on bounding the conditional density of X_1 by a Gaussian density. E.g.

$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(X_1) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\right)^2 \Big] \lesssim \frac{1}{\tau^{1/2}} \left(\int_{-\delta}^{\delta} \mathrm{d}x \right) \times \mathbb{E}[\, \mathbb{P}[\, d_{\partial K}(X_1) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]] \lesssim \frac{\delta^2}{\tau^{1/2}}$$

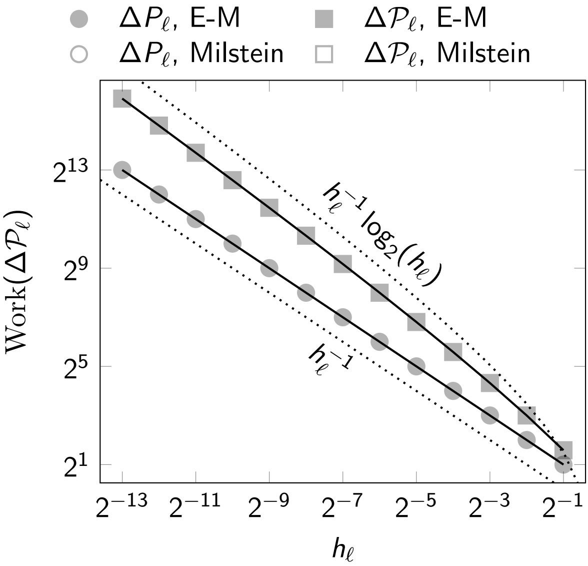
Numerical Results on GBM



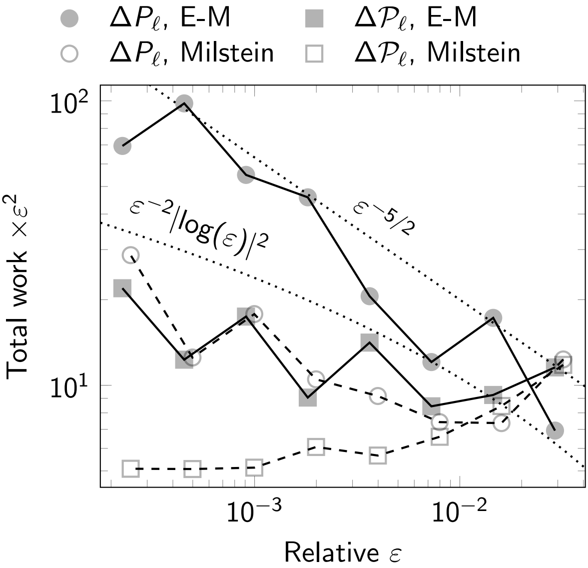
Numerical Results on GBM



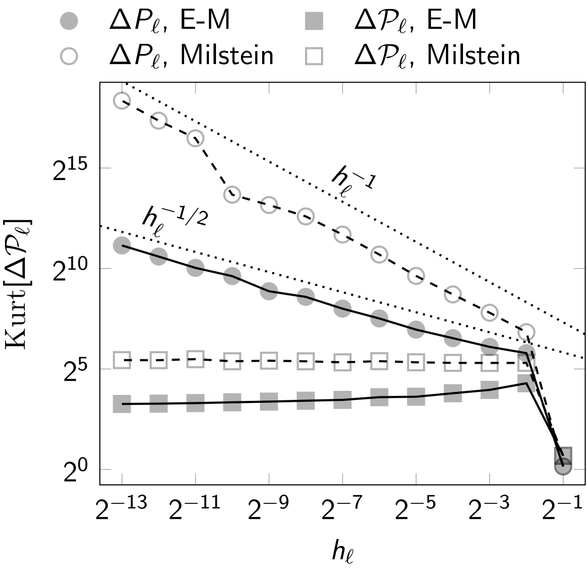
Numerical Results on GBM



Numerical Results on GBM



Numerical Results on GBM



What's done

- We also consider a sequence $\tau_{\ell'} = 2^{-\eta\ell'}$ for some $\eta > 0$. For $\eta > 1$, this reduces the work of $\Delta\mathcal{P}_\ell$ to $\mathcal{O}(2^\ell)$.
- More theoretical and numerical analysis for antithetic estimators.
- Manuscript "Multilevel Path Branching for Digital Options" coming soon.

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?
- Particle systems and Multi-index Monte Carlo.
- Approximate CDFs.
- Parabolic SPDEs with MLMC or MIMC. Method extends naturally, but analysis could be more challenging.

