Concentration inequalities for the Multilevel Monte Carlo Euler method applied to SDEs

Ahmed Kebaier
Université Paris 13

joint work with Benjamin Jourdain

ADVANCES IN FINANCIAL MATHEMATICS

January 12th, 2016

Outline of The Talk

1. CI for the Monte Carlo Euler method
   - Model
   - The concentration inequality
   - Used approach
2. CI for Multilevel Monte Carlo Euler method
   - Framework
   - Main results
3. Sketch of the proof
   - CI with the Clark-Ocone representation formula
   - Moment generating function of the error terms

Outline

1. CI for the Monte Carlo Euler method
   - Model
   - The concentration inequality
   - Used approach
2. CI for Multilevel Monte Carlo Euler method
   - Framework
   - Main results
3. Sketch of the proof
   - CI with the Clark-Ocone representation formula
   - Moment generating function of the error terms

The Model

- Let $W = (W^1, \ldots, W^q) \in \mathbb{R}^q$ be a B.M. and X solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \text{with}$$

$$(\mathcal{H}_{b,\sigma})\ \forall x, y \in \mathbb{R}^d\ |b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq C_{b,\sigma}|x - y|,$$

- Let X^n be the Euler scheme with time step $\delta = T/n$

$$dX_t^n = b(X_{\eta_n(t)})dt + \sigma_j(X_{\eta_n(t)})dW_t, \quad \eta_n(t) = [t/\delta]\delta.$$

- In this context

$$\forall p \geq 1,\ \mathbb{E}^{1/p}[\sup_{0\leq t\leq T} |X_t - X_t^n|^p] \leq K_{b,\sigma,T}\sqrt{\delta},\ \text{with } K_{b,\sigma,T} < \infty.$$

CI for Monte Carlo Euler method

The following result is due to Frikha and Menozzi (2012)

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a continuous Lipschitz function satisfying

$$|f(x) - f(y)| \leq [f]_{\mathrm{Lip.}} |x - y| \text{ for all } (x, y) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Assume we have condition $(\mathcal{H}_{b,\sigma})$ with uniformly bounded $\sigma(\cdot)$.

If $(X^n_{T,i})_{1 \leq i \leq N}$ denote independent copies of the Euler scheme X^n_T with time step $\delta = T/n$, Then for $N \in \mathbb{N}^*$ and $\alpha \geq 0$

$$\mathbb{P}\left(\left|\frac{1}{N} \sum_{i=1}^{N} f(X^n_{T,i}) - \mathbb{E} f(X^n_T)\right| \geq \alpha \right) \leq 2 \exp\left(\frac{-N\alpha^2}{2 C_{b,\sigma,T} [f]^2_{\mathrm{Lip.}}} \right),$$

where $C_{b,\sigma,T}$ is an explicit positive constant depending on b, σ, d and T.

Method used for the proof

- Recall the Gaussian concentration (GC) property

For any Lipschitz function ϕ with constant $[\phi]_{\mathrm{Lip.}}$ and for $G \sim \mathcal{N}(0, I_d)$ we have

$$\mathbb{E}\left(\exp(\lambda[\phi(G) - \mathbb{E}\phi(G)])\right) \leq \exp\left(\frac{1}{2}\lambda^2[\phi]^2_{\mathrm{Lip.}}\right)$$

- Conditionally to $X^n_{\frac{(k-1)T}{n}}$, $1 \leq k \leq n$, write

$$X^n_{\frac{kT}{n}} \overset{law}{=} \phi_{k,n}(G), \text{ where}$$

$$\phi_{k,n}({\color{magenta}x}) = X^n_{\frac{(k-1)T}{n}} + b(X^n_{\frac{(k-1)T}{n}})\frac{T}{n} + \sqrt{\frac{T}{n}}\sigma(X^n_{\frac{(k-1)T}{n}}){\color{magenta}x}, \forall x \in \mathbb{R}^d.$$

- Prove that $\phi_{k,n}$ is Lipschitz with a suitable explicit constant depending on k and n and apply the GC property.

Outline

Our setting

We consider a SDE with constant diffusion coefficient:

- Let $W = (W^1, \ldots, W^d) \in \mathbb{R}^d$ be a B.M. and X solution to

$$dX_t = b(X_t)dt + dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \text{with}$$

$(\mathcal{H}_b)$ $\forall x, y \in \mathbb{R}^d \; |b(x) - b(y)| \leq C_b|x - y|,$

- Let X^n be the Euler scheme with time step $\delta = T/n$

$$dX^n_t = b(X_{\eta_n(t)})dt + dW_t, \quad \eta_n(t) = [t/\delta]\delta.$$

- In this context

$$\forall p \geq 1, \; \mathbb{E}^{1/p}[\sup_{0\leq t\leq T} |X_t - X^n_t|^p] \leq K_{b,T}\delta, \text{ with } K_{b,T} < \infty.$$

The Euler Multilevel Monte Carlo scheme [Giles, 2008]

- Use $L+1$ Euler schemes with time steps $\frac{T}{m^\ell}$ for $\ell = 0, \ldots, L$ such that $m^L = n$, so that

$$\mathbb{E}(f(X_T^n)) = \mathbb{E}\left(f(X_T^{m^0})\right) + \sum_{\ell=1}^{L} \mathbb{E}(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}))$$

- The Multilevel method for Euler scheme estimator of $\mathbb{E}(f(X_T^n))$

$$\hat{Q} = \frac{1}{N_0}\sum_{k=0}^{N_0} f(X_{T,k}^1) + \sum_{\ell=0}^{L} \frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}})\right).$$

$$\mathrm{Var}(\hat{Q}) = \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell^{-1} m^{-2\ell}\right) \text{ and } \mathrm{Bias}^2(\hat{Q}, \mathbb{E}f(X_T)) = \mathcal{O}\left(m^{-2L}\right)$$

- For a given precison ε, by the complexity Theorem of Giles

$$N_\ell = \mathcal{O}\left(\varepsilon^{-2} m^{-\frac{3}{2}\ell}\right) \rightsquigarrow C_{MMC}^\star = \mathcal{O}\left(\varepsilon^{-2}\right) \ll C_{MC} = \mathcal{O}\left(\varepsilon^{-3}\right)$$

Assumptions

Assumption (R1)

The function $b \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ and there exist finite constants $[\dot{b}]_\infty$ and $a_{\Delta b}$ such that

$$\forall x \in \mathbb{R}^d, \quad \|\nabla b(x)\| \leq [\dot{b}]_\infty,$$
$$\forall x \in \mathbb{R}^d, \; |\Delta b(x) - \Delta b(x_0)| \leq 2a_{\Delta b}(1 + |x - x_0|).$$

Assumption (R2)

1. **Assumption (R1)** is satisfied.
2. Moreover the function $b \in \mathcal{C}^3(\mathbb{R}^d, \mathbb{R}^d)$ and there exist positive constants $[\ddot{b}]_\infty$ and $a_{\nabla \Delta b}$ such that

$$\forall x \in \mathbb{R}^d, \quad \|\nabla^2 b(x)\| \leq [\ddot{b}]_\infty,$$
$$\forall x \in \mathbb{R}^d, \; |\nabla[\Delta b(x)] - \nabla[\Delta b(x_0)]| \leq a_{\nabla \Delta b}(1 + |x - x_0|).$$

Main result

Theorem 1

- *Let $f \in C_b^1(\mathbb{R}^d, \mathbb{R})$ such that ∇f is a bounded Lipschitz function with Lipchitz constant $[\dot{f}]_{\mathrm{Lip.}}$ and such that*

$$|\nabla f| \leq [\dot{f}]_\infty.$$

- *If* **(R2)** *holds, then* $\forall\, 0 \leq \alpha \leq \mathcal{C}\mathcal{C}' \min_{1\leq\ell\leq L}\{m^\ell N_\ell\} \sum_{\ell=0}^{L} m^{-2\ell} N_\ell^{-1}$,

$$\mathbb{P}\left(|\hat{Q} - \mathbb{E}f(X_T^{m^L})| \geq \alpha\right) \leq 2\exp\left(-\frac{\alpha^2}{4\mathcal{C}\sum_{\ell=0}^{L} m^{-2\ell} N_\ell^{-1}}\right)$$

- Recall that for the MMC $\mathrm{Var}(\hat{Q}) = \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell^{-1} m^{-2\ell}\right)$

Restriction

- However, for $\forall\, \alpha \geq \mathcal{C}\mathcal{C}' \min_{1\leq \ell \leq L}\{m^{\ell} N_{\ell}\} \sum_{\ell=0}^{L} m^{-2\ell} N_{\ell}^{-1}$ we only have

$$\mathbb{P}\left(|\hat{Q} - \mathbb{E}f(X_T^{m^L})| \geq \alpha \right) \leq 2\exp\left(-\frac{\mathcal{C}}{2} \min_{1\leq \ell \leq L} m^{\ell} N_{\ell} \alpha \right).$$

- This restriction on α is influenced by the choice of the sample sizes N_{ℓ}.

