Weighted multilevel Monte Carlo

Tony Ware
Joint work with Yu Li

ICCF 2-5 April 2024, Amsterdam

Multilevel Monte Carlo

Multilevel Monte Carlo

- **Scenario:** we want to estimate a quantity $\overline{P} = \mathbb{E}[P]$.
- We cannot sample P directly. But we do have access to a sequence of ***approximate*** estimators P_l^l for $l = 0, 1, 2, \ldots$.
- $\overline{P}_l = \mathbb{E}[P_l^l] \to \overline{P}$ as $l \to \infty$, but the samples become increasing costly to generate.
- Typically, $\left|\overline{P}_l - \overline{P}\right|$ is inversely proportional to the cost, while the sampling error is inversely proportional to the square of the number of samples, so that the cost of generating an overall error of ϵ is proportional to $1/\epsilon^3$.

MLMC (Giles, 2008) exploits multiple levels of estimates to reduce the cost to - under certain conditions - as low as $1/\epsilon^2$.

Multilevel Monte Carlo

- Suppose that, whenever we generate a sample of P_l^l, we are able to generate a **correlated** sample of P_{l-1}^l, with mean $\overline{P}_{l-1}$.
- We collect these together as basic estimators:
$$Y_0 = P_0^0, \text{ and } Y_l = P_l^l - P_{l-1}^l \text{ for } l = 1, 2, \ldots$$
- A MLMC estimator at level L takes the form $\boxed{\mathcal{P}_L = \sum_{l=0}^{L} \alpha_l^L Y_l}$.
- The numbers of samples α_l^L are chosen to minimise the cost per unit variance W_L^2:
$$\alpha_l^L = \frac{\Delta_l}{\eta_l} W_L, \text{ with } W_L = \sum_{l=0}^{L} \Delta_l \eta_l,$$
where Δ_l^2 is the variance of Y_l and η_l^2 is the cost of computing Y_l.

Multilevel Monte Carlo

Two ideas

First idea

The ratio of successive numbers of samples is ***independent of*** L, so we can write the MLMC estimator **recursively**:

$$\boxed{\mathcal{P}_0 = {}_{\Delta_0^2}Y_0,} \text{ and, for } l > 0, \boxed{\mathcal{P}_l = \sum_{l'=0}^{l} {}_{\alpha_{l'}^l} Y_{l'} = {}_{\alpha_l} Y_l + {}_{\beta_l}\mathcal{P}_{l-1},}$$

where $\alpha_l = \alpha_l^l = \frac{\Delta_l W_l}{\eta_l}$ and $\beta_l = \frac{W_l}{W_{l-1}}$.

Expanding $\mathcal{P}_l = {}_{\alpha_l}P_l^l - \left({}_{\alpha_l}P_{l-1}^l - {}_{\beta_l}\mathcal{P}_{l-1}\right)$, and noticing that $\mathbb{E}[{}_{\alpha_l}P_{l-1}^l - {}_{\beta_l}\mathcal{P}_{l-1}] = 0$, leads to …

Multilevel Monte Carlo

Two ideas

Second idea

…adding **weights** (in the spirit of control variates (Kebaier 2005)):

$$\mathcal{P}_l = {}_{\alpha_l}P_l^l - \theta_{l-1}\left({}_{\alpha_l}P_{l-1}^l - {}_{\beta_l}\mathcal{P}_{l-1}\right) = {}_{\alpha_l}Y_l^{\theta_{l-1}} + \theta_{l-1}\,{}_{\beta_l}\mathcal{P}_{l-1}$$

where $Y_l^\theta = P_l^l - \theta P_{l-1}^l$.
This results in the ***weighted MLMC estimate***:

$$\boxed{\mathcal{P}_L = \sum_{l=0}^{L} \Theta_l^L\, {}_{\left(\frac{\alpha_l \beta_L}{\beta_l}\right)}Y_l^{\theta_l},}$$

where $\Theta_l^L = \prod_{l'=l}^{L-1} \theta_{l'}$.

Weighted Multilevel Monte Carlo

- We optimise both the numbers of samples and the weights. It is most natural to do this recursively, minimising the cost of $\mathcal{P}_l$ subject to unit variance.
- We start with $\mathcal{P}_0 = {}_{\alpha_0}Y_0 = {}_{\alpha_0}P_0^0$.
  In this case, $\alpha_0 = \Delta_0^2$, and $W_0 = \Delta_0 \eta_0$, so that $\mathbb{V}[\mathcal{P}_0] = 1$.
- (Note that we can set the variances to some value v^2 by multiplying all the values of α by v^2.)
- Suppose we have an optimised WMLMC estimator $\mathcal{P}_{l-1}$. For fixed θ_{l-1}, we can minimise the cost by setting

$$\alpha_l = \frac{\Delta_l^{\theta_{l-1}} W_l}{\eta_l}, \quad \beta_l = |\theta_{l-1}| \frac{W_l}{W_{l-1}},$$
$$W_l = \Delta_l^{\theta_{l-1}} \eta_l + |\theta_{l-1}| W_{l-1}.$$

Weighted Multilevel Monte Carlo

- If $\rho_l = \text{Corr}[P_l^l, P_{l-1}^l] \leq \frac{W_{l-1}}{\eta_l \sigma_{l-1}}$, we take $\theta_{l-1} = 0$, and l becomes our new coarsest level.
- Otherwise, the optimal θ_{l-1} is positive, and we have

$$\theta_{l-1} = \frac{\rho_l \sigma_l}{\sigma_{l-1}} - \frac{\Delta_l^{\theta_{l-1}} W_{l-1}}{\eta_l \sigma_l^2}, \text{ with } \Delta_l^{\theta_{l-1}} = \sqrt{\frac{\sigma_l^2(1-\rho_l^2)}{1-\frac{W_{l-1}^2}{\eta_l^2 \sigma_l^2}}}$$

- This results in

$$\frac{W_l}{\sigma_l} = \rho_l \frac{W_{l-1}}{\sigma_{l-1}} + \sqrt{(1-\rho_l^2)\left(\eta_l^2 - \frac{W_{l-1}^2}{\sigma_{l-1}^2}\right)}$$

which allows us to determine the optimal weights θ_l and the efforts α_l and β_l.

Weighted Multilevel Monte Carlo

Comparison with MLMC

- In order to gain insight into the comparative the costs associated with MLMC and WMLMC it is convenient to assume that $\sigma_l \equiv \sigma$ and that $\eta_l^2 = M\eta_{l-1}^2$, and to focus on the two-level case. We have that $W_0 = \sigma\eta_0 = \widehat{W}_0$ (where $\widehat{W}_l$ is the cost of a single-level estimator at level l).
- We have, for the optimal weighted version

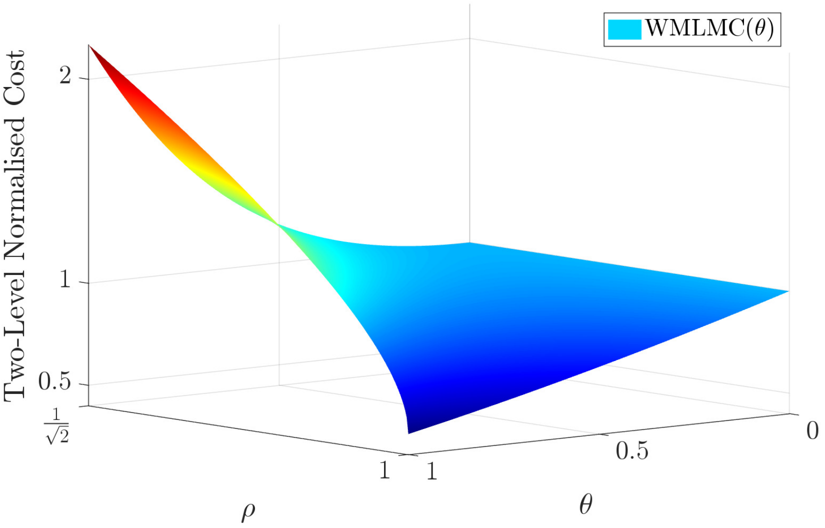
$$W_1 = \begin{cases} \widehat{W}_0 \left(\rho_1 + \sqrt{(1-\rho_1^2)(M-1)}\right) & \text{if } \rho_1 > \frac{1}{\sqrt{M}} \\ \widehat{W}_1 \ (= \sigma\eta_1) & \text{otherwise.} \end{cases}$$

