SDE Numerical Analysis for Multilevel Function Approximation

Filippo De Angelis, Mike Giles, Christoph Reisinger

Mathematical Institute, University of Oxford

MCQMC'24, University of Waterloo, Canada

August 19-23, 2024

Outline

- Stefan Heinrich's MLMC for parametric integration
- MLFA for SDEs – extension of Heinrich's approach
  - numerical analysis for integrable SDEs
  - MIMC decomposition for SDE approximations
  - numerical analysis for smooth and non-smooth "payoffs"
  - strong convergence for pathwise sensitivities
- conclusions and references

MLMC for parametric integration

Stefan Heinrich's original MLMC research (2001) concerned the approximation of

$$f(\theta) = \mathbb{E}[g(\theta;\omega)],$$

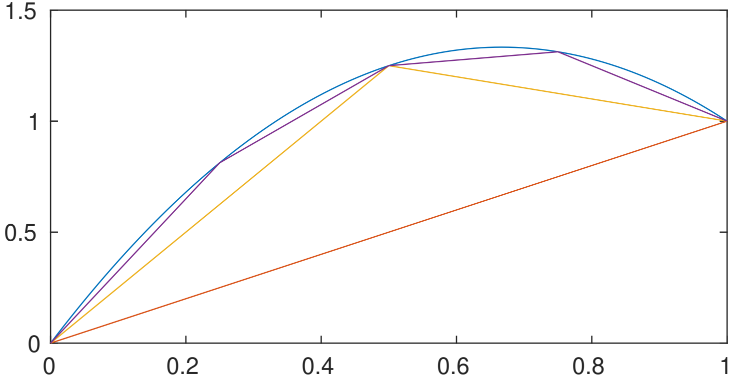
given exact sampling of $g(\theta;\omega)$ at unit cost, with $\theta \in [0,1]^d$.

In his formulation, the MLMC telescoping sum is

$$f \approx I_L[f] = I_0[f] + \sum_{\ell=1}^{L} I_\ell[f] - I_{\ell-1}[f]$$

where $I_\ell[f]$ represents a level ℓ interpolation, e.g. piecewise linear interpolation in 1D with spacing $2^{-\ell}$, and tensor product multilinear interpolation in higher dimensions.

MLMC for parametric integration



Here we see 3 levels of approximation, with the difference $I_\ell[f] - I_{\ell-1}[f]$ getting progressively smaller as ℓ increases.

MLMC for parametric integration

Heinrich then approximates $(I_\ell - I_{\ell-1})[f]$ through Monte Carlo sampling at required values of θ:

$$(I_\ell - I_{\ell-1})[f] \approx \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} (I_\ell - I_{\ell-1})[g(\,\cdot\,;\omega^{\ell,m})]$$

As $\ell \to \infty$, $(I_\ell - I_{\ell-1})[g] \to 0$, and therefore

$$\mathbb{V}\left[\,(I_\ell - I_{\ell-1})[g]\,\right] \;\equiv\; \mathbb{E}\left[\,\left\|\,(I_\ell - I_{\ell-1})[g] - \mathbb{E}[(I_\ell - I_{\ell-1})[g]]\,\right\|_2^2\,\right] \;\to\; 0$$

so fewer MC samples needed on finer levels.

Analysis assumes the number of θ points increases exponentially with dimension so the resulting complexity for linear interpolation is of order

$$\begin{array}{cl} \varepsilon^{-2}, & d < 2r \\ \varepsilon^{-2}\,|\log\varepsilon|^2, & d = 2r \\ \varepsilon^{-d/r}, & d > 2r \end{array}$$

where $r \in \{1,2\}$ is the degree of smoothness of f and g with respect to θ.

MLMC for parametric integration

Heinrich's work is the key starting point for our current work which extends it in several directions:

- PDEs with appropriate numerical approximation (Filippo)
- sparse grid interpolation to address curse of dimensionality (Filippo)
- replace $g(\theta;\omega)$ by $P(\theta;\omega) = g(S_T(\theta;\omega))$ where S_T is the solution of an SDE
- weaker assumptions on smoothness of g and hence P
- numerical approximation of $f(\theta) \equiv \mathbb{E}[P(\theta;\omega)]$ in cases without a finite variance, finite expected cost unbiased estimator
- numerical analysis for PDEs (Filippo) and SDEs