Discussion on the choice of the sample sizes $(N_\ell)_{0\leq\ell\leq L}$

- Let us fix the precison $\varepsilon = \mathcal{O}\left(m^{-L}\right) = \mathcal{O}\left(n^{-1}\right)$

Optimization rule

$$\mathrm{Var}(\hat{Q}) = \mathcal{O}\left(\sum_{\ell=0}^{L} N_\ell^{-1} m^{-2\ell}\right) \sim \mathrm{Bias}^2(\hat{Q}, \mathbb{E}f(X_T)) = \mathcal{O}\left(m^{-2L}\right)$$

- As seen before, this leads to the choice

$$N_\ell = m^{2L-\frac{3}{2}\ell} \text{ for } 0 \leq \ell \leq L.$$

- This choice achieves an optimal complexity

$$C^\star_{MMC} = \mathcal{O}\left(\varepsilon^{-2}\right) = \mathcal{O}\left(m^{2L}\right) = \mathcal{O}\left(n^2\right)$$

- For this choice, our constraint on α rewrites

$$\alpha \leq CC' \min_{1\leq\ell\leq L}\{m^\ell N_\ell\} \sum_{\ell=0}^{L} m^{-2\ell} N_\ell^{-1} = CC' m^{-\frac{1}{2}L} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

- This should be compared with the precision $\varepsilon = \mathcal{O}\left(\frac{1}{n}\right)$.

Maximizing the range of α

- Question: Can we maximize the range of values for α while keeping the same precision $\varepsilon = \mathcal{O}(1/n)$?
- At first, note that $\alpha \propto \min_{1\leq \ell \leq L} m^{\ell} N_{\ell}$

1. Then a natural choice is the sequence

$$N_{\ell} = \frac{1}{L} m^{2L-\ell}$$

  - This yields a complexity $C^{\star}_{MMC} = \mathcal{O}\left(m^{2L}\right) = \mathcal{O}\left(n^2\right)$
  - But $\mathrm{Var}(\hat{Q}) = \mathcal{O}\left(Lm^{-2L}\right) = \mathcal{O}\left(\log(n)/n^2\right)$

  $\Rightarrow$ This leads to $\alpha \leq CC'/L = CC'/\log(n)$

2. Another possible choice is given by

$$N_{\ell} = m^{2L-\frac{3}{2}\ell}\mathbf{1}_{\{0\leq \ell\leq \beta L\}} + m^{(2-\frac{\beta}{2})L-\ell}\mathbf{1}_{\{\beta L\leq \ell\leq L\}}, \beta \in (0,1]$$

  - This leads to $C^{\star}_{MMC} = \mathcal{O}\left(n^2\right)$ and $\mathrm{Var}(\hat{Q}) = \mathcal{O}\left(1/n^2\right)$

  $\Rightarrow$ This leads to $\alpha \leq CC'm^{-\frac{\beta}{2}L} = CC'/n^{\beta/2}$

Outline

Sketch of the proof

Recall That $\hat{Q} = \hat{Q}_1 + \hat{Q}_2$ where

$$\hat{Q}_1 := \frac{1}{N_0} \sum_{k=0}^{N_0} f(X^1_{T,k}) - \mathbb{E}f(X^1_T)$$

and

$$\hat{Q}_2 := \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(f(X^{m^\ell}_{T,k}) - f(X^{m^{\ell-1}}_{T,k}) - \mathbb{E}[f(X^{m^\ell}_{T,k}) - f(X^{m^{\ell-1}}_{T,k})] \right).$$

Then by Markov inequality we have

$$\begin{aligned} \mathbb{P}\left(\hat{Q} - \mathbb{E}f(X^{m^L}_T) \geq \alpha\right) &\leq e^{-\lambda\alpha} \mathbb{E}\left[\exp\left(\lambda[\hat{Q} - \mathbb{E}f(X^{m^L}_T)]\right)\right] \\ &\leq e^{-\lambda\alpha} \mathbb{E}\left[\exp(\lambda\hat{Q}_1)\right] \mathbb{E}\left[\exp(\lambda\hat{Q}_2)\right] \end{aligned}$$

- **First term.**
By Frikha and Menozzi (2012), we have the existence of an explicit positive constant C depending on b, d, T and f such that

$$\mathbb{E}\left[\exp(\lambda \hat{Q}_1)\right] \leq \exp\left\{\frac{\lambda^2 C}{N_0}\right\}.$$

- **Second term.**
We use the independence property to write

$$\mathbb{E}\left[\exp(\lambda \hat{Q}_2)\right] \leq$$

$$\prod_{\ell=1}^{L}\left(\mathbb{E}\left[\exp\left\{\frac{\lambda}{N_\ell}\left(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) - \mathbb{E}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})]\right)\right\}\right]\right)^{N_\ell}$$

Houdré and Privault (2002) approach

Let $F \in \mathbb{D}^{1,2}$ be an $\mathcal{F}_T$-measurable s.t. $\mathbb{E}[e^{\lambda|F|}] < \infty$, $\forall \lambda > 0$.

