Multilevel Sequential Monte Carlo samplers & particle filters

Kody Law
& A. Beskos (UCL), A. Jasra (NUS), K. Kamatani (Osaka), R. Tempone (KAUST), Y. Zhou (NUS)

Sensitivity, Error and Uncertainty Quantification for Atomic, Plasma, and Material Data, Stony Brook University, New York, USA

November 13, 2015
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
MLMC BIP SMC MLSMC MLPF Summary

Orientation

Aim: Approximately sample from probability distribution $\eta_\infty$, which needs to be approximated by some $\eta_L$, and can be evaluated only up to a normalizing constant.

Solution: The multilevel Monte Carlo (MLMC) framework is extended to Sequential Monte Carlo (SMC) samplers, yielding the MLSMC method for Bayesian inference problems.

- MLMC methods reduce cost to error $= O(\varepsilon)$, can be used in the case that $\eta_L$ can be sampled from directly [G08].
- Here it is assumed that $\eta_\infty$ and $\eta_L$ cannot be sampled from directly, but can be evaluated up to a normalizing constant (e.g. Bayesian inference problems).
- SMC Samplers are a general framework for sampling from such distributions [DDJ06].
- Hot off the press: multilevel particle filter (MLPF) for the sequential data case.
Inverse Problems

Data

\[ y = G(u) + e \]

forward model (PDE)

\[ y \in \mathbb{R}^M \]
\[ u \in X \]
\[ G : X \rightarrow \mathbb{R}^M. \]

Data \( y \) are limited in number, noisy, and indirect.

Parameter \( u \) often a function, discretized.

Continuous, bounded, and 1st order differentiable.
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
Single level Monte Carlo

Aim: Approximate \( \eta(g) := \mathbb{E}_\eta(g) \) for \( g : E \to \mathbb{R} \).

Monte Carlo approach

- Discretize the space \( \Rightarrow \) approximate distribution \( \eta_L \).
- Sample \( U_L^{(i)} \sim \eta_L \) i.i.d., and approximate

\[
\eta_L(g) := \mathbb{E}_{\eta_L}(g) \approx \hat{Y}_L^{N_L} := \frac{1}{N_L} \sum_{i=1}^{N_L} g(U_L^{(i)}).
\]

Mean square error (MSE) \( \mathbb{E}\{\hat{Y}_L - \mathbb{E}_{\eta_\infty}[g(U)]\}^2 \) splits into

\[
\mathbb{E}\{\hat{Y}_L - \mathbb{E}_{\eta_L}[g(U)]\}^2 + \mathbb{E}\{\mathbb{E}_{\eta_L}[g(U)] - \mathbb{E}_{\eta_\infty}[g(U)]\}^2
\]

- Variance = \( O(N_L^{-1}) \)
- Bias

Cost to achieve MSE = \( O(\varepsilon^2) \) is \( \text{Cost}(U_L^{(i)}) \times \varepsilon^{-2} \).
Multilevel Monte Carlo I

Introduce a hierarchy of discretization levels $\{\eta_l\}_{l=1}^L$ and define

$$Y_l = \{ \mathbb{E}_{\eta_l}[g(U)] - \mathbb{E}_{\eta_{l-1}}[g(U)] \},$$

with $\eta_{-1} := 0$.

Observe the telescopic sum

$$\mathbb{E}_{\eta_L}[g(U)] = \sum_{l=0}^{L} Y_l.$$

Each term can be unbiasedly approximated by

$$Y_l^{N_l} = \frac{1}{N_l} \sum_{i=1}^{N_l} \{ g(U_l^{(i)}) - g(U_{l-1}^{(i)}) \}$$

where $g(U_{-1}^{(i)}) := 0$. 
Multilevel Monte Carlo approach:

- Sample i.i.d. $(U_l, U_{l-1})^{(i)} \sim \tilde{\eta}^l$, such that
  \[ \int \tilde{\eta}^l du_{l-1} = \eta_{l-1}, \]
  and approximate
  \[ \eta_L(g) \approx \hat{Y}_{L, \text{Multi}} := \sum_{l=0}^{L} Y_l^{N_l}. \]

