MLMC analysis of the stochastic heat equation

Mike Giles

Mathematical Institute, University of Oxford

Motivation

New research interest – MLMC for parabolic SPDEs:

- 1D stochastic heat equation is the simplest example driven by space-time white noise (cylindrical Wiener process)
- focus on
  - three different noise representations: spectral, mass-lumped finite element, finite volume
  - three different quantities of interest (QoI): squared amplitude of a single Fourier mode, energy $\|u\|_2^2$, and $\langle \varphi, u \rangle^2$

Key messages:

- finite volume treatment has worst MLMC variance
- finite element is as good as spectral, and both benefit from Richardson extrapolation to overcome MLMC oddity ($\beta > 2\alpha$)

Multilevel Monte Carlo

Given a sequence $P_0, P_1, P_2, \ldots \to P$

$$\mathbb{E}[P] \;\approx\; \mathbb{E}[P_L] \;=\; \mathbb{E}[P_0] + \sum_{\ell=1}^{L} \mathbb{E}[P_\ell - P_{\ell-1}]$$

so we can use the estimator

$$N_0^{-1} \sum_{n=1}^{N_0} P_0^{(0,n)} \;+\; \sum_{\ell=1}^{L} \left\{ N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(P_\ell^{(\ell,n)} - P_{\ell-1}^{(\ell,n)} \right) \right\}$$

with independent estimation for each level of correction.

$\mathbb{V}[P_\ell - P_{\ell-1}] \to 0$ as $\ell \to \infty$ means we don't need many samples on finer levels.

MLMC Meta Theorem

(Slight generalisation of version in 2008 *Operations Research* paper)

If there exist independent estimators Y_ℓ based on N_ℓ Monte Carlo samples, each costing C_ℓ, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$ and

i) $|\mathbb{E}[P_\ell - P]| \leq c_1\, 2^{-\alpha\,\ell}$

ii) $\mathbb{E}[Y_\ell] = \begin{cases} \mathbb{E}[P_0], & \ell = 0 \\ \mathbb{E}[P_\ell - P_{\ell-1}], & \ell > 0 \end{cases}$

iii) $\mathbb{V}[Y_\ell] \leq c_2\, N_\ell^{-1} 2^{-\beta\,\ell}$

iv) $\mathbb{E}[C_\ell] \leq c_3\, 2^{\gamma\,\ell}$

MLMC Theorem

then there exists a positive constant c_4 such that for any $\varepsilon < 1$ there exist L and N_ℓ for which the multilevel estimator

$$Y = \sum_{\ell=0}^{L} Y_\ell,$$

has a mean-square-error with bound $\mathbb{E}\left[\left(Y - \mathbb{E}[P]\right)^2\right] < \varepsilon^2$

with an expected computational cost C with bound

$$C \leq \begin{cases} c_4\, \varepsilon^{-2}, & \beta > \gamma, \\ c_4\, \varepsilon^{-2}(\log \varepsilon)^2, & \beta = \gamma, \\ c_4\, \varepsilon^{-2-(\gamma-\beta)/\alpha}, & 0 < \beta < \gamma. \end{cases}$$

Note: the MLMC parameters α, β, γ determine the asymptotic cost.

Stochastic heat equation

Stochastic heat equation:

$$\mathrm{d}u - u_{xx}\,\mathrm{d}t = \mathrm{d}W, \qquad 0 < x < 1,\ t > 0,$$

subject to homogeneous initial data and b.c.'s.