Clark Ocone Formula

$$F - \mathbb{E}F = \int_0^T \mathbb{E}(D_s F|\mathcal{F}_s)dW_s$$

Markov inequality

$$\mathbb{P}\left(F - \mathbb{E}F \geq \alpha\right) \leq e^{-\lambda\alpha}\mathbb{E}\left(e^{\lambda(F-\mathbb{E}F)}\right) = e^{-\lambda\alpha}\mathbb{E}\left(e^{\lambda\int_0^T \mathbb{E}(D_s F|\mathcal{F}_s)dW_s}\right)$$

Martingale property

If moreover $\|\mathcal{D}F\|_{L^\infty} \leq K$, $K > 0$, then for some $p > 1$ we have

$$\mathbb{P}\left(F - \mathbb{E}F \geq \alpha\right) \leq \exp\left(\frac{1}{2}pK^2T\lambda^2 - \lambda\alpha\right).$$

Optimizing in λ yields $\mathbb{P}\left(F - \mathbb{E}F \geq \alpha\right) \leq \exp\left(-\frac{\alpha^2}{2pK^2T}\right)$

Then, by Clarck's Ocone formula we have

$$f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}}) - \mathbb{E}[f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})] = \int_0^T K_r^\ell \cdot dW_r,$$

where

$$K_{r,j}^\ell := \mathbb{E}\left[\mathcal{D}_r^j f(X_T^{m^\ell}) - \mathcal{D}_r^j f(X_T^{m^{\ell-1}}) | \mathcal{F}_r\right].$$

Therefore,

$$\mathbb{E}\left[\exp(\lambda \hat{Q}_2)\right] \leq$$

$$\prod_{\ell=1}^{L}\left(\mathbb{E}\left[\exp\left\{\frac{p\lambda}{N_\ell}\int_0^T K_r^\ell \cdot dW_r - \frac{p^2\lambda^2}{2N_\ell^2}\int_0^T |K_r^\ell|^2 dr\right\}\right]\right)^{\frac{N_\ell}{p}}$$

$$\times\left(\mathbb{E}\left[\exp\left\{\frac{p^2\lambda^2}{2(p-1)N_\ell^2}\int_0^T |K_r^\ell|^2 dr\right\}\right]\right)^{\frac{(p-1)N_\ell}{p}}.$$

Now, by the Malliavin chain rule we have

$$|K_r^\ell|^2 = \left|\mathbb{E}\left[\mathcal{D}_r X_T^{m^\ell} \nabla f(X_T^{m^\ell}) - \mathcal{D}_r X_T^{m^{\ell-1}} \nabla f(X_T^{m^{\ell-1}}) | \mathcal{F}_r\right]\right|^2,$$

Under our assumption, we also have

Lemma

For all $1 \leq j \leq d$ we have

$$\Big(\sup_{r\in[0,T]} \sup_{t\in[r,T]} |\mathcal{D}_r^j X_t|\Big) \vee \Big(\sup_{r\in[0,T]} \sup_{t\in[r,T]} \sup_{n\in\mathbb{N}^*} |\mathcal{D}_r^j X_t^n|\Big) \leq e^{T[\dot{b}]_\infty}.$$

Then, the process $(\exp\{\frac{p\lambda}{N_\ell}\int_0^t K_r^\ell \cdot dW_r - \frac{p^2\lambda^2}{2N_\ell^2}\int_0^t |K_r^\ell|^2 dr\})_{0\leq t\leq T}$ is a martingale, which leads us to

$$\mathbb{E}\left[\exp(\lambda \hat{Q}_2)\right] \leq \prod_{\ell=1}^{L} \mathbb{E}^{N_\ell \frac{(p-1)}{p}}\left[\exp\left\{\frac{p^2\lambda^2}{2(p-1)N_\ell^2}\int_0^T |K_r^\ell|^2 dr\right\}\right].$$

