Theoretical Properties of Quasistationary Monte Carlo Methods

Andi Wang

University of Oxford

a.wang@stats.ox.ac.uk

Joint with Divakar Kumar, Gareth Roberts and David Steinsaltz

11 July 2018
Let $X = (X_t)$ be an ant undergoing a diffusion on $\mathbb{R}^d$. Introduce killing rate

$$\kappa : \mathbb{R}^d \rightarrow [0, \infty).$$

At rate $\kappa(X_t)$ the ant is killed; call this time $\tau_\partial$. 
Let $X = (X_t)$ be an ant undergoing a diffusion on $\mathbb{R}^d$. Introduce killing rate

$$\kappa : \mathbb{R}^d \rightarrow [0, \infty).$$

At rate $\kappa(X_t)$ the ant is killed; call this time $\tau_\partial$.

We will consider $\mathbb{P}_x(X_t \in \cdot | \tau_\partial > t)$ for large $t$. 
Let $X = (X_t)$ be an ant undergoing a diffusion on $\mathbb{R}^d$. Introduce killing rate

$$\kappa : \mathbb{R}^d \to [0, \infty).$$

At rate $\kappa(X_t)$ the ant is killed; call this time $\tau_\partial$.

We will consider $\mathbb{P}_x(X_t \in \cdot | \tau_\partial > t)$ for large $t$.

If these converge to $\pi$ as $t \to \infty$, $\pi$ is an example of a quasistationary distribution.
Example

Take $X$ to be a standard Brownian motion on $\mathbb{R}^2$, $\kappa(y) = \|y\|^2$. 
Example

Take $X$ to be a standard Brownian motion on $\mathbb{R}^2$, $\kappa(y) = \|y\|^2$. 

What can be said about $P(X_t \in \cdot | \tau_{\partial} > t)$ for large $t$?
Example

Take $X$ to be a standard Brownian motion on $\mathbb{R}^2$, $\kappa(y) = \|y\|^2$.

What can be said about $P(X_t \in \cdot | \tau_\partial > t)$ for large $t$?
Quasistationary Monte Carlo methods aim to sample from a target distribution $\pi$, where $\pi$ is a quasistationary distribution.

Quasistationary Monte Carlo methods aim to sample from a target distribution $\pi$, where $\pi$ is a quasistationary distribution.

The quasistationary framework enables the principled use of subsampling techniques to give exact Bayesian inference with a sub-linear cost in the number of observations\(^1\).

---

Foundational Results

\[
dX_t = \nabla A(X_t) \, dt + dW_t, \quad X_0 = x \in \mathbb{R}^d.
\]

Theorem (Convergence to Quasistationarity)

Under certain assumptions, the diffusion \( X \) killed at rate \( \kappa \) has quasilimiting distribution \( \pi \). That is, for each measurable \( E \subset \mathbb{R}^d \) we have as \( t \to \infty \),

\[
P_x(X_t \in E | \tau_\partial > t) \to \pi(E).
\]
Foundational Results

$$dX_t = \nabla A(X_t) \, dt + dW_t, \quad X_0 = x \in \mathbb{R}^d.$$ 

**Theorem (Convergence to Quasistationarity)**

Under certain assumptions, the diffusion $X$ killed at rate $\kappa$ has quasilimiting distribution $\pi$. That is, for each measurable $E \subset \mathbb{R}^d$ we have as $t \to \infty$,

$$\mathbb{P}_x(X_t \in E | \tau_\partial > t) \to \pi(E).$$

**Theorem (Rates of convergence)**

Additionally, $X$ converges to quasistationarity $\pi$ at the same rate as the Langevin diffusion targeting $\pi^2 / 2A$ converges to stationarity.
Suppose we have a killed diffusion $X$ with quasilimiting distribution $\pi$. So $\mathbb{P}(X_t \in \cdot | \tau_\partial > t) \to \pi$. How might we simulate try to simulate $\pi$?
Suppose we have a killed diffusion $X$ with quasilimiting distribution $\pi$. So $\mathbb{P}(X_t \in \cdot \mid \tau_\partial > t) \to \pi$. How might we simulate try to simulate $\pi$?

1. Rejection sampling. Infeasible.
Simulating from QSDs

Suppose we have a killed diffusion $X$ with quasilimiting distribution $\pi$. So $\mathbb{P}(X_t \in \cdot \mid \tau_\theta > t) \rightarrow \pi$. How might we simulate try to simulate $\pi$?

1. Rejection sampling. Infeasible.
2. Continuous-time sequential Monte Carlo. Feasible but involved.
Suppose we have a killed diffusion $X$ with quasilitmiting distribution $\pi$. So $P(X_t \in \cdot | \tau_\partial > t) \to \pi$. How might we simulate try to simulate $\pi$?

1. Rejection sampling. Infeasible.
2. Continuous-time sequential Monte Carlo. Feasible but involved.
An Example ReScaLE Trajectory
Does this algorithm converge to the quasistationary distribution $\pi$?
Does this algorithm converge to the quasistationary distribution $\pi$?

**Theorem (Convergence in compact setting)**

When the state space is compact, we have that (after time-changing) $(\mu_t)$ is an *asymptotic pseudo-trajectory* for a deterministic semiflow $\Phi$ almost surely.

It follows that $\mu_t$ converges to $\pi$ almost surely.
Convergence Result

Does this algorithm converge to the quasistationary distribution $\pi$?

**Theorem (Convergence in compact setting)**

When the state space is compact, we have that (after time-changing) $(\mu_t)$ is an asymptotic pseudo-trajectory for a deterministic semiflow $\Phi$ almost surely.

It follows that $\mu_t$ converges to $\pi$ almost surely.

**Conjecