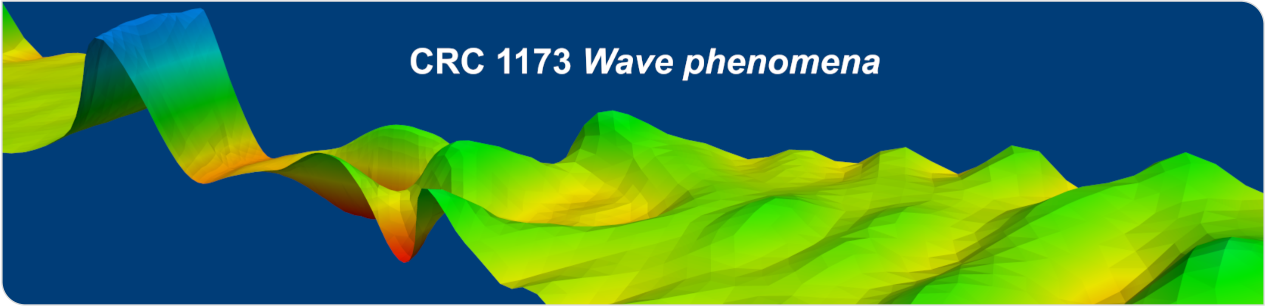
Karlsruhe Institute of Technology

A Fully Parallelized and Budgeted Multi-level Monte Carlo Method

Applications to Acoustic Wave Equations - Parallelization Limits - Automated High-Performance Computing

Niklas Baumgarten, Christian Wieners, Sebastian Krumscheid | 23.10.2023



Model Problem and Investigation Goals

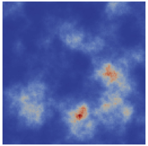
Acoustic wave: Search $(\mathbf{v}, p)\colon \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

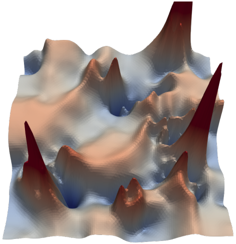
$$\text{such that} \quad \begin{cases} \rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) &= \mathbf{f} & \mathcal{D} \times (0, T] \\ \partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) &= g & \mathcal{D} \times (0, T] \\ \mathbf{v} \cdot \mathbf{n} &= 0 & \Gamma \times (0, T] \\ \mathbf{v}(0) &= \mathbf{v}_0 & \mathcal{D} \\ p(0) &= p_0 & \mathcal{D} \end{cases}$$

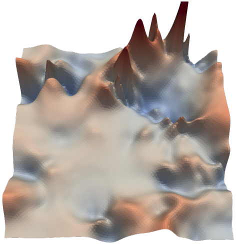
Model Problem and Investigation Goals

Acoustic wave: Search $(\mathbf{v}, p)\colon \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

such that
$$\begin{cases} \rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) = \mathbf{f} & \mathcal{D} \times (0, T] \\ \partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) = g & \mathcal{D} \times (0, T] \\ \mathbf{v} \cdot \mathbf{n} = 0 & \Gamma \times (0, T] \\ \mathbf{v}(0) = \mathbf{v}_0 & \mathcal{D} \\ p(0) = p_0 & \mathcal{D} \end{cases}$$



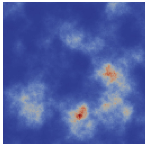


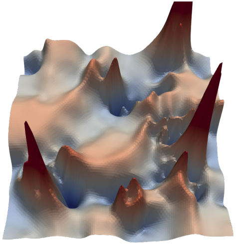


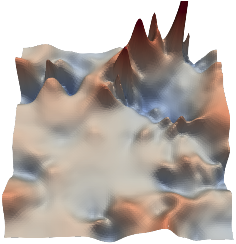
Model Problem and Investigation Goals

Acoustic wave: Search $(\mathbf{v}, p)\colon \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

such that
$$
\begin{cases}
\rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) = \mathbf{f} & \mathcal{D} \times (0, T] \\
\partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) = g & \mathcal{D} \times (0, T] \\
\mathbf{v} \cdot \mathbf{n} = 0 & \Gamma \times (0, T] \\
\mathbf{v}(0) = \mathbf{v}_0 & \mathcal{D} \\
p(0) = p_0 & \mathcal{D}
\end{cases}
$$





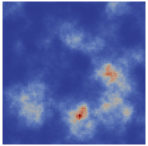


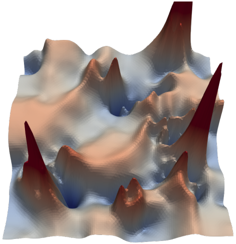
Determine:
$$
\mathbb{E}[\mathrm{Q}] := \int_\Omega \mathrm{Q}(\omega) \mathrm{d}\mathbb{P} \approx M^{-1} \sum_{m=1}^{M} \mathrm{Q}_\ell(\mathbf{y}^{(m)}) =: \widehat{\mathrm{Q}}^{\mathrm{MC}}_{\ell,M}
$$

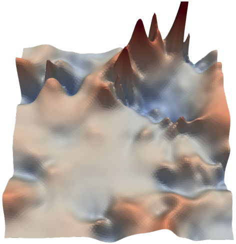
Model Problem and Investigation Goals

Acoustic wave: Search $(\mathbf{v}, p) \colon \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

such that
$$\begin{cases} \rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) &= \mathbf{f} & \mathcal{D} \times (0, T] \\ \partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) &= g & \mathcal{D} \times (0, T] \\ \mathbf{v} \cdot \mathbf{n} &= 0 & \Gamma \times (0, T] \\ \mathbf{v}(0) &= \mathbf{v}_0 & \mathcal{D} \\ p(0) &= p_0 & \mathcal{D} \end{cases}$$







Determine:
$$\mathbb{E}[\mathrm{Q}] := \int_\Omega \mathrm{Q}(\omega) \mathrm{d}\mathbb{P} \approx M^{-1} \sum_{m=1}^{M} \mathrm{Q}_\ell(\mathbf{y}^{(m)}) =: \widehat{\mathrm{Q}}_{\ell,M}^{\mathrm{MC}}$$

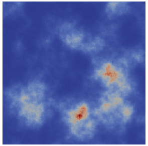
Goal: Find combination of methods to minimize total error

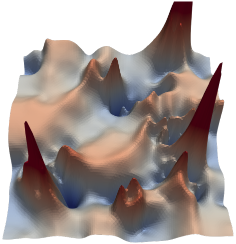
$$\mathrm{err}_{\text{total}} = \mathrm{err}_{\text{input}} + \mathrm{err}_{\text{disc}} + \mathrm{err}_{\text{model}} + \mathrm{err}_{\text{solve}} + \mathrm{err}_{\text{float}} + \mathrm{err}_{\text{bug}} + \dots$$

Model Problem and Investigation Goals

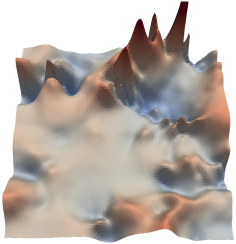
Acoustic wave: Search $(\mathbf{v}, p) : \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

such that
$$
\begin{cases}
\rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) = \mathbf{f} & \mathcal{D} \times (0, T] \\
\partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) = g & \mathcal{D} \times (0, T] \\
\mathbf{v} \cdot \mathbf{n} = 0 & \Gamma \times (0, T] \\
\mathbf{v}(0) = \mathbf{v}_0 & \mathcal{D} \\
p(0) = p_0 & \mathcal{D}
\end{cases}
$$





Determine: $$\mathbb{E}[\mathrm{Q}] := \int_\Omega \mathrm{Q}(\omega) \mathrm{d}\mathbb{P} \approx M^{-1} \sum_{m=1}^{M} \mathrm{Q}_\ell(\mathbf{y}^{(m)}) =: \widehat{\mathrm{Q}}_{\ell,M}^{\mathrm{MC}}$$



Goal: Find combination of methods to minimize total error

$$\mathrm{err}_{\text{total}} = \mathrm{err}_{\text{input}} + \mathrm{err}_{\text{disc}} + \mathrm{err}_{\text{model}} + \mathrm{err}_{\text{solve}} + \mathrm{err}_{\text{float}} + \mathrm{err}_{\text{bug}} + \dots$$

Constraint: Finite computational capacities (CPUs, time, memory)

$\Rightarrow$ Introduce budget for error minimization and utilize effective parallelization

Multi-level Monte Carlo - Introduction

Assumptions: Let $\alpha, \beta, \gamma > 0$ and

$$\left|\mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}]\right| \lesssim h_\ell^{\alpha}$$

$$\mathbb{V}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] \lesssim h_\ell^{\beta}$$

$$\mathrm{C}\left(\mathrm{Q}_\ell(\mathbf{y}^{(m)})\right) \lesssim h_\ell^{-\gamma}$$

Multi-level Monte Carlo - Introduction

Assumptions: Let $\alpha, \beta, \gamma > 0$ and

$$
\begin{aligned}
|\mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}]| &\lesssim h_\ell^{\alpha} \\
\mathbb{V}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] &\lesssim h_\ell^{\beta} \\
\mathrm{C}\left(\mathrm{Q}_\ell(\mathbf{y}^{(m)})\right) &\lesssim h_\ell^{-\gamma}
\end{aligned}
$$

Idea: Telescoping sum with $\mathrm{Y}_0 := \mathrm{Q}_0, \mathrm{Y}_\ell := \mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}$

$$\mathbb{E}[\mathrm{Q}_L] = \mathbb{E}[\mathrm{Q}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] = \sum_{\ell=0}^{L} \mathbb{E}[\mathrm{Y}_\ell]$$

Multi-level Monte Carlo - Introduction

Assumptions: Let $\alpha, \beta, \gamma > 0$ and

$$|\mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}]| \lesssim h_\ell^\alpha$$

$$\mathbb{V}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] \lesssim h_\ell^\beta$$

$$\mathrm{C}\left(\mathrm{Q}_\ell(\mathbf{y}^{(m)})\right) \lesssim h_\ell^{-\gamma}$$

Idea: Telescoping sum with $\mathrm{Y}_0 := \mathrm{Q}_0, \mathrm{Y}_\ell := \mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}$

$$\mathbb{E}[\mathrm{Q}_L] = \mathbb{E}[\mathrm{Q}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] = \sum_{\ell=0}^{L} \mathbb{E}[\mathrm{Y}_\ell]$$

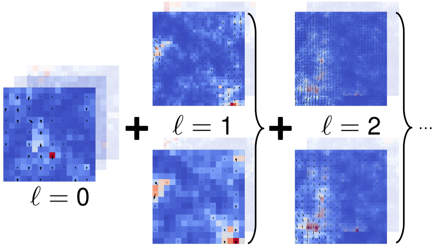
Multi-level Estimator:

$$\widehat{\mathrm{Q}}^{\mathrm{MLMC}}_{\{M_\ell\}_{\ell=0}^L} = \sum_{\ell=0}^{L} \widehat{\mathrm{Y}}^{\mathrm{MC}}_{\ell,M_\ell} = \sum_{\ell=0}^{L} M_\ell^{-1} \sum_{m=1}^{M_\ell} \mathrm{Y}_\ell(\mathbf{y}^{(m)})$$

Multi-level Monte Carlo - Introduction

Assumptions: Let $\alpha, \beta, \gamma > 0$ and

$$\left|\mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}]\right| \lesssim h_\ell^\alpha$$
$$\mathbb{V}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] \lesssim h_\ell^\beta$$
$$\mathrm{C}\left(\mathrm{Q}_\ell(\mathbf{y}^{(m)})\right) \lesssim h_\ell^{-\gamma}$$



Idea: Telescoping sum with $\mathrm{Y}_0 := \mathrm{Q}_0$, $\mathrm{Y}_\ell := \mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}$

$$\mathbb{E}[\mathrm{Q}_L] = \mathbb{E}[\mathrm{Q}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] = \sum_{\ell=0}^{L} \mathbb{E}[\mathrm{Y}_\ell]$$

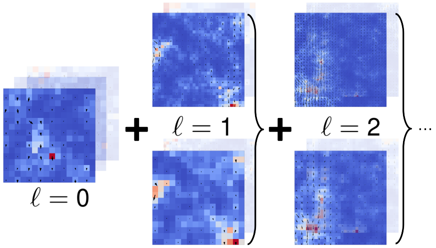
Multi-level Estimator:

$$\widehat{\mathrm{Q}}^{\mathrm{MLMC}}_{\{M_\ell\}_{\ell=0}^L} = \sum_{\ell=0}^{L} \widehat{\mathrm{Y}}^{\mathrm{MC}}_{\ell, M_\ell} = \sum_{\ell=0}^{L} M_\ell^{-1} \sum_{m=1}^{M_\ell} \mathrm{Y}_\ell(\mathbf{y}^{(m)})$$

Multi-level Monte Carlo - Introduction

Assumptions: Let $\alpha, \beta, \gamma > 0$ and

$$|\mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}]| \lesssim h_\ell^\alpha$$
$$\mathbb{V}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] \lesssim h_\ell^\beta$$
$$\mathrm{C}\left(\mathrm{Q}_\ell(\mathbf{y}^{(m)})\right) \lesssim h_\ell^{-\gamma}$$



Idea: Telescoping sum with $\mathrm{Y}_0 := \mathrm{Q}_0$, $\mathrm{Y}_\ell := \mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}$

$$\mathbb{E}[\mathrm{Q}_L] = \mathbb{E}[\mathrm{Q}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\mathrm{Q}_\ell - \mathrm{Q}_{\ell-1}] = \sum_{\ell=0}^{L} \mathbb{E}[\mathrm{Y}_\ell]$$

Multi-level Estimator:

$$\widehat{\mathrm{Q}}^{\mathrm{MLMC}}_{\{M_\ell\}_{\ell=0}^L} = \sum_{\ell=0}^{L} \widehat{\mathrm{Y}}^{\mathrm{MC}}_{\ell, M_\ell} = \sum_{\ell=0}^{L} M_\ell^{-1} \sum_{m=1}^{M_\ell} \mathrm{Y}_\ell(\mathbf{y}^{(m)})$$

Epsilon-Cost Theorem: $\exists \{M_\ell\}_{\ell=0}^L$, such that

$$\mathrm{err}_{\mathrm{MSE}} = \sum_{\ell=0}^{L} M_\ell^{-1} \mathbb{V}[\mathrm{Y}_\ell] + (\mathbb{E}[\mathrm{Q}_L - \mathrm{Q}])^2 < \epsilon^2$$

and $\quad \mathrm{T}_\epsilon := \mathrm{C}_\epsilon \lesssim \begin{cases} \epsilon^{-2} & \beta > \gamma \\ \epsilon^{-2-(\gamma-\beta)/\alpha} & \beta < \gamma \end{cases}$

M. Giles. *Multilevel Monte Carlo path simulation.* (2008)

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T}_\mathrm{B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#$CPU $\cdot$ hours

Budgeted Multi-level Monte Carlo - Introduction

Karlsruhe Institute of Technology

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T_B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#$CPU $\cdot$ hours

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and $\mathrm{T_B}$ have to be reserved for HPC
- Empirical study of algorithms

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T_B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#\text{CPU} \cdot \text{hours}$

