Hierarchical Monte Carlo Fusion

Ryan Chan

Murray Pollock [1], Gareth Roberts [1], Petros Dellaportas [2]

[1]University of Warwick

[2]University College London

8 March, 2019

Warwick Statistics

The Alan Turing Institute

Outline

- ‘Fork-and-Join’ Monte Carlo Fusion
  - Fusion Problem
  - Fusion Idea
  - An extended target density
  - Rejection sampling
  - Double Langevin Approach
- Example - Tempering
- ‘Hierarchical’ Monte Carlo Fusion

Fusion Problem

- Target of interest:

$$\pi(\boldsymbol{x}) \propto f_1(\boldsymbol{x}) \cdots f_C(\boldsymbol{x}) = \prod_{c=1}^{C} f_c(\boldsymbol{x})$$

  - C: number of cores / experts / ‘views’
  - f_c are *sub-posteriors*
- Applications:
  - Combining views of multiple experts on a topic
  - Combining views in a privacy setting
  - Big Data (by construction)
  - Tempering & sampling from multi-modal densities (by construction)

Fusion Problem

- Target of interest:

$$\pi(\boldsymbol{x}) \propto f_1(\boldsymbol{x}) \cdots f_C(\boldsymbol{x}) = \prod_{c=1}^{C} f_c(\boldsymbol{x})$$

  - C: number of cores / experts / ‘views’
  - f_c are *sub-posteriors*
- Applications:
  - Combining views of multiple experts on a topic
  - Combining views in a privacy setting
  - Big Data (by construction)
  - Tempering & sampling from multi-modal densities (by construction)

Fusion Idea

- Target of interest:

$$\pi(\boldsymbol{x}) \propto f_1(\boldsymbol{x}) \cdots f_C(\boldsymbol{x}) = \prod_{c=1}^{C} f_c(\boldsymbol{x})$$

  - Assume we can sample from $f_c(\boldsymbol{x})$
- Based on a simple (A)BC idea:
  1. Simulate $\boldsymbol{x}_1 \sim f_1, \boldsymbol{x}_2 \sim f_2, \ldots, \boldsymbol{x}_c \sim f_c$
  2. Accept if $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_c$, else go to 1
  3. Return $\boldsymbol{x} := \boldsymbol{x}_1 \sim \prod_{i=1}^{C} f_c \propto \pi$

Fusion Idea

- Target of interest:

$$\pi(\boldsymbol{x}) \propto f_1(\boldsymbol{x}) \cdots f_C(\boldsymbol{x}) = \prod_{c=1}^{C} f_c(\boldsymbol{x})$$

  - Assume we can sample from $f_c(\boldsymbol{x})$
- Based on a simple (A)BC idea:
  1. Simulate $\boldsymbol{x}_1 \sim f_1, \boldsymbol{x}_2 \sim f_2, \ldots, \boldsymbol{x}_c \sim f_c$
  2. Accept if $\boldsymbol{x}_1 = \boldsymbol{x}_2 = \cdots = \boldsymbol{x}_c$, else go to 1
  3. Return $\boldsymbol{x} := \boldsymbol{x}_1 \sim \prod_{i=1}^{C} f_c \propto \pi$

Langevin Diffusion

- Consider,

$$\mathrm{d}\boldsymbol{X}_t = \frac{1}{2}\nabla \log \pi(\boldsymbol{X}_t)\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{X}_0 = \boldsymbol{x} \in \mathbb{R}^d$$

  - Invariant distribution π
  - If $\boldsymbol{X}_0 \sim \pi$, then for all $t > 0$, $\boldsymbol{X}_t \sim \pi$

Langevin Diffusion

- Consider,

$$\mathrm{d}\boldsymbol{X}_t = \frac{1}{2}\nabla \log \pi(\boldsymbol{X}_t)\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{X}_0 = \boldsymbol{x} \in \mathbb{R}^d$$

  - Invariant distribution π
  - If $\boldsymbol{X}_0 \sim \pi$, then for all $t > 0$, $\boldsymbol{X}_t \sim \pi$

