Sampling with Time-Changed Markov Processes

Giorgos Vasdekis
Newcastle University, U.K.

Joint work with A. Bertazzi

4 December 2025,
Lancaster CSML

Outline

Goal and Motivation

- **Goal:** MCMC to target the density π on $\mathbb{R}^d$.

Goal and Motivation

- **Goal:** MCMC to target the density π on $\mathbb{R}^d$.
- Consider continuous time Markov processes.

Goal and Motivation

- **Goal:** MCMC to target the density π on $\mathbb{R}^d$.
- Consider continuous time Markov processes.
- Most local algorithms have problems on heavy tailed and multi-modal targets.

Goal and Motivation

- **Goal:** MCMC to target the density π on $\mathbb{R}^d$.
- Consider continuous time Markov processes.
- Most local algorithms have problems on heavy tailed and multi-modal targets.
- We suggest a general framework for algorithms to speed up or slow down in various areas.

Goal and Motivation

- **Goal:** MCMC to target the density π on $\mathbb{R}^d$.
- Consider continuous time Markov processes.
- Most local algorithms have problems on heavy tailed and multi-modal targets.
- We suggest a general framework for algorithms to speed up or slow down in various areas.
- Follows up work from (Vasdekis and Roberts 2023).

Challenges with Traditional MCMC

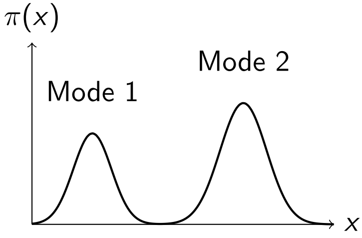
Multimodal Distributions

- Processes get trapped in local modes
- Difficult to traverse low-density regions
- Poor mixing between modes

Challenges with Traditional MCMC

Multimodal Distributions

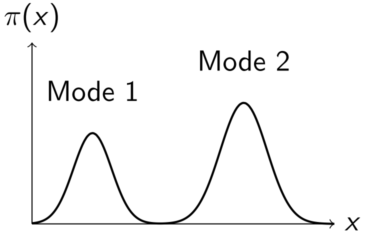
- Processes get trapped in local modes
- Difficult to traverse low-density regions
- Poor mixing between modes



Challenges with Traditional MCMC

Multimodal Distributions

- Processes get trapped in local modes
- Difficult to traverse low-density regions
- Poor mixing between modes



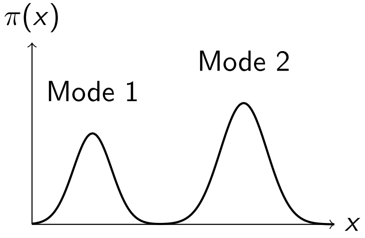
Heavy-Tailed Distributions

- Slow exploration of tails
- Inefficient sampling of rare events

Challenges with Traditional MCMC

Multimodal Distributions

- Processes get trapped in local modes
- Difficult to traverse low-density regions
- Poor mixing between modes



Heavy-Tailed Distributions

- Slow exploration of tails
- Inefficient sampling of rare events

Need for More Efficient MCMC Methods

Traditional MCMC methods often struggle with these challenging distributions, leading to slow convergence and poor mixing.

Core Idea of Time-Changed Markov Processes

Time-changed process

A time-changed Markov process adjusts the speed of time of a base Markov process via a user-chosen, state-dependent function.

Core Idea of Time-Changed Markov Processes

Time-changed process

A time-changed Markov process adjusts the speed of time of a base Markov process via a user-chosen, state-dependent function.

- Start with a well-understood (time-homogeneous) **base process** $(Y_t)_{t \geq 0}$.

Core Idea of Time-Changed Markov Processes

Time-changed process

A time-changed Markov process adjusts the speed of time of a base Markov process via a user-chosen, state-dependent function.

- Start with a well-understood (time-homogeneous) **base process** $(Y_t)_{t\geq 0}$.
- Apply a time transformation to get a new process X_t.

Core Idea of Time-Changed Markov Processes

Time-changed process

A time-changed Markov process adjusts the speed of time of a base Markov process via a user-chosen, state-dependent function.

- Start with a well-understood (time-homogeneous) **base process** $(Y_t)_{t \geq 0}$.
- Apply a time transformation to get a new process X_t.
- The transformation is regulated by a **speed function** $s : \mathbb{R}^d \to (0, +\infty)$.

Core Idea of Time-Changed Markov Processes

Time-changed process

A time-changed Markov process adjusts the speed of time of a base Markov process via a user-chosen, state-dependent function.

- Start with a well-understood (time-homogeneous) **base process** $(Y_t)_{t \geq 0}$.
- Apply a time transformation to get a new process X_t.
- The transformation is regulated by a **speed function** $s : \mathbb{R}^d \to (0, +\infty)$.
- Intuitively:
  - When $s(x)$ is large, time accelerates
  - When $s(x)$ is small, time decelerates

Mathematical Formulation

Time-Changed Process Definition

$$X_t = Y_{r(t)}$$

where

$$r(t) = \int_0^t s(X_u)du$$

Mathematical Formulation

Time-Changed Process Definition

$$X_t = Y_{r(t)}$$

where

$$r(t) = \int_0^t s(X_u)du$$

Intuition:

$$\frac{dX_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot \frac{dr_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot s(X_t)$$

Mathematical Formulation

Time-Changed Process Definition

$$X_t = Y_{r(t)}$$

where

$$r(t) = \int_0^t s(X_u)du$$

Intuition:

$$\frac{dX_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot \frac{dr_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot s(X_t)$$

At time $t = 0$:

$$X_0' = Y_0' \cdot s(X_0)$$

Mathematical Formulation

Time-Changed Process Definition

$$X_t = Y_{r(t)}$$

where

$$r(t) = \int_0^t s(X_u)du$$

Intuition:

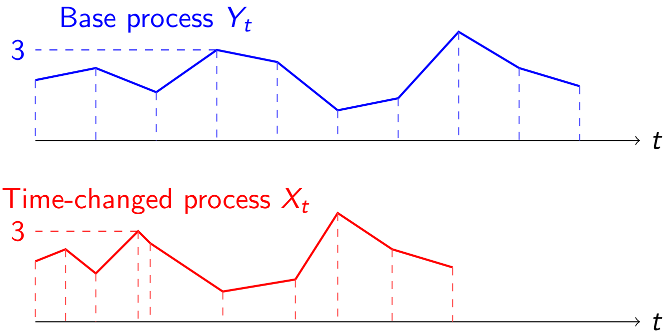
$$\frac{dX_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot \frac{dr_t}{dt} = \frac{dY_{r(t)}}{dr(t)} \cdot s(X_t)$$

At time $t = 0$:

$$X_0' = Y_0' \cdot s(X_0)$$

"X follows the path of Y, but s times faster."

Visualizing Time-Changed Processes



Assumptions on s

Assumptions on s

- Continuous function.

Assumptions on s

Assumptions on s

- Continuous function.
- $\pi(s) := \int s(x)\pi(dx) < \infty$.

Assumptions on s

Assumptions on s

- Continuous function.
- $\pi(s) := \int s(x)\pi(dx) < \infty$.
- Lower bounded: $s(x) \geq s_0 > 0$.

Connection between $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$

- X_t follows the same paths as Y_t, but with adjusted speed.

Connection between $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$

- X_t follows the same paths as Y_t, but with adjusted speed.

- X_t and Y_t spend different amount of time in any given set.

Connection between $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$

- X_t follows the same paths as Y_t, but with adjusted speed.
- X_t and Y_t spend different amount of time in any given set.
- X_t and Y_t have different stationary distributions.

Connection between $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$

- X_t follows the same paths as Y_t, but with adjusted speed.
- X_t and Y_t spend different amount of time in any given set.
- X_t and Y_t have different stationary distributions.

Informal Result: Invariant distributions

Let $\tilde{\pi}(dx) = \frac{1}{\pi(s)} s(x)\pi(dx)$.

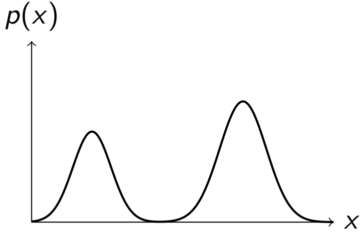
- Y targets $\tilde{\pi} \iff X$ targets π.

where $\pi(s) = \int_{\mathbb{R}^d} s(x)\pi(dx)$

Example: Multimodal Distribution

Challenge:

- Base process gets trapped in local modes
- Rarely crosses low-density regions



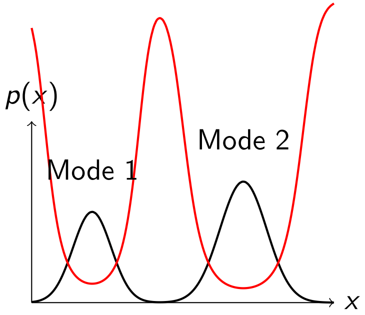
Example: Multimodal Distribution

Challenge:

- Base process gets trapped in local modes
- Rarely crosses low-density regions

Solution with Time-Change:

- Set high $s(x)$ in low-density regions
- Process spends less real time there
- But visits these regions more frequently
- Improves mode-hopping behavior



Key Assumptions

Assumption 1 (Speed Function)

The speed function $s : \mathbb{R}^d \to \mathbb{R}_+$ is continuous, satisfies $\int s(y)\pi(dy) < \infty$, and there exists $s_0 > 0$ such that $s(x) \geq s_0$ for all $x \in \mathbb{R}^d$.

Assumption 2 (LLN for Base Process)

For any $f \in L^1(\tilde{\pi})$ and any initial condition $x \in E$:

$$\frac{1}{T}\int_0^T f(Y_u)du \xrightarrow{T\to\infty} \int_E f(y)\tilde{\pi}(dy) \quad \text{a.s.} \tag{1}$$

where $\tilde{\pi}(dx) = \frac{1}{\pi(s)}s(x)\pi(dx)$.

Law of Large Numbers for X_t

Theorem (Invariance and LLN for Time-Changed Process)

Under Assumptions 1 and 2, the process X has π as unique stationary distribution.

Law of Large Numbers for X_t

Theorem (Invariance and LLN for Time-Changed Process)

Under Assumptions 1 and 2, the process X has π as unique stationary distribution. Furthermore, for any $f \in L^1(\pi)$ and all initial conditions $x \in \mathbb{R}^d$:

$$\frac{1}{T}\int_0^T f(X_t)dt \xrightarrow{T\to\infty} \int_{\mathbb{R}^d} f(x)\pi(dx) \quad \text{a.s.} \tag{2}$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$.

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

“Proof”

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt$$

“Proof”

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$
$$= \frac{r(T)}{T} \cdot \underbrace{\frac{1}{r(T)}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du}$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$
$$= \frac{r(T)}{T} \cdot \underbrace{\frac{1}{r(T)}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du}_{\text{LLN for } Y}$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$
$$= \frac{r(T)}{T}\cdot \underbrace{\frac{1}{r(T)}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du}_{\text{LLN for } Y} \xrightarrow{T\to\infty} C\int_{\mathbb{R}^d}\frac{f(y)}{s(y)}\tilde{\pi}(dy)$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$

$$= \frac{r(T)}{T} \cdot \underbrace{\frac{1}{r(T)}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du}_{\text{LLN for } Y} \xrightarrow{T\to\infty} C\int_{\mathbb{R}^d} \frac{f(y)}{s(y)}\tilde{\pi}(dy)$$

$$= C\int_{\mathbb{R}^d} \frac{f(y)}{s(y)}s(y)\pi(dy)$$

"Proof"

Use a change of variables $u = r(t) = \int_0^t s(X_v)dv$. So

$$du = r^{'}(t)dt = s(X_t)dt = s(Y_{r(t)})dt = s(Y_u)dt.$$

Write,

$$\frac{1}{T}\int_0^T f(X_t)dt = \frac{1}{T}\int_0^T f(Y_{r(t)})dt = \frac{1}{T}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du$$
$$= \frac{r(T)}{T}\cdot \underbrace{\frac{1}{r(T)}\int_0^{r(T)} f(Y_u)\frac{1}{s(Y_u)}du}_{\text{LLN for } Y} \xrightarrow{T\to\infty} C\int_{\mathbb{R}^d}\frac{f(y)}{s(y)}\tilde{\pi}(dy)$$
$$= C\int_{\mathbb{R}^d}\frac{f(y)}{s(y)}s(y)\pi(dy) = \int_{\mathbb{R}^d} f(y)\pi(dy).$$

Convergence Properties

Geometric Ergodicity

Under suitable conditions on the base process Y and speed function s, the time-changed process X_t is geometrically ergodic, even if the base process Y is not:

$$\|\mathbb{P}(X_t \in \cdot) - \pi(\cdot)\|_{TV} \leq M(x) \exp\{-\lambda t\}, t \geq 0$$

Convergence Properties

Geometric Ergodicity

Under suitable conditions on the base process Y and speed function s, the time-changed process X_t is geometrically ergodic, even if the base process Y is not:

$$\|\mathbb{P}(X_t \in \cdot) - \pi(\cdot)\|_{TV} \leq M(x)\exp\{-\lambda t\}, t \geq 0$$

Uniform Ergodicity

Under suitable conditions on Y and assuming that $s(x)$ grows sufficiently fast as $\|x\| \to \infty$, the time-changed process X can be uniformly ergodic even when the base process Y is not:

$$\|\mathbb{P}(X_t \in \cdot) - \pi(\cdot)\|_{TV} \leq M\exp\{-\lambda t\}, t \geq 0$$

Convergence Properties

Central Limit Theorem

Under appropriate conditions, the time-changed process satisfies a central limit theorem, with asymptotic variance expressed in terms of the base process.

Example 1: SDEs

- *Diffusions remain diffusions under time-change.*

Example 1: SDEs

- *Diffusions remain diffusions under time-change.*
- Original diffusion:
$$dY_t = \tilde{b}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t$$

Example 1: SDEs

- *Diffusions remain diffusions under time-change.*
- Original diffusion:

$$dY_t = \tilde{b}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t$$

- Time-changed diffusion:

$$dX_t = s(X_t)\tilde{b}(X_t)dt + s^{1/2}(X_t)\sigma(\tilde{X}_t)dW_t.$$

Example 1: SDEs

- *Diffusions remain diffusions under time-change.*
- Original diffusion:

$$dY_t = \tilde{b}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t$$

- Time-changed diffusion:

$$dX_t = s(X_t)\tilde{b}(X_t)dt + s^{1/2}(X_t)\sigma(\tilde{X}_t)dW_t.$$

- Original targets $\tilde{\pi} \Rightarrow$ Time-Changed targets π.