- And for MLMLC (which corresponds to taking $\theta_0 = 1$)

$$W_1 = \begin{cases} \widehat{W}_0 \left(1 + \sqrt{2M(1-\rho_1)}\right) & \text{if } \rho_1 > 1 - \frac{1}{2}\left(1 - \frac{1}{\sqrt{M}}\right)^2 \\ \widehat{W}_1 & \text{otherwise.} \end{cases}$$

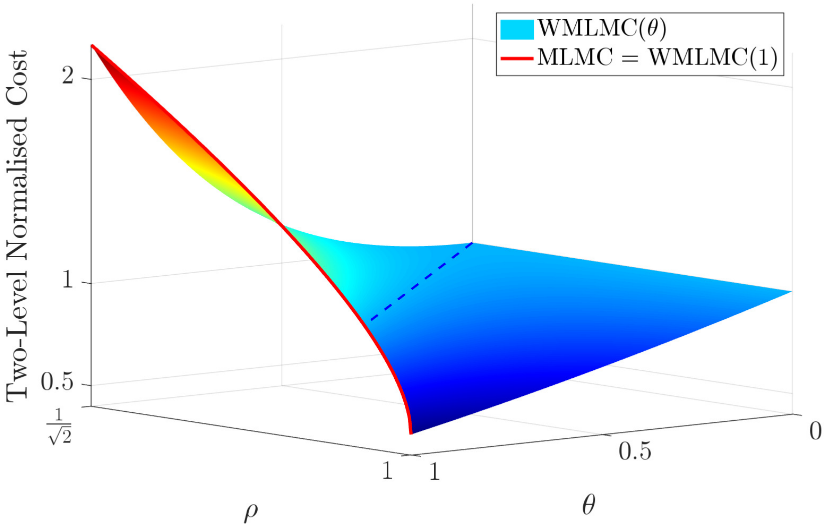
Weighted Multilevel Monte Carlo

Two-level normalised cost (i.e. $W_1^2/\widehat{W}_1^2$)



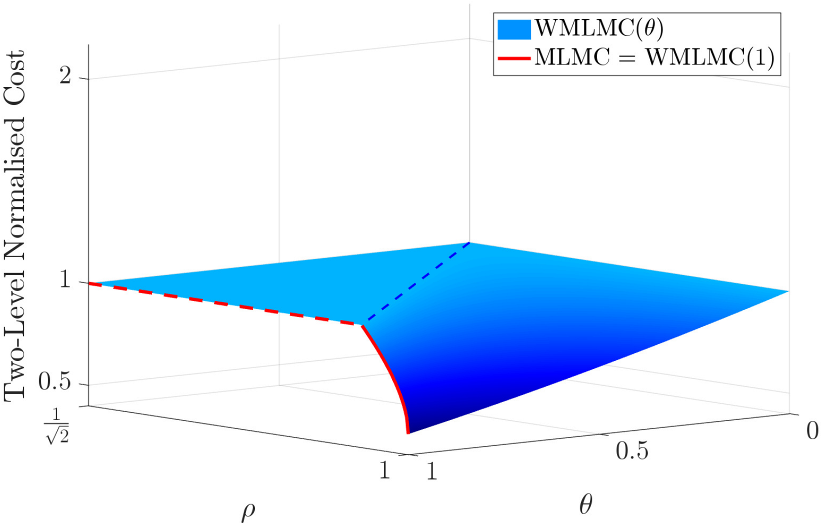
Weighted Multilevel Monte Carlo

Two-level normalised cost (i.e. $W_1^2/\widehat{W}_1^2$)