- Mean square error (MSE) given by
  \[
  \mathbb{E}\left\{ \hat{Y}_{L, \text{Multi}} - \mathbb{E}_{\eta_\infty}[g(U)] \right\}^2 = \]
  \[
  \mathbb{E}\left\{ \hat{Y}_{L, \text{Multi}} - \mathbb{E}_{\eta_L}[g(U)] \right\}^2 + \mathbb{E}\left\{ \mathbb{E}_{\eta_L}[g(U)] - \mathbb{E}_{\eta_\infty}[g(U)] \right\}^2.
  \]
  \(\text{variance} = \sum_{l=0}^{L} V_l/N_l\)
  \(\text{bias}\)

- Fix bias by choosing \(L\). Minimize cost \(C = \sum_{l=0}^{L} C_l N_l\) as a function of \(\{N_l\}_{l=0}^{L}\) for fixed variance \(\Rightarrow N_l \propto \sqrt{V_l/C_l}.\)
Multilevel vs. Single level

Assume $h_l = M^{-l}$ and there are $\alpha$, and $\beta > \zeta$ such that

(i) weak error $|E[g(U_l) - g(U)]| = O(h_i^\alpha)$.

(ii) strong error $E|g(U_l) - g(U)|^2 = O(h_i^\beta) \Rightarrow V_l = O(h_i^\beta)$,

(iii) computational cost for a realisation of $g(U_l) - g(U_{l-1})$, $C_l = O(h_i^{-\zeta})$.

In both cases, require $h_L^\alpha = O(\varepsilon) \Rightarrow L \propto |\log \varepsilon|$.

- **Single level cost** $C = O(\varepsilon^{-\zeta/\alpha-2})$ : cost per sample is $C_L \propto \varepsilon^{-\zeta/\alpha}$, and fixed $V \propto \varepsilon^2 \Rightarrow N_L \propto \varepsilon^{-2}$.

- **Multilevel cost** $C_{ML} = O(\varepsilon^{-2})$ : $N_l \propto \varepsilon^{-2} h_i^{(\beta+\zeta)/2}$, so $V \propto \varepsilon^2 K_L$ and $C \propto \varepsilon^{-2} K_L$ for $K_L = \sum_{l=0}^{L} h_i^{(\beta-\zeta)/2} = O(1)$ [G08] – cost of simulating a scalar random variable.

- **Example**: Milstein solution of SDE

\[ C = O(\varepsilon^{-3}) \quad \text{vs.} \quad C_{ML} = O(\varepsilon^{-2}). \]
Inverse Problems

Data

\[ y = G(u) + e \]

Data \( y \) are limited in number, noisy, and indirect.

Parameter

\( y \in \mathbb{R}^M \)

Parameter \( u \) often a function, discretized.

\( u \in X \)

Continuous, bounded, and 1st order differentiable.

\( G : X \rightarrow \mathbb{R}^M. \)
Forward problem

Let \( V := H^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega) =: V^* \), \( \Omega \subset \mathbb{R}^d \) with \( \partial\Omega \in C^1 \) convex, and \( f \in V^* \). Consider

\[
-\nabla \cdot (\hat{u}\nabla p) = f, \quad \text{on } \Omega \\
p = 0, \quad \text{on } \partial\Omega,
\]

where

\[
\hat{u}(x) = \bar{u}(x) + \sum_{k=1}^{K} u_k \sigma_k \Phi_k(x).
\]

Define \( u = \{u_k\}_{k=1}^{K} \in E := \prod_{k=1}^{K} [-1, 1] \), with \( u_k \sim U[-1, 1] \) i.i.d. This determines the prior distribution for \( u \). Assume \( \bar{u}, \Phi_k \in C^\infty \), and \( \|\Phi_k\|_\infty = 1 \) for all \( k \). Also require

\[
\inf_{x} \hat{u}(x) \geq \inf_{x} \bar{u}(x) - \sum_{k=1}^{K} \sigma_k \geq u_* > 0.
\]
Bayesian inference problem

Let $p(\cdot; u)$ denote the weak solution for parameter value $u$, and define

$$\mathcal{G}(p) = [g_1(p), \cdots, g_M(p)]^\top,$$

where $g_m \in V^*$ for $m = 1, \ldots, M$.