Example 1: SDEs

- *Diffusions remain diffusions under time-change.*
- Original diffusion:

$$dY_t = \tilde{b}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t$$

- Time-changed diffusion:

$$dX_t = s(X_t)\tilde{b}(X_t)dt + s^{1/2}(X_t)\sigma(\tilde{X}_t)dW_t.$$

- Original targets $\tilde{\pi} \Rightarrow$ Time-Changed targets π.
- (Roberts and Stramer 2002)

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.
- It is a time-change of the process $dY_t = \nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}dW_t$, which targets $\tilde{\pi}(dy) \propto s(y)\pi(dy)$.

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right)dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.
- It is a time-change of the process $dY_t = \nabla \log\left(s(Y_t)\pi(Y_t)\right)dt + \sqrt{2}dW_t$, which targets $\tilde{\pi}(dy) \propto s(y)\pi(dy)$.
- Special case: $s(x) = \pi^{-\alpha}(x)$ for $\alpha \in (0,1)$

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.
- It is a time-change of the process $dY_t = \nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}dW_t$, which targets $\tilde{\pi}(dy) \propto s(y)\pi(dy)$.
- Special case: $s(x) = \pi^{-\alpha}(x)$ for $\alpha \in (0,1)$
- The base process becomes $dY_t = \nabla \log\left(\pi(Y_t)^{1-\alpha}\right) dt + \sqrt{2}dW_t$.

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.

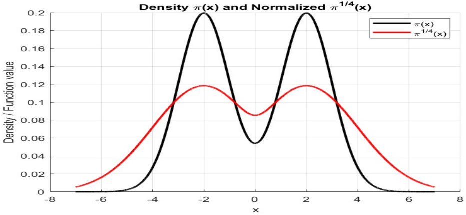
- It is a time-change of the process $dY_t = \nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}dW_t$, which targets $\tilde{\pi}(dy) \propto s(y)\pi(dy)$.

- Special case: $s(x) = \pi^{-\alpha}(x)$ for $\alpha \in (0, 1)$

- The base process becomes $dY_t = \nabla \log\left(\pi(Y_t)^{1-\alpha}\right) dt + \sqrt{2}dW_t$. Targets $\pi^{1-\alpha}$. "Less multi-modal".

Overdamped Langevin Dynamics

- The process $dX_t = s(X_t)\nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}s^{1/2}(X_t)dW_t$ targets π.
- It is a time-change of the process $dY_t = \nabla \log\left(s(Y_t)\pi(Y_t)\right) dt + \sqrt{2}dW_t$, which targets $\tilde{\pi}(dy) \propto s(y)\pi(dy)$.
- Special case: $s(x) = \pi^{-\alpha}(x)$ for $\alpha \in (0, 1)$
- The base process becomes $dY_t = \nabla \log\left(\pi(Y_t)^{1-\alpha}\right) dt + \sqrt{2}dW_t$. Targets $\pi^{1-\alpha}$. "Less multi-modal".



Example 2: Markov Jump Processes

- *Markov Jump processes remain Markov Jump processes under time-change*

Example 2: Markov Jump Processes

- *Markov Jump processes remain Markov Jump processes under time-change*
- Original Markov Jump process: Rate $\lambda : \mathbb{R}^d \to (0, +\infty)$, Jump kernel Q.

Example 2: Markov Jump Processes

- *Markov Jump processes remain Markov Jump processes under time-change*
- Original Markov Jump process: Rate $\lambda : \mathbb{R}^d \to (0, +\infty)$, Jump kernel Q.

  That is, if the process is at y, it stays there for time $\tau \sim exp(\lambda(y))$. Then it jumps according to $Q(y, \cdot)$.

Example 2: Markov Jump Processes

- *Markov Jump processes remain Markov Jump processes under time-change*
- Original Markov Jump process: Rate $\lambda : \mathbb{R}^d \to (0, +\infty)$, Jump kernel Q.

