# Multi-index Sequential Monte Carlo ratio estimators for Bayesian inverse problems

**Shangda Yang**
Kody J.H. Law, Neil Walton and Ajay Jasra

ESI Workshop 1

May 5, 2022

MANCHESTER 1824
The University of Manchester

# Aim

We are aiming to find

$$\pi_\infty(\varphi(x)) := \mathbb{E}_{\pi_\infty}[\varphi(x)],$$

where

- $x \sim \pi_\infty$
- $\varphi(x) : \mathbb{R} \to \mathbb{R}$.

However, we can only **approximate** $\pi_\infty$ by $\pi_\alpha$ with $\alpha \in \mathbb{Z}_+^D$ being the index of refinement.

## Bayesian Inference

Under Bayesian context, the target distribution can only be evaluated up to normalising constant

$$\pi_\alpha(dx) := \pi_\alpha(dx|y) \propto L_\alpha(x)\pi_0(dx),$$

where

- $x \in \mathbb{R}^{d_x}$ is the unknown parameter
- $y \in \mathbb{R}^{d_y}$ is the observations
- $\pi_0(dx)$ is the prior distribution
- $L_\alpha(dx) \propto \pi_\alpha(y|x)$ is the approximated likelihood.

# Notations

| **Exact** | **Approximate** |
|---|---|
| $L(x) \propto \pi(y\|x)$ | $L_\alpha(x) \propto \pi_\alpha(y\|x)$ |
| $f(dx) := L(x)\pi_0(dx)$ | $f_\alpha(dx) := L_\alpha(x)\pi_0(dx)$ |
| $Z := \int_X f(x)dx$ | $Z_\alpha := \int_X f_\alpha(x)dx$ |
| $\pi := \pi(dx\|y) = f(dx)/Z$ | $\pi_\alpha := \pi_\alpha(dx\|y) = f_\alpha(dx)/Z_\alpha$ |
| $\pi(\varphi) := \mathbb{E}_\pi[\varphi(x)]$ | $\pi_\alpha(\varphi_\alpha) := \mathbb{E}_{\pi_\alpha}[\varphi_\alpha(x)]$ |

## Monte Carlo

Assuming can simulate samples from $\pi_\alpha(dx|y)$, we can construct the single level Monte Carlo estimator for $\mathbb{E}_{\pi_\alpha}[\varphi(x)]$ as

$$N^{-1} \sum_{i=1}^{N} \varphi_\alpha(x^i),$$

where $N$ is the number of samples.

Complexity: $\text{MSE}^{-\xi}$ ($\varepsilon^{-2\xi}$) for $\xi > 1$ (e.g. SDE/PDE)

Canonical: $\text{MSE}^{-1}$ if no discretisation bias (e.g. integral of an arbitrary function on [0,1])

# Multilevel Monte Carlo (MLMC)

With **MLMC**, we apply the approximation

$$\pi(\varphi) \approx \sum_{\ell=0}^{L} \mathbb{E}[\varphi_\ell(x) - \varphi_{\ell-1}(x)], \tag{1}$$

where $L > 0$ is the finest level of discretisation and $\varphi_\ell \equiv 0$ if $\ell < 0$.

# Multi-index Monte Carlo (MIMC)

With **MIMC**, we apply the approximation

$$\pi(\varphi) \approx \sum_{\alpha \in \mathcal{I}} \mathbb{E}[\Delta\varphi_\alpha(x)], \tag{2}$$

where

- $\mathcal{I} \subset \mathbb{Z}_+^{\mathbb{D}}$
- the difference of difference operator is defined as $\Delta\varphi_\alpha := \Delta_D \circ \cdots \circ \Delta_1 \varphi_\alpha$ with $\Delta_i \varphi_\alpha := \varphi_\alpha - \varphi_{\alpha - e_i}$
- $\varphi_\alpha \equiv 0$ if $\alpha_i < 0$ for any $i$.

E.g. $\mathcal{D} = 2$:

$$\Delta\varphi_\alpha = \varphi(u_\alpha(x)) - \varphi(u_{\alpha_1,\alpha_2-1}(x)) - \varphi(u_{\alpha_1-1,\alpha_2}(x)) + \varphi(u_{\alpha_1-1,\alpha_2-1}(x))$$

# MIMC

Define the vector of multi-indices:

$$\boldsymbol{\alpha}(\alpha) := (\boldsymbol{\alpha}_1(\alpha), \dots, \boldsymbol{\alpha}_{2^D}(\alpha)) \in \mathbb{Z}_+^{D \times 2^D},$$

where

- $\boldsymbol{\alpha}_1(\alpha) = \alpha$
- $\boldsymbol{\alpha}_{2^D}(\alpha) = \alpha - \sum_{i=1}^{D} e_i$
- $e_i$ is the $i^{\text{th}}$ standard basis vector in $\mathbb{R}^D$.