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and $\mathrm{T_B}$ have to be reserved for HPC
- Empirical study of algorithms

Knapsack Problem:

$$\min_{(L, \{M_\ell\}_{\ell=0}^L)} \mathrm{err}_{\text{MSE}} = \sum_{\ell=0}^{L} M_\ell^{-1} \mathbb{V}[\mathrm{Y}_\ell] + (\mathbb{E}[\mathrm{Q}_L - \mathrm{Q}])^2$$

such that $$\sum_{\ell=0}^{L} \sum_{m=1}^{M_\ell} \mathrm{C}_\ell(\mathbf{y}^{(m)}) \leq \mathrm{B}$$

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T}_\mathrm{B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#\mathrm{CPU} \cdot \mathrm{hours}$

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and T_B have to be reserved for HPC
- Empirical study of algorithms

Knapsack Problem:

$$\min_{(L,\{M_\ell\}_{\ell=0}^{L})} \mathrm{err}_{\mathrm{MSE}} = \sum_{\ell=0}^{L} M_\ell^{-1} \mathbb{V}[\mathrm{Y}_\ell] + (\mathbb{E}[\mathrm{Q}_L - \mathrm{Q}])^2$$

such that $$\sum_{\ell=0}^{L} \sum_{m=1}^{M_\ell} \mathrm{C}_\ell(\mathbf{y}^{(m)}) \leq \mathrm{B}$$

Conjecture: For a feasible and parallel execution, it is

$$\epsilon \lesssim \underbrace{(1-\lambda_\mathrm{p}) \cdot \mathrm{T}_\mathrm{B}^{-\delta}}_{=:\epsilon_\mathrm{s}} + \underbrace{\lambda_\mathrm{p}(|\mathcal{P}| \cdot \mathrm{T}_\mathrm{B})^{-\delta}}_{=:\epsilon_\mathrm{p}}$$

with $\delta \in \left\{ \frac{1}{2}, \frac{\alpha}{(2\alpha + (\gamma - \beta))} \right\}$ and $\lambda_\mathrm{p} \in [0, 1]$.

Baumgarten et al. *A Fully Parallelized and Budgeted MLMC Method.* arXiv Preprint (2023)

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T_B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#\mathrm{CPU} \cdot \mathrm{hours}$

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and $\mathrm{T_B}$ have to be reserved for HPC
- Empirical study of algorithms

Knapsack Problem:

$$\min_{(L,\{M_\ell\}_{\ell=0}^L)} \mathrm{err}_{\mathrm{MSE}} = \sum_{\ell=0}^{L} M_\ell^{-1} \mathbb{V}[\mathrm{Y}_\ell] + (\mathbb{E}[\mathrm{Q}_L - \mathrm{Q}])^2$$

such that $\sum_{\ell=0}^{L} \sum_{m=1}^{M_\ell} \mathrm{C}_\ell(\mathbf{y}^{(m)}) \leq \mathrm{B}$

Conjecture: For a feasible and parallel execution, it is

$$\epsilon \lesssim \underbrace{(1-\lambda_\mathrm{p}) \cdot \mathrm{T_B}^{-\delta}}_{=:\epsilon_\mathrm{s}} + \underbrace{\lambda_\mathrm{p}(|\mathcal{P}| \cdot \mathrm{T_B})^{-\delta}}_{=:\epsilon_\mathrm{p}}$$

with $\delta \in \left\{\frac{1}{2}, \frac{\alpha}{(2\alpha+(\gamma-\beta))}\right\}$ and $\lambda_\mathrm{p} \in [0, 1]$.

Baumgarten et al. *A Fully Parallelized and Budgeted MLMC Method.* arXiv Preprint (2023)

Dynamic Programming (DP):

- Decomposition in overlapping subproblems
- Solve subproblems with optimal strategy
- Reutilization of preexisting results

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T_B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#\mathrm{CPU} \cdot \mathrm{hours}$

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and $\mathrm{T_B}$ have to be reserved for HPC
- Empirical study of algorithms

Conjecture: For a feasible and parallel execution, it is

$$\epsilon \lesssim \underbrace{(1 - \lambda_\mathrm{p}) \cdot \mathrm{T_B}^{-\delta}}_{=:\epsilon_\mathrm{s}} + \underbrace{\lambda_\mathrm{p}(|\mathcal{P}| \cdot \mathrm{T_B})^{-\delta}}_{=:\epsilon_\mathrm{p}}$$

with $\delta \in \left\{ \frac{1}{2}, \frac{\alpha}{(2\alpha + (\gamma - \beta))} \right\}$ and $\lambda_\mathrm{p} \in [0, 1]$.

Baumgarten et al. *A Fully Parallelized and Budgeted MLMC Method.* arXiv Preprint (2023)

Knapsack Problem:

$$\min_{(L, \{M_\ell\}_{\ell=0}^{L})} \mathrm{err}_{\mathrm{MSE}} = \sum_{\ell=0}^{L} M_\ell^{-1} \mathbb{V}[\mathrm{Y}_\ell] + (\mathbb{E}[\mathrm{Q}_L - \mathrm{Q}])^2$$

such that $\displaystyle \sum_{\ell=0}^{L} \sum_{m=1}^{M_\ell} \mathrm{C}_\ell(\mathbf{y}^{(m)}) \leq \mathrm{B}$

Dynamic Programming (DP):

- Decomposition in overlapping subproblems
- Solve subproblems with optimal strategy
- Reutilization of preexisting results

$\Rightarrow$ Use DP for approximated knapsack problem

Budgeted Multi-level Monte Carlo - Introduction

Goal: Replace accuracy ϵ by budget $|\mathcal{P}| \cdot \mathrm{T_B} =: \mathrm{B} > 0$ measured in $[\mathrm{B}] = \#\mathrm{CPU} \cdot \mathrm{hours}$

Motivation:

- Often no a priori knowledge about α, β and γ
- $\mathcal{P}$ and $\mathrm{T_B}$ have to be reserved for HPC
- Empirical study of algorithms

Approximated Knapsack Problem:

$$\min_{(L, \{M_\ell\}_{\ell=0}^L)} \quad \sum_{\ell=0}^{L} M_\ell^{-1} s_{\mathrm{Y}_\ell}^2 + \left(\widehat{\mathrm{err}}_{\mathrm{disc}}(\widehat{\mathrm{Y}}_\ell, \widehat{\alpha}) \right)^2$$

$$\text{such that} \quad \sum_{\ell=0}^{L} M_\ell \widehat{\mathrm{C}}_\ell \leq \mathrm{B}$$

Conjecture: For a feasible and parallel execution, it is

$$\epsilon \lesssim \underbrace{(1 - \lambda_\mathrm{p}) \cdot \mathrm{T_B}^{-\delta}}_{=:\epsilon_\mathrm{s}} + \underbrace{\lambda_\mathrm{p}(|\mathcal{P}| \cdot \mathrm{T_B})^{-\delta}}_{=:\epsilon_\mathrm{p}}$$

with $\delta \in \left\{ \frac{1}{2}, \frac{\alpha}{(2\alpha + (\gamma - \beta))} \right\}$ and $\lambda_\mathrm{p} \in [0, 1]$.