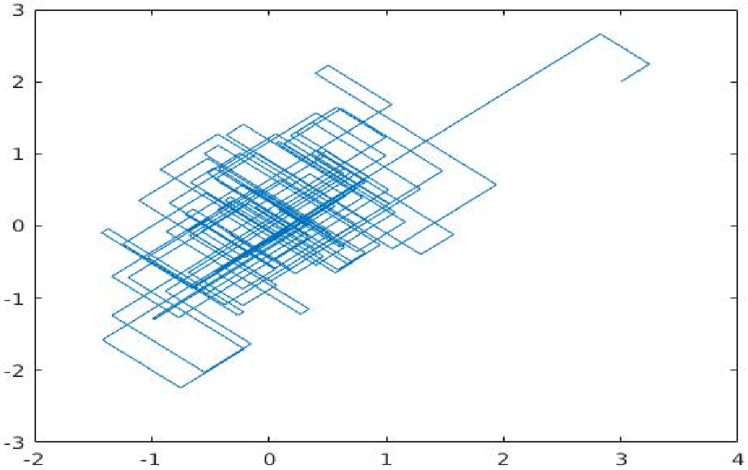
  That is, if the process is at y, it stays there for time $\tau \sim exp(\lambda(y))$. Then it jumps according to $Q(y, \cdot)$.
- Time-changed Markov Jump process: Rate $s \cdot \lambda$, Jump kernel Q.

Example 2: Markov Jump Processes

- *Markov Jump processes remain Markov Jump processes under time-change*
- Original Markov Jump process: Rate $\lambda : \mathbb{R}^d \to (0, +\infty)$, Jump kernel Q.

  That is, if the process is at y, it stays there for time $\tau \sim exp(\lambda(y))$. Then it jumps according to $Q(y, \cdot)$.
- Time-changed Markov Jump process: Rate $s \cdot \lambda$, Jump kernel Q.
- Original process targets $\tilde{\pi} \Rightarrow$ Time-changed targets π.

Running Example: Zig-Zag Process (ZZP)



Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(Y_t, v), \quad t \geq 0$$

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(Y_t, v), \quad t \geq 0 \quad \text{, where} \quad \lambda_i(y, v) = [-v_i \partial_i \log \pi(y)]^+$$

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(Y_t, v), \quad t \geq 0 \ \text{, where} \quad \lambda_i(y, v) = [-v_i \partial_i \log \pi(y)]^+$$

Suppose that the first arrival times are T_1, T_2.

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(Y_t, v), \quad t \geq 0 \text{ , where } \quad \lambda_i(y, v) = [-v_i \partial_i \log \pi(y)]^+$$

   Suppose that the first arrival times are T_1, T_2.
4. Let $T = \min\{T_1, T_2\}$ and

$$j = \operatorname{argmin}\{T_1, T_2\}.$$

5. Set $y = Y_T$ and flip the velocity: $v_j = -v_j$.

Running Example: Zig-Zag Process (ZZP)

Algorithm (in 2-d) targeting π

1. Start from point $(y, v) = (y_1, y_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (Y_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} Y_t = v, \quad t \geq 0, \quad Y_0 = y, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(Y_t, v), \quad t \geq 0 \quad , \text{where} \quad \lambda_i(y, v) = [-v_i \partial_i \log \pi(y)]^+$$

   Suppose that the first arrival times are T_1, T_2.
4. Let $T = \min\{T_1, T_2\}$ and

$$j = \text{argmin}\{T_1, T_2\}.$$

5. Set $y = Y_T$ and flip the velocity: $v_j = -v_j$.
6. Repeat from step 1.

Time-Changed Zig-Zag Process

Algorithm (in 2-d) targeting π

1. Start from point $(x, v) = (x_1, x_2; v_1, v_2) \in \mathbb{R}^2 \times \{-1, 1\}^2$.
2. The process (X_t, V_t) moves according to the deterministic dynamics:

$$\frac{d}{dt} X_t = s(X_t) \cdot v, \quad t \geq 0, \quad X_0 = x, \quad \text{and} \quad V_t = v, \quad t \geq 0.$$

3. For all $i = 1, 2$, consider a non-homogeneous Poisson process with intensity

$$m_i(t) = \lambda_i(X_t, v), \quad t \geq 0, \quad \text{where} \quad \lambda_i(x, v) = s(x) \left[-v_i \partial_i \log\left(s(x) \pi(x) \right) \right]^+ .$$

   Suppose that the first arrival times are T_1, T_2.
4. Let $T = \min\{T_1, T_2\}$ and

$$j = \text{argmin}\{T_1, T_2\}.$$

5. Set $x = X_T$ and flip the velocity: $v_j = -v_j$.
6. Repeat from step 1.

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.
Finding where the process will be after time $t \to$ Solve an easy one-dimensional ODE.

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.
Finding where the process will be after time $t \to$ Solve an easy one-dimensional ODE.
Starting position x_i, direction v_i.
For $i = 2, \ldots, d$,

$$X^i(t) = x_i + v_1 v_i \left(X^1(t) - x_1 \right)$$

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.
Finding where the process will be after time $t \to$ Solve an easy one-dimensional ODE.
Starting position x_i, direction v_i.
For $i = 2, \ldots, d$,

$$X^i(t) = x_i + v_1 v_i \left(X^1(t) - x_1\right) = y_2 + v_1 v_2 X^1(t),$$

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.
Finding where the process will be after time $t \to$ Solve an easy one-dimensional ODE.
Starting position x_i, direction v_i.
For $i = 2, \ldots, d$,

$$X^i(t) = x_i + v_1 v_i \left(X^1(t) - x_1\right) = y_2 + v_1 v_2 X^1(t),$$

$$X^1(t) = -c_1 + c_2 \tan\left(\arctan\left(\frac{x_1 + c_1}{c_2}\right) + v_1 c_2 d\ t\right), \quad t \in [0, t^*(x, v)),$$

Why do we care about this process?