Elliptic SDEs

Definiton ((Si) *sets*)

We say that a set $S \subset \mathbb{R}^d$ is an (Si) set if there exists an orthonormal matrix A and a Lipschitz function f such that $S = A\widetilde{S}$ for the set

$$\widetilde{S} = \{x \in \mathbb{R}^d \;:\; f(x_{-1}) = x_1\},$$

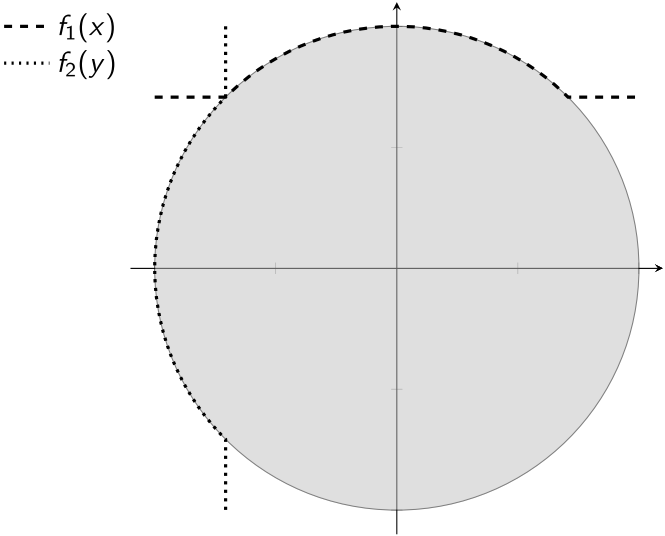
and $A\widetilde{S}$ denoting the image of $\widetilde{S}$ under the transformation $x \to Ax$.

Lemma

For $K \subset \mathbb{R}^d$ assume that $\partial K \subseteq \bigcup_{j=1}^n S_j$ for some finite n and (Si) sets $\{S_j\}_{j=1}^n$. Assume further that a and σ are bounded and uniformly Hölder continuous and σ is uniformly elliptic then

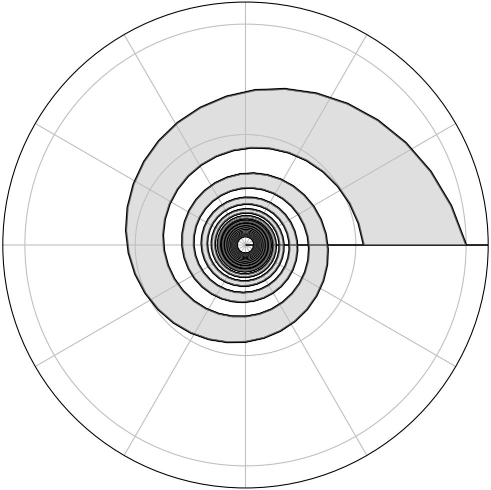
$$\mathbb{E}\Big[\left(\mathbb{P}[\, d_{\partial K}(X_1) \leq \delta \,|\, \mathcal{F}_{1-\tau}\,]\right)^2 \Big] \leq C\,\frac{\delta^2}{\tau^{1/2}}.$$

A nice set



$$\delta K = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$$

A not-so-nice set



$$\partial K = \{(r,\theta) \in \mathbb{R}_+ \times [0, 2\pi] \; : \; r = (n + \theta/\pi)^{-\frac{1}{2}}, n \in \mathbb{N}\}$$

Exponentials of Elliptic SDEs

What about a Geometric Brownian Motion $Y_t = \exp(X_t)$?

$$\mathrm{d}Y_t = aY_t\,\mathrm{d}t + \sigma Y_t\,\mathrm{d}W_t$$
$$\mathrm{d}X_t = a\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$

Exponentials of Elliptic SDEs

What about a Geometric Brownian Motion $Y_t = \exp(X_t)$?

$$\mathrm{d}Y_t = aY_t\,\mathrm{d}t + \sigma Y_t\,\mathrm{d}W_t$$
$$\mathrm{d}X_t = a\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$$

Lemma

For $K \subset \mathbb{R}^d$ assume that $\partial K \subseteq \bigcup_{j=1}^n \exp(S_j)$ for some finite n and (Si) sets $\{S_j\}_{j=1}^n$. Assume further that a and σ are bounded and uniformly Hölder continuous and σ is uniformly elliptic then

$$\mathbb{E}\Big[\big(\mathbb{P}[\, d_{\partial K}(\exp(X_1)) \le \delta \,|\, \mathcal{F}_{1-\tau} \,]\big)^2 \Big] \le C\, \frac{\delta^2}{\tau^{1/2}}.$$