MLFA for parametric integration

If $Y(\theta;\omega)$ is an unbiased estimator for $f(\theta) \equiv \mathbb{E}[P(\theta;\omega)]$, with

$$\begin{aligned} \| (I_\ell - I)\,[f]\, \| &< c_1\, 2^{-r\ell} \\ \mathbb{V}\left[(I_\ell - I_{\ell-1})\,[Y]\right] &< c_2\, 2^{-s\ell} \end{aligned}$$

and the total expected cost is bounded by $c_3 \displaystyle\sum_0^L 2^{d\ell} M_\ell$, for M_ℓ samples per level, then ε r.m.s. accuracy can be achieved with cost of order

$$\begin{array}{ll} \varepsilon^{-2}, & d < s \\ \varepsilon^{-2} |\log \varepsilon|^2, & d = s \\ \varepsilon^{-2-(d-s)/r}, & d > s \end{array}$$

This is a slight generalisation of Heinrich's original result which corresponds to $s = 2r$.

Numerical analysis

In 1D, using piecewise linear interpolation, the maximum value of $(I_\ell - I_{\ell-1})\,[Y]$ is at a midpoint of a coarse θ interval, so the numerical analysis involves bounding

$$\mathbb{E}\left[\left(\delta^2 Y(\theta;\omega)\right)^2\right]$$

where $\delta^2 Y(\theta_0;\omega) = Y(\theta_0{+}\Delta\theta;\omega) - 2\,Y(\theta_0;\omega) + Y(\theta_0{-}\Delta\theta;\omega)$

We are concerned with applications in mathematical finance for which

$$P(\theta;\omega) = g(S_T(\theta;\omega))$$

with S_T being the final value for an SDE solution with ω representing the driving Brownian motion.

Numerical analysis for integrable SDEs

For an integrable SDE, we use $Y(\theta;\omega) = P(\theta;\omega) = g(S_T(\theta;\omega))$.

We assume the SDE satisfies the usual conditions and therefore for each $p>0$ there exists $c^{(p)}$ such that

$$\mathbb{E}\left[\|S_T\|^p\right] \leq c^{(p)}$$

Furthermore, we assume the drift and diffusion coefficients are smooth w.r.t. θ, S and therefore for integer $q>0$, and any $p>0$, there exists $c^{(p,q)}$ such that

$$\mathbb{E}\left[\left\|\frac{\partial^q S_T}{\partial \theta^q}\right\|^p\right] \leq c^{(p,q)}$$

This can be proved given bounded derivatives for the drift and diffusion coefficients, but I haven't yet found a reference for it.

Numerical analysis for integrable SDEs

For twice-differentiable payoff functions,

$$\begin{aligned}\dot{Y}(\theta,\omega) &= g'(S_T(\theta,\omega))\ \dot{S}_T(\theta,\omega),\\ \ddot{Y}(\theta,\omega) &= g''(S_T(\theta,\omega))\ (\dot{S}_T(\theta,\omega))^2 + g'(S_T(\theta,\omega))\ \ddot{S}_T(\theta,\omega).\end{aligned}$$

where $\dot{Y} \equiv \partial Y/\partial\theta$, and $g' \equiv \mathrm{d}g/\mathrm{d}S$. We then have

$$\delta^2 Y = \int_{\theta_0-\Delta\theta}^{\theta_0+\Delta\theta} (\Delta\theta - |\theta-\theta_0|)\ \ddot{Y}(\theta,\omega)\,\mathrm{d}\theta,$$

and hence $\delta^2 Y = O(\Delta\theta^2)$ and $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta^4)$, giving $s{=}4$ as well as $r{=}2$ in the meta-theorem.

This corresponds to the smooth case analysed by Stefan Heinrich. However, in mathematical finance the payoff function is rarely twice-differentiable.

Numerical analysis for non-smooth payoffs

At the other extreme, consider a digital option for which the payoff is an indicator function $g(S_T) = \mathbb{1}_{S_T \in K}$

For this, we follow previous research in assuming that there exists a constant c such that for all θ, and all $\delta > 0$,

$$\mathbb{P}[d(S_T, \partial K) < \delta] < c\,\delta$$

Heuristically, this corresponds to S_T having a bounded density, but it also requires the set K to not be pathological.

G., Haji-Ali (2024/5) give conditions under which this assumption is satisfied, and also examples of pathological K for which it is not.

Numerical analysis for non-smooth payoffs

Heuristic analysis:

- $O(\Delta\theta)$ probability of $S_T(\theta_0;\omega)$ being within $O(\Delta\theta)$ of ∂K
- $\Longrightarrow$ $O(\Delta\theta)$ probability of $S_T(\theta;\omega)$ for $\theta_0 - \Delta\theta < \theta < \theta_0 + \Delta\theta$ crossing ∂K, giving $\delta^2 Y = O(1)$
- otherwise, $\delta^2 Y = 0$
- hence, $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta)$

The rigorous version of this gives

$$\mathbb{E}[(\delta^2 Y)^2] = o(\Delta\theta^{1-\delta})$$

for any $\delta > 0$, so $s \approx 1$, but $r = 2$.

