Multi-level Monte Carlo Methods for SPDEs

Klaus Ritter

TU Darmstadt

Joint work with Thomas Müller-Gronbach

Uni Passau

I. Computational Problem

Stochastic Heat Equation

$$
\begin{aligned}
dX(t) &= \Delta X(t)\, dt + B(t, X(t))\, dW(t), \\
X(0) &= x_0.
\end{aligned}
$$

Assumptions

(i) Δ Dirichlet Laplacian on $D =]0,1[^d$, $x_0 \in H = L_2(D)$,

I. Computational Problem

Stochastic Heat Equation

$$dX(t) = \Delta X(t)\,dt + B(t, X(t))\,dW(t),$$
$$X(0) = x_0.$$

Assumptions

(i) Δ Dirichlet Laplacian on $D =]0,1[^d$, $x_0 \in H = L_2(D)$,

(ii) multiplication operators $B(t,x)h = G(t,x) \cdot h$ for $G : [0,T] \times H \to H$ satisfying Lipschitz/Hölder conditions,

I. Computational Problem

Stochastic Heat Equation

$$dX(t) = \Delta X(t)\,dt + B(t, X(t))\,dW(t),$$
$$X(0) = x_0.$$

Assumptions

(i) Δ Dirichlet Laplacian on $D =]0,1[^d$, $x_0 \in H = L_2(D)$,

(ii) multiplication operators $B(t,x)h = G(t,x) \cdot h$ for
$G : [0,T] \times H \to H$ satisfying Lipschitz/Hölder conditions,

(iii) Brownian motion

$$W(t) = \sum_{\boldsymbol{i} \in \mathbb{N}^d} |\boldsymbol{i}|_2^{-\gamma/2} \cdot \beta_{\boldsymbol{i}}(t) \cdot h_{\boldsymbol{i}}$$

with eigenfunctions $h_{\boldsymbol{i}}$ of Δ, independent scalar Bms $(\beta_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{N}^d}$,
either $\gamma = 0$ and $d = 1$ or $\gamma > d \in \mathbb{N}$.

Quadrature Problem: compute

$$S(f) = \mathrm{E}(f(V)) = \int_{\mathfrak{V}} f \, dP_V,$$

where

$$V = X \quad \text{and} \quad f : \mathfrak{V} \to \mathbb{R} \quad \text{for} \quad \mathfrak{V} = L_2([0, T], H) \tag{1}$$

Quadrature Problem: compute

$$S(f) = \mathrm{E}(f(V)) = \int_{\mathfrak{V}} f \, dP_V,$$

where

$$V = X \quad \text{and} \quad f : \mathfrak{V} \to \mathbb{R} \quad \text{for} \quad \mathfrak{V} = L_2([0,T], H) \tag{1}$$

or

$$V = X(T) \quad \text{and} \quad f : \mathfrak{V} \to \mathbb{R} \quad \text{for} \quad \mathfrak{V} = H. \tag{2}$$

Quadrature Problem: compute

$$S(f) = \mathrm{E}(f(V)) = \int_{\mathfrak{V}} f \, dP_V,$$

where

$$V = X \quad \text{and} \quad f : \mathfrak{V} \to \mathbb{R} \quad \text{for} \quad \mathfrak{V} = L_2([0,T], H) \tag{1}$$

or

$$V = X(T) \quad \text{and} \quad f : \mathfrak{V} \to \mathbb{R} \quad \text{for} \quad \mathfrak{V} = H. \tag{2}$$

Classical Monte Carlo Approach: construct a (weak) approximation $\widehat{V}$ to V and sample from $P_{\widehat{V}}$.

For weak approximation, see Hausenblas (2003), Debusche, Printems (2007), Geissert, Kovacs, Larsson (2008), Lindner (2009).

II. Algorithms, Error, and Cost

Randomized algorithm to compute $S(f) = \mathrm{E}(f(V))$ for $f \in F$

$$\widehat{S} : F \times \Omega \to \mathbb{R}, \quad \widehat{S}(f,\omega) = \varphi(f(V_1(\omega)), \ldots, f(V_\nu(\omega)))$$

with evaluation of f at random knots

$$V_1(\omega), \ldots, V_\nu(\omega) \in \bigcup_{m=1}^{\infty} \mathfrak{V}_m$$

for a fixed scale of finite dimensional subspaces

$$\mathfrak{V}_1 \subset \mathfrak{V}_2 \subset \ldots \subset \mathfrak{V}.$$

II. Algorithms, Error, and Cost

Randomized algorithm to compute $S(f) = \mathrm{E}(f(V))$ for $f \in F$

$$\widehat{S} : F \times \Omega \to \mathbb{R}, \quad \widehat{S}(f, \omega) = \varphi(f(V_1(\omega)), \ldots, f(V_\nu(\omega)))$$

with evaluation of f at random knots

$$V_1(\omega), \ldots, V_\nu(\omega) \in \bigcup_{m=1}^{\infty} \mathfrak{V}_m$$

for a fixed scale of finite dimensional subspaces

$$\mathfrak{V}_1 \subset \mathfrak{V}_2 \subset \ldots \subset \mathfrak{V}.$$

Cost and Error of $\widehat{S}$

$$\mathrm{cost}(\widehat{S}) = E\Big(\sum_{i=1}^{\nu} \inf\{\dim \mathfrak{V}_m : V_i \in \mathfrak{V}_m\} + \#\,\text{arith.op.} + \#\,\text{rand.numbers}\Big),$$

$$e(\widehat{S}) = \sup_{f \in F} e(\widehat{S}(f)) = \sup_{f \in F} \big(E|S(f) - \widehat{S}(f)|^2\big)^{1/2}.$$

III. Multi-level Monte Carlo Algorithms

Assumptions

- F is the class of Lip(1)-functionals on $\mathfrak{V}$,

$$|f(v) - f(w)| \leq \|v - w\|, \qquad v, w \in \mathfrak{V}.$$

III. Multi-level Monte Carlo Algorithms

Assumptions

- F is the class of Lip(1)-functionals on $\mathfrak{V}$,

$$|f(v) - f(w)| \leq \|v - w\|, \qquad v, w \in \mathfrak{V}.$$

- $(\widehat{V}_n)_n$ is a sequence of approximations $\widehat{V}_n = \psi_n(W)$ of $V = \psi(W)$ such that

$$\big(\mathrm{E}\,\|\widehat{V}_n - V\|^2\big)^{1/2} \preceq n^{-\alpha} \quad \text{with } \alpha \in \,]0, 1/2]\,,$$

(C(α))

$$\mathrm{cost}(f(\widehat{V}_n), f(\widehat{V}_m)) \preceq \max(n, m).$$

III. Multi-level Monte Carlo Algorithms

Assumptions

- F is the class of Lip(1)-functionals on $\mathfrak{V}$,

$$|f(v) - f(w)| \le \|v - w\|, \qquad v, w \in \mathfrak{V}.$$

- $(\widehat{V}_n)_n$ is a sequence of approximations $\widehat{V}_n = \psi_n(W)$ of $V = \psi(W)$ such that

$$\left(\mathrm{E}\,\|\widehat{V}_n - V\|^2\right)^{1/2} \preceq n^{-\alpha} \quad \text{with } \alpha \in \,]0, 1/2]\,,$$

(C(α))

$$\mathrm{cost}(f(\widehat{V}_n), f(\widehat{V}_m)) \preceq \max(n, m).$$

Standard Monte Carlo with k independent replications of $\widehat{V}_n$

$$\widehat{S}_{k,n}(f) = \frac{1}{k}\sum_{i=1}^{k} f(\widehat{V}_n^{(i)}).$$

Multi-level Monte Carlo Basic idea: For $\boldsymbol{n} \in \mathbb{N}^L$ with $n_1 < \ldots < n_L$

$$\mathrm{E}(f(\widehat{V}_{n_L})) = \sum_{\ell=2}^{L} \mathrm{E}(\underbrace{f(\widehat{V}_{n_\ell}) - f(\widehat{V}_{n_{\ell-1}})}_{\Delta_\ell(f)}) + \mathrm{E}(\underbrace{f(\widehat{V}_{n_1})}_{\Delta_1(f)}).$$