# MIMC

Two cases:

- can sample from a coupling of $(\pi_{\boldsymbol{\alpha}_1(\alpha)}, \ldots, \pi_{\boldsymbol{\alpha}_{2D}(\alpha)})$ (a joint distribution with the marginal $\pi_{\boldsymbol{\alpha}_j(\alpha)}$) $\rightarrow$ construct a standard multi-index estimator [5]
- cannot $\rightarrow$ approximate coupling and construct different estimators [7, 13]

# Approximate Coupling

Let

- $\Pi_0(d\boldsymbol{x}) = \pi_0(d\boldsymbol{x}_1)\prod_{i=2}^{2^D}\delta_{\boldsymbol{x}_1}(d\boldsymbol{x}_i)$ ($\delta_{\boldsymbol{x}_1}$ the Dirac delta function)
- $\mathbf{L}_\alpha(\boldsymbol{x}) = \max\{L_{\boldsymbol{\alpha}_1(\alpha)}(\boldsymbol{x}_1), \ldots, L_{\boldsymbol{\alpha}_{2^D}(\alpha)}(\boldsymbol{x}_{2^D})\}$.

Approximate coupling:

$$F_\alpha(d\boldsymbol{x}) = \mathbf{L}_\alpha(\boldsymbol{x})\Pi_0(d\boldsymbol{x}), \qquad \Pi_\alpha(d\boldsymbol{x}) = \frac{1}{F_\alpha(1)}F_\alpha(d\boldsymbol{x}). \tag{3}$$

# Markov Chain Monte Carlo (MCMC)

MCMC:

- Construct an ergodic (aperiodic and positive recurrent) Markov chain such that it is $\pi-$invariant
- Simulate a trajectory and construct an estimator

Metropolis-Hastings method:

- Design a proposal Markov kernel $\mathcal{M}$
- Sample $x' \sim \mathcal{M}(x^{(i)}, \cdot)$
- Let $x^{(i+1)} = x'$ with probability $\min\left\{1, \frac{f(x')\mathcal{M}(x', x^{(i)})}{f(x^{(i)})\mathcal{M}(x^{(i)}, x')}\right\}$
- Let $x^{(i+1)} = x^{(i)}$, otherwise

# Sequential Monte Carlo Sampler (SMC sampler)

Sequential Monte Carlo sampler is the combination of (sequential) importance (re)sampling and Markov chain Monte Carlo.

- **Sample** from the target $\pi$ by sampling from a simple distribution (importance sampling)
- **Mutate/Propagate** by Markov kernels to decorrelate the samples, along a sequence of distributions converging to the target (sequential importance sampling and Markov chain Monte Carlo)
- **Resample** when the effective sample size of the samples drops down to certain values (Resampling)

## Sequential Monte Carlo Sampler (SMC sampler)

Let for $j = 1, ..., J$

- $h_j = f_{j+1}/f_j$ where $f_1 = \pi_0$, $f_J = f$ and $f_j$ for $j = 2, ..., J-1$ are some intermediate distributions (e.g. $f_j(x) = L(x)^{\tau_j}\pi_0(x)$, where $\tau_1 = 0$, $\tau_j < \tau_{j+1}$, and $\tau_J = 1$)
- $\pi_j(x) = \frac{1}{f_j(1)} f_j(x)$
- $\mathcal{M}_i$ be a Markov transition kernel such that $(\pi_{\alpha,j}\mathcal{M}_j)(dx) = \pi_j(dx)$

# Sequential Monte Carlo Sampler (SMC sampler)

**Algorithm 1** SMC sampler

Let $x_1^{(i)} \sim \pi_1$ for $i = 1, \ldots, N$, and $Z_1^N = 1$. For $j = 2, \ldots, J$, repeat the following steps:

(0) Store $Z_j^N = Z_{j-1}^N \frac{1}{N} \sum_{k=1}^N h_{j-1}({}_{j-1}^{(k)})$.

For $i = 1, ..., N$:

(i) Define $w_j^i = h_{j-1}({}_{j-1}^{(i)}) / \sum_{k=1}^N h_{j-1}({}_{j-1}^{(k)})$.

(ii) Resample. Select $I_j^i \sim \{w_j^1, \ldots, w_j^N\}$, and let $\hat{x}_j^{(i)} = x_{j-1}^{(I_j^i)}$.

(iii) Mutate. Draw $x_j^{(i)} \sim \mathcal{M}_j(\hat{x}_j^{(i)}, \cdot)$.