$\mathrm{d}W$ is the increment of space-time white noise (cylindrical Wiener process) so that for arbitrary f, g,

$$\langle f, \mathrm{d}W \rangle$$

is Normally-distributed with zero mean and variance $\|f\|^2\,\mathrm{d}t$, and

$$\mathbb{E}[\langle f, \mathrm{d}W\rangle \langle g, \mathrm{d}W\rangle] = \langle f, g\rangle$$

Stochastic heat equation

Solution has expansion in orthonormal modes $e_k(x) = \sqrt{2}\sin(k\pi x)$:

$$u(x,t) = \sum_{k=1}^{\infty} \widehat{u}_k(t)\, e_k(x)$$

Substituting gives Ornstein-Uhlenbeck SDE for each mode:

$$\mathrm{d}\widehat{u}_k = -\lambda_k \hat{u}_k \,\mathrm{d}t + \mathrm{d}\widehat{W}_k, \quad \lambda_k = k^2\pi^2,$$

where $\widehat{W}_k$ are independent Brownian motions, due to orthonormality of eigenmodes.

Stochastic heat equation

The O-U solution is

$$\widehat{u}_k(t) = \int_0^t e^{-\lambda_k(t-s)} \, \mathrm{d}\widehat{W}_k(s),$$

and hence

$$\mathbb{E}[\hat{u}_k^2(t)] = \frac{1}{2\lambda_k}(1 - e^{-2\lambda_k t}) \equiv \sigma_k^2 = O(k^{-2})$$

and $\mathbb{V}[\widehat{u}_k^2] = 2\sigma_k^4 = O(k^{-4})$.

As well as square amplitude of a single Fourier mode, other two QoIs are:

- energy $P = \|u\|_2^2 = \sum_{k=1}^{\infty} \widehat{u}_k^2$
- squared functional $P = \langle \varphi, u \rangle^2 = \sum_{k=1}^{\infty} \widehat{\varphi}_k^2 \widehat{u}_k^2$

Stochastic heat equation

Insight comes from truncating expansion to give:

$$u_K(x,t) = \sum_{k=1}^{K-1} \widehat{u}_k(t)\, e_k(x),$$

where K is roughly equivalent to $1/\Delta x$ in a numerical approximation.

For energy,

$$P - P_K = \sum_{k=K}^{\infty} \widehat{u}_k^2,$$

so $\mathbb{E}[P{-}P_K] = O(K^{-1})$ and $\mathbb{V}[P{-}P_K] = O(K^{-3})$.

If $\widehat{\varphi}_k = O(k^{-p})$ then for squared functional $\mathbb{E}[P{-}P_K]$ and $\mathbb{V}[P{-}P_K]$ are both $O(K^{-2p-1})$.

Spectral noise representation

Finite difference approximation to PDE, and spectral (K-L) representation of noise gives semi-discretisation

$$\mathrm{d}U_j - \frac{1}{\Delta x^2}(U_{j+1} - 2U_j + U_{j-1})\,\mathrm{d}t = \sum_{k=1}^{J-1} \mathrm{d}\widehat{W}_k\, e_k(x_j).$$

The semi-discrete solution is then

$$U_j(t) = \sum_{k=1}^{J-1} \widehat{U}_k(t)\, e_k(x_j),$$

in which $\widehat{U}_k$ satisfy the modified O-U SDE

$$\mathrm{d}\widehat{U}_k = -\widetilde{\lambda}_k\, \widehat{U}_k\,\mathrm{d}t + \mathrm{d}\widehat{W}_k,$$

with $\widetilde{\lambda}_k = (4/\Delta x^2)\sin^2(k\pi\Delta x/2) = \lambda_k + O(k^4\Delta x^2)$.

Spectral noise representation

The fully-discrete equations use Euler-Maruyama time discretisation

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) + \sum_{k=1}^{J-1} \Delta \widehat{W}_k^n \, e_k(x_j).$$

For MLMC use $\Delta x_\ell \propto 2^{-\ell}$, $\Delta t_\ell \propto 4^{-\ell}$ with fixed $\Delta t/\Delta x^2 < 1/2$ for stability.