Multi-level Monte Carlo Basic idea: For $\boldsymbol{n} \in \mathbb{N}^L$ with $n_1 < \ldots < n_L$

$$\mathrm{E}(f(\widehat{V}_{n_L})) = \sum_{\ell=2}^{L} \mathrm{E}(\underbrace{f(\widehat{V}_{n_\ell}) - f(\widehat{V}_{n_{\ell-1}})}_{\Delta_\ell(f)}) + \mathrm{E}(\underbrace{f(\widehat{V}_{n_1})}_{\Delta_1(f)}).$$

Choose $\boldsymbol{k} \in \mathbb{N}^L$ with $k_1 \geq \ldots \geq k_L$ and define

$$\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}(f) = \sum_{\ell=1}^{L} \frac{1}{k_\ell} \sum_{i=1}^{k_\ell} \Delta_\ell^{(i)}(f)$$

with independent

$$\underbrace{\Delta_1^{(1)}(f), \ldots, \Delta_1^{(k_1)}(f)}_{\text{i.i.d. as } \Delta_1(f)}, \ldots, \underbrace{\Delta_L^{(1)}(f), \ldots, \Delta_L^{(k_L)}(f)}_{\text{i.i.d. as } \Delta_L(f)}.$$

Theorem

For the standard Monte Carlo method with $k \asymp n^{2\alpha}$

$$e(\widehat{S}_{k,n}) \preceq \left(\mathrm{cost}(\widehat{S}_{k,n})\right)^{-\alpha/(1+2\alpha)}.$$

For $f \in F$ with Var$(f(X)) > 0$ and $\mathrm{cost}(f(\widehat{X}_n)) \asymp n$

$$e(\widehat{S}_{k,n}(f)) \asymp \left(\mathrm{cost}(\widehat{S}_{k,n})\right)^{-\alpha/(1+2\alpha)}.$$

Theorem

For the standard Monte Carlo method with $k \asymp n^{2\alpha}$

$$e(\widehat{S}_{k,n}) \preceq \big(\mathrm{cost}(\widehat{S}_{k,n})\big)^{-\alpha/(1+2\alpha)}.$$

For $f \in F$ with Var$(f(X)) > 0$ and $\mathrm{cost}(f(\widehat{X}_n)) \asymp n$

$$e(\widehat{S}_{k,n}(f)) \asymp \big(\mathrm{cost}(\widehat{S}_{k,n})\big)^{-\alpha/(1+2\alpha)}.$$

For the multi-level Monte Carlo method with $n_\ell = 2^\ell, k_\ell = 2^{L-\ell}$

$$e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq \big(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}})\big)^{-\alpha} \cdot \big(\log(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))\big)^{\lfloor 2\alpha \rfloor}.$$

Theorem

For the standard Monte Carlo method with $k \asymp n^{2\alpha}$

$$e(\widehat{S}_{k,n}) \preceq \left(\mathrm{cost}(\widehat{S}_{k,n})\right)^{-\alpha/(1+2\alpha)}.$$

For $f \in F$ with $\mathrm{Var}(f(X)) > 0$ and $\mathrm{cost}(f(\widehat{X}_n)) \asymp n$

$$e(\widehat{S}_{k,n}(f)) \asymp \left(\mathrm{cost}(\widehat{S}_{k,n})\right)^{-\alpha/(1+2\alpha)}.$$

For the multi-level Monte Carlo method with $n_\ell = 2^\ell, k_\ell = 2^{L-\ell}$

$$e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq \left(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}})\right)^{-\alpha} \cdot \left(\log(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))\right)^{\lfloor 2\alpha \rfloor}.$$

Remark Better bounds if bias estimates are also available.

Multi-level Monte Carlo

- *integral equations, parametric integration,* see Heinrich (1998), Heinrich, Sindambiwe (1999),

Multi-level Monte Carlo

- *integral equations, parametric integration,* see Heinrich (1998), Heinrich, Sindambiwe (1999),
- *SDEs,* see Giles (2008, . . .), Giles, Higham, Mao (2009), Avikainen (2009), Dereich (2009), Dereich, Heidenreich (2009), Hoel, v. Schwerin, Szepessy, Tempone (2009),

Multi-level Monte Carlo

- *integral equations, parametric integration,* see Heinrich (1998), Heinrich, Sindambiwe (1999),
- *SDEs,* see Giles (2008, . . .), Giles, Higham, Mao (2009), Avikainen (2009), Dereich (2009), Dereich, Heidenreich (2009), Hoel, v. Schwerin, Szepessy, Tempone (2009),
- *measures on sep. Banach spaces,* optimality results for quadrature problem in terms of quantization numbers and Kolmogorov widths, see Creutzig, Dereich, Müller-Gronbach, R (2009),

Multi-level Monte Carlo

- *integral equations, parametric integration,* see Heinrich (1998), Heinrich, Sindambiwe (1999),
- *SDEs,* see Giles (2008, . . .), Giles, Higham, Mao (2009), Avikainen (2009), Dereich (2009), Dereich, Heidenreich (2009), Hoel, v. Schwerin, Szepessy, Tempone (2009),
- *measures on sep. Banach spaces,* optimality results for quadrature problem in terms of quantization numbers and Kolmogorov widths, see Creutzig, Dereich, Müller-Gronbach, R (2009),
- *integration on sequence spaces,* see Hickernell, Müller-Gronbach, Niu, R (2009). See also Kuo, Sloan, Wasilkowski, Woźniakowski (2009) for a deterministic counterpart.

Recall: multi-level yields order α if

$$\begin{aligned} &\left(\mathrm{E}\,\|\widehat{V}_n - V\|^2\right)^{1/2} \preceq n^{-\alpha} \quad \text{with } \alpha \in \left]0, 1/2\right], \\ &\mathrm{cost}(f(\widehat{V}_n), f(\widehat{V}_m)) \preceq \max(n, m). \end{aligned} \tag{C(α)}$$

Theorem Müller-Gronbach, R (2007)

Consider problem (1). Suppose that

(i) $G(t,x) = G(t)$ and $G : [0,T] \to H$ is smooth, or

(ii) $G(t,x) = g \circ x$ and $g : \mathbb{R} \to \mathbb{R}$ is smooth and $d = 1$.

Theorem Müller-Gronbach, R (2007)

Consider problem (1). Suppose that

(i) $G(t,x) = G(t)$ and $G : [0,T] \to H$ is smooth, or

(ii) $G(t,x) = g \circ x$ and $g : \mathbb{R} \to \mathbb{R}$ is smooth and $d = 1$.

Then C(α) holds with

$$\alpha = \frac{\min(\gamma - d, d) + 2}{2(d+2)}.$$

Theorem Müller-Gronbach, R (2007)

Consider problem (1). Suppose that

(i) $G(t,x) = G(t)$ and $G : [0,T] \to H$ is smooth, or

(ii) $G(t,x) = g \circ x$ and $g : \mathbb{R} \to \mathbb{R}$ is smooth and $d = 1$.

Then C(α) holds with

$$\alpha = \frac{\min(\gamma - d, d) + 2}{2(d+2)}.$$

Remarks

- Optimality of α.
- Key: use of non-uniform time discretization.

IV. Numerical Experiments

See Graubner (2008).

