UPPSALA UNIVERSITET

Multilevel Monte Carlo Methods for Rare Event Probabilities (and Quantiles)

Daniel Elfverson
daniel.elfverson@it.uu.se

Division of Scientific Computing
Uppsala University
Sweden

Joint work with F. Hellman (Uppsala) and A. Målqvist (Gothenburg)

Outline

Introduction: **Rare Event Probabilities**

- X – a stochastic variable

Definition: Rare event/failure probability

The failure probability p given y is:

$$p = \Pr(X \leq y) \quad \text{or}$$
$$p = F(y),$$

where $F(\cdot)$ is the cdf associated with X.

Assumption

For the failure probability p to be unique we assume the following Lipschitz continuity of $F(\cdot)$

$$|F(x) - F(y)| \leq C|x - y|, \quad \text{for} \quad x, y \in \mathbb{R}.$$

- The failure probability p given y is:

$$p = F(y) = \Pr(X \leq y)$$

- **Goal** – Estimate the probability $p \approx \widehat{Q}$ to a given root mean square error (RMSE), $e(\widehat{Q}) \leq \varepsilon$, using minimal computational cost

Introduction: **Problem formulation**

Model problem

- $\mathcal{M}$ – model
- V – a function space
- $(\Omega, \Sigma, \mathbb{P})$ – a probability space

We assume that there exists a unique solution $u \in V$ given any $\omega \in \Omega$ $\mathbb{P}$-almost surely: that is

$$\mathcal{M}(\omega, u) = 0 \quad \text{a.s.}$$

- $X(u) : V \to \mathbb{R}$ – A quantity of interest (functional) of the solution u
- The solution u is uniquely determined by the data ω, $X(\omega) := X(u(\omega))$

Spatial discretization

Assumption: Numerical error for samples

- For each sample $\omega_i \in \Omega$ the numerical approximation $X_\ell^\epsilon(\omega_i)$ of $X(\omega_i)$ satisfies

$$|X(\omega_i) - X_\ell^\epsilon(\omega_i)| \leq \epsilon_\ell,$$

  for any $\epsilon_\ell > 0$

- Further, the work W for computing $X_\ell^\epsilon(\omega_i)$ depends on the error tolerances and satisfies

$$C\epsilon_\ell^{-q} \leq W(X^\epsilon(\omega_i)) \leq \epsilon_\ell^{-q},$$

  where $C \leq 1$ and $q > 0$ are independent of ω_i

- Let $Q(\omega) = \mathbb{1}(X(\omega) < y)$ and $Q_\ell^\epsilon(\omega) = \mathbb{1}(X_\ell^\epsilon(\omega) < y)$ be binomial distributed random variables

Lemma

Given the previous assumption the following statements

M1 $\left|\mathbb{E}\left[Q_\ell^\epsilon(\omega) - Q(\omega)\right]\right| \leq C_1\epsilon_\ell$,

M2 $\mathbb{V}\left[Q_\ell^\epsilon(\omega) - Q_{\ell-1}^\epsilon(\omega)\right] \leq C_2\epsilon_\ell$ *for* $\ell \geq 1$,

M3 $\mathbb{E}\left[W(Q_\ell^\epsilon(\omega))\right] = C_3\epsilon_\ell^{-q}$,

are satisfied where C_1, C_2, *and* C_3 *do not depend on the sample or the underlying discretization*

MLMC method

- Given $\epsilon_0 > \epsilon_1 > \cdots > \epsilon_\ell$ and $\{N_\ell\}_{\ell=0}^{L}$
- Let $Y_0^\epsilon(\omega) = Q_0^\epsilon(\omega)$ and $Y_\ell^\epsilon(\omega) = Q_\ell^\epsilon(\omega) - Q_{\ell-1}^\epsilon(\omega)$ for $\ell \geq 1$, the MLMC estimator is

$$\widehat{Q}_{\{N_\ell\},\epsilon}^{ML} = \sum_{\ell=0}^{L} N_\ell^{-1} \sum_{i=1}^{N_\ell} Y_\ell^\epsilon(\omega_i)$$

- The computational cost for the MLMC estimator is

$$\mathcal{C}_q\left(\widehat{Q}_{\{N_\ell\},\epsilon}^{ML}\right) = \sum_{\ell=0}^{L} N_\ell \mathcal{C}_q\left(Y_\ell^\epsilon(\omega_i)\right) \sim \sum_{\ell=0}^{L} N_\ell \epsilon_\ell^{-q}$$

Theorem

Then there exist a constant L and a sequence $\{N_\ell\}$ such that the RMSE is less then ε, with the required work in terms of ε,

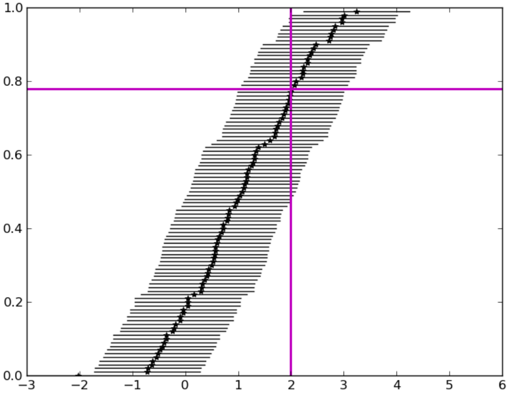
$$\mathbb{E}\left[\mathcal{C}_q\left(\widehat{Q}^{ML}_{\{N_\ell\},\epsilon}\right)\right] \lesssim \begin{cases} \varepsilon^{-2} & q < 1 \\ \varepsilon^{-2}(\log \varepsilon^{-1})^2 & q = 1 \\ \varepsilon^{-1-q} & q > 1 \end{cases} .$$

Proof.

See Giles 08. □

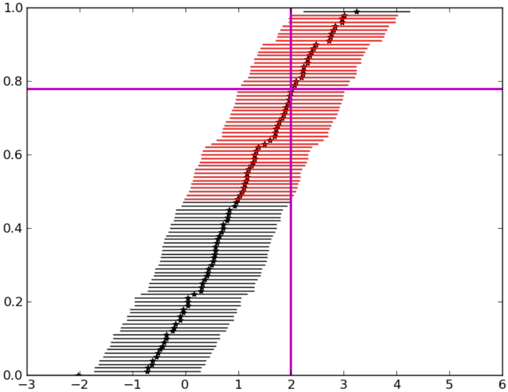
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_0 = 1$
- $\# I_0 = 100$



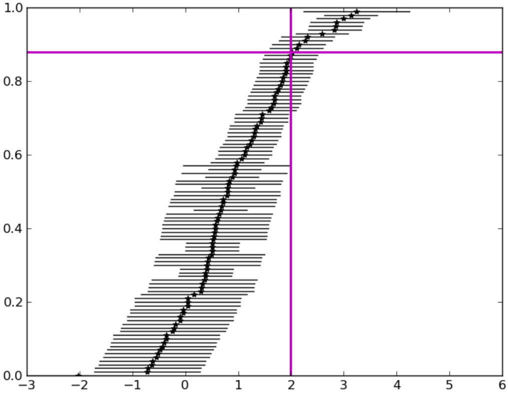
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_0 = 1$
- $\#I_1 = 51$



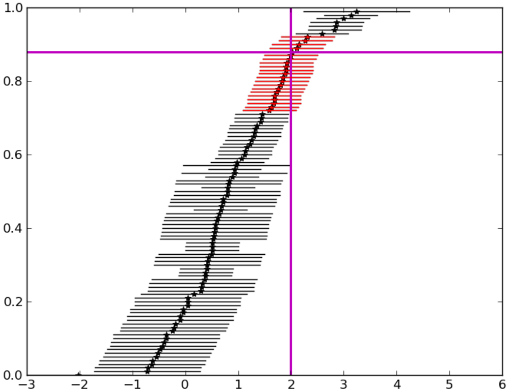
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_1 = 0.5$
- $\#I_1 = 51$