**DATA :**

$$y = \mathcal{G}(p(\cdot; u)) + e, \quad e \sim N(0, \Gamma), \quad e \perp u.$$

The *unnormalized* density of $u|y$ is given by

$$\gamma(u) = e^{-\Phi[\mathcal{G}(p(\cdot; u))]}; \quad \Phi(\mathcal{G}) = \frac{1}{2} |\mathcal{G} - y|_\Gamma^2.$$

**TARGET :**

$$\eta(u) = \frac{\gamma(u)}{Z}, \quad Z = \int_{E} \gamma(u) du.$$
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
Let \((E, \mathcal{E})\) be a measurable space, and \(\mathcal{B}_b(E)\) the class of bounded and measurable real-valued functions, with \(\|f\|_\infty := \sup_{u \in E} |f(u)|\).

Consider \(K : E \times \mathcal{E} \to \mathbb{R}^+\). For a finite measure \(\mu\) on \((E, \mathcal{E})\) and \(f : E \to \mathbb{R}\) use notations

\[
\mu K : A \mapsto \int K(u, A) \, \mu(du) ; \quad Kf : u \mapsto \int f(v) \, K(u, dv).
\]

Also write \(\mu(f) = \int f(u) \mu(du)\).
Algorithm

Distributions $\eta_l$ dictated by an accuracy parameter $h_l$ (here FEM mesh diameter) $\infty > h_0 > h_1 \cdots > h_\infty = 0$. Approximate $\mathbb{E}_{\eta_L}[g(U)] = \eta_L(g) = \int_E g(u)\eta_L(u)du$.

Idea: interlace sequential importance resampling (selection) along the hierarchy, and mutation by MCMC kernels.

- Initialize i.i.d. $U_0^i \sim \eta_0, i = 1, \ldots, N$. For $l \in \{0, \ldots, L - 1\}$:
  - Resample $\{\hat{U}_l^i\}_{i=1}^N$ according to the weights $\{G_l(U_l^i) = (\gamma_{l+1}/\gamma_l)(U_l^i)\}_{i=1}^N$.
  - Draw $U_{l+1}^i \sim M_{l+1}(\hat{U}_l^i, \cdot)$, where $M_{l+1}$ is an MCMC kernel such that $\eta_{l+1}M_{l+1} = \eta_{l+1}$.

For $\phi : E \to \mathbb{R}, l \in \{0, \ldots, L\}$, we have the following estimators

$$\mathbb{E}_{\eta_l}[\phi(U)] \approx \eta_l^N(\phi) := \frac{1}{N} \sum_{i=1}^N \phi(U_l^i) .$$
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
Preliminaries

\[
\mathbb{E}_{\eta_L}[g(U)] = \mathbb{E}_{\eta_0}[g(U)] + \sum_{l=1}^{L} \left\{ \mathbb{E}_{\eta_l}[g(U)] - \mathbb{E}_{\eta_{l-1}}[g(U)] \right\}
\]

\[
= \mathbb{E}_{\eta_0}[g(U)] + \sum_{l=1}^{L} \mathbb{E}_{\eta_{l-1}} \left[ \left( \frac{\gamma_l(U)Z_{l-1}}{\gamma_{l-1}(U)Z_l} - 1 \right) g(U) \right].
\]

Sample \((U_0^{1:N_0}, \ldots, U_{L-1}^{1:N_{L-1}})\) as in single level SMC, but with \(+\infty > N_0 \geq N_1 \cdots \geq N_{L-1} \geq 1\). The joint probability distribution is

\[
\prod_{i=1}^{N_0} \eta_0(du^i_0) \prod_{l=1}^{L-1} \prod_{i=1}^{N_l} \frac{\eta_{l-1}^{N_l-1}(G_{l-1}M_l(du^i_l))}{\eta_{l-1}^{N_l-1}(G_{l-1})},
\]

where \(\eta_{l-1}^{N_l-1}(G_{l-1}M_l(du_l)) := \frac{1}{N_l} \sum_{i=1}^{N_l} G_{l-1}(U^i_{l-1})M_l(U^i_{l-1}, du_l)\).