Take for example $s(x) = 1 + \|x\|^2$.
Until a change of direction, the process moves in a straight line.
Finding where the process will be after time $t \to$ Solve an easy one-dimensional ODE.
Starting position x_i, direction v_i.
For $i = 2, \dots, d$,

$$X^i(t) = x_i + v_1 v_i \left(X^1(t) - x_1\right) = y_2 + v_1 v_2 X^1(t),$$

$$X^1(t) = -c_1 + c_2 \tan\left(\arctan\left(\frac{x_1 + c_1}{c_2}\right) + v_1 c_2 d \; t\right), \quad t \in [0, t^*(x, v)),$$

and

$$c_2 = \sqrt{\frac{1 + y_2^2}{d} - c_1^2}, \quad c_1 = \frac{(y \cdot v)}{d} v_1, \quad y_2 = x_2 - v_1 v_i x_1.$$

Comparison with Related Concepts

Diffeomorphic Transformations

Tempering/Annealing

- Modifies target distribution → Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$

Similarities

Comparison with Related Concepts

Diffeomorphic Transformations

Tempering/Annealing

- Modifies target distribution $\rightarrow$ Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Similarities

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution $\rightarrow$ Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t \geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \left|\det\left(\nabla f(y)\right)\right|$.

Similarities

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution → Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t\geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \left|\det\left(\nabla f(y)\right)\right|$.
- Map to samples back to $X_t = f(Y_t) \rightarrow$ Approximate samples from π.

Similarities

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution $\rightarrow$ Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t \geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \left|\det\left(\nabla f(y)\right)\right|$.
- Map to samples back to $X_t = f(Y_t) \rightarrow$ Approximate samples from π.
- Change paths, keep timing.

Similarities

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution $\rightarrow$ Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t \geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \, |\det(\nabla f(y))|$.
- Map to samples back to $X_t = f(Y_t) \rightarrow$ Approximate samples from π.
- Change paths, keep timing.

Similarities

- Both can help with multimodal/heavy-tailed distributions.

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution $\rightarrow$ Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t \geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \left|\det\left(\nabla f(y)\right)\right|$.
- Map to samples back to $X_t = f(Y_t) \rightarrow$ Approximate samples from π.
- Change paths, keep timing.

Similarities

- Both can help with multimodal/heavy-tailed distributions.
- Speed function plays similar role to Jacobian determinant

Comparison with Related Concepts

Tempering/Annealing

- Modifies target distribution → Targets $\tilde{\pi}(x) = \pi^{\alpha}(x)$, $\alpha \in (0, 1)$
- Requires careful tuning of temperatures α.

Diffeomorphic Transformations

- Applies diffeomorphism f^{-1} to state space.
- Run a process $(Y_t)_{t\geq 0}$ that targets the push-forward measure $\tilde{\pi}(y) = \pi(f(y)) \left|\det(\nabla f(y))\right|$.
- Map to samples back to $X_t = f(Y_t) \rightarrow$ Approximate samples from π.
- Change paths, keep timing.

Similarities

- Both can help with multimodal/heavy-tailed distributions.
- Speed function plays similar role to Jacobian determinant
- Both time and diffeomorphic transformations can achieve uniform ergodicity under suitable conditions

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

- If the process can be implemented directly, use

$$F(T) := \frac{1}{T}\int_0^T f(X_t)dt.$$

.

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

- If the process can be implemented directly, use

$$F(T) := \frac{1}{T}\int_0^T f(X_t)dt.$$

- Sample the process at times $X_\delta, X_{2\delta}, X_{3\delta}, \dots$.

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

- If the process can be implemented directly, use

$$F(T) := \frac{1}{T}\int_0^T f(X_t)dt.$$

- Sample the process at times $X_\delta, X_{2\delta}, X_{3\delta}, \dots$. Use the Monte Carlo estimator

$$\frac{1}{N}\sum_{k=1}^N f(X_k).$$

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

- If the process can be implemented directly, use

$$F(T) := \frac{1}{T}\int_0^T f(X_t)dt.$$

- Sample the process at times $X_\delta, X_{2\delta}, X_{3\delta}, \dots$. Use the Monte Carlo estimator

$$\frac{1}{N}\sum_{k=1}^{N} f(X_k).$$

- Generally this would work easier with PDMPs.

Implementation of Time-Changed processes (Approach 1)

Goal: Estimate $\int_{\mathbb{R}^d} f(x)\pi(dx)$.

- If the process can be implemented directly, use

$$F(T) := \frac{1}{T}\int_0^T f(X_t)dt.$$

- Sample the process at times $X_\delta, X_{2\delta}, X_{3\delta}, \ldots$. Use the Monte Carlo estimator

$$\frac{1}{N}\sum_{k=1}^{N} f(X_k).$$

- Generally this would work easier with PDMPs.
- *For example:* Time-changed Zig-Zag process with speed

$$s(x) = \left(1 + \|x\|^2\right)^{1+k}, k = 0, 1, 2, \ldots$$

Implementation of Time-Changed processes (Approach 2)

- Target $\tilde{\pi}$. Simulate $(Y_t)_{t\geq 0}$. Use

$$F(r^{-1}(T)) := \frac{\int_0^T \frac{f(Y_t)}{s(Y_t)} dt}{\int_0^T \frac{1}{s(Y_t)} dt}$$

Implementation of Time-Changed processes (Approach 2)

- Target $\tilde{\pi}$. Simulate $(Y_t)_{t\geq 0}$. Use

$$F(r^{-1}(T)) := \frac{\int_0^T \frac{f(Y_t)}{s(Y_t)} dt}{\int_0^T \frac{1}{s(Y_t)} dt}$$

Connections **Self-normalised Importance Sampling**

Implementation of Time-Changed processes (Approach 2)

- Target $\tilde{\pi}$. Simulate $(Y_t)_{t\geq 0}$. Use

$$F(r^{-1}(T)) := \frac{\int_0^T \frac{f(Y_t)}{s(Y_t)} dt}{\int_0^T \frac{1}{s(Y_t)} dt}$$

  Connections **Self-normalised Importance Sampling**

- Discretise Y. Take the points $Y_\delta, Y_{2\delta}, Y_{3\delta}, \ldots$. Use

$$\frac{\sum_{k=0}^{N} \frac{f(Y_{k\delta})}{s(Y_{k\delta})}}{\sum_{k=0}^{N} \frac{1}{s(Y_{k\delta})}}$$

Implementation of Time-Changed processes (Approach 2)

- Target $\tilde{\pi}$. Simulate $(Y_t)_{t\geq 0}$. Use

$$F(r^{-1}(T)) := \frac{\int_0^T \frac{f(Y_t)}{s(Y_t)} dt}{\int_0^T \frac{1}{s(Y_t)} dt}$$

  Connections **Self-normalised Importance Sampling**

- Discretise Y. Take the points $Y_\delta, Y_{2\delta}, Y_{3\delta}, \ldots$. Use

$$\frac{\sum_{k=0}^N \frac{f(Y_{k\delta})}{s(Y_{k\delta})}}{\sum_{k=0}^N \frac{1}{s(Y_{k\delta})}}$$

Connection with **Importance Markov Chain.**

Implementation of Time-Changed processes (Approach 3)

- Use Markov jump process with rate s.

Implementation of Time-Changed processes (Approach 3)

- Use Markov jump process with rate s.
- Jump kernel $\tilde{Q}$ with stationary distribution $\tilde{\pi}$.

Implementation of Time-Changed processes (Approach 3)

- Use Markov jump process with rate s.
- Jump kernel $\tilde{Q}$ with stationary distribution $\tilde{\pi}$.
- Easier to implement in some cases.

Implementation of Time-Changed processes (Approach 3)

- Use Markov jump process with rate s.
- Jump kernel $\tilde{Q}$ with stationary distribution $\tilde{\pi}$.
- Easier to implement in some cases.
- $\tilde{Q}$ can be a discretisation of your favourite process.

Simulations: Heavy-Tailed Targets (Scenario 1)

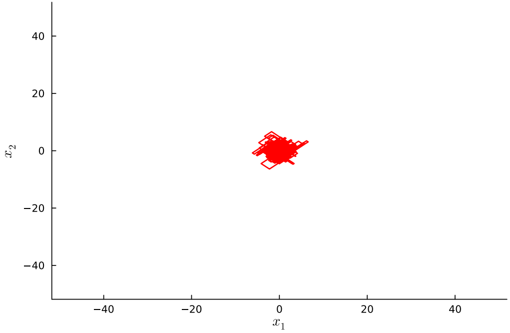


Figure: $s(x) = 1$

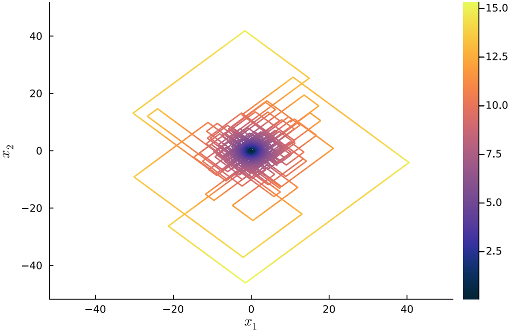


Figure: $s(x) = (1 + |x|^2)^2$

Figure: Student(5) target.