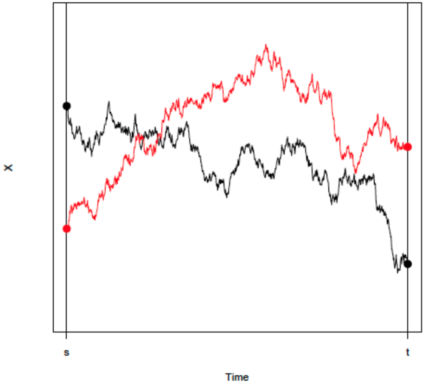
Langevin Diffusion

- Consider,

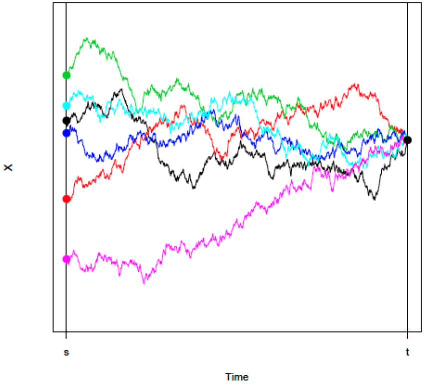
$$\mathrm{d}\boldsymbol{X}_t = \frac{1}{2}\nabla \log \pi(\boldsymbol{X}_t)\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{X}_0 = \boldsymbol{x} \in \mathbb{R}^d$$

  - Invariant distribution π
  - If $\boldsymbol{X}_0 \sim \pi$, then for all $t > 0$, $\boldsymbol{X}_t \sim \pi$

Fusion Idea



Fusion Idea with Langevin Diffusion



An Extended Target

Proposition

Suppose that $p_c(\boldsymbol{y} \mid \boldsymbol{x}^{(c)})$ is the transition density of a stochastic process with stationary distribution $f_c^2(\boldsymbol{x})$. The $(C+1)d$-dimensional (fusion) density proportional to the integrable function

$$g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c^2(\boldsymbol{x}^{(c)}) p_c(\boldsymbol{y} \mid \boldsymbol{x}^{(c)}) \cdot \frac{1}{f_c(\boldsymbol{y})} \right]$$

admits the marginal density π for $\boldsymbol{y}$.

An Extended Target

- Idea: construct a rejection sampler for g
- If we can sample from g, then we can obtain a draw from the fusion target density ($\boldsymbol{y} \sim \pi$)

Rejection Sampling

- Let $p_c(\boldsymbol{y} \mid \boldsymbol{x}) := p_{T,c}(\boldsymbol{y} \mid \boldsymbol{x})$, the transition density of the d-dimensional (double) Langevin (DL) diffusion processes $\boldsymbol{X}_t^{(c)}$ for $c = 1, \ldots, C$, from $\boldsymbol{x}$ to $\boldsymbol{y}$ for a pre-defined time $T > 0$ given by

$$\mathrm{d}\boldsymbol{X}_t^{(c)} = \nabla \log f_c(\boldsymbol{X}_t^{(c)})\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t^c,$$

  - $\boldsymbol{W}_t^{(c)}$ is d-dimensional Brownian motion
  - ∇ is the gradient operator over $\boldsymbol{x}$

Rejection Sampling

- Extended Target Density:

$$g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c^2(\boldsymbol{x}^{(c)}) p_{T,c}(\boldsymbol{y} \mid \boldsymbol{x}^{(c)}) \cdot \frac{1}{f_c(\boldsymbol{y})} \right]$$

- Consider the proposal density h for the extended target g:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right) \right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T} \right)$$

  - $\bar{\boldsymbol{x}} = \frac{1}{C} \sum_{c=1}^{C} \boldsymbol{x}^{(c)}$; T is an arbitrary positive constant

Warwick Statistics
The Alan Turing Institute

Rejection Sampling

- Extended Target Density:

$$g(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c^2(\boldsymbol{x}^{(c)}) p_{T,c}(\boldsymbol{y} \mid \boldsymbol{x}^{(c)}) \cdot \frac{1}{f_c(\boldsymbol{y})} \right]$$

- Consider the proposal density h for the extended target g:

$$h(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right) \right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T} \right)$$