Idea: Approximate \(\dagger\) using SMC sample hierarchy.
The MLSMC consistent estimator of $\eta_L(g)$ is given by

$$\hat{Y} := \eta_0^0(g) + \sum_{l=1}^{L} \left\{ \frac{\eta_{l-1}^{N_{l-1}}(gG_l-1)}{\eta_{l-1}^{N_{l-1}}(G_l-1)} - \eta_{l-1}^{N_{l-1}}(g) \right\}.$$ 

i) the $L + 1$ terms above are not unbiased estimates of $\mathbb{E}_{\eta_l}[g(U)] - \mathbb{E}_{\eta_{l-1}}[g(U)]$, so decompose MSE as:

$$\mathbb{E}\left[\{\hat{Y} - \mathbb{E}_{\eta_{\infty}}[g(U)]\}^2\right] \leq 2 \mathbb{E}\left[\{\hat{Y} - \mathbb{E}_{\eta_L}[g(U)]\}^2\right] + 2 \left\{\mathbb{E}_{\eta_L}[g(U)] - \mathbb{E}_{\eta_{\infty}}[g(U)]\right\}^2.$$ 

ii) the same $L + 1$ estimates are not independent, so a more complex error analysis will be required to characterise $\mathbb{E}\left[\{\hat{Y} - \mathbb{E}_{\eta_L}[g(U)]\}^2\right]$. 
(A1) There exist $0 < \underline{C} < \bar{C} < +\infty$ such that

$$\sup_{1 \leq l \leq L} \sup_{u \in E} G_l(u) \leq \bar{C}$$

$$\inf_{1 \leq l \leq L} \inf_{u \in E} G_l(u) \geq \underline{C}.$$ 

(A2) There exist a $\rho \in (0, 1)$ such that for any $1 \leq p \leq L - 1$, $(u, v) \in E^2$, $A \in \mathcal{E}$

$$\int_A M_p(u, du') \geq \rho \int_A M_p(v, du').$$
Main result

Define

\[ V_l := \| \frac{Z_{l-1}}{Z_l} G_{l-1} - 1 \|_\infty^2 = O(h_l^\beta) . \]

**Theorem (BJLTZ15)**

Assume (A1-2). For any \( g \in B_b(E) \)

\[
E \left[ \left\{ \hat{Y} - E_{\eta_l}[g(U)] \right\}^2 \right] = \frac{V}{2} \\
\lesssim \frac{1}{N_0} + \sum_{l=1}^{L} \left( \frac{V_l}{N_l} + \left( \frac{V_l}{N_l} \right)^{1/2} \sum_{q=l+1}^{L} \frac{V_q^{1/2}}{N_q} \right) .
\]

In particular, for \( \beta > \zeta \), \( L \) and \( \{N_l\}_{l=0}^{L} \) can be chosen such that \( \text{MSE} = O(\varepsilon^2) \) for computational cost \( = O(\varepsilon^{-2}) \).
Error as a function of runtime

Algorithm

\[ \text{MSE} = \epsilon^2 \]

\[ \text{Cost} \propto \sum_{l=1}^{L} N_l h_l^{-1} \]

MLMC BIP SMC MLSMC MLPF Summary

mlsmc smc
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
Change of setup: filtering

- Filtering involves a sequence of Bayesian inversions separated by propagation in time (in this case through an SDE).
- Let \( \hat{\eta}_{0,m}, \ldots, \hat{\eta}_{L,m}, \ldots, \hat{\eta}_{\infty,m} \) denote the time \( m \) filtering distribution at a hierarchy of levels (time discretization).
- Mutation \( M^\ell \) is now coupled propagation of a pair of initial conditions through an SDE discretized at two successive mesh-refinements, for \( \ell = 0, \ldots, L \).
- Likelihood for \( m^{th} \) observation is denoted \( G_m \).
- Selection is performed by novel pairwise coupled resampling which preserves marginals.
- MLMC results carry over with somewhat weaker rate.
Particle filter

\( Q^L(x, \cdot) \) : kernel associated to Euler-Marayuma discretization of SDE at level \( L \) with initial condition \( x \).

Generate \( \hat{\eta}^N_L = \sum_{i=1}^{N_L} \delta \hat{U}_m^L i \sim \hat{\eta}_L, m \) using

**Particle filter algorithm:**

For \( i = 1, \ldots, N_L \), draw \( \hat{U}_0^L i \sim \mu_0 \).

Initialize \( m = 1 \). Do

(i) For \( i = 1, \ldots, N_L \), draw \( U_m^L i \sim Q^L(\hat{U}_{m-1}^L, \cdot) \);

(ii) For \( k = 1, \ldots, N_L \), draw \( I_{m}^{L,k} \) according to multinomial distribution \( \{ w_{m}^i \}_{i=1}^{N_L} \), where

\[
    w_{m}^i := G_m(U_m^L i) / \sum_{j=1}^{N_L} G_m(U_m^L j).
\]

(iii) \( \hat{U}_m^{L,k} \leftarrow U_m^L I_{m}^{L,k} \).