If the fine grid uses

$$\sum_{k=1}^{J-1} \Delta \widehat{W}_k^n \, e_k(x_j),$$

then the coarse grid uses

$$\sum_{m=n}^{n+3} \sum_{k=1}^{J/2-1} \Delta \widehat{W}_k^m \, e_k(x_j).$$

Key lemma for E-M time discretisation

Lemma

For the Ornstein-Uhlenbeck SDE $\quad \mathrm{d}u_t = -\lambda\, u_t\, \mathrm{d}t + \mathrm{d}W_t,$

with Euler-Maruyama approximation $\quad U_{n+1} = (1-\lambda\Delta t)\, U_n + \Delta W_n,$

there exists a constant C such that for any fixed t and integer $n = t/\Delta t$ with $\lambda\,\Delta t \leq 1$,

$$\begin{aligned} \left|\mathbb{E}[u^2(t) - U_n^2]\,\right| &< C\,\Delta t, \\ \mathbb{V}[u(t) - U_n] &< C\,\lambda\,\Delta t^2, \\ \mathbb{V}[u^2(t) - U_n^2] &< C\,\Delta t^2. \end{aligned}$$

Proof.

True for $\lambda{=}1$, then follows for all λ through re-scaling. □

Spectral noise representation

Lemma

For fixed t and $k < 1/\Delta x$:

$$\begin{aligned} \left| \mathbb{E}[\widehat{U}_k(t)^2 - \hat{u}_k^2(t)] \right| &\lesssim \Delta x^2, \\ \mathbb{V}[\widehat{U}_k(t) - \hat{u}_k(t)] &\lesssim k^2 \Delta x^4, \\ \mathbb{V}[\widehat{U}_k(t)^2 - \hat{u}_k^2(t)] &\lesssim \Delta x^4. \end{aligned}$$

Corollary

For fixed t, $\Delta t/\Delta x^2 \leq 1/4$, integer $n = t/\Delta t$, and $k < 1/\Delta x$,

$$\begin{aligned} \left| \mathbb{E}[\,(\widehat{U}_k^n)^2 - \widehat{u}_k^2(t)\,] \right| &\lesssim \Delta x^2, \\ \mathbb{V}[\,\widehat{U}_k^n - \widehat{u}_k(t)\,] &\lesssim k^2 \Delta x^4, \\ \mathbb{V}[\,(\widehat{U}_k^n)^2 - \widehat{u}_k^2(t)\,] &\lesssim \Delta x^4. \end{aligned}$$

Spectral noise representation

Consequences for QoIs:

- single mode: $\alpha = 2,\ \beta = 4$
- energy: $\alpha = 1,\ \beta = 3$ (Note: $\beta > 2\alpha$ highly unusual in MLMC)
- functional: $\alpha = 2,\ \beta = 4$ if p=2

Note that γ=3 since $\Delta x_\ell \propto 2^{-\ell}$, $\Delta t_\ell \propto 4^{-\ell}$.

Finite element noise

Using the usual "hat" piecewise linear basis functions ϕ_j, and mass-lumping, the Galerkin semi-discrete approximation is

$$\mathrm{d}U_j - \frac{1}{\Delta x^2}(U_{j+1} - 2U_j + U_{j-1})\,\mathrm{d}t = \frac{1}{\Delta x}\langle \mathrm{d}W, \phi_j\rangle, \quad j = 1, \ldots, J{-}1.$$

Note that $\mathbb{E}\left[\langle \mathrm{d}W, \phi_{j_1}\rangle\langle \mathrm{d}W, \phi_{j_2}\rangle\right] = 0$ when $|j_1{-}j_2| > 1$, and

$$\mathbb{E}\left[\langle \mathrm{d}W, \phi_j\rangle^2\right] = \tfrac{2}{3}\Delta x\,\mathrm{d}t, \quad \mathbb{E}\left[\langle \mathrm{d}W, \phi_j\rangle\langle \mathrm{d}W, \phi_{j\pm1}\rangle\right] = \tfrac{1}{6}\Delta x\,\mathrm{d}t,$$

The Euler-Maruyama discretisation gives

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) + \Delta W_j^n,$$

where ΔW_j^n can be simulated as

$$\Delta W_j^n = \sqrt{\Delta t/\Delta x}\,\left(Z_j^n/\sqrt{3} + Z_{j-1/2}^n/\sqrt{6} + Z_{j+1/2}^n/\sqrt{6}\right)$$

using iid standard Normals Z_j^n, $Z_{j\pm1/2}^n$.