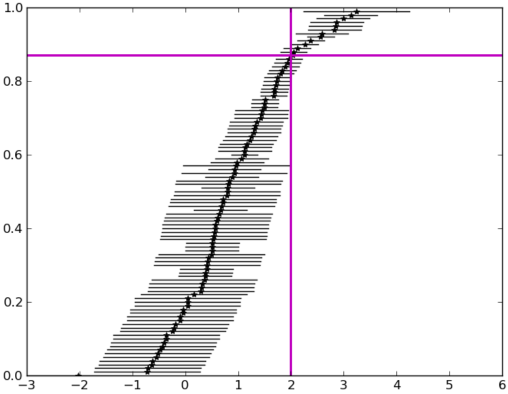
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_1 = 0.5$
- $\# I_2 = 21$



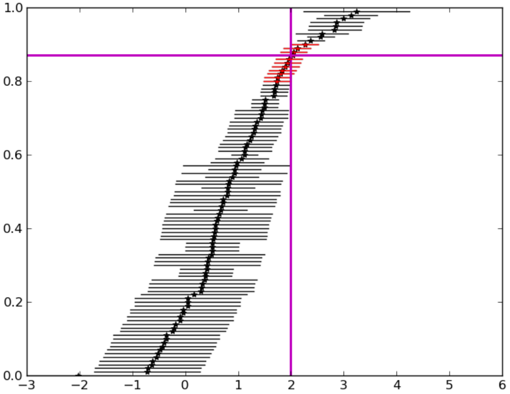
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_2 = 0.25$
- $\#I_2 = 21$



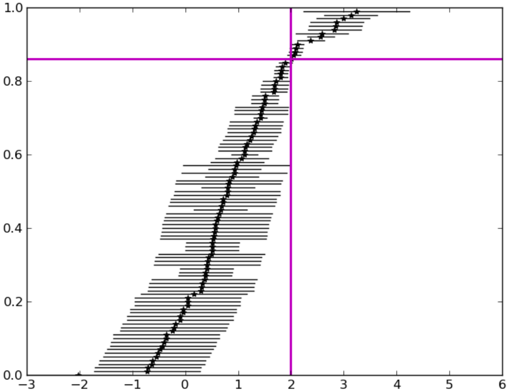
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_2 = 0.25$
- $\#I_3 = 11$



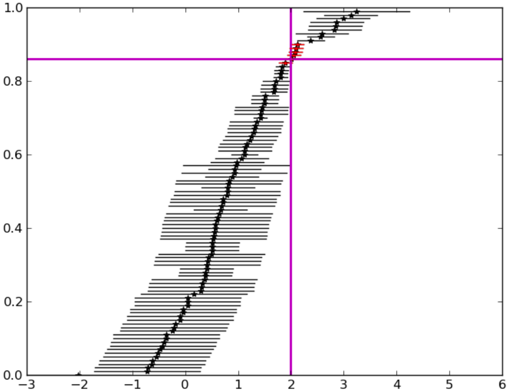
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_3 = 0.0125$
- $\# I_3 = 11$



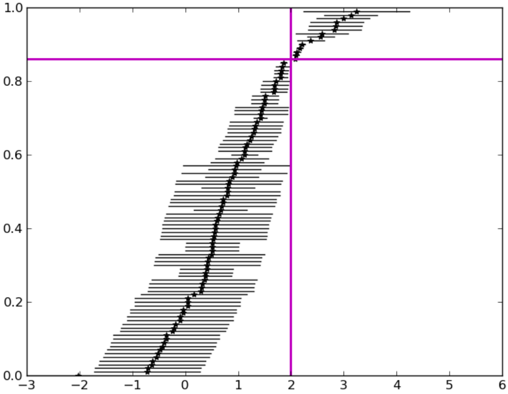
Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_3 = 0.0125$
- $\# I_4 = 6$



Selective algorithm

- $N = 100$
- $y = 2$
- $\epsilon_4 = 0.00625$
- $\# I_4 = 6$



Lemma

The MC estimator is equivalent to the MC estimator using selective refinement.