Baumgarten et al. *A Fully Parallelized and Budgeted MLMC Method.* arXiv Preprint (2023)

Dynamic Programming (DP):

- Decomposition in overlapping subproblems
- Solve subproblems with optimal strategy
- Reutilization of preexisting results

$\Rightarrow$ Use DP for approximated knapsack problem

Budgeted Multi-level Monte Carlo - Algorithm

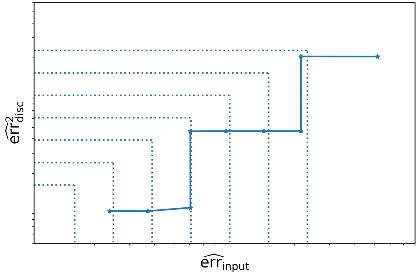
$$\texttt{data} = \Big\{ \mathrm{i} \mapsto \Big\{ \mathrm{err}_\mathrm{i}, \{M_{\mathrm{i},\ell}\}_{\ell=0}^{L_\mathrm{i}}, \{\widehat{\mathrm{Q}}_{\mathrm{i},\ell}\}_{\ell=0}^{L_\mathrm{i}}, \{\widehat{\mathrm{C}}_{\mathrm{i},\ell}\}_{\ell=0}^{L_\mathrm{i}}, \{\widehat{\mathrm{Y}}_{\mathrm{i},\ell}\}_{\ell=0}^{L_\mathrm{i}}, \dots \Big\} \Big\}$$

$$\texttt{function BMLMC}(\mathrm{B}_0, \{M_{0,\ell}^{\text{init}}\}_{\ell=0}^{L_0}):$$

$$\begin{cases} \texttt{for } \ell = L_0, \dots, 0: & \Delta\texttt{data}_{0,\ell} \leftarrow \texttt{MS-FEM}(M_{0,\ell}^{\text{init}}, \mathcal{P}) \\ \texttt{data}_0 \leftarrow \texttt{Welford}(\texttt{data}_{-1}, \Delta\texttt{data}_0) & \texttt{return BMLMC}(\mathrm{B}_0 - \sum_{\ell=0}^{L_0} \mathrm{C}_\ell, \eta \cdot \mathrm{err}_0) \end{cases}$$

$$\texttt{function BMLMC}(\mathrm{B}_\mathrm{i}, \epsilon_\mathrm{i}):$$

$$\begin{cases} \texttt{if } \mathrm{B}_\mathrm{i} \approx 0: & \texttt{return err}_{\mathrm{i}-1} \\ \texttt{if } \widehat{\mathrm{err}}_{\texttt{disc}}(\texttt{data}_{\mathrm{i}-1}) \geq \sqrt{1-\theta}\epsilon_\mathrm{i}: & L_\mathrm{i} \leftarrow L_\mathrm{i} + 1 \\ \texttt{if } \widehat{\mathrm{err}}_{\texttt{input}}(\texttt{data}_{\mathrm{i}-1}) \geq \theta\epsilon_\mathrm{i}^2: & \widehat{M}_{\mathrm{i},\ell}^{\text{opt}} \sim \left\lceil (\sqrt{\theta}\epsilon)^{-2} \sqrt{s_{\mathrm{Y}_\ell}^2 / \widehat{\mathrm{C}}_\ell} \right\rceil \\ \texttt{for } \ell = L_\mathrm{i}, \dots, 0: & \Delta M_{\mathrm{i},\ell} \leftarrow \max\left\{ \widehat{M}_{\mathrm{i},\ell}^{\text{opt}} - M_{\mathrm{i}-1,\ell}, 0 \right\} \\ \widehat{\mathrm{C}}_\mathrm{i} \leftarrow \sum_{\ell=0}^{L_\mathrm{i}} \Delta M_{\mathrm{i},\ell} \widehat{\mathrm{C}}_{\mathrm{i}-1,\ell} & \\ \texttt{if } \widehat{\mathrm{C}}_\mathrm{i} = 0: & \texttt{return BMLMC}(\mathrm{B}_\mathrm{i}, \eta \cdot \epsilon_\mathrm{i}) \\ \texttt{if } \widehat{\mathrm{C}}_\mathrm{i} > \mathrm{B}_\mathrm{i}: & \texttt{return BMLMC}(\mathrm{B}_\mathrm{i}, 0.5 \cdot (\epsilon_\mathrm{i} + \epsilon_{\mathrm{i}-1})) \\ \texttt{for } \ell = L_\mathrm{i}, \dots, 0: & \Delta\texttt{data}_{\mathrm{i},\ell} \leftarrow \texttt{MS-FEM}(\Delta M_{\mathrm{i},\ell}, \mathcal{P}) \\ \texttt{data}_\mathrm{i} \leftarrow \texttt{Welford}(\texttt{data}_{\mathrm{i}-1}, \Delta\texttt{data}_\mathrm{i}) & \texttt{return BMLMC}(\mathrm{B}_\mathrm{i} - \sum_{\ell=0}^{L_\mathrm{i}} \mathrm{C}_\ell, \epsilon_\mathrm{i}) \end{cases}$$



M. Giles. *Multilevel Monte Carlo methods.* (2015)

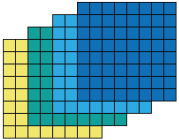
Collier et al. *A continuation multilevel Monte Carlo algorithm.* (2015)

Multi-sample Finite Element Method (MS-FEM)

Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$

Multi-sample Finite Element Method (MS-FEM)

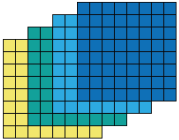
Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$



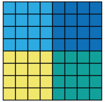
$1 < |\mathcal{P}| \leq M_\ell$

Multi-sample Finite Element Method (MS-FEM)

Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$



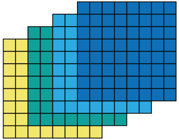
$1 < |\mathcal{P}| \leq M_\ell$



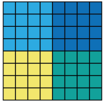
$1 < |\mathcal{P}|$ & $M_\ell = 1$

Multi-sample Finite Element Method (MS-FEM)

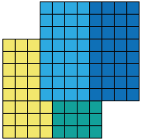
Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$



$1 < |\mathcal{P}| \leq M_\ell$



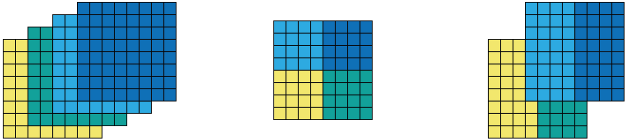
$1 < |\mathcal{P}|$ & $M_\ell = 1$



$1 < M_\ell < |\mathcal{P}|$

Multi-sample Finite Element Method (MS-FEM)

Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$



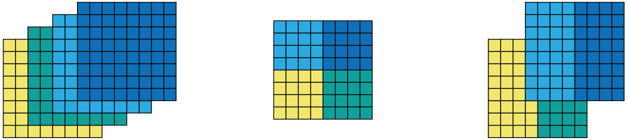
$1 < |\mathcal{P}| \leq M_\ell$ $\qquad$ $1 < |\mathcal{P}|$ & $M_\ell = 1$ $\qquad$ $1 < M_\ell < |\mathcal{P}|$

Minimize Communication: Search $k \in \mathbb{N}_0$, such that

$$2^k \leq \frac{|\mathcal{P}|}{M_\ell} < 2^{k+1} \;\Rightarrow\; \mathcal{P} = \bigsqcup_{m=1}^{M_\ell} \mathcal{P}_k^{(m)} \text{ with } \left|\mathcal{P}_k^{(m)}\right| = 2^k$$

Multi-sample Finite Element Method (MS-FEM)

Problem: Approximate M_ℓ-times a PDE on discretization level ℓ on a fixed set of CPUs $\mathcal{P}$



$1 < |\mathcal{P}| \leq M_\ell$ $\quad$ $1 < |\mathcal{P}|$ & $M_\ell = 1$ $\quad$ $1 < M_\ell < |\mathcal{P}|$

Define: Set of FE meshes

$$\mathcal{M}_\mathcal{P} := \left\{ \mathcal{M}^{(m)}_{\mathcal{P}_k} \right\}_{m=1}^{M_\ell}$$