Finite element noise

The coarse grid basis functions can be expressed in terms of fine grid ones:

$$\phi_{\ell-1,j}(x) = \tfrac{1}{2}\phi_{\ell,j-1}(x) + \phi_{\ell,j}(x) + \tfrac{1}{2}\phi_{\ell,j+1}(x).$$

so there is a natural MLMC coupling with

$$\Delta W^n_{\ell-1,j} = \sum_{m=n}^{n+3} \left(\tfrac{1}{4}\,\Delta W^m_{\ell,j-1} + \tfrac{1}{2}\,\Delta W^m_{\ell,j} + \tfrac{1}{4}\,\Delta W^m_{\ell,j+1} \right).$$

One of the takeaways from this talk is that this is easy to implement, very natural, and comparable to the spectral treatment in accuracy.

Finite element noise

The Fourier modes from the semi-discrete equations are given by

$$\mathrm{d}\widehat{U}_k = -\widetilde{\lambda}_k \, \widehat{U}_k \, \mathrm{d}t + \mathrm{d}\widetilde{W}_k$$

where

$$\begin{aligned}
\mathrm{d}\widetilde{W}_k &= \left\langle \sum_{j=1}^{J-1} e_k(x_j)\, \phi_j, \mathrm{d}W \right\rangle \\
&= \frac{4\sin^2(k\pi\Delta x/2)}{k^2\pi^2\Delta x^2} \mathrm{d}\widehat{W}_k \\
&+ \sum_{l=1}^{\infty} \frac{4\sin^2(k\pi\Delta x/2)}{(2lJ{+}k)^2\pi^2\Delta x^2} \mathrm{d}\widehat{W}_{2lJ+k} - \sum_{l=1}^{\infty} \frac{4\sin^2(k\pi\Delta x/2)}{(2lJ{-}k)^2\pi^2\Delta x^2} \mathrm{d}\widehat{W}_{2lJ-k}
\end{aligned}$$

Note: aliasing coefficient proportional to $(2lJ{\pm}k)^{-2} = O(\Delta x^2)$

Finite element noise

By bounding the difference from the spectral solution, can establish very similar lemmas concerning the accuracy of the semi-discrete and fully-discrete solutions.

Hence, obtain the same α, β, γ for the different output QoIs.

Finite volume noise

For the finite volume treatment

$$\phi_j(x) = \mathbf{1}_{[x_j-\Delta x/2,\, x_j+\Delta x/2]}(x)$$

so

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n) + \Delta W_j^n,$$

where $\Delta W_j^n = \sqrt{\Delta t/\Delta x}\ Z_j^n$, the Z_j^n again iid standard Normal.

However, this doesn't give a natural MLMC coupling.

Finite volume noise

Instead split the interval into two halves,

$$\phi_j(x) = \mathbf{1}_{[x_j-\Delta x/2,\, x_j]}(x) + \mathbf{1}_{[x_j,\, x_j+\Delta x/2]}(x),$$

$$\Delta W_j^n = \sqrt{\frac{\Delta t}{2\Delta x}}\,(Z^n_{j-1/4} + Z^n_{j+1/4})$$

The MLMC coupling is then

$$\Delta W^n_{\ell-1,j} = \sqrt{\frac{\Delta t}{8\Delta x}}\, \sum_{m=n}^{n+3} \left(Z^m_{j-3/4} + Z^m_{j-1/4} + Z^m_{j+1/4} + Z^m_{j+3/4} \right),$$

corresponding to integrating over the coarse grid interval $[x_{j-1}, x_{j+1}]$.