Lemma

Given N samples in a MC method, the expected number of samples, $\mathbb{E}[\# I_\ell]$, on level $\ell = 0, \ldots, L$ can be bounded as

$$\mathbb{E}[\# I_\ell] \lesssim N\epsilon_\ell.$$

MLMC method using selective refinement

- Gives exactly the same estimator as the (standard) MLMC (previous lemma)
- The cost for the MLMC estimator using selective refinement

$$\mathcal{C}_q\left(\widehat{Q}^{MLS}_{\{N_\ell\},\epsilon}\right) = \sum_{\ell=0}^{L} N_\ell \mathcal{C}_q^\ell,$$

where $\mathcal{C}_q^\ell$ is the "effective" cost for one sample on level ℓ

$$\mathcal{C}_q\left(\widehat{Q}^{MLS}_{\{N_\ell\},\epsilon}\right) = \sum_{\ell=0}^{L} N_\ell \sum_{j=0}^{\ell} \mathcal{C}_q\left(Y_j^\epsilon(\omega_i)\right) \# I_{(j)}/N_\ell$$

$$\sim \sum_{\ell=0}^{L} N_\ell \sum_{j=0}^{\ell} \epsilon_j^{-q+1}$$

Theorem (Computable complexity for the Multilevel Monte Carlo method with selective refinement)

There exist a constant L and a sequence $\{N_\ell\}$ such that the RMSE is less then ε, with the required work in terms of ε,

$$\mathbb{E}\left[\mathcal{C}_q\left(\widehat{Q}^{MLS}_{\{N_\ell\},\epsilon}\right)\right] \lesssim \begin{cases} \varepsilon^{-2} & q < 2 \\ \varepsilon^{-2}(\log \epsilon^{-1})^2 & q = 2 \\ \varepsilon^{-q} & q > 2 \end{cases}.$$

The method is optimal in the sense:

- ($q < 2$) same as the standard MC method on level $= 0$
- ($q > 2$) same complexity as one sample on the finest level L

$$\mathbb{E}\left[\mathcal{C}_q\left(\widehat{Q}^{MLS}_{\{N_\ell\},\epsilon}\right)\right] \lesssim \begin{cases} N & q < 2 \\ \mathcal{C}_q\left(Q^{\epsilon}_L(\omega)\right) & q > 2 \end{cases}$$

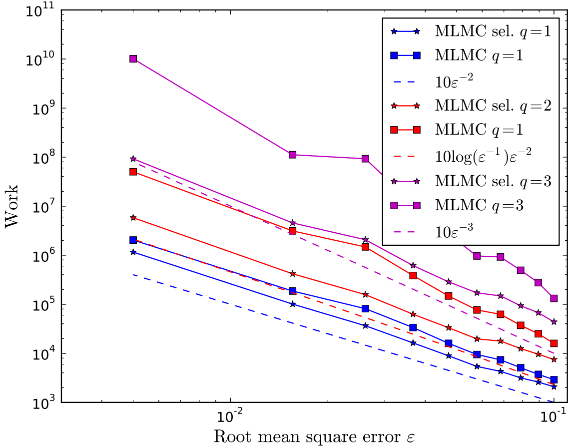
Recall the work for the standard MLMC (without selective refinement)

$$\mathbb{E}\left[\mathcal{C}_q\left(\widehat{Q}^{ML}_{\{N_\ell\},\epsilon}\right)\right] \lesssim \begin{cases} N & q < 1 \\ N^{1/2}\mathcal{C}_q\left(Q^{\epsilon}_L(\omega)\right) & q > 1 \end{cases}$$

Numerical verification: **Demonstrational problem**

- The algorithm proposed in Giles 08 is used to compute the MLMC estimator
- For each ε the algorithm is computed 1000 times to compute the expected work

- Estimate $p = F(y)$ for $q = \{1, 2, 3\}$



Example 1:

Solving a PDE in $2D$ to accuracy ε, on a uniform mesh, using a numerical method with convergence rate $p = 1$, and using multigrid to solve the linear system. The computational cost is $\sim \epsilon^{-2}$.

Example 2:

Solving a PDE in $3D$ to accuracy ε, on a uniform mesh, using a numerical method with convergence rate $p = 1$, and using multigrid to solve the linear system. The computational cost is $\sim \epsilon^{-3}$.