Simulations: Heavy-Tailed Targets (Scenario 2)

Target: $\pi(x) \sim \exp\left\{-\|x\|^{1/2}\right\}, \quad d = 20$

Simulations: Heavy-Tailed Targets (Scenario 2)

Target: $\pi(x) \sim \exp\left\{-\|x\|^{1/2}\right\}, \quad d = 20$

Speed used: $s(x) = \left(1 + \|x\|^2\right)^{(1+a)/2}$

Simulations: Heavy-Tailed Targets (Scenario 2)

Target: $\pi(x) \sim \exp\left\{-\|x\|^{1/2}\right\}, \quad d = 20$

Speed used: $s(x) = \left(1 + \|x\|^2\right)^{(1+a)/2}$

Algorithmic Efficiency

Algorithms	**ESS/Lik.Eval.**	**ESS/min**
Zig-Zag	$0.3 \cdot 10^{-3}$	124.9
Time-changed ZZ ($a = 0$)	$3.9 \cdot 10^{-3}$	2847.7
Time-changed ZZ ($a = 1$)	$\mathbf{6.3 \cdot 10^{-3}}$	**4471.3**
Space Transformed RWM	$2.8 \cdot 10^{-3}$	1966.8

ESS for Zig-Zag, Time-transformed Zig-Zag, and Space Transformed Random Walk Metropolis

Simulations: Heavy-Tailed Targets (Scenario 3)

Speed function: $s(x) = \pi(x)^{-a}, \quad a \in (0, 1/3).$

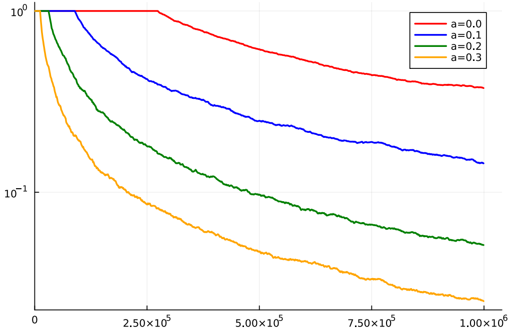


Figure: Median of the relative square error (y-axis) vs number of jumps of the base process (x-axis).

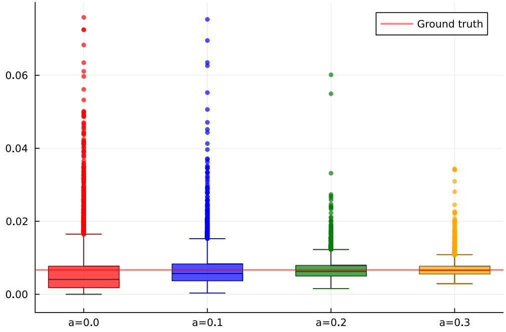


Figure: Estimates of $\mathbb{P}(\|x\| > 150)$ for different values of a.

Figure: Target: Student(1) distribution in $\mathbb{R}^2$.

Simulations: Multi-Modal Targets

Mixture of Normals

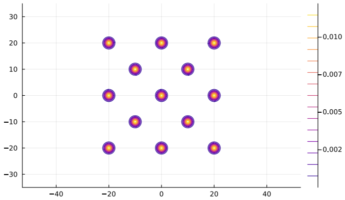


Figure: Level curves of π.

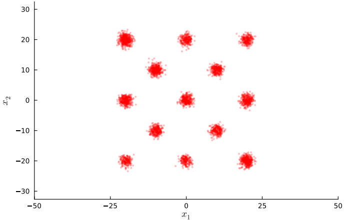


Figure: Time-changed: $s(x) = \pi(x)^{-0.9}$.

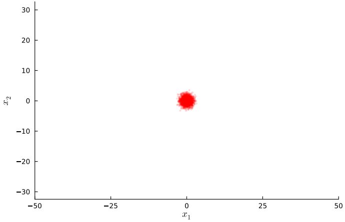


Figure: Without time-change: $s(x) = 1$.

Summary and Future Work

Future Research Directions

- Optimal choice of speed function for specific targets.
- Adaptive methods to learn optimal speed functions.
- Study the high-dimensional behaviour/ scaling limits.

References

- Bertazzi, A., & Vasdekis, G. (2025). Sampling with time-changed Markov processes. *arXiv preprint arXiv:2501.15155*. Under Revision.
- Roberts, G. O., & Stramer, O. (2002). Langevin diffusions and Metropolis-Hastings algorithms. *Methodology and computing in applied probability*, 4(4), 337-357.
- Bierkens, J., Fearnhead, P., & Roberts, G. (2019). The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data. *The Annals of Statistics*, 47(3), 1288-1320.
- Vasdekis, G., & Roberts, G. O. (2023). Speed-up Zig-Zag. *The Annals of Applied Probability*, 33(6A):4693 – 4746, 2023.

Thank you for your attention!

Paper: Sampling with time-changed Markov processes
Available at: `https://arxiv.org/abs/2501.15155`