Minimize Communication: Search $k \in \mathbb{N}_0$, such that

$$2^k \leq \frac{|\mathcal{P}|}{M_\ell} < 2^{k+1} \;\Rightarrow\; \mathcal{P} = \bigsqcup_{m=1}^{M_\ell} \mathcal{P}_k^{(m)} \text{ with } \left|\mathcal{P}_k^{(m)}\right| = 2^k$$

MS-FEM: Search for representation of

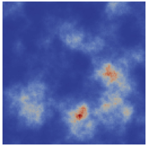
$$(\mathbf{u}_\ell)_{m=1}^{M_\ell} \in \prod_{m=1}^{M_\ell} V_\ell^{(m)}$$

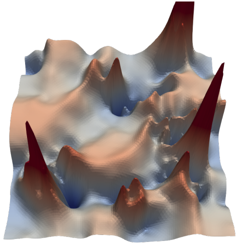
defined on $\mathcal{M}_\mathcal{P}$

Model Problem and Discretization

Acoustic wave: Search $(\mathbf{v}, p) \colon \Omega \times \mathcal{D} \times [0, T] \to \mathbb{R}^{D+1}$,

such that
$$\begin{cases} \rho(\omega)\partial_t \mathbf{v}(\omega) - \nabla p(\omega) = \mathbf{f} & \mathcal{D} \times (0, T] \\ \partial_t p(\omega) - \operatorname{div}(\mathbf{v}(\omega)) = g & \mathcal{D} \times (0, T] \\ \mathbf{v} \cdot \mathbf{n} = 0 & \Gamma \times (0, T] \\ \mathbf{v}(0) = \mathbf{v}_0 & \mathcal{D} \\ p(0) = p_0 & \mathcal{D} \end{cases}$$



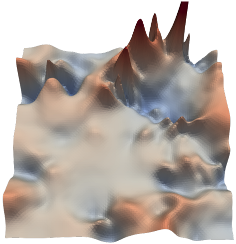


Discontinuous Galerkin (dG): Search for $(\mathbf{v}, p)^\top =: \mathbf{u}_\ell \in V_{\ell,\mathbf{p}}^{\mathrm{dG}}$

$$\mathrm{M}_\ell \partial_t \mathbf{u}_\ell + \mathrm{A}_\ell \mathbf{u}_\ell = \mathbf{b}_\ell \quad \text{and} \quad \mathbf{u}_\ell(0) = \mathbf{u}_{\ell,0}$$

Implicit midpoint-rule (IMPR): Solve for $t_n = n\tau_\ell$ with $\tau_\ell = T/N_\ell^\mathcal{T}$

$$\left(\mathrm{M}_\ell + \frac{\tau_\ell}{2}\mathrm{A}_\ell\right)\mathbf{u}_\ell(t_n) = \left(\mathrm{M}_\ell - \frac{\tau_\ell}{2}\mathrm{A}_\ell\right)\mathbf{u}_\ell(t_{n-1}) + \tau_\ell \mathbf{b}_\ell(t_{n-1/2})$$



Circulant Embedding: Sample from log-normally distributed material density ρ

Experimental Setup via Continuous Delivery (CD)

Goal: Find combination of methods for minimal error

$$\text{err}_{\text{total}} = \text{err}_{\text{input}} + \text{err}_{\text{disc}} + \text{err}_{\text{model}} + \text{err}_{\text{bug}} + \dots$$

under computational constraint

Experimental Setup via Continuous Delivery (CD)

Goal: Find combination of methods for minimal error

$$\mathrm{err}_{\text{total}} = \mathrm{err}_{\text{input}} + \mathrm{err}_{\text{disc}} + \mathrm{err}_{\text{model}} + \mathrm{err}_{\text{bug}} + \dots$$

under computational constraint

Default Budget: $\mathrm{B} = |\mathcal{P}| \cdot \mathrm{T}_{\mathrm{B}} = 1024 \cdot 6\text{h}$

Experimental Setup via Continuous Delivery (CD)

Goal: Find combination of methods for minimal error

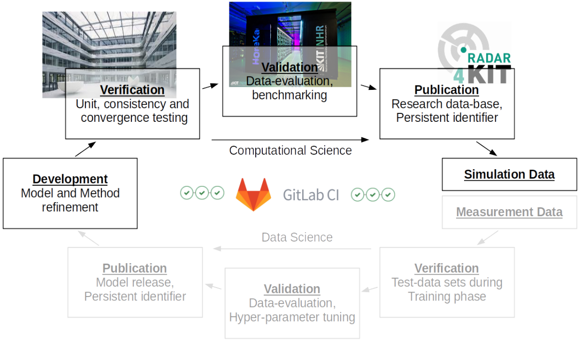
$$\text{err}_{\text{total}} = \text{err}_{\text{input}} + \text{err}_{\text{disc}} + \text{err}_{\text{model}} + \text{err}_{\text{bug}} + \dots$$

under computational constraint

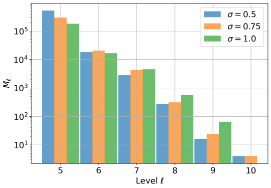
Default Budget: $\mathrm{B} = |\mathcal{P}| \cdot \mathrm{T}_{\mathrm{B}} = 1024 \cdot 6\text{h}$

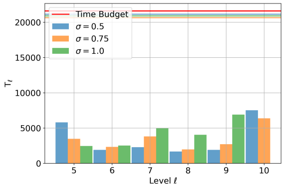
Automated High-Performance Computing:

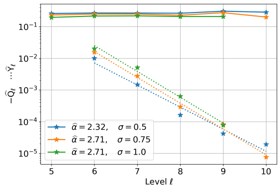
- Code verification
- Model and method comparison
- Data collection and postprocessing

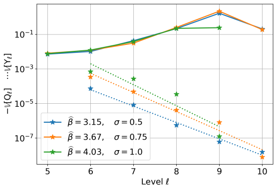


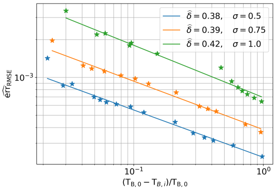
Numerical Experiments - Covariance Function

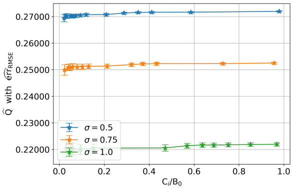












Covariance Function:

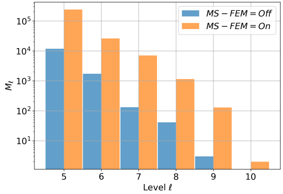
$$\mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \exp\left(- \left\| \left(\frac{\mathbf{x}_1 - \mathbf{x}_2}{\lambda} \right) \right\|_2^\nu \right)$$

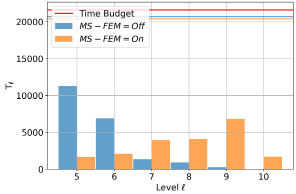
with $\lambda = 0.15$, $\nu = 1.8$ and $\sigma \in \{0.5, 0.75, 1.0\}$

Numerical Experiments - Parallelization

1st Experiment:

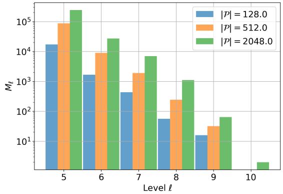
- Solver parallelization vs. MS-FEM with fixed $\mathrm{B} = 2048 \cdot 6\mathrm{h}$

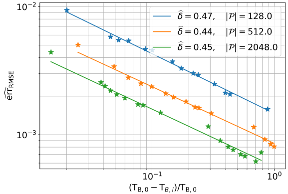




2nd Experiment:

- Weak scaling measurement with $\mathrm{T_B} = 6\mathrm{h}$ on $|\mathcal{P}| \in \{128, 512, 2048\}$