  - $\bar{\boldsymbol{x}} = \frac{1}{C} \sum_{c=1}^{C} \boldsymbol{x}^{(c)}$; T is an arbitrary positive constant

Rejection Sampling

- Simulation from h is easy:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right)\right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T}\right)$$

1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ independently
2. Simulate $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T\mathbb{I}_d}{C})$

Rejection Sampling

- Simulation from h is easy:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right) \right] \cdot \exp\left(-\frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T} \right)$$

1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ independently
2. Simulate $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T\mathbb{I}_d}{C})$

Rejection Sampling

- Simulation from h is easy:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right)\right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T}\right)$$

1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ independently
2. Simulate $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T\mathbb{I}_d}{C})$

Rejection Sampling - acceptance probability

- Rejection sampling:

$$\frac{g(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(C)}, \mathbf{y})}{h(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(C)}, \mathbf{y})} \propto \rho \cdot Q$$

$$\begin{cases} \rho := \rho(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(C)}) = e^{-\frac{C\sigma^2}{2T}} \\ \sigma^2 = \frac{1}{C}\sum_{c=1}^{C} \left\| \mathbf{x}^{(c)} - \bar{\mathbf{x}} \right\|^2 \end{cases}$$

$$\begin{cases} Q = \mathbb{E}_{\bar{\mathbb{W}}}(E) \\ E := \prod_{c=1}^{C} \left[\exp\left\{ -\int_0^T \left(\phi_c(\mathbf{x}_t^{(c)}) - \Phi_c \right) \mathrm{d}t \right\} \right] \end{cases}$$

where $\bar{\mathbb{W}}$ denotes the law of C independent Brownian bridges $\mathbf{x}_t^{(1)}, \ldots, \mathbf{x}_t^{(C)}$ with $\mathbf{x}_0 = \mathbf{x}^{(c)}$ and $\mathbf{x}_T^{(c)} = \mathbf{y}$

Rejection Sampling - acceptance probability

- Rejection sampling:

$$\frac{g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

$$\begin{cases} \rho := \rho(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}) = e^{-\frac{C\sigma^2}{2T}} \\ \sigma^2 = \frac{1}{C} \sum_{c=1}^{C} \left\| \boldsymbol{x}^{(c)} - \bar{\boldsymbol{x}} \right\|^2 \end{cases}$$

$$\begin{cases} Q = \mathbb{E}_{\bar{\mathbb{W}}}(E) \\ E := \prod_{c=1}^{C} \left[\exp\left\{ -\int_0^T \left(\phi_c(\boldsymbol{x}_t^{(c)}) - \Phi_c \right) \mathrm{d}t \right\} \right] \end{cases}$$

where $\bar{\mathbb{W}}$ denotes the law of C independent Brownian bridges $\boldsymbol{x}_t^{(1)}, \ldots, \boldsymbol{x}_t^{(C)}$ with $\boldsymbol{x}_0 = \boldsymbol{x}^{(c)}$ and $\boldsymbol{x}_T^{(c)} = \boldsymbol{y}$

Rejection Sampling - acceptance probability

- Rejection sampling:

$$\frac{g(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

$$\begin{cases} \rho := \rho(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}) = e^{-\frac{C\sigma^2}{2T}} \\ \sigma^2 = \frac{1}{C} \sum_{c=1}^{C} \left\| \boldsymbol{x}^{(c)} - \bar{\boldsymbol{x}} \right\|^2 \end{cases}$$

$$\begin{cases} Q = \mathbb{E}_{\bar{\mathbb{W}}}(E) \\ E := \prod_{c=1}^{C} \left[\exp\left\{ -\int_0^T \left(\phi_c(\boldsymbol{x}_t^{(c)}) - \Phi_c \right) \mathrm{d}t \right\} \right] \end{cases}$$

where $\bar{\mathbb{W}}$ denotes the law of C independent Brownian bridges $\boldsymbol{x}_t^{(1)}, \dots, \boldsymbol{x}_t^{(C)}$ with $\boldsymbol{x}_0 = \boldsymbol{x}^{(c)}$ and $\boldsymbol{x}_T^{(c)} = \boldsymbol{y}$

Double Langevin Approach - Summary

- Proposal:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right) \right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T} \right)$$

- Accept $\boldsymbol{y}$ as a draw from fusion density π with probability:

$$\frac{g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

- Monte Carlo Fusion Algorithm:
  1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ and $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T \mathbb{I}_d}{C})$
  2. Accept $\boldsymbol{y}$ with probability $\rho \cdot Q$

Warwick Statistics The Alan Turing Institute

Double Langevin Approach - Summary

- Proposal:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right)\right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T}\right)$$

- Accept $\boldsymbol{y}$ as a draw from fusion density π with probability:

$$\frac{g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

- Monte Carlo Fusion Algorithm:
  1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ and $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T \mathbb{I}_d}{C})$
  2. Accept $\boldsymbol{y}$ with probability $\rho \cdot Q$

Warwick Statistics The Alan Turing Institute

Double Langevin Approach - Summary

- Proposal:

$$h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y}) \propto \prod_{c=1}^{C} \left[f_c\left(\boldsymbol{x}^{(c)}\right)\right] \cdot \exp\left(- \frac{C \cdot \|\boldsymbol{y} - \bar{\boldsymbol{x}}\|^2}{2T} \right)$$

- Accept $\boldsymbol{y}$ as a draw from fusion density π with probability:

$$\frac{g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

- Monte Carlo Fusion Algorithm:
  1. Simulate $\boldsymbol{x}^{(c)} \sim f_c(\boldsymbol{x})$ and $\boldsymbol{y} \sim \mathrm{N}(\bar{\boldsymbol{x}}, \frac{T\mathbb{I}_d}{C})$
  2. Accept $\boldsymbol{y}$ with probability $\rho \cdot Q$

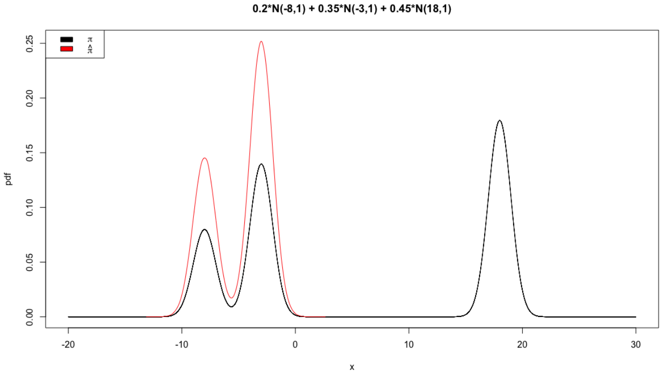
Example - Tempering

- Consider the target $\pi(\boldsymbol{x})$ is mulit-modal
- Using MCMC can be computationally expensive

Example - Tempering

- Consider the target $\pi(\boldsymbol{x})$ is mulit-modal
- Using MCMC can be computationally expensive

Example - Tempering



Example - Tempering

- Consider the power-tempered target distribution ($\beta \in (0,1]$)

$$\pi_\beta(\boldsymbol{x}) = [\pi(\boldsymbol{x})]^\beta$$

- Choose β such that $\frac{1}{\beta} \in \mathbb{Z}$ and Markov chain sampling from the tempered target, $\pi_\beta(\boldsymbol{x})$ can mix well across the entire sample space, then

$$\pi(x) = \pi(x)^{\frac{1}{\beta} \cdot \beta} = \prod_{i=1}^{\frac{1}{\beta}} \pi_\beta(x) \ \left(= \prod_{c=1}^{C} f_c(\boldsymbol{x}) \right)$$

Warwick Statistics The Alan Turing Institute

Example - Tempering

- Consider the power-tempered target distribution ($\beta \in (0, 1]$)

$$\pi_\beta(\boldsymbol{x}) = [\pi(\boldsymbol{x})]^\beta$$

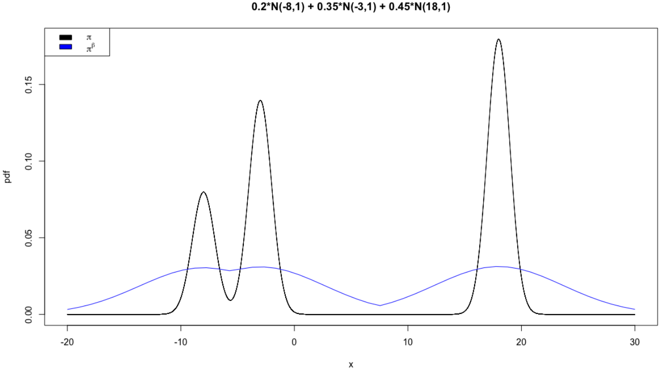
- Choose β such that $\frac{1}{\beta} \in \mathbb{Z}$ and Markov chain sampling from the tempered target, $\pi_\beta(\boldsymbol{x})$ can mix well across the entire sample space, then

$$\pi(x) = \pi(x)^{\frac{1}{\beta} \cdot \beta} = \prod_{i=1}^{\frac{1}{\beta}} \pi_\beta(x) \ \left(= \prod_{c=1}^{C} f_c(\boldsymbol{x}) \right)$$

Warwick Statistics — The Alan Turing Institute

Example - Tempering

- Problem: for many examples, β must be very small for Markov chain sampling for $\pi_{\beta}(\boldsymbol{x})$ to mix well.



Example - Tempering

- Recall the acceptance probability:

$$\frac{g(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

$$\begin{cases} \rho := \rho(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(C)}) = e^{-\frac{C\sigma^2}{2T}} \\ \sigma^2 = \frac{1}{C} \sum_{c=1}^{C} \left\| \boldsymbol{x}^{(c)} - \bar{\boldsymbol{x}} \right\|^2 \end{cases}$$

$$\begin{cases} Q = \mathbb{E}_{\bar{\mathbb{W}}}(E) \\ E := \prod_{c=1}^{C} \left[\exp\left\{ -\int_0^T \left(\phi_c(\boldsymbol{x}_t^{(c)}) - \Phi_c \right) \mathrm{d}t \right\} \right] \end{cases}$$

- If C is large, then ρ and Q are very small

Example - Tempering

- Recall the acceptance probability:

$$\frac{g(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})}{h(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}, \boldsymbol{y})} \propto \rho \cdot Q$$

$$\begin{cases} \rho := \rho(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(C)}) = e^{-\frac{C\sigma^2}{2T}} \\ \sigma^2 = \frac{1}{C} \sum_{c=1}^{C} \left\| \boldsymbol{x}^{(c)} - \bar{\boldsymbol{x}} \right\|^2 \end{cases}$$

$$\begin{cases} Q = \mathbb{E}_{\bar{\mathbb{W}}}(E) \\ E := \prod_{c=1}^{C} \left[\exp\left\{ -\int_0^T \left(\phi_c(\boldsymbol{x}_t^{(c)}) - \Phi_c \right) \mathrm{d}t \right\} \right] \end{cases}$$

- If C is large, then ρ and Q are very small

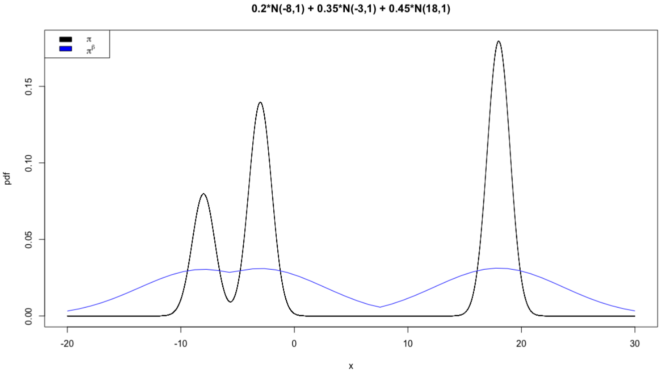
Warwick Statistics | The Alan Turing Institute