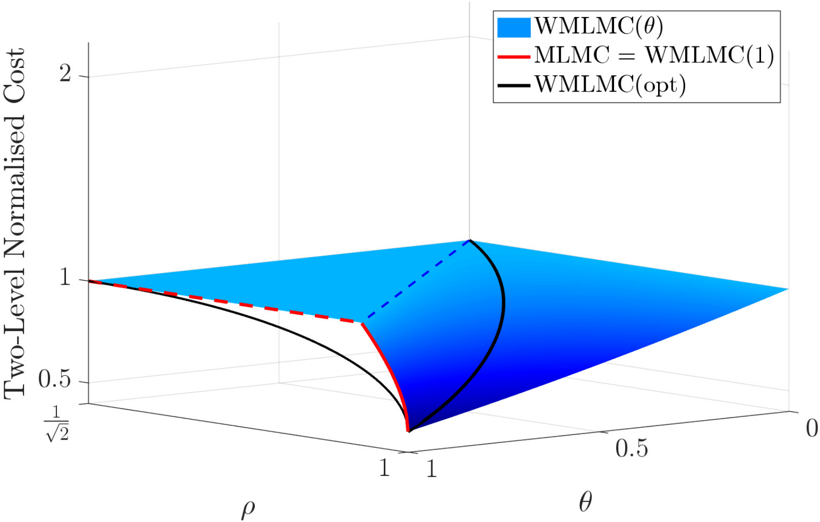
Weighted Multilevel Monte Carlo

Two-level normalised cost (i.e. $W_1^2/\widehat{W}_1^2$)



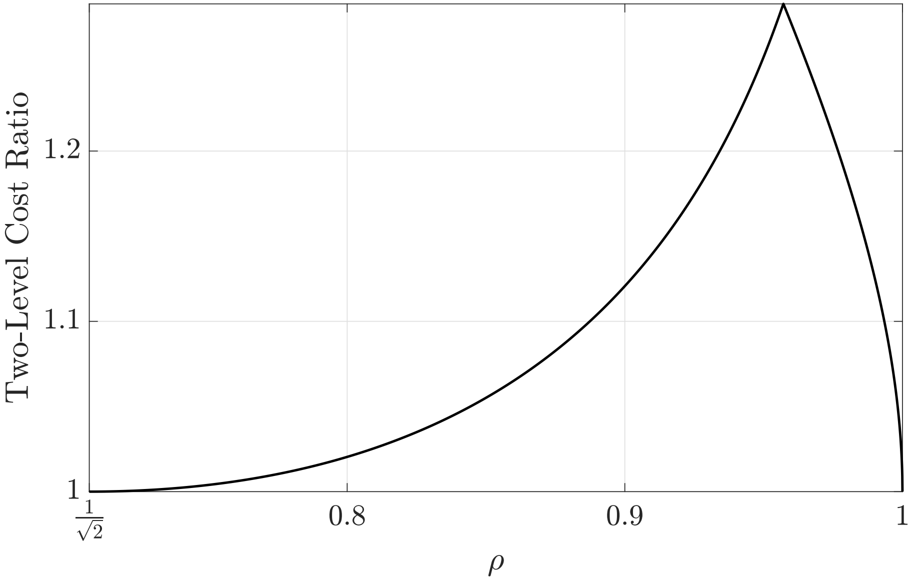
Weighted Multilevel Monte Carlo

Two-level normalised cost (i.e. $W_1^2/\widehat{W}_1^2$)