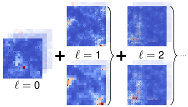


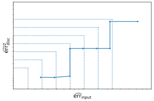


Conclusion and Outlook

Conclusion:

- Budgeted Multi-level Monte Carlo (BMLMC)

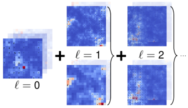


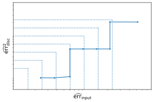


Conclusion and Outlook

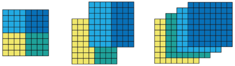
Conclusion:

- Budgeted Multi-level Monte Carlo (BMLMC)





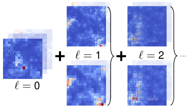
- Multi-Sample Finite Element Method (MS-FEM)

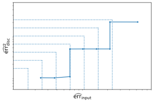


Conclusion and Outlook

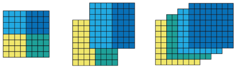
Conclusion:

- Budgeted Multi-level Monte Carlo (BMLMC)

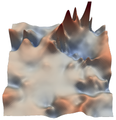


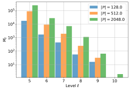


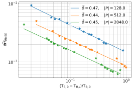
- Multi-Sample Finite Element Method (MS-FEM)



- Acoustic Wave Simulations in Random Media



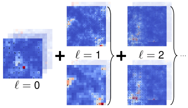


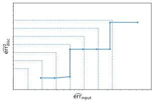


Conclusion and Outlook

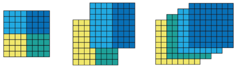
Conclusion:

- Budgeted Multi-level Monte Carlo (BMLMC)

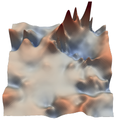


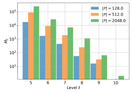


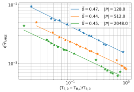
- Multi-Sample Finite Element Method (MS-FEM)



- Acoustic Wave Simulations in Random Media







Further Work:

- Other PDEs with various FE&UQ methods

N. Baumgarten. *A Fully Parallelized and Budgeted Multi-level Monte Carlo Framework for Partial Differential Equations: From Mathematical Theory to Automated Large-Scale Computations.* (2023)

Baumgarten, Wieners. *The parallel finite element system M++ with integrated multilevel preconditioning and multilevel Monte Carlo methods.* (2021)

- FEM-Software M++ & Main Source of Talk

Wieners, Corallo, Schneiderhan, Stengel, Lindner, Rheinbay, Baumgarten. *Mpp 3.3.0.* (2023)

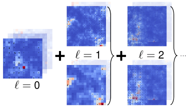
Baumgarten et al. *A Fully Parallelized and Budgeted Multi-level Monte Carlo Method and the Application to Acoustic Waves.* arXiv Preprint (2023)

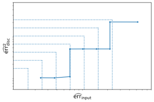
Karlsruhe Institute of Technology

Conclusion and Outlook

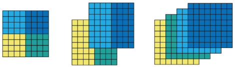
Conclusion:

- Budgeted Multi-level Monte Carlo (BMLMC)

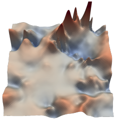


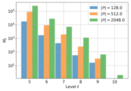


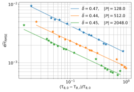
- Multi-Sample Finite Element Method (MS-FEM)



- Acoustic Wave Simulations in Random Media







Further Work:

- Other PDEs with various FE&UQ methods

  N. Baumgarten. *A Fully Parallelized and Budgeted Multi-level Monte Carlo Framework for Partial Differential Equations: From Mathematical Theory to Automated Large-Scale Computations.* (2023)

  Baumgarten, Wieners. *The parallel finite element system M++ with integrated multilevel preconditioning and multilevel Monte Carlo methods.* (2021)

- FEM-Software M++ & Main Source of Talk

  Wieners, Corallo, Schneiderhan, Stengel, Lindner, Rheinbay, Baumgarten. *Mpp 3.3.0.* (2023)

  Baumgarten et al. *A Fully Parallelized and Budgeted Multi-level Monte Carlo Method and the Application to Acoustic Waves.* arXiv Preprint (2023)

In Progress / Outlook / Interests:

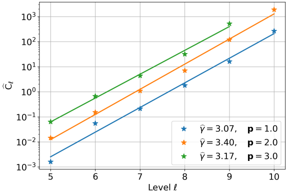
- Interfaces: Umbridge & Ginkgo
- (B)MLSC, (B)MLQMC, (B)MIMC
- Implementation of $\widehat{\mathbf{u}}_L$ in Mpp 3.3.1
- SGD/ADAM for Optimal Control
- Bayesian Inverse UQ via SMC

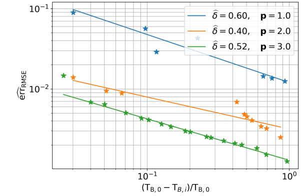
Numerical Experiments - Method Comparison

Karlsruhe Institute of Technology

1st Experiment:

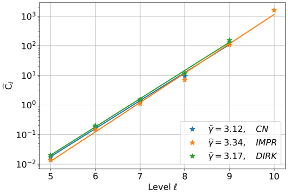
- Polynomial degree $\mathbf{p} \in \{1, 2, 3\}$ of $V_{\ell,\mathbf{p}}^{\mathrm{dG}}$

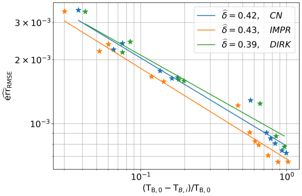




2nd Experiment:

- Diagonal implicit Runge-Kutta (DIRK)
- Implicit midpoint-rule (IMPR)
- Crank-Nicolson (CN)





Numerical Experiments - Verification of Conjecture

Weak Scaling Experiments:

- Fixed time budget $\mathrm{T_B} = 6\mathrm{h}$
- Variable amount of processing units

$$|\mathcal{P}| \in \left\{2^{-k}\,|\mathcal{P}_{\mathsf{max}}| \,:\, |\mathcal{P}_{\mathsf{max}}| = 8192,\; k = 0, \dots, 7\right\}$$

Numerical Experiments - Verification of Conjecture

Weak Scaling Experiments:

- Fixed time budget $\mathrm{T_B} = 6\mathrm{h}$
- Variable amount of processing units

$$|\mathcal{P}| \in \left\{2^{-k} \, |\mathcal{P}_{\mathsf{max}}| : \, |\mathcal{P}_{\mathsf{max}}| = 8192, \, k = 0, \dots, 7\right\}$$

- Reduction factor $\eta \in \{0.7, 0.8, 0.9\}$ for $\epsilon_{\mathtt{i}} := \eta \epsilon_{\mathtt{i}-1}$

Numerical Experiments - Verification of Conjecture

Weak Scaling Experiments:

- Fixed time budget $\mathrm{T_B} = 6\mathrm{h}$
- Variable amount of processing units

$$|\mathcal{P}| \in \left\{2^{-k}\,|\mathcal{P}_{\max}| \,:\, |\mathcal{P}_{\max}| = 8192,\ k = 0, \dots, 7\right\}$$

- Reduction factor $\eta \in \{0.7, 0.8, 0.9\}$ for $\epsilon_{\mathtt{i}} := \eta\epsilon_{\mathtt{i}-1}$
- Assume λ_{p} is fixed for a fixed η, then

$$\epsilon \lesssim (1 - \lambda_{\mathrm{p}}) \cdot \mathrm{T_B}^{-\delta} + \lambda_{\mathrm{p}}(|\mathcal{P}| \cdot \mathrm{T_B})^{-\delta}$$

  can be estimated with

$$\widehat{\mathrm{err}}_{\mathrm{RMSE},k} = \widehat{\mathrm{err}}_{\mathrm{RMSE,s}} + \widehat{\mathrm{err}}_{\mathrm{RMSE,p}} \cdot 2^{k\widehat{\delta}}$$