Finite volume noise

The Fourier modes from the semi-discrete equations are given by

$$\mathrm{d}\widehat{U}_k = -\widetilde{\lambda}_k \, \widehat{U}_k \, \mathrm{d}t + \mathrm{d}\widetilde{W}_k$$

where now

$$\begin{aligned} \mathrm{d}\widetilde{W}_k &= \left\langle \sum_{j=1}^{J-1} e_k(x_j) \, \phi_j, \mathrm{d}W \right\rangle \\ &= \frac{2\sin(k\pi\Delta x/2)}{k\pi\Delta x} \, \mathrm{d}\widehat{W}_k \\ &+ \sum_{l=1}^{\infty} \frac{2\sin(k\pi\Delta x/2)}{(2lJ{+}k)\pi\Delta x} \, \mathrm{d}\widehat{W}_{2lJ+k} - \sum_{l=1}^{\infty} \frac{2\sin(k\pi\Delta x/2)}{(2lJ{-}k)\pi\Delta x} \, \mathrm{d}\widehat{W}_{2lJ-k}. \end{aligned}$$

Note: aliasing coefficient now proportional to $(2lJ{\pm}k)^{-1} = O(\Delta x)$ which leads to a larger difference from the spectral solution.

Finite volume noise

Corollary

For fixed t, $\Delta t/\Delta x^2 \leq 1/4$, integer $n = t/\Delta t$, and $k < 1/\Delta x$,

$$\begin{aligned}
\left| \mathbb{E}[\,(\widehat{U}_k^n)^2 - \widehat{u}_k^2(t)\,] \right| &\lesssim \Delta x^2, \\
\mathbb{V}[\,\widehat{U}_k^n - \widehat{u}_k(t)\,] &\lesssim \Delta x^2, \\
\mathbb{V}[\,(\widehat{U}_k^n)^2 - \widehat{u}_k^2(t)\,] &\lesssim k^{-2}\Delta x^2.
\end{aligned}$$

Consequences for QoIs:

- single mode: $\alpha = 2,\ \beta = 2$
- energy: $\alpha = 1,\ \beta = 2$
- functional: $\alpha = 2,\ \beta = 2$ if p=2

Richardson extrapolation

If

$$P_\ell = P + a\, 2^{-\ell} + O(2^{-2\ell}),$$

then

$$P_\ell^{\rm ex} = 2\, P_\ell - P_{\ell-1} = P + O(2^{-2\ell}).$$

Doubling weak order α eliminates anomalous $\beta > 2\alpha$ situation.

With extrapolation, MLMC estimator becomes

$$\begin{aligned} 2\, P_\ell - P_{\ell-1} - (2\, P_{\ell-1} - P_{\ell-2}) &= 2\,(P_\ell - P_{\ell-1}) - (P_{\ell-1} - P_{\ell-2}) \\ &= 2\, P_\ell - 3\, P_{\ell-1} + P_{\ell-2} \end{aligned}$$

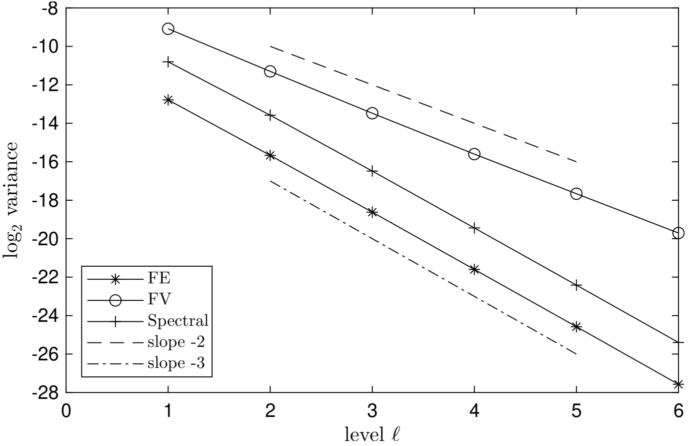
Extrapolation increases the variance, but it greatly reduces the finest level L required for weak convergence, so overall gives big savings.