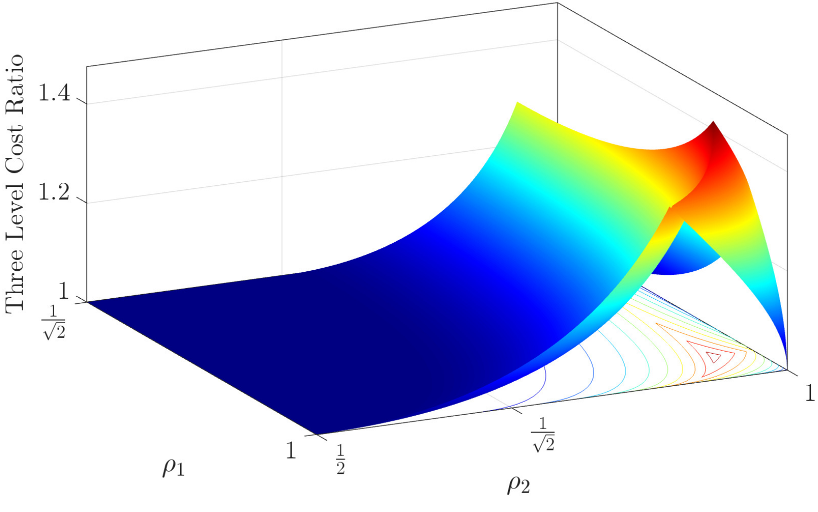
Weighted Multilevel Monte Carlo

Two-level cost ratio (unweighted/weighted)



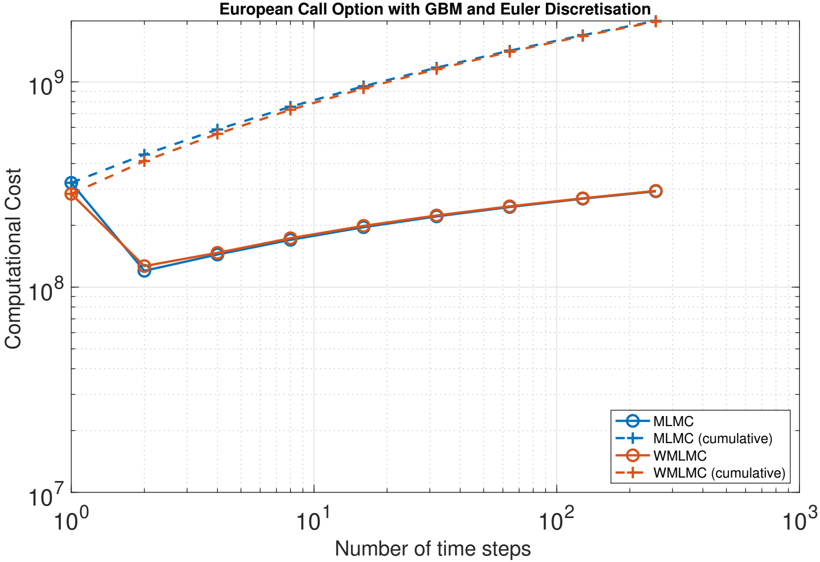
Weighted Multilevel Monte Carlo

Three-level cost ratio (unweighted/weighted)