Numerical analysis for non-smooth payoffs

Similarly, for Lipschitz functions with a bounded second derivative except on ∂K (e.g. European put/call functions), the heuristic analysis is:

- $O(\Delta\theta)$ probability of $S_T(\theta_0;\omega)$ being within $O(\Delta\theta)$ of ∂K
- $\Longrightarrow$ $O(\Delta\theta)$ probability of $S_T(\theta;\omega)$ for $\theta_0 - \Delta\theta < \theta < \theta_0 + \Delta\theta$ crossing ∂K, giving $\delta^2 Y = O(\Delta\theta)$
- otherwise, $\delta^2 Y = O(\Delta\theta^2)$
- hence, $\mathbb{E}[(\delta^2 Y)^2] = O(\Delta\theta^3)$

The rigorous version of this gives

$$\mathbb{E}[(\delta^2 Y)^2] = o(\Delta\theta^{3-\delta})$$

for any $\delta > 0$, so $s \approx 3$, but $r = 2$.

MIMC for SDE approximations

Almost all SDEs in mathematical finance need to be approximated, so we use MIMC approach of Haji-Ali, Nobile & Tempone (2016), starting from

$$f \approx \sum_{\ell=0}^{L} (I_\ell - I_{\ell-1})[f], \qquad I_{-1}[f] \equiv 0,$$

where I_ℓ uses a tensor product interpolation with spacing proportional to $2^{-\ell}$ in each direction.

We then replace f with a timestep approximation (e.g. Euler-Maruyama) expansion

$$f \approx \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \Delta I_\ell[\Delta f_{\ell'}], \qquad \Delta I_\ell[\Delta f_{\ell'}] \equiv (I_\ell - I_{\ell-1})[f_{\ell'} - f_{\ell'-1}]$$

in which L'_ℓ is a decreasing function of ℓ, since $\Delta I_\ell[f]$ becomes smaller as ℓ increases and so less relative accuracy is required in its approximation.

MIMC for SDE approximations

The final step is to replace $\Delta I_\ell[\Delta f_{\ell'}]$ by a Monte Carlo estimate, giving the MIMC estimator

$$\widetilde{f} = \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \left(\frac{1}{M_{\ell,\ell'}} \sum_{m=1}^{M_{\ell,\ell'}} \Delta I_\ell [\Delta P_{\ell'}(\cdot;\omega^{\ell,\ell',m})] \right)$$

We now need to choose $L, L'_\ell, M_{\ell,\ell'}$ to achieve the desired accuracy at the minimum cost.

$$\begin{aligned} \mathbb{E}[\widetilde{f}-f] &= (I_L - I)[f] + \sum_{\ell=0}^{L} (I_\ell - I_{\ell-1})[f_{L'(\ell)} - f] \\ \Longrightarrow \quad \left\| \mathbb{E}[\widetilde{f}-f] \right\| &\leq \|(I_L - I)[f]\| + \sum_{\ell=0}^{L} \|(I_\ell - I_{\ell-1})[f_{L'(\ell)} - f]\| \end{aligned}$$

and

$$\mathbb{V}[\widetilde{f}] = \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} \left(\frac{1}{M_{\ell,\ell'}} \mathbb{V}\left[\Delta I_\ell[\Delta P_{\ell'}(\cdot;\omega)]\right] \right)$$

MIMC for SDE approximations

If we have

$$\| \Delta I_\ell [\Delta f_{\ell'}] \| \;<\; c_1 \, 2^{-r\ell - \alpha \ell'}$$

$$\mathbb{V}\left[\Delta I_\ell [\Delta P_{\ell'}]\right] \;<\; c_2 \, 2^{-s\ell - \beta \ell'}$$

and the total cost is bounded by

$$c_3 \sum_{\ell=0}^{L} \sum_{\ell'=0}^{L'_\ell} 2^{d\ell + \gamma \ell'} M_{\ell,\ell},$$

then ε RMS accuracy can be achieved at a computational cost of order

$$\begin{array}{ll} \varepsilon^{-2}, & \eta < 0 \\ \varepsilon^{-2-\eta} \, |\log \varepsilon|^p, & \eta \geq 0 \end{array}$$

for some p (see MIMC analysis by Haji-Ali *et al* (2016)), where

$$\eta = \max\left(\frac{\gamma - \beta}{\alpha}, \frac{d - s}{r} \right).$$

Numerical analysis

The challenge now is to bound $\mathbb{V}\left[\Delta I_\ell[\Delta P_{\ell'}]\right]$