1. Introduction
2. MLMC and MIMC
3. Samplers
4. MISMC
5. Numerical Results

# MISMC

**Multi-index sequential Monte Carlo (MISMC):**

**MIMC + SMC sampler + Estimators**

**Estimators** $\begin{cases} \text{self-normalised (SN) increments estimator} \\ \text{ratio estimator} \end{cases}$

# Self-normalised increments estimator

Self-normalised increments estimator (lack of rigorous convergence results for realistic assumptions):

$$\pi(\varphi) \approx \sum_{\alpha \in \mathcal{I}} \Delta(\pi_\alpha(\varphi_\alpha)) = \sum_{\alpha \in \mathcal{I}} \Delta\left(\frac{1}{Z_\alpha} f_\alpha(\varphi_\alpha)\right) \tag{4}$$

- Sample from the approximate coupling (3) using SMC sampler
- Construct the self-normalized importance sampling estimators

$$\Delta\left(\frac{\sum_{i=1}^N \varphi_\alpha(x^{(i)}) \frac{\pi_\alpha(x^{(i)})}{\Pi_\alpha(x^{(i)})}}{\sum_{i=1}^N \frac{\pi_\alpha(x^{(i)})}{\Pi_\alpha(x^{(i)})}}\right), \quad x^{(i)} \sim \Pi_\alpha$$

for each individual summands of (4)

# Self-normalised increments estimator assumption

(A2) For every $n \geq 0$, $\varphi : \mathbb{N}_0^d \times \mathsf{X}^{n+1} \times \Theta \to \mathbb{R}$ bounded, every $\alpha \in \mathcal{I}$, there exist a $C(\alpha(1:k_\alpha))$, with $\lim_{\min_{1\leq i\leq d}\alpha_i \to +\infty} C(\alpha(1:k_\alpha)) = 0$, such that for any collection of scalar, bounded random variables $\beta(\alpha(1:k_\alpha), 2i, 2i-1)$, $i \in \{1, \dots, k'_\alpha\}$ we have almost surely

$$\sup_{(x_{0:n}(1:k_\alpha),\theta)\in(\bigotimes_{i=1}^{k_\alpha} \mathsf{X}_{\alpha(i)}^{n+1})\times\Theta} \Big|\Big\{\sum_{i=1}^{k'_\alpha} \tau_{i,\alpha}\beta(\alpha(1:k_\alpha), 2i, 2i-1)\Big[\varphi_{\alpha(2i)}(x_{0:n}(1:k_\alpha),\theta) - \varphi_{\alpha(2i-1)}(x_{0:n}(1:k_\alpha),\theta)\Big]\Big\}\Big| \leq C(\alpha(1:k_\alpha)) \sum_{i=1}^{k'_\alpha} |\beta(\alpha(1:k_\alpha), 2i, 2i-1)|^2.$$

# Ratio estimator

Ratio estimator (by ratio decomposition)

$$\pi(\varphi) = \frac{f(\varphi)}{f(1)} \approx \frac{\sum_{\alpha \in \mathcal{I}} \Delta(f_\alpha(\varphi))}{\sum_{\alpha \in \mathcal{I}} \Delta(f_\alpha(1))} \tag{5}$$

- Sample from the approximate coupling (3) using SMC sampler
- Construct estimators for the unnormalised increments $\Delta(f_\alpha(\zeta_\alpha))$ for $\zeta_\alpha = \varphi_\alpha$ and $\zeta_\alpha = 1$

# Estimators for the unnormalised increments of increments

Define for $j = 1, \ldots, J$, and for random variables $\mathbf{x}_j^{(i)}$, $i = 1, \ldots, N$

- $\Pi_{\alpha,j}^N(d\mathbf{x}) := \frac{1}{N}\sum_{i=1}^N \delta_{\mathbf{x}_j^{(i)}}(d\mathbf{x})$
- $H_{\alpha,j} = F_{\alpha,j+1}/F_{\alpha,j}$
- $Z_\alpha^N := \prod_{j=1}^{J-1} \Pi_{\alpha,j}^N(H_{\alpha,j})$
- $F_\alpha^N(d\mathbf{x}) := Z_\alpha^N \Pi_{\alpha,J}^N(d\mathbf{x})$
- $\psi_{\zeta_\alpha}(\mathbf{x}) := \sum_{k=1}^{2^D} \iota_k \omega_k(\mathbf{x})\zeta_{\boldsymbol{\alpha}_k(\alpha)}(\mathbf{x}_k)\,, \qquad \omega_k(\mathbf{x}) := \frac{L_{k(\alpha)}(\mathbf{x}_k)}{\mathbf{L}_\alpha(\mathbf{x})}$,
  where $\iota_k \in \{-1, 1\}$ is the sign of the $k^{\text{th}}$ term in $\Delta f_\alpha$.