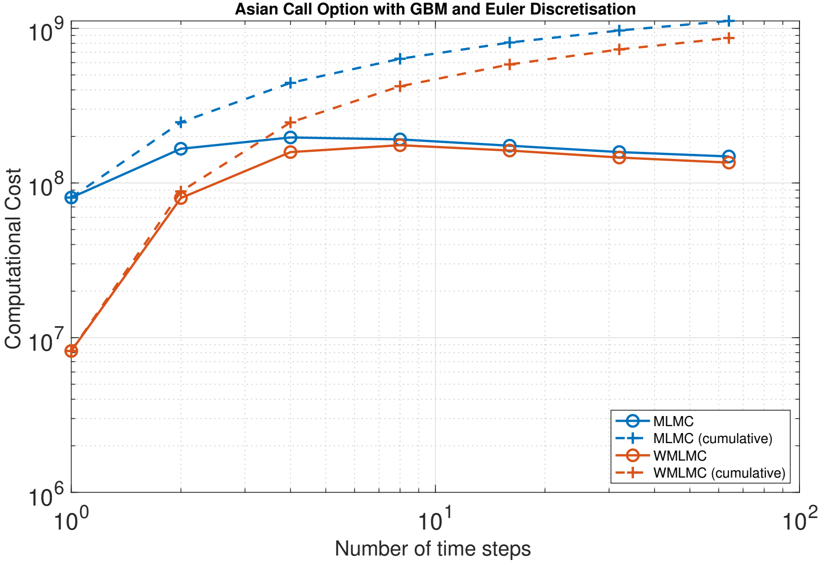
Weighted Multilevel Monte Carlo

Numerical Comparisons



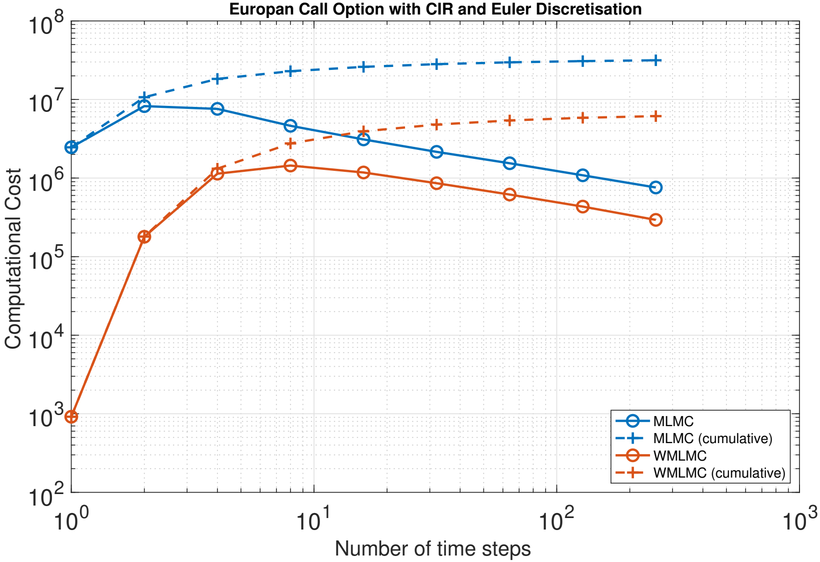
Weighted Multilevel Monte Carlo

Numerical Comparisons



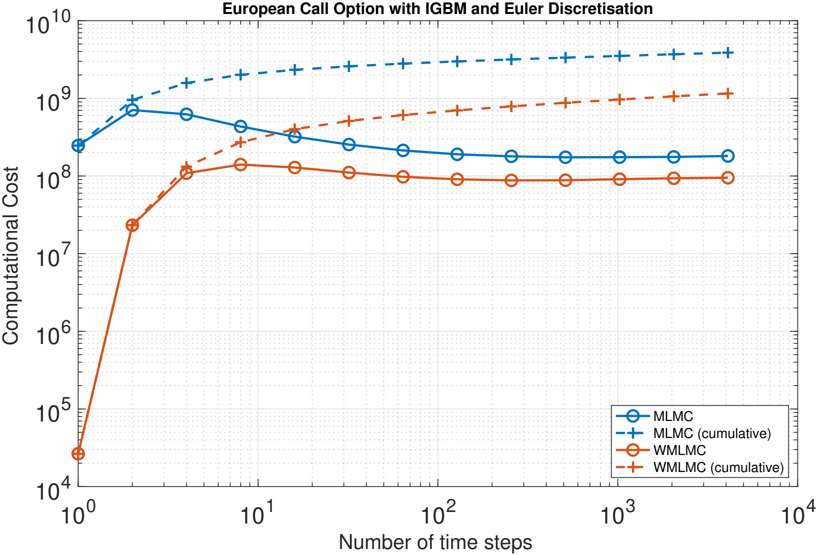
Weighted Multilevel Monte Carlo

Numerical Comparisons



Weighted Multilevel Monte Carlo

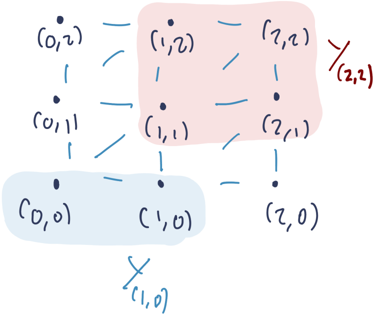
Numerical Comparisons



Weighted Multi-Index MLMC

Weighted Multi-Index Multilevel Monte Carlo

- Our estimates are now indexed by multi-indices $\lambda \in \mathbb{N}^2$.
- The MIMLMC method uses alternating signs to define Y_λ.
- It can (again) be written in a recursive form, and weights can be added.
- There is no longer an explicit expression for the optimal weights, but they can be determined numerically at each node.