\( m \leftarrow m + 1 \)
Multilevel particle filter

\( M^\ell([x, y], \cdot) \): coupled kernel with marginals \( Q^\ell(x, \cdot) \) and \( Q^{\ell-1}(y, \cdot) \).

Generate \( \hat{\eta}^{ML}_{L,m} = \sum_{\ell=0}^{L} \sum_{i=1}^{N_\ell} (\delta_{\hat{U}_{m,1}^{\ell,i}} - \delta_{\hat{U}_{m,2}^{\ell,i}}) \approx \hat{\eta}_{L,m} \), with

\( \delta_{\hat{U}_{m,2}^{0,i}} := 0 \), using

**Multilevel particle filter algorithm:**

For \( \ell = 0, 1, \ldots, L \) and \( i = 1, \ldots, N_\ell \), draw \( \hat{U}_{0,1}^{\ell,i} \sim \mu_0 \), and let \( \hat{U}_{0,2}^{\ell,i} = \hat{U}_{0,1}^{\ell,i} \).

Initialize \( m = 1 \). Do

(i) For \( \ell = 0, 1, \ldots, L \) and \( i = 1, \ldots, N_\ell \), draw

\( (U_{m,1}^{\ell,i}, U_{m,2}^{\ell,i}) \sim M^\ell((\hat{U}_{m-1,1}^{\ell,i}, \hat{U}_{m-1,2}^{\ell,i}), \cdot) \);

(ii) For \( \ell = 0, 1, \ldots, L \) and \( k = 1, \ldots, N_\ell \), draw \( (I_{m,1}^{\ell,k}, I_{m,2}^{\ell,k}) \) according to the coupled resampling procedure;

(iii) \( (\hat{U}_{m,1}^{\ell,k}, \hat{U}_{m,2}^{\ell,k}) \leftarrow (U_{m,1}^{\ell,k}, U_{m,2}^{\ell,k}) \) for \( k = 1, \ldots, N_\ell \).

\( m \leftarrow m + 1 \)
Given \( \{\{U_{m,1}^{\ell,i}, U_{m,2}^{\ell,i}\}_{i=1}^{N_\ell}\}_{\ell=0}^{L} \),

For \( \ell = 0, 1, \ldots, L \) define

\[
W_{m,1}^{\ell,i} = \frac{G_m(U_{m,1}^{\ell,i})}{\sum_{j=1}^{N_\ell} G_m(U_{m,1}^{\ell,j})} \quad \text{and} \quad W_{m,2}^{\ell,i} = \frac{G_m(U_{m,2}^{\ell,i})}{\sum_{j=1}^{N_\ell} G_m(U_{m,2}^{\ell,j})}.
\]
Coupled resampling procedure:

a. with probability \( \alpha_m^\ell = \sum_{i=1}^{N_\ell} w_{m,1}^{\ell,i} \land w_{m,2}^{\ell,i} \), draw \( I_{m,1}^{\ell,k} \) according to

\[
P(I_{m,1}^{\ell} = i) = \frac{w_{m,1}^{\ell,i} \land w_{m,2}^{\ell,i}}{\sum_{j=1}^{N_\ell} w_{m,1}^{\ell,j} \land w_{m,2}^{\ell,j}}, \quad i = 1, \ldots, N_\ell.
\]

and let \( I_{m,2}^{\ell,k} = I_{m,1}^{\ell,k} \).

b. with probability \( 1 - \alpha_m^\ell \), draw \((I_{m,1}^{\ell,k}, I_{m,2}^{\ell,k})\) independently according to the probabilities

\[
P(I_{m,1}^{\ell} = i) = \frac{w_{m,1}^{\ell,i} - w_{m,1}^{\ell,i} \land w_{m,2}^{\ell,i}}{\sum_{j=1}^{N_\ell} w_{m,1}^{\ell,j} - w_{m,1}^{\ell,j} \land w_{m,2}^{\ell,j}}; \quad i = 1, \ldots, N_\ell.
\]

\[
P(I_{m,2}^{\ell} = i) = \frac{w_{m,2}^{\ell,i} - w_{m,1}^{\ell,i} \land w_{m,2}^{\ell,i}}{\sum_{j=1}^{N_\ell} w_{m,2}^{\ell,j} - w_{m,1}^{\ell,j} \land w_{m,2}^{\ell,j}}, \quad i = 1, \ldots, N_\ell.
\]
Assuming 1-step strong error order $\beta$, weak error order $\alpha$, and cost $\zeta$ (for Euler-Marayuma $\alpha = \beta = \zeta = 1$), the following theorem holds:

**Theorem (JKLZ15)**

Under suitable regularity assumptions on $M^\ell$ and $G$, for any $\varphi \in \mathcal{B}_b(\mathbb{R}^d) \cap \text{Lip}(\mathbb{R}^d)$

$$
\mathbb{E} \left[ \left\{ \hat{\eta}^{ML}_m(\varphi) - \hat{\eta}^L_m(\varphi) \right\}^2 \right] \lesssim \sum_{\ell=1}^L \frac{h^{\beta/2}_\ell}{N_\ell}
$$

In particular, for $\beta/2 > \zeta$, $L$ and $\{N_\ell\}_{\ell=0}^L$ can be chosen such that $\text{MSE} = \mathcal{O}(\varepsilon^2)$ for computational cost $= \mathcal{O}(\varepsilon^{-2})$.

Note that the coupled resampling effectively reduces rate $\beta \to \beta/2$. 
Numerical examples

- \( dX_t = a(X_t)dt + b(X_t)dW_t, \quad X_0 = x_0, \)
  with \( X_t \in \mathbb{R}^d, \ t \geq 0 \) and \( \{W_t\}_{t \in [0,t]} \) a Brownian motion of appropriate dimension.

- Partial observations \( \{y_1, \ldots, y_n\} \) available and \( Y_k|X_k \) has a density function \( G(y_k, x_k) =: G_k(x_k) \).

- Euler Marayuma discretization with \( h_\ell = 2^{-\ell} \). For constant diffusion \( \beta = 2 \), non-constant \( \beta = 1 \).

<table>
<thead>
<tr>
<th>Example</th>
<th>( a(x) )</th>
<th>( b(x) )</th>
<th>( G(y; x) )</th>
<th>( \varphi(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>OU</td>
<td>( \theta(\mu - x) )</td>
<td>( \sigma )</td>
<td>( \mathcal{N}(x, \tau^2) )</td>
<td>( x )</td>
</tr>
<tr>
<td>GBM</td>
<td>( \mu x )</td>
<td>( \sigma x )</td>
<td>( \mathcal{N}(\log x, \tau^2) )</td>
<td>( x )</td>
</tr>
<tr>
<td>Langevin</td>
<td>( \frac{1}{2} \nabla \log \pi(x) )</td>
<td>( \sigma )</td>
<td>( \mathcal{N}(0, \tau^2 e^x) )</td>
<td>( \tau^2 e^x )</td>
</tr>
<tr>
<td>NLM</td>
<td>( \theta(\mu - x) )</td>
<td>( \frac{\sigma}{\sqrt{1+x^2}} )</td>
<td>( \mathcal{L}(x, s) )</td>
<td>( x )</td>
</tr>
</tbody>
</table>
Numerical examples: rates
Numerical examples: cost

![Graph showing cost and error with different algorithms](image-url)
Outline

1. Multilevel Monte Carlo Sampling
2. Bayesian inference problem
3. Sequential Monte Carlo Samplers
4. Multilevel Sequential Monte Carlo Samplers
5. Multilevel Particle Filter
6. Summary
Summary

- Multilevel Sequential Monte Carlo sampler (MLSMC), and particle filter (MLPF) can perform asymptotically as well as MLMC.
- Cost-to-$\epsilon$ can be asymptotically the same as for a scalar random variable!
- MLPF strong error is effectively reduced by coupled resampling $\beta \rightarrow \beta/2$.
- For MLSMC example $\beta > \zeta$. If $\beta \leq \zeta$, cost is somewhat higher, analogous to standard MLMC.
- For MLPF examples $\beta/2 \leq \zeta$. For $\beta/2 > \zeta$ optimal results can be obtained again.
- If $\zeta > 2\alpha$ then the optimal cost is $\epsilon^{-\zeta/\alpha}$, the cost of a single simulation at the finest level.
- Many improvements and more complicated models are forthcoming.
References

Thank you!
New book!