Hierarchical Monte Carlo Fusion

- Suppose we want to sample from

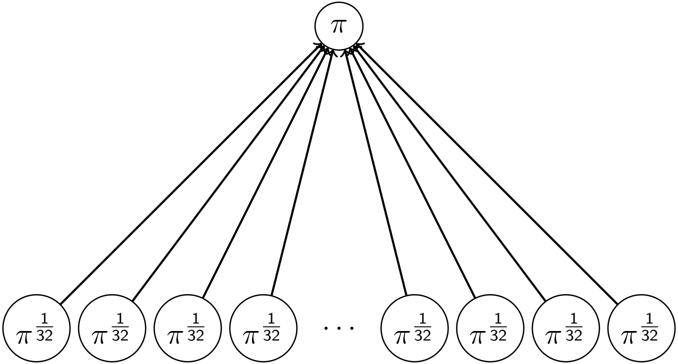
$$\pi(x) = 0.2\mathrm{N}(-8, 1) + 0.35\mathrm{N}(-3, 1) + 0.45\mathrm{N}(18, 1)$$

and $\beta = \frac{1}{32}$



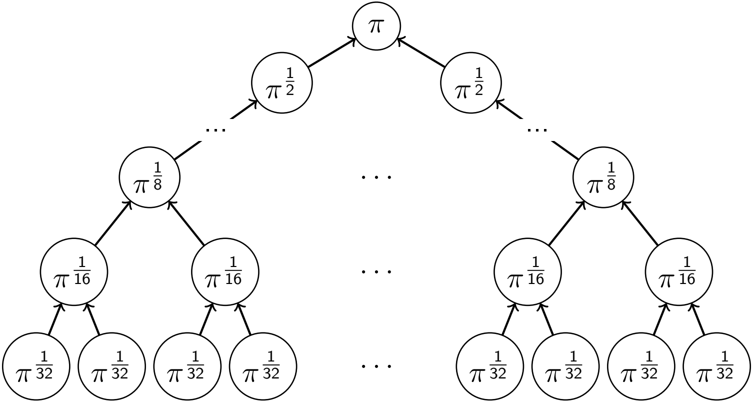
Hierarchical Monte Carlo Fusion

- Instead of a 'fork-and-join' approach:

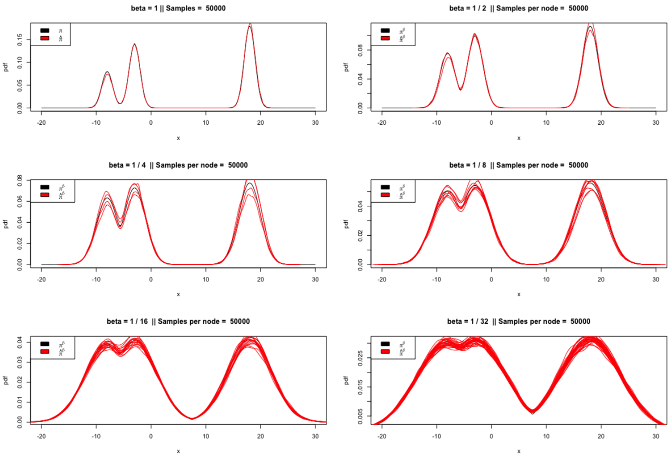


Hierarchical Monte Carlo Fusion

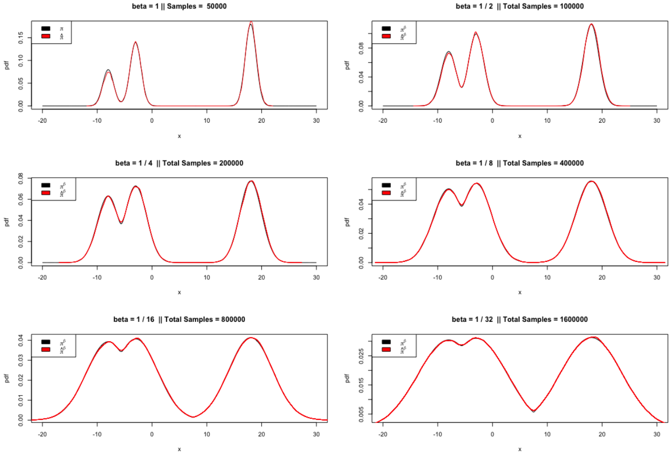
- We could perform fusion in a hierarchical fashion:



Hierarchical Monte Carlo Fusion - Example



Hierarchical Monte Carlo Fusion - Example



Hierarchical Monte Carlo Fusion - Further work

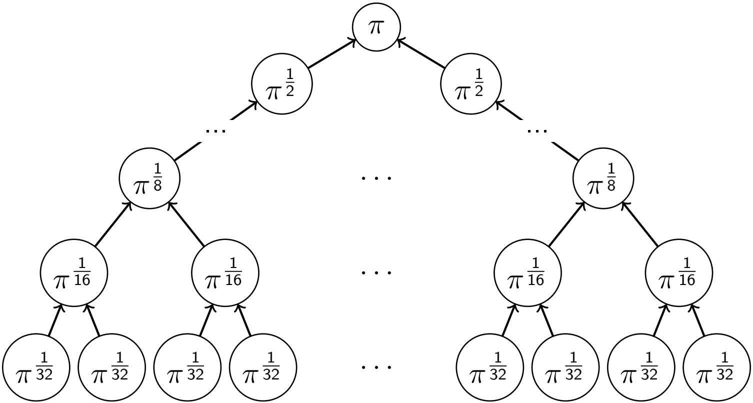
- Quantifying the propagation of error arising from the hierarchy
- In the more general case, how do we choose which sub-posteriors to combine?
- What is the effect of increasing dimensionality in the sub-posteriors?
- Are there other hierarchy structures that make more sense?

Hierarchical Monte Carlo Fusion - Further work

- Quantifying the propagation of error arising from the hierarchy
- In the more general case, how do we choose which sub-posteriors to combine?
- What is the effect of increasing dimensionality in the sub-posteriors?
- Are there other hierarchy structures that make more sense?

Hierarchical Monte Carlo Fusion

- We could perform fusion in a hierarchical fashion:



Hierarchical Monte Carlo Fusion - Further work

- Quantifying the propagation of error arising from the hierarchy
- In the more general case, how do we choose which sub-posteriors to combine?
- What is the effect of increasing dimensionality in the sub-posteriors?
- Are there other hierarchy structures that make more sense?

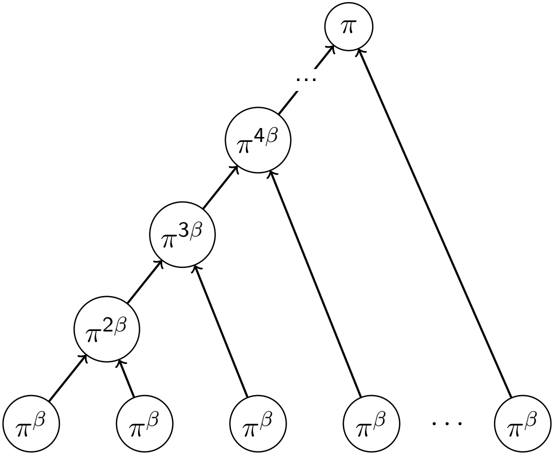
Hierarchical Monte Carlo Fusion - Further work

- Quantifying the propagation of error arising from the hierarchy
- In the more general case, how do we choose which sub-posteriors to combine?
- What is the effect of increasing dimensionality in the sub-posteriors?
- Are there other hierarchy structures that make more sense?

Hierarchical Monte Carlo Fusion - Further work

- Quantifying the propagation of error arising from the hierarchy
- In the more general case, how do we choose which sub-posteriors to combine?
- What is the effect of increasing dimensionality in the sub-posteriors?
- Are there other hierarchy structures that make more sense?

Hierarchical Monte Carlo Fusion



Monte Carlo Fusion - Further work in my PhD

- Making Fusion more efficient
- Fusion in a privacy setting (CONfidential Fusion / Confusion)
- Dimensionally mis-matched sub-posteriors setting

Monte Carlo Fusion - Further work in my PhD

- Making Fusion more efficient
- Fusion in a privacy setting (CONfidential Fusion / Confusion)
- Dimensionally mis-matched sub-posteriors setting

Monte Carlo Fusion - Further work in my PhD

- Making Fusion more efficient
- Fusion in a privacy setting (CONfidential Fusion / Confusion)
- Dimensionally mis-matched sub-posteriors setting

Any questions?