$$\mathcal{P}_\lambda = \alpha_\lambda Y_\lambda^{t_\lambda} + \sum_{\mu \in \square_\lambda^-} \theta_\mu^\lambda \left(\frac{\beta_\lambda}{\beta_\mu}\right) \mathcal{P}_\mu$$

$$= \sum_{\mu \leq \lambda} \Theta_\mu^\lambda \left(\frac{\alpha_\mu \beta_\lambda}{\beta_\mu}\right) Y_\mu .$$

Weighted Multi-Index Multilevel Monte Carlo

Some more details

We write

$$\mathcal{P}_\lambda = {}_{\alpha_\lambda}Y_\lambda^{t_\lambda} + \sum_{\mu \in \square_\lambda^-} \theta_\mu^\lambda \ {}_{\left(\frac{\beta_\lambda}{\beta_\mu}\right)}\mathcal{P}_\mu$$

$$= \sum_{\mu \leq \lambda} \Theta_\mu^\lambda \ {}_{\left(\frac{\alpha_\mu \beta_\lambda}{\beta_\mu}\right)}Y_\mu.$$

- The index set $\square_\lambda^- = \{\mu \geq (0,0) | \max_i \lambda_i - \mu_i = 1\}$.
- Θ_μ^λ is the sum over all paths from μ to λ of the product of the values of $\theta_{\mu'}^{\lambda'}$ along each path.
- The effort $\frac{\alpha_\mu \beta_\lambda}{\beta_\mu}$ is independent of the path from μ to λ by virtue of the recursive construction.

Weighted Multi-Index Multilevel Monte Carlo

Some *more* details

- Given the recursive relation

$$\mathcal{P}_\lambda = {}_{\alpha_\lambda}Y_\lambda^{t_\lambda} + \sum_{\mu \in \square_\lambda^-} \theta_\mu^\lambda \; {}_{\left(\frac{\beta_\lambda}{\beta_\mu}\right)}\mathcal{P}_\mu,$$

  we can, given $t_\lambda = [\theta_\mu^\lambda]_{\mu \in \square_\lambda^-}$, determine α_λ and β_λ so as to minimise the effort needed to ensure $\mathbb{V}[\mathcal{P}_\lambda] = 1$.
- It remains to optimise over t_λ, and this involves minimising the sum of the square roots of two quadratic forms defined by t_λ. This minimisation is performed numerically at each node.
- The unweighted MIMLMC method consists in setting $\theta_\mu^\lambda = (-1)^{1+|\lambda-\mu|}$.

Weighted Multi-Index Multilevel Monte Carlo

Numerical comparisons: Zakai SPDE

We consider the Zakai SPDE (Reisinger/Wang 2018):

$$\mathrm{d}v = -\mu\frac{\partial v}{\partial x}\mathrm{d}t + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}\mathrm{d}t - \sqrt{\rho}\frac{\partial v}{\partial x}\mathrm{d}M_t.$$

The quantity of interest is $L_t = 1 - \int_0^\infty v(t,x)\mathrm{d}x$. We compute with a timestep $k = 4^{-m+1}$ and a grid spacing in x of $h = 2^{-n}$, and compare the effort required to achieve a fixed variance using the unweighted and weighted MIMLMC:

$n\backslash m$	0	1	2	3
0	1.0	1.0	1.0	1.0
1	1.6	1.5	1.5	1.7
2	1.8	1.9	1.8	2.0
3	2.0	1.9	1.9	1.9