Numerical Experiments - Verification of Conjecture

Weak Scaling Experiments:

- Fixed time budget $\mathrm{T_B} = 6\mathrm{h}$
- Variable amount of processing units

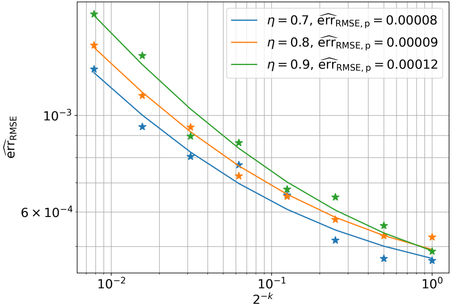
$$|\mathcal{P}| \in \left\{2^{-k}\,|\mathcal{P}_{\max}| \,:\, |\mathcal{P}_{\max}| = 8192,\ k = 0, \dots, 7\right\}$$

- Reduction factor $\eta \in \{0.7, 0.8, 0.9\}$ for $\epsilon_{\mathrm{i}} := \eta\epsilon_{\mathrm{i}-1}$
- Assume λ_{p} is fixed for a fixed η, then

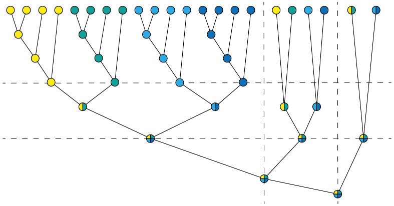
$$\epsilon \lesssim (1 - \lambda_{\mathrm{p}}) \cdot \mathrm{T_B}^{-\delta} + \lambda_{\mathrm{p}}(|\mathcal{P}| \cdot \mathrm{T_B})^{-\delta}$$

can be estimated with

$$\widehat{\mathrm{err}}_{\mathrm{RMSE},k} = \widehat{\mathrm{err}}_{\mathrm{RMSE,s}} + \widehat{\mathrm{err}}_{\mathrm{RMSE,p}} \cdot 2^{k\widehat{\delta}}$$



Budgeted Multi-level Monte Carlo - Optimality Principle

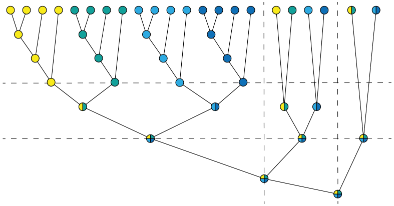


Welford's Update Algorithm:

function Welford$(\mathrm{data}_A, \mathrm{data}_B)$:

$$
\begin{cases}
M_{AB,\ell} & \leftarrow \quad M_{B,\ell} + M_{A,\ell} \\
\delta_{AB,\ell} & \leftarrow \quad \widehat{\mathrm{Q}}_{B,\ell} - \widehat{\mathrm{Q}}_{A,\ell} \\
\widehat{\mathrm{Q}}_{AB,\ell} & \leftarrow \quad \widehat{\mathrm{Q}}_{A,\ell} + \frac{M_{B,\ell}}{M_{AB,\ell}} \delta_{AB,\ell} \\
S_{\mathrm{Q}_2,AB,\ell} & \leftarrow \quad S_{\mathrm{Q}_2,A,\ell} + S_{\mathrm{Q}_2,B,\ell} + \frac{M_{A,\ell} M_{B,\ell}}{M_{AB,\ell}} \delta^2_{AB,\ell} \\
s^2_{\mathrm{Q}_{AB,\ell}} & \leftarrow \quad (M_{AB,\ell} - 1)^{-1} S_{\mathrm{Q}_2,AB,\ell} \\
\texttt{return} & \quad \{ M_{AB,\ell}, \widehat{\mathrm{Q}}_{AB,\ell}, S_{\mathrm{Q}_2,AB,\ell}, \dots \}_{\ell=0}^{L_{\mathrm{i}}}
\end{cases}
$$

Budgeted Multi-level Monte Carlo - Optimality Principle



Reward Function: Compute with Welford's Update Algorithm

$$\Delta\mathrm{err}_{\mathrm{MSE}} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$$

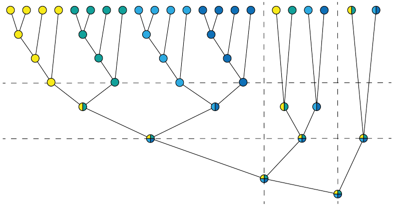
with $\mathtt{data}_{\mathtt{i}-1} \in \mathcal{S}$ and $\{\Delta M_{\mathtt{i},\ell}\}_{\ell=0}^{L_{\mathtt{i}}} \in \mathcal{A}$.

Welford's Update Algorithm:

function Welford($\mathrm{data}_A, \mathrm{data}_B$):

$$\begin{cases} M_{AB,\ell} & \leftarrow \; M_{B,\ell} + M_{A,\ell} \\ \delta_{AB,\ell} & \leftarrow \; \widehat{\mathrm{Q}}_{B,\ell} - \widehat{\mathrm{Q}}_{A,\ell} \\ \widehat{\mathrm{Q}}_{AB,\ell} & \leftarrow \; \widehat{\mathrm{Q}}_{A,\ell} + \frac{M_{B,\ell}}{M_{AB,\ell}} \delta_{AB,\ell} \\ S_{\mathrm{Q}_{2},AB,\ell} & \leftarrow \; S_{\mathrm{Q}_{2},A,\ell} + S_{\mathrm{Q}_{2},B,\ell} + \frac{M_{A,\ell} M_{B,\ell}}{M_{AB,\ell}} \delta^2_{AB,\ell} \\ s^2_{\mathrm{Q}_{AB,\ell}} & \leftarrow \; (M_{AB,\ell} - 1)^{-1} S_{\mathrm{Q}_{2},AB,\ell} \\ \mathtt{return} & \{M_{AB,\ell}, \widehat{\mathrm{Q}}_{AB,\ell}, S_{\mathrm{Q}_{2},AB,\ell}, \dots\}_{\ell=0}^{L_{\mathtt{i}}} \end{cases}$$

Budgeted Multi-level Monte Carlo - Optimality Principle



Reward Function: Compute with Welford's Update Algorithm

$$\Delta \mathrm{err}_{\mathsf{MSE}} \colon \mathcal{S} \times \mathcal{A} \to \mathbb{R}$$

with $\mathtt{data}_{\mathtt{i}-1} \in \mathcal{S}$ and $\{\Delta M_{\mathtt{i},\ell}\}_{\ell=0}^{L_{\mathtt{i}}} \in \mathcal{A}$.