# Ratio Estimator

The ratio estimator is

$$\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}} = \frac{\sum_{\alpha\in\mathcal{I}} F_{\alpha}^{N_\alpha}(\psi_{\varphi_\alpha})}{\max\{\sum_{\alpha\in\mathcal{I}} F_{\alpha}^{N_\alpha}(\psi_1), Z_{\min}\}},$$

where

- $\mathcal{I} \subseteq \mathbb{Z}_+^D$ is the index set
- $\{N_\alpha\}_{\alpha\in\mathcal{I}}$ are the number of samples
- $\varphi : X \to \mathbb{R}$ is the quantity of interest
- $Z_{\min}$ is a lower bound on $Z$.

## Estimators for the unnormalised increments of increments

Note:

The estimator $F_\alpha^N(\psi_{\zeta_\alpha})$ is unbiased

$$\mathbb{E}[F_\alpha^N(\psi_{\zeta_\alpha})] = F_\alpha(\psi_{\zeta_\alpha}) = \Delta f_\alpha(\zeta_\alpha)$$

.

# Estimators for the unnormalised increments of increments

## Assumption 4.1

*Let $J \in \mathbb{N}$ be given. For each $j \in \{1, \dots, J\}$ there exists some $C > 0$ such that for all $(\alpha, x) \in \mathbb{Z}_+^D \times X$*

$$C^{-1} < Z, H_{\alpha,j}(x) \le C \,.$$

## Assumption 4.2

*For any $\zeta :\to$ bounded and Lipschitz, there exist $C, \beta_i, s_i, \gamma_i > 0$ for $i = 1, \dots, D$ such that for resolution vector $(2^{-\alpha_1}, \dots, 2^{-\alpha_D})$, i.e. resolution $2^{-\alpha_i}$ in the $i^{\text{th}}$ direction, the following holds*

(B) $|f_\alpha(\zeta) - f(\zeta)| =: B_\alpha \le C 2^{-\langle \alpha, s \rangle}$;

(V) $\int_X (\Delta(L_\alpha(x)\zeta_\alpha(x)))^2 \pi_0(dx) =: V_\alpha \le C 2^{-\langle \alpha, \beta \rangle}$;

(C) $\mathrm{COST}(F_\alpha(\psi_\varphi)) =: C_\alpha \propto 2^{\langle \alpha, \gamma \rangle}$.

# Multi-index sequential Monte Carlo ratio estimator

## Theorem 1

*Assume 4.1. Then for any $J \in \mathbb{N}$ there exists a $C > 0$ such that for any $N \in \mathbb{N}$, $\psi : X^{2^D} \to \mathbb{R}$ bounded and measurable and $\alpha \in \mathbb{Z}_+^D$*

$$\mathbb{E}\left[\left|F_\alpha^N(\psi) - F_\alpha(\psi)\right|^2\right] \leq \frac{C}{N} F_\alpha(\psi^2).$$

## Lemma 4.1

*Assume 4.1. Then there exist a $C > 0$ such that for any $\alpha \in \mathbb{Z}_+^D$*

$$F_\alpha(\psi_{\zeta_\alpha}^2) \leq C \int_X (\Delta(L_\alpha(x)\zeta_\alpha(x)))^2 \pi_0(dx)\,.$$

# MISMCRE with the tensor product index set

## Theorem 2

*Assume 4.1 and 4.2, with* $\sum_{j=1}^{D} \frac{\gamma_j}{s_j} \leq 2$ *and* $\beta_i > \gamma_i$ *for* $i = 1, \ldots, D$. *Then for any* $\varepsilon > 0$ *and suitable* $\varphi : X \to \mathbb{R}$, *it is possible to choose a tensor product (TP) index set* $\mathcal{I}_{L_1:L_d} := \{\alpha \in \mathbb{N}^D : \alpha_1 \in \{0, ..., L_1\}, ..., \alpha_D \in \{0, ..., L_D\}\}$ *and* $\{N_\alpha\}_{\alpha \in \mathcal{I}_{L_1:L_D}}$, *such that for some* $C > 0$

$$\mathbb{E}[(\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}} - \pi(\varphi))^2] \leq C\varepsilon^2 ,$$

*and* $\mathrm{COST}(\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}}) \leq C\varepsilon^{-2}$, *the canonical rate.*