Weighted Multi-Index Multilevel Monte Carlo

Numerical comparisons: a hybrid COS-Monte-Carlo basket option valuation

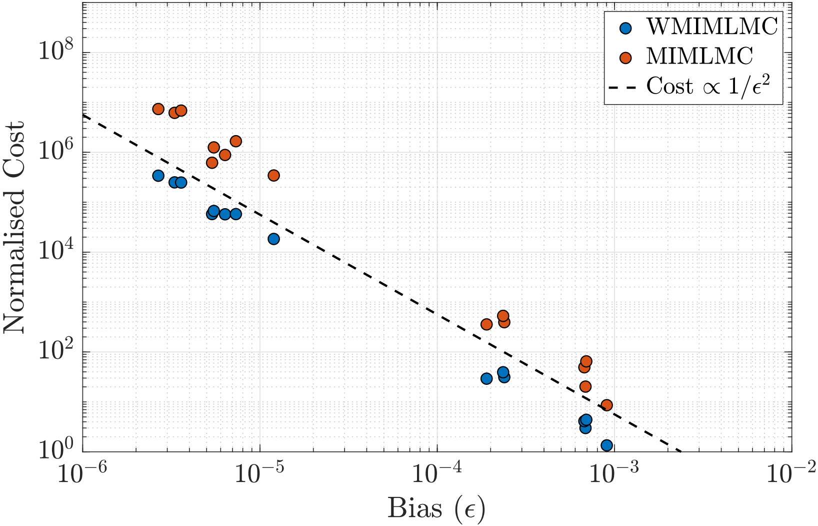
- Here we express a $(d+2)$-asset basket option payoff as a function $\Lambda(Z_1, Z_2)$ of independent normal random variables $Z_1 \in \mathbb{R}^2$ and $Z_2 \in R^d$. The option value can be written as a nested expectation:

$$\mathbb{E}[\Lambda_2(Z_2)], \quad \text{with } \Lambda_2(z_2) = \mathbb{E}[\Lambda(Z_1, z_2)].$$

- For each sample of Z_2, the inner expectation is computed using a two-dimensional COS method, with $(2^{1+L_i} - 1)$ modes used in dimension i (for $i = 1, 2$).
- We again compare MIMC with WMIMC. For some combinations of L_1 and L_2, MIMC provides no advantage over a single-level estimate.

Weighted Multi-Index Multilevel Monte Carlo

Numerical comparisons: a hybrid COS-Monte-Carlo basket option valuation



Weighted Multi-Index Multilevel Monte Carlo

Numerical comparisons: a hybrid COS-Monte-Carlo basket option valuation

The ratio between the unweighted and weighted MIMLMC costs to achieve a given variance

$L_1 \backslash L_2$	0	1	2	3	4	5	6
0							
1	1.49	1.59	1.57	1.56	1.75	1.71	1.70
2	1.46	1.61	1.86	2.85	3.51	3.17	4.26
3	1.59	1.81	1.93	6.18	6.57	11.78	14.58
4	1.58	2.17	4.25	11.93	10.44	15.22	18.47
5	1.61	2.36	5.81	12.39	18.51	24.46	27.37
6	1.64	2.35	5.82	13.25	21.41	28.56	31.48

Conclusions

- MLMC and MIMLMC can be formulated recursively, and in that way are naturally viewed as nested control variate variance reduction techniques.
- As such, weights can be added, and the optimal weights computed at each node.
- The addition of weights does not change the asymptotic *rate* of convergence, but it does allow for more efficient use of estimates at coarser resolutions (and lower correlations) than with unweighted MLMC, resulting in potentially significant gains in performance.
- The gains are relatively insensitive to the optimal choice of weights.
- The multi-index version offers potentially even greater relative improvement.

References

References I

Amri, M. R. E., Mycek, P., Ricci, S., and De Lozzo, M. (2023). Multilevel surrogate-based control variates. *arXiv preprint arXiv:2306.10800*.

Giles, M. B. (2008). Multilevel Monte Carlo path simulation. *Operations Research-Baltimore*, 56(3):607–617.

Giles, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328.

Giorgi, D., Lemaire, V., and Pagès, G. (2020). Weak error for nested multilevel Monte Carlo. *Methodology and Computing in Applied Probability*, 22(3):1325–1348.

Lemaire, V. and Pagès, G. (2017). Multilevel Richardson-Romberg extrapolation.

Pagès, G. (2007). Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity.

Reisinger, C. and Wang, Z. (2018). Analysis of Multi-Index Monte Carlo estimators for a Zakai SPDE. *Journal of Computational Mathematics*, 36(2):202–236.

Ruijter, M. and Oosterlee, C. (2012). Two-dimensional Fourier Cosine series expansion method for pricing financial options. *SIAM Journal on Scientific Computing*, 34(5):B642–B671.