We consider

- a heat equation and
- a Burgers equation

with

- $d = 1$,
- additive space-time white noise,
- initial value zero,
- and Dirichlet boundary conditions.

Heat equation

$$dX(t) = \Delta X(t)\, dt + dW(t).$$

Heat equation

$$dX(t) = \Delta X(t)\,dt + dW(t).$$

Strong approximations $\widehat{V}_n$ of

$$V = X \qquad \text{or} \qquad V = X(1)$$

with $C(1/6)$ via finite differences, implicit Euler scheme and bilinear interpolation, see Gyöngy (1999). Then

$$e(\widehat{S}_{k,n}) \asymp (\mathrm{cost}(\widehat{S}_{k,n}))^{-1/8},$$

$$e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq (\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))^{-1/6}.$$

Heat equation

$$dX(t) = \Delta X(t)\, dt + dW(t).$$

Strong approximations $\widehat{V}_n$ of

$$V = X(1)$$

with $C(1/2)$ by spectral Galerkin with non-uniform time discretization and implicit Euler scheme, see Müller-Gronbach, R, Wagner (2008). Then

$$e(\widehat{S}_{k,n}) \asymp (\mathrm{cost}(\widehat{S}_{k,n}))^{-1/4},$$

$$e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq (\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))^{-1/2} \cdot \log(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}})).$$

Heat equation

$$dX(t) = \Delta X(t)\,dt + dW(t).$$

Strong approximations $\widehat{V}_n$ of

$$V = X(1)$$

with $C(1/2)$ by spectral Galerkin with non-uniform time discretization and implicit Euler scheme, see Müller-Gronbach, R, Wagner (2008). Then

$$e(\widehat{S}_{k,n}) \asymp (\mathrm{cost}(\widehat{S}_{k,n}))^{-1/4},$$

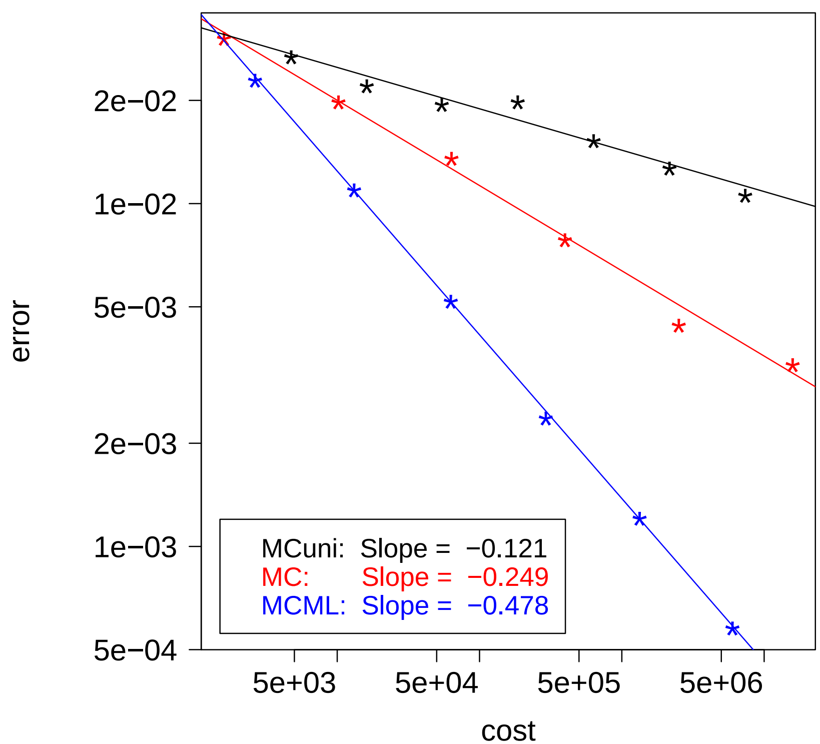
$$e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq (\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))^{-1/2} \cdot \log(\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}})).$$