# MISMCRE with the total degree index set

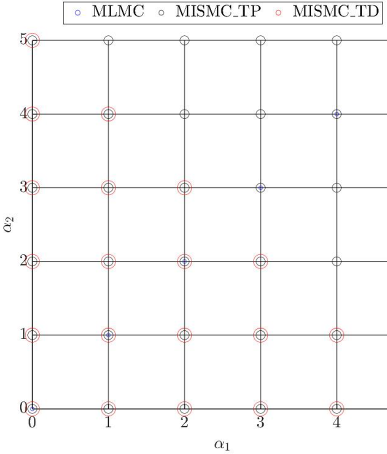
## Theorem 3

*Assume 4.1 and 4.2, with* $\beta_i > \gamma_i$ *for* $i = 1, \ldots, D$. *Then for any* $\varepsilon > 0$ *and suitable* $\varphi : X \to \mathbb{R}$, *it is possible to choose a total degree (TD) index set* $\mathcal{I}_L := \{\alpha \in \mathbb{N}^D : \sum_{i=1}^D \delta_i \alpha_i \leq L, \sum_{i=1}^D \delta_i = 1\}$, $\delta_i \in (0,1]$ *and* $\{N_\alpha\}_{\alpha \in \mathcal{I}_L}$, *such that for some* $C > 0$

$$\mathbb{E}[(\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}} - \pi(\varphi))^2] \leq C\varepsilon^2 ,$$

*and* $\mathrm{COST}(\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}}) \leq C\varepsilon^{-2}$, *the canonical rate.*

# Index sets



## Crucial point for theoretical results

### Lemma 4.2

*For the estimator* $\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}} = \frac{\sum_{\alpha\in\mathcal{I}} F_\alpha^{N_\alpha}(\psi_{\varphi_\alpha})}{\max\{\sum_{\alpha\in\mathcal{I}} F_\alpha^{N_\alpha}(\psi_1), Z_{\min}\}}$, *the following inequality holds*

$$\mathbb{E}[(\widehat{\varphi}_{\mathcal{I}}^{\mathrm{MI}} - \pi(\varphi))^2] \leq C \max_{\zeta\in\{\varphi,1\}} \left( \sum_{\alpha\in\mathcal{I}} \mathbb{E}\left[\left(F_\alpha^{N_\alpha}(\psi_{\zeta_\alpha}) - F_\alpha(\psi_{\zeta_\alpha})\right)^2\right] + \left(\sum_{\alpha\notin\mathcal{I}} F_\alpha(\psi_{\zeta_\alpha})\right)^2\right),$$

*for some* $C > 0$.

# Sub-canonical and Canonical

MISMC with TP index set:

- relies on the assumption that $\sum_{j=1}^{D} \frac{\gamma_j}{s_j} \leq 2$ (the samples at the finest index do not dominate the cost)
- if assumption is violated, only sub-canonical complexity
- sub-canonical rate may often be D-dependent (curse-of-dimensionality)

MISMC with TD index set:

- releases this constraint
- improves the computational complexity for many problems from sub-canonical to canonical

1. Introduction
2. MLMC and MIMC
3. Samplers
4. MISMC
5. Numerical Results

# Elliptic PDEs

We consider the following elliptic PDE,

$$-\nabla \cdot (a(x)\nabla u(x)) = \mathrm{f}, \qquad \text{on } \Omega \,, \tag{6}$$

$$u(x) = 0, \qquad \text{on } \partial\Omega \,, \tag{7}$$

where

- $\Omega \subset \mathbb{R}^D$ is a convex domain
- $\partial\Omega \in C^0$ is the boundary
- $f : \Omega \to \mathbb{R}$
- $a(x) : \Omega \to \mathbb{R}_+$, parameterised by a random variable $x \in X$.

We solve the PDE by the **finite element method**.

# Bayesian settings

- Define
$$\mathcal{G}(u(x)) = [v_1(u(x)), \ldots, v_m(u(x))]^\top,$$
where $v_i \in L^2$ and $v_i(u(x)) = \int v_i(z)u(x)(z)dz$ for $i = 1, \ldots, m$, for some $m \geq 1$.
- Assume the observations take the form
$$y = \mathcal{G}(u(x)) + \nu,$$
where $\nu \sim N(0, \Xi)$, and $x$ and $\nu$ are independent.
- Define
$$L(x) := \exp\left( -\frac{1}{2}|y - \mathcal{G}(u(x))|_\Xi^2 \right) \propto \pi(y|x),$$
where $|w|_\Xi := (w^\top \Xi^{-1} w)^{1/2}$

## Bayesian Settings

Denote the weak approximation of (6)-(7) at resolution multi-index $\alpha$ by $u_\alpha(x)$, the approximated likelihood is given by

$$L_\alpha(x) := \exp(-\frac{1}{2}|y - \mathcal{G}(u_\alpha(x))|_\Xi^2),$$

and the associated target is

$$\pi_\alpha(dx) \propto L_\alpha(x)\pi_0(dx)\,. \tag{8}$$

# 1D Toy Example

PDE:

- $\Omega = [0, 1]$
- $a(x) = 1$
- f $= x$, where $x$ is a random input with a uniform prior such that $x \sim U[-1, 1]$
- analytical solution $u(x) = -0.5x(z^2 - z)$

Observations:

- ten observations in the interval $(0, 1)$ with a step size $1/10$
- generated by $y = -0.5x^*(z^2 - z) + \nu$, where $y = [y_1, \ldots, y_{10}]$, $z = [z_1, \ldots, z_{10}]$, $x^* = 0.2581$ drawn from $U[-1, 1]$ and $\nu \sim N(0, 0.2^2)$.

# 1D Toy Example

Regularity:

- $s = 2$, $\beta = 4$, $\gamma = 1$

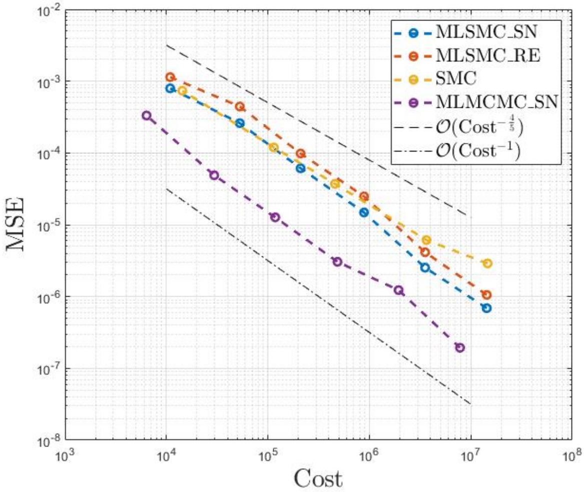
Quantity of interest:

- $\varphi(x) = x^2$

Complexity:

- Multilevel methods: $\text{MSE}^{-1}$
- Single level method: $\text{MSE}^{-4/5}$

# 1D Toy Example



# 2D Elliptic PDE with random diffusion coefficient

PDE:

- $\Omega = [0,1]^2$
- $a(x) = 3 + x_1 \cos(3z_1)\sin(3z_2) + x_2 \cos(z_1)\sin(z_2)$, where $x$ is a random input with a uniform prior such that $x \sim U[-1,1]^2$
- $f = 100$

Observations:

- four observations {(0.25,0.25), (0.25,0.75), (0.75,0.25), (0.75,0.75)}
- generated by $y = u_\alpha(x^*) + \nu$, where $u_\alpha(x^*)$ is the approximate solution of the PDE at $\alpha = [10, 10]$ with $x^* = [-0.4836, -0.5806]$ drawn from $U[-1,1]^2$, and $\nu \sim N(0, 0.5^2)$.

# 2D Elliptic PDE with random diffusion coefficient

Regularity:

- $s_1 = s_2 = 2$, $\beta_1 = \beta_2 = 4$, $\gamma_1 = \gamma_2 = 1$

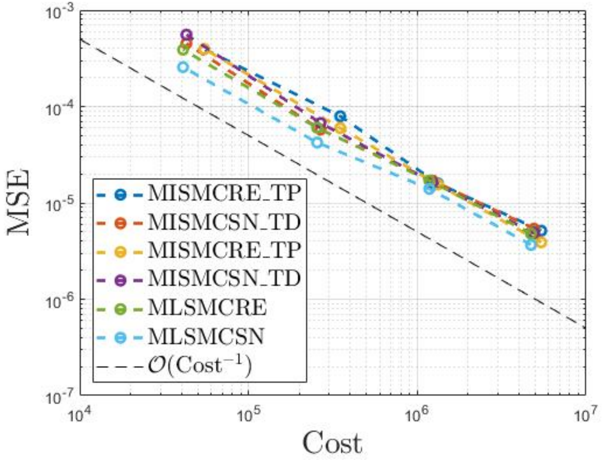
Quantity of interest:

- $\varphi(x) = x_1^2 + x_2^2$

Complexity:

- Multilevel and multi-index methods: $\mathrm{MSE}^{-1}$
- Single level method: $\mathrm{MSE}^{-2/3}$

## 2D Elliptic PDE with random diffusion coefficient



# Log Gaussian (Cox) Process models [6, 10][12]

Observations: the location of $n = 126$ Scots pine saplings in a natural forest in Finland, denoted $z_1, \ldots, z_n \in [0,1]^2$.