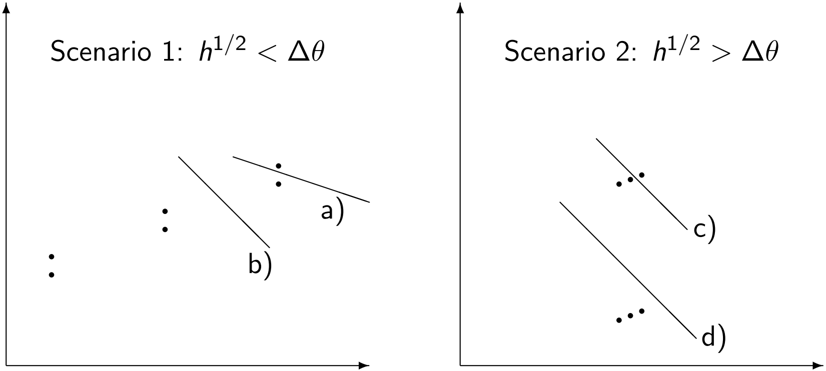
On a level ℓ with spacing $\Delta\theta$, and level ℓ' with timestep h, this involves bounding

$$\mathbb{V}\Big[\ \Big(g(\widehat{S}(\theta_0{-}\Delta\theta,h,\omega))\ -\ 2\,g(\widehat{S}(\theta_0,h,\omega))\ +\ g(\widehat{S}(\theta_0{+}\Delta\theta,h,\omega))\Big) \\ -\Big(g(\widehat{S}(\theta_0{-}\Delta\theta,2h,\omega))-2\,g(\widehat{S}(\theta_0,2h,\omega))+g(\widehat{S}(\theta_0{+}\Delta\theta,2h,\omega))\Big)\ \Big]$$

In the smooth case, this variance is $O(\Delta\theta^4 h)$ for the E-M discretisation, and $O(\Delta\theta^4 h^2)$ for Milstein.

Numerical analysis

In the non-smooth case, there are a number of scenarios to consider for the E-M discretisation regarding the position of ∂K:



Numerical analysis

Eventually, the conclusion is that the variance for the Euler-Maruyama discretisation is approximately

$O(\min(h^{1/2}, \Delta\theta))$ for the digital case, and

$O(\min(\Delta\theta\, h, \Delta\theta^3))$ for the Lipschitz case.

Modifying the meta-theorem, we obtain complexity which is approximately

$O(\varepsilon^{-2-\max((d+1)/4),\ (d-1)/2}$ for the digital case, and

$O(\varepsilon^{-2-\max((d-1)/4,\ (d-3)/2)})$ for the Lipschitz case.

When $d=1$, we get the usual MLMC complexity; as d increases we hit the curse of dimensionality

Strong convergence for pathwise sensitivities

The numerical analysis requires the following strong convergence result for the Euler-Maruyama discretisation.

For any $p>0$ there exists $c^{(p)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\dot{S}}_t - \dot{S}_t\|^p\right] \quad \leq \quad c^{(p)} h^{p/2}$$
$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\ddot{S}}_t - \ddot{S}_t\|^p\right] \quad \leq \quad c^{(p)} h^{p/2}$$

This can also be proved given bounded derivatives for the drift and diffusion coefficients, but I haven't yet found a reference for it.

Conclusions and future work

Conclusions:

- excellent asymptotic efficiency in approximating parametric functions arising from SDEs – nearly optimal in some cases
- initial numerical results support numerical analysis

Future work:

- numerical results
- extension to sparse interpolation, giving complexity $\lesssim\ O(\varepsilon^{-5/2})$
- investigate use of path-branching and conditional expectation for improved variance for non-smooth cases

References

F. De Angelis. PhD thesis, in preparation, Oxford University, 2024.

M.B. Giles, A.-L. Haji-Ali. 'Multilevel path branching for digital options'. Annals of Applied Probability, to appear, 2024/5.

M.B. Giles, O. Sheridan-Methven, 'Analysis of nested multilevel Monte Carlo using approximate Normal random variables'. SIAM/ASA Journal on Uncertainty Quantification 10(1):200-226, 2022.

A.-L. Haji-Ali, F. Nobile, R. Tempone. 'Multi-index Monte Carlo: when sparsity meets sampling'. Numerische Mathematik, 132:767-806, 2016.

S. Heinrich. 'Multilevel Monte Carlo methods'. Lecture Notes in Computer Science, 2179:58-67, 2001.