Welford's Update Algorithm:

function Welford$(\mathtt{data}_A, \mathtt{data}_B)$:

$$\begin{cases} M_{AB,\ell} & \leftarrow \quad M_{B,\ell} + M_{A,\ell} \\ \delta_{AB,\ell} & \leftarrow \quad \widehat{\mathrm{Q}}_{B,\ell} - \widehat{\mathrm{Q}}_{A,\ell} \\ \widehat{\mathrm{Q}}_{AB,\ell} & \leftarrow \quad \widehat{\mathrm{Q}}_{A,\ell} + \frac{M_{B,\ell}}{M_{AB,\ell}} \delta_{AB,\ell} \\ S_{\mathrm{Q}_2,AB,\ell} & \leftarrow \quad S_{\mathrm{Q}_2,A,\ell} + S_{\mathrm{Q}_2,B,\ell} + \frac{M_{A,\ell} M_{B,\ell}}{M_{AB,\ell}} \delta_{AB,\ell}^2 \\ s^2_{\mathrm{Q}_{AB,\ell}} & \leftarrow \quad (M_{AB,\ell} - 1)^{-1} S_{\mathrm{Q}_2,AB,\ell} \\ \mathtt{return} & \quad \{M_{AB,\ell}, \widehat{\mathrm{Q}}_{AB,\ell}, S_{\mathrm{Q}_2,AB,\ell}, \dots\}_{\ell=0}^{L_{\mathtt{i}}} \end{cases}$$

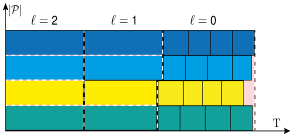
Bellman Equation: Estimation of final MSE

$$\widehat{\mathrm{err}}_{\mathsf{MSE}}^{\mathsf{final}} = \widehat{\mathrm{err}}_{\mathsf{MSE}}^{\mathsf{init}} - \widehat{\mathrm{err}}_{\mathsf{MSE}}(\mathrm{B}_{\mathtt{i}}, \epsilon_0)$$

for $\eta \in (0,1)$ with

$$\widehat{\mathrm{err}}_{\mathsf{MSE}}(\mathrm{B}_{\mathtt{i}}, \epsilon_{\mathtt{i}}) = \max_{\substack{\{\Delta M_{\mathtt{i},\ell}\}_{\ell=0}^{L_{\mathtt{i}}} \\ \text{s.t. } \widehat{\mathrm{C}}_{\mathtt{i}} < \mathrm{B}_{\mathtt{i}}}} \Big\{ \Delta \mathrm{err}_{\mathsf{MSE}}(\mathtt{data}_{\mathtt{i}-1}, \{\Delta M_{\mathtt{i},\ell}\}_{\ell=0}^{L_{\mathtt{i}}}) + \widehat{\mathrm{err}}_{\mathsf{MSE}}\Big(\mathrm{B}_{\mathtt{i}} - \sum_{\ell=0}^{L_{\mathtt{i}}} \mathrm{C}_{\mathtt{i},\ell}, \, \eta \cdot \epsilon_{\mathtt{i}}\Big) \Big\}$$

Budgeted Multi-level Monte Carlo - Parallelization Bias

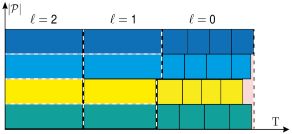


Recall: For a parallel BMLMC execution, it holds

$$\epsilon \lesssim \underbrace{(1-\lambda_{\mathrm{p}}) \cdot \mathrm{T}_{\mathrm{B}}^{-\delta}}_{=:\epsilon_{\mathrm{s}}} + \underbrace{\lambda_{\mathrm{p}}(|\mathcal{P}| \cdot \mathrm{T}_{\mathrm{B}})^{-\delta}}_{=:\epsilon_{\mathrm{p}}}$$

with $\delta \in \left\{ \frac{1}{2}, \frac{\alpha}{(2\alpha+(\gamma-\beta))} \right\}$ and $\lambda_{\mathrm{p}} \in [0, 1]$.

Budgeted Multi-level Monte Carlo - Parallelization Bias



Recall: For a parallel BMLMC execution, it holds

$$\epsilon \lesssim \underbrace{(1-\lambda_{\mathrm{p}}) \cdot \mathrm{T}_{\mathrm{B}}^{-\delta}}_{=:\epsilon_{\mathrm{s}}} + \underbrace{\lambda_{\mathrm{p}}(|\mathcal{P}| \cdot \mathrm{T}_{\mathrm{B}})^{-\delta}}_{=:\epsilon_{\mathrm{p}}}$$

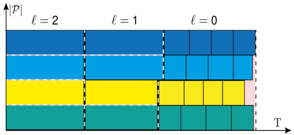
with $\delta \in \left\{\frac{1}{2}, \frac{\alpha}{(2\alpha+(\gamma-\beta))}\right\}$ and $\lambda_{\mathrm{p}} \in [0,1]$.

Ansatz: Inverted ϵ-Cost for feasible execution

$$\lambda_{\mathrm{p}} = 0: \quad \epsilon \lesssim \mathrm{T}_{\mathrm{B}}^{-\delta} = \mathrm{C}_{\epsilon}^{-\delta}$$

$$\lambda_{\mathrm{p}} = 1: \quad \epsilon \lesssim (|\mathcal{P}| \cdot \mathrm{T}_{\mathrm{B}})^{-\delta} = \mathrm{C}_{\epsilon}^{-\delta}$$

Budgeted Multi-level Monte Carlo - Parallelization Bias



Recall: For a parallel BMLMC execution, it holds

$$\epsilon \lesssim \underbrace{(1-\lambda_p)\cdot T_B^{-\delta}}_{=:\epsilon_s} + \underbrace{\lambda_p(|\mathcal{P}|\cdot T_B)^{-\delta}}_{=:\epsilon_p}$$

with $\delta \in \left\{\frac{1}{2}, \frac{\alpha}{(2\alpha+(\gamma-\beta))}\right\}$ and $\lambda_p \in [0,1]$.

Ansatz: Inverted ϵ-Cost for feasible execution

$$\lambda_p = 0: \ \epsilon \lesssim T_B^{-\delta} = C_\epsilon^{-\delta}$$

$$\lambda_p = 1: \ \epsilon \lesssim (|\mathcal{P}|\cdot T_B)^{-\delta} = C_\epsilon^{-\delta}$$

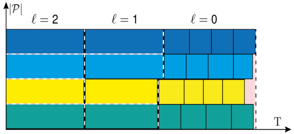
Gustafson's Law: Consider hypothetical speedup

$$S := \frac{T_{\epsilon,s}}{T_{\epsilon,p}} = \frac{\lambda_s + \lambda_p|\mathcal{P}|^{-\delta}}{\lambda_s + \lambda_p} = (1-\lambda_p) + \lambda_p|\mathcal{P}|^{-\delta}$$

Additional error reduction by utilizing $|\mathcal{P}|$ units

$$\epsilon \lesssim S \cdot T_B^{-\delta}$$

Budgeted Multi-level Monte Carlo - Parallelization Bias



Recall: For a parallel BMLMC execution, it holds

$$\epsilon \lesssim \underbrace{(1 - \lambda_p) \cdot T_B^{-\delta}}_{=:\epsilon_s} + \underbrace{\lambda_p(|\mathcal{P}| \cdot T_B)^{-\delta}}_{=:\epsilon_p}$$

with $\delta \in \left\{ \frac{1}{2}, \frac{\alpha}{(2\alpha + (\gamma - \beta))} \right\}$ and $\lambda_p \in [0, 1]$.

Ansatz: Inverted ϵ-Cost for feasible execution

$$\lambda_p = 0: \quad \epsilon \lesssim T_B^{-\delta} = C_\epsilon^{-\delta}$$
$$\lambda_p = 1: \quad \epsilon \lesssim (|\mathcal{P}| \cdot T_B)^{-\delta} = C_\epsilon^{-\delta}$$

Gustafson's Law: Consider hypothetical speedup

$$S := \frac{T_{\epsilon,s}}{T_{\epsilon,p}} = \frac{\lambda_s + \lambda_p |\mathcal{P}|^{-\delta}}{\lambda_s + \lambda_p} = (1 - \lambda_p) + \lambda_p |\mathcal{P}|^{-\delta}$$

Additional error reduction by utilizing $|\mathcal{P}|$ units

$$\epsilon \lesssim S \cdot T_B^{-\delta}$$

Algorithmic Requirements:

Large L_i $\longrightarrow$ FE parallelization
Large $M_{i,\ell}$ $\longrightarrow$ Sample parallelization
Unknown $M_{i,\ell}$ $\longrightarrow$ Dynamic load distribution
Usage of `data` $\longrightarrow$ Distributed state machine