- For $\ell \in \{1,\ldots,L\}$, we denote

$$U_T^\ell := X_T^{m^\ell} - X_T^{m^{\ell-1}}.$$

- Hence, using the above Lemma and the fact that ∇f is a Lipschitz continuous and bounded function, we get

$$|K_r^\ell|^2 \le 2e^{2T[\dot{b}]_\infty}[\dot{f}]_{\mathrm{lip}}^2 \mathbb{E}\left[|U_T^\ell|^2|\mathcal{F}_r\right] + 2[\dot{f}]_\infty^2 \mathbb{E}\left[\|\mathcal{D}_r U_T^\ell\|^2|\mathcal{F}_r\right],$$

- By Cuachy Schwarz and Jensen inequalities we get

$$\mathbb{E}\left[\exp\left\{\frac{p^2\lambda^2}{2(p-1)N_\ell^2}\int_0^T |K_r^\ell|^2 dr\right\}\right] \le$$
$$\mathbb{E}^{\frac{1}{2}}\left[\exp\left\{\frac{2p^2\lambda^2 e^{2T[\dot{b}]_\infty}[\dot{f}]_{\mathrm{lip}}^2}{(p-1)N_\ell^2}\int_0^T \mathbb{E}\left[|U_T^\ell|^2|\mathcal{F}_r\right] dr\right\}\right] \times$$
$$\left(\frac{1}{d}\sum_{i=1}^d \mathbb{E}\left[\exp\left\{\frac{2dp^2\lambda^2[\dot{f}]_\infty^2}{(p-1)N_\ell^2}\int_0^T \mathbb{E}\left[|\mathcal{D}_r^i U_T^\ell|^2|\mathcal{F}_r\right] dr\right\}\right]\right)^{\frac{1}{2}}.$$

Theorem 2

Let $m, n \in \mathbb{N}^*$, $U_T = X_T^n - X_T^{mn}$ *and* ρ *be a constant satisfying*

$$0 \leq \rho \leq \frac{(mn)^2}{C_1 T (m-1)^2}$$

where C_1 *is an explicit positive constant. Then, Under* (**R1**), *we have the existence of an explicit positive constant* C_2 *such that*

$$\mathbb{E}\left[\exp\left\{\rho \int_0^T \mathbb{E}\left[|U_T|^2 | \mathcal{F}_r\right] dr\right\}\right] \leq \exp\left\{\rho C_2 \left(\frac{(m-1)T}{mn}\right)^2\right\}.$$

Theorem 3

Let ρ be a constant satisfying

$$0 \leq \rho \leq \frac{(mn)^2}{C_3 T(m-1)^2}$$

where C_3 is an explicit positive constant. Then, Under (**R2**)*, we have the existence of an explicit positive constant C_4 such that for all $1 \leq j \leq d$ we have*

$$\mathbb{E}\left[\exp\left\{\rho \int_0^T \mathbb{E}\left[|\mathcal{D}_r^j U_T^\ell|^2 |\mathcal{F}_r\right] dr\right\}\right] \leq \exp\left\{\rho C_4 \left(\frac{(m-1)T}{mn}\right)^2\right\}.$$

- Then, according to theorems 2 and 3 combined with the above inequalities we deduce that for

$$\lambda \leq \mathcal{C}' \min_{1\leq \ell \leq L} m^{\ell} N_{\ell}$$

with $\mathcal{C}'$ an explicit positive constant

$$\mathbb{P}\left(\hat{Q} - \mathbb{E}f(X_T^{m^L}) \geq \alpha\right) \leq \exp\left\{\psi_\alpha(\lambda)\right\},$$

where $\mathcal{C}$ is an explicit positive constant and

$$\psi_\alpha(\lambda) := \lambda^2 \mathcal{C} \sum_{\ell=0}^{L} m^{-2\ell} N_\ell^{-1} - \lambda\alpha,$$

- When $\alpha \leq \mathcal{C}\mathcal{C}' \min_{1\leq\ell\leq L}\{m^{\ell}N_{\ell}\}\sum_{\ell=0}^{L} m^{-2\ell}N_{\ell}^{-1}$,

$$\min_{\lambda\in[0,\mathcal{C}\min_{1\leq\ell\leq L} m^{\ell}N_{\ell}]} \psi_{\alpha}(\lambda) = -\frac{\alpha^2}{4\mathcal{C}\sum_{\ell=0}^{L} m^{-2\ell}N_{\ell}^{-1}}$$

- Which leads to the

Concentration Inequality

$$\mathbb{P}\left(\hat{Q} - \mathbb{E}f(X_T^{m^L}) \geq \alpha\right) \leq \exp\left(-\frac{\alpha^2}{4\mathcal{C}\sum_{\ell=0}^{L} m^{-2\ell}N_{\ell}^{-1}}\right).$$

Thank you !