Process applied: $\Lambda = \exp(x)$ where $x$ is a Gaussian process (a priori), for $z \in [0,2]^2$,

$$x(z) = \theta_1 + \sum_{k \in \mathbb{Z} \times \mathbb{Z}_+ \cup \mathbb{Z}_+ \times 0} \zeta_k(\theta)(\xi_k \phi_k(z) + \xi_k^* \phi_{-k}(z)),$$

where

- $\xi_k \sim \mathcal{CN}(0,1)$ i.i.d. and $\xi_k^*$ is the complex conjugate of $\xi_k$
- $\mathcal{CN}(0,1)$ denotes a standard complex Normal distribution
- $\phi_k(z) \propto \exp[\pi i z \cdot k]$ are Fourier series basis functions
- $\zeta_k^2(\theta) = \theta_2/((\theta_3 + k_1^2)(\theta_3 + k_2^2))^{\frac{\beta+1}{2}}$.

## Likelihoods

$$
\text{(LGC)} \qquad \frac{d\pi}{d\pi_0}(x) \propto \exp\left[\sum_{j=1}^{n} x(z_j) - \int_{[0,1]^2} \exp(x(z))dz\right],
$$

$$
\text{(LGP)} \qquad \frac{d\pi}{d\pi_0}(x) \propto \exp\left[\sum_{j=1}^{n} x(z_j) - n \log \int_{[0,1]^2} \exp(x(z))dz\right].
$$

# Finite Approximation

Finite approximation of $x(z)$:

$$x_\alpha(z) = \theta_1 + \sum_{k \in \mathcal{A}_\alpha} \zeta_k(\theta)(\xi_k \phi_k(z) + \xi_k^* \phi_{-k}(z)), \quad \xi_k \sim \mathcal{CN}(0,1) \text{ i.i.d.},$$

where
$\mathcal{A}_\alpha := \{-2^{\alpha_1/2}, \dots, 2^{\alpha_1/2}\} \times \{1, \dots, 2^{\alpha_2/2}\} \cup \{1, \dots, 2^{\alpha_2/2}\} \times 0$
and can be approximated on a grid
$\{0, 2^{-\alpha_1}, \dots 1 - 2^{-\alpha_1}\} \times \{0, 2^{-\alpha_2}, \dots 1 - 2^{-\alpha_2}\}$ using the FFT
with a cost $\mathcal{O}((\alpha_1 + \alpha_2) 2^{\alpha_1 + \alpha_2})$.

# Finite Approximation

Finite approximation of the likelihood:

$$\text{(LGC)} \qquad \frac{d\pi_\alpha}{d\pi_{0,\alpha}}(x_\alpha) \propto \exp\left[\sum_{j=1}^{n} \hat{x}_\alpha(z_j) - Q(\exp(x_\alpha))\right],$$

$$\text{(LGP)} \qquad \frac{d\pi_\alpha}{d\pi_{0,\alpha}}(x_\alpha) \propto \exp\left[\sum_{j=1}^{n} \hat{x}_\alpha(z_j) - n\log Q(\exp(x_\alpha))\right],$$

where $\hat{x}_\alpha(z)$ is defined as an interpolant over the grid output from FFT and $Q$ denotes a quadrature rule.

# Log Gaussian (Cox) Process Models

Parameters for LGC:

- $\beta = 1.6$, $\theta = (\theta_1, \theta_2, \theta_3) = (0, 1, 110.339)$

Parameters for LGP:

- $\beta = 1.6$, $\theta = (\theta_1, \theta_2, \theta_3) = (0, 1, 27.585)$

Regularity:

- $s_1 = s_2 = 0.8$, $\beta_1 = \beta_2 = 1.6$, $\gamma_1 = \gamma_2 = 1 + \omega$ for $\omega > 0$

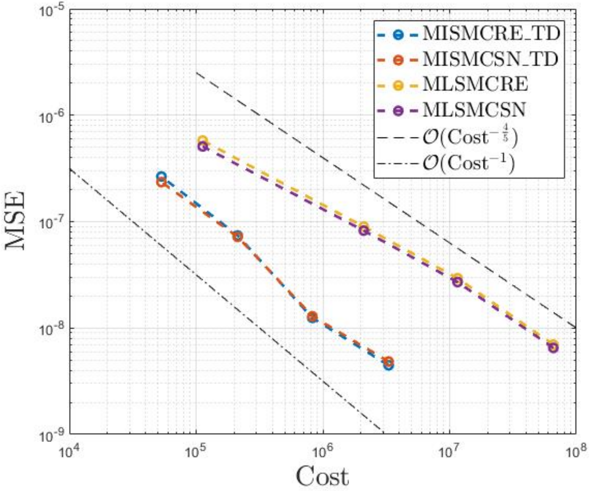
Quantity of interest:

- $\varphi(x) = \int_{[0,1]^2} \exp(x(z))dz$

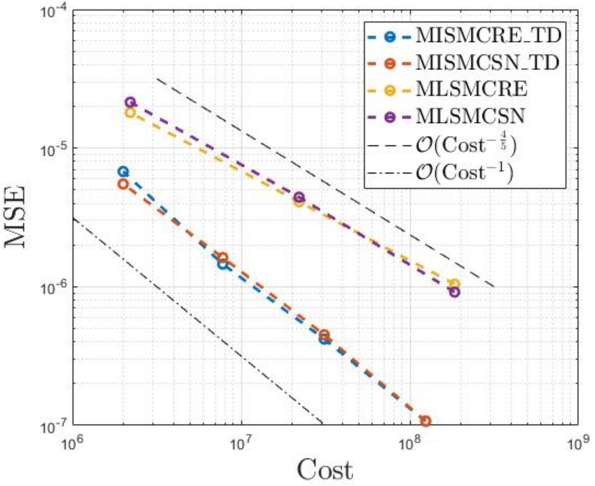
Complexity:

- Grid-based approach: approximately $\text{MSE}^{-19/4}$
- Single level method with circulant embedding: approximately $\text{MSE}^{-9/4}$
- MLSMC or MISMC with TP set: $\text{MSE}^{-5/4-\omega}$
- MISMC with TD set: $\text{MSE}^{-1}$

# Log Gaussian Cox Process Model



# Log Gaussian Process Model



## Summary

- MISMCRE has rigorous theoretical results under realistic assumptions due to the unbiased estimation of the unnormalised increments of increments.
- MISMCRE with TD set can improve the complexity of MLSMCRE and MISMCRE with TP set from subcanonical to canonical for some problems.

# References I

[1] Alexandros Beskos, Ajay Jasra, Kody J. H. Law, Raul Tempone, and Yan Zhou. "Multilevel sequential Monte Carlo samplers". In: *Stochastic Processes and their Applications* 127.5 (2017), pp. 1417–1440.

[2] T. Cui, Ajay Jasra, and Kody J. H. Law. "Multi-Index Sequential Monte Carlo methods". In: *Preprint* (2018).

[3] Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. "Sequential Monte Carlo samplers". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68.3 (2006), pp. 411–436.

[4] Michael B Giles. "Multilevel Monte Carlo methods". In: *Acta Numerica* 24 (2015), p. 259.

[5] Abdul-Lateef Haji-Ali, Fabio Nobile, and Raúl Tempone. "Multi-index Monte Carlo: when sparsity meets sampling". In: *Numerische Mathematik* 132.4 (2016), pp. 767–806.

# References II


[6] Jeremy Heng, Adrian N Bishop, George Deligiannidis, and Arnaud Doucet. "Controlled sequential monte carlo". In: *The Annals of Statistics* 48.5 (2020), pp. 2904–2929.

[7] Ajay Jasra, Kengo Kamatani, Kody J. H. Law, and Yan Zhou. "A multi-index Markov chain Monte Carlo method". In: *International Journal for Uncertainty Quantification* 8.1 (2018).

[8] Ajay Jasra, Kengo Kamatani, Kody J. H. Law, and Yan Zhou. "Bayesian static parameter estimation for partially observed diffusions via multilevel Monte Carlo". In: *SIAM Journal on Scientific Computing* 40.2 (2018), A887–A902.

[9] Ajay Jasra, Kengo Kamatani, Kody JH Law, and Yan Zhou. "Multilevel particle filters". In: *SIAM Journal on Numerical Analysis* 55.6 (2017), pp. 3068–3096.

# References III

[10] Jesper Møller, Anne Randi Syversveen, and Rasmus Plenge Waagepetersen. "Log gaussian cox processes". In: *Scandinavian journal of statistics* 25.3 (1998), pp. 451–482.

[11] Andrew M Stuart. "Inverse problems: a Bayesian perspective". In: *Acta numerica* 19 (2010), pp. 451–559.

[12] Surya T Tokdar and Jayanta K Ghosh. "Posterior consistency of logistic Gaussian process priors in density estimation". In: *Journal of statistical planning and inference* 137.1 (2007), pp. 34–42.

[13] Yaxian Xu, Ajay Jasra, and Kody JH Law. "Multi-Index Sequential Monte Carlo Methods for partially observed Stochastic Partial Differential Equations". In: *arXiv preprint arXiv:1805.00415* (2018).

# Thank You!

Paper: https://arxiv.org/abs/2203.05351

Codes: https://github.com/Shangda-Yang/MISMCRE.git