Numerical results

- $T = 1/8$
- $\Delta x_\ell = 2^{-\ell-2}$, $\Delta t_\ell = 2^{-2\ell-6} \implies \Delta t_\ell / \Delta x_\ell^2 = 1/4$.
- 10^5 samples for spectral and finite element (FE),
  4×10^5 for finite volume (FV) due to poorer variance
- $\log_2 \mathbb{V}[P_\ell - P_{\ell-1}]$ and $\log_2 |\mathbb{E}[P_\ell - P_{\ell-1}]|$ plotted versus level ℓ
- functional weighting is $\varphi(x) = 2\sqrt{3}\,\min(x, 1-x)$ so $p{=}2$

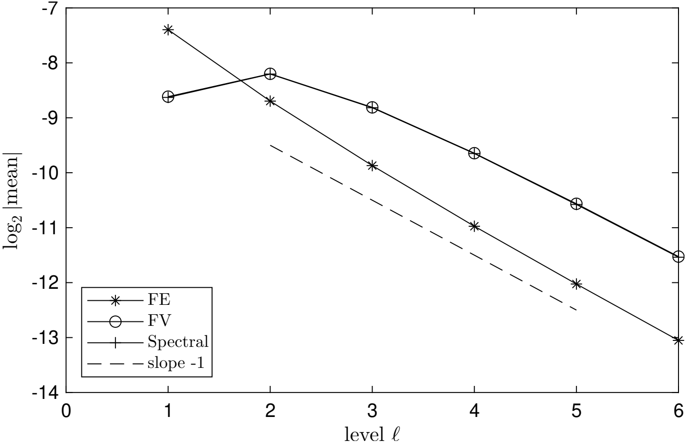
Numerical results

$\mathbb{V}[P_\ell - P_{\ell-1}]$ for energy QOI:



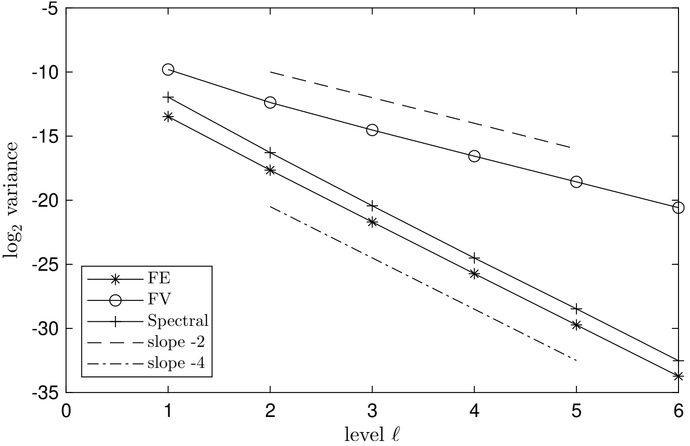
Numerical results

$\mathbb{E}[P_\ell - P_{\ell-1}]$ for energy QOI:



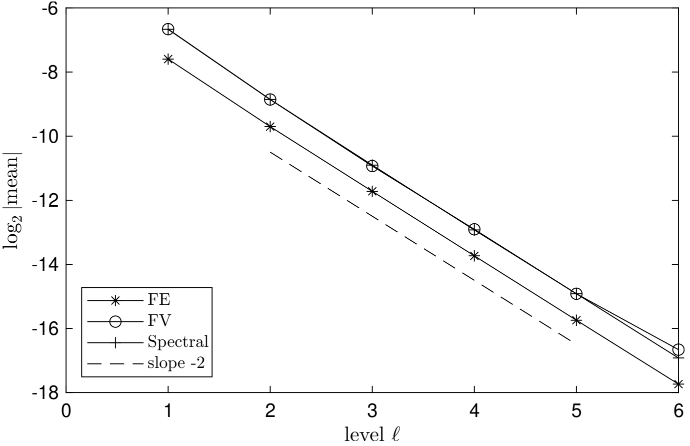
Numerical results

$\mathbb{V}[P_\ell - P_{\ell-1}]$ for functional QOI:



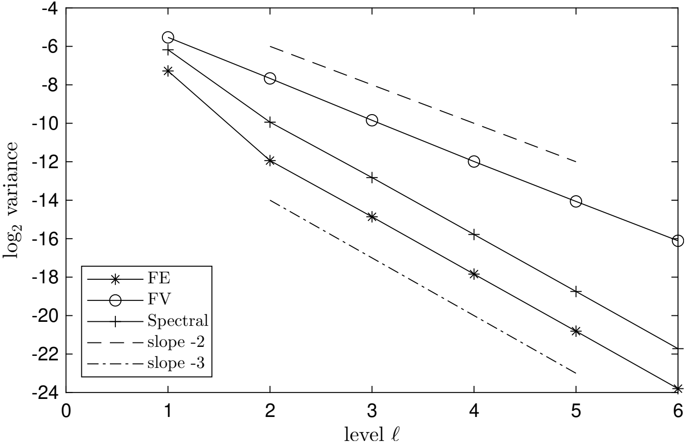
Numerical results

$\mathbb{E}[P_\ell - P_{\ell-1}]$ for functional QOI:



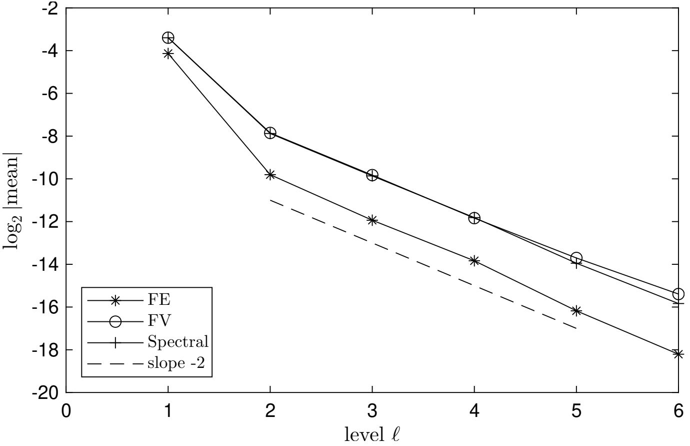
Numerical results

$\mathbb{V}[P_\ell - P_{\ell-1}]$ for energy QOI with extrapolation:



Numerical results

$\mathbb{E}[P_\ell - P_{\ell-1}]$ for energy QOI with extrapolation:



Conclusions

- Fourier analysis provides complete analysis of MLMC variance for stochastic heat equation
- 3 different white noise treatments, 3 different output QoI's
- finite volume treatment is clearly the worst due to aliasing errors
- finite element treatment as good as spectral treatment – both need Richardson extrapolation to get full benefits

- now ready to move on to more interesting SPDEs

References

A. Barth, A. Lang, and C. Schwab.
Multilevel Monte Carlo method for parabolic stochastic partial differential equations.
BIT Numerical Mathematics, 53(1):3–27, 2013.

F. Cornalba and J. Fischer.
Multilevel Monte Carlo methods for the Dean-Kawasaki equation from fluctuating hydrodynamics.
SIAM Journal on Numerical Analysis, 63(1):262–287, 2025.

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

G.J. Lord, C.E. Powell, and T. Shardlow.
An Introduction to Computational Stochastic PDEs.
Cambridge Texts in Applied Mathematics. CUP, 2014.