Remark Uniform time-discretization or finite difference approach yields

$$e(\widehat{S}_{k,n}) \asymp (\mathrm{cost}(\widehat{S}_{k,n}))^{-1/8}, \qquad e(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}) \preceq (\mathrm{cost}(\widehat{S}^{\mathsf{ML}}_{\boldsymbol{k},\boldsymbol{n}}))^{-1/6}.$$

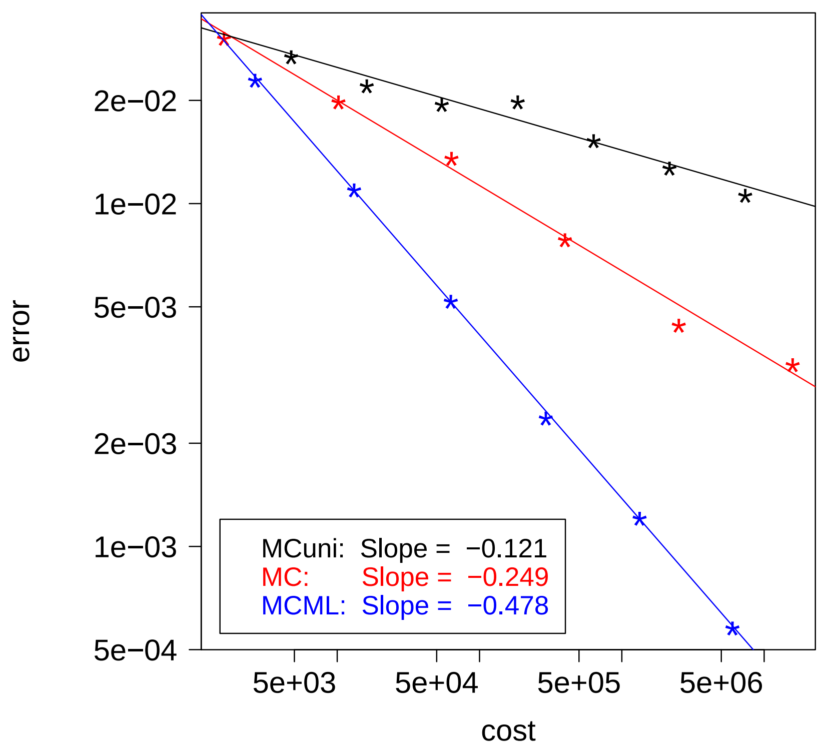
Numerical results

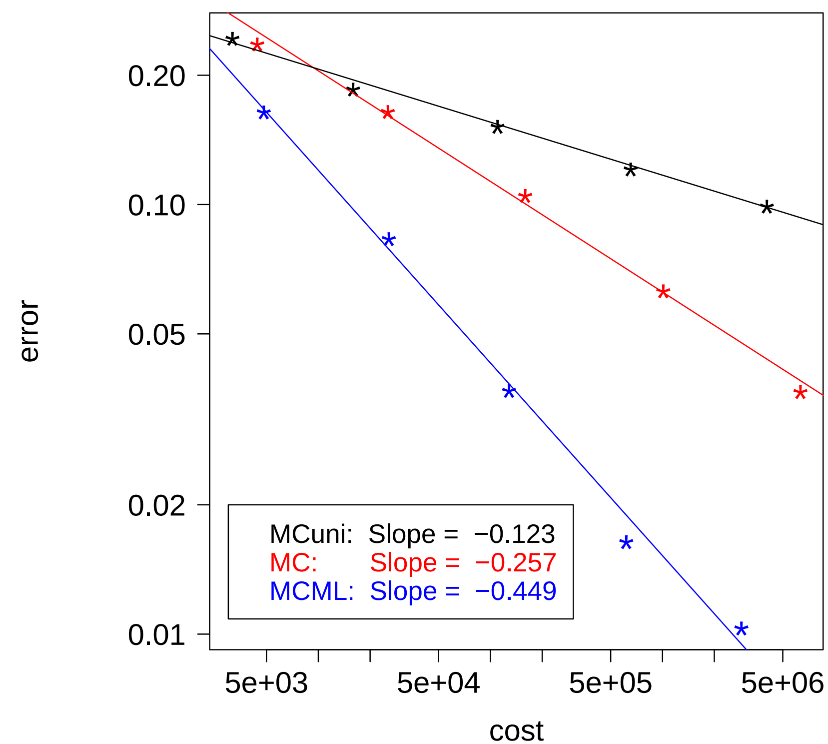
$$\int_0^1 X(1,u)\,du$$



Numerical results

$$\int_0^1 X(1,u)\,du \qquad\qquad \max_{0\le u\le 1} X(1,u).$$





Burgers equation

$$dX(t) = \left(\frac{\partial^2}{\partial u^2} X(t) + X(t) \cdot \frac{\partial}{\partial u} X(t) \right) dt + dW(t).$$

Burgers equation

$$dX(t) = \left(\frac{\partial^2}{\partial u^2} X(t) + X(t) \cdot \frac{\partial}{\partial u} X(t) \right) dt + dW(t).$$

Error bound for semi-discretization in space via finite differences, see Alabert, Gyöngy (2006).

Burgers equation

$$dX(t) = \left(\frac{\partial^2}{\partial u^2} X(t) + X(t) \cdot \frac{\partial}{\partial u} X(t) \right) \, dt + dW(t).$$

Error bound for semi-discretization in space via finite differences, see Alabert, Gyöngy (2006).

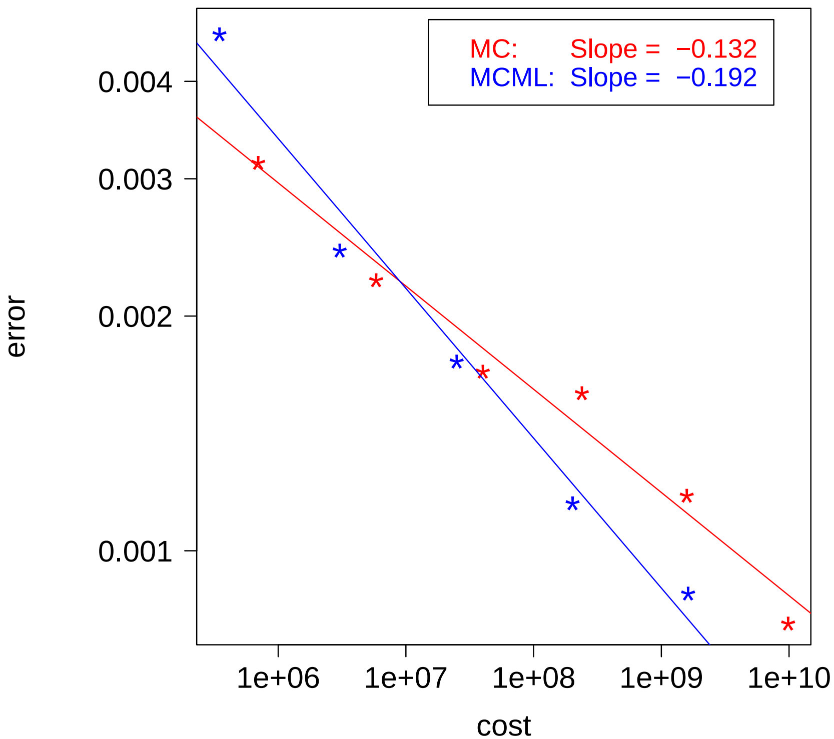
Strong approximations $\widehat{V}_n$ of

$$V = X(1)$$

via finite difference, implicit Euler scheme, and linear interpolation, supposing that $\alpha = 1/6$.

Numerical results

$$\int_0^1 X(1, u)\, du$$



Numerical results

$$\int_0^1 X(1,u)\,du \qquad\qquad \max_{0\le u\le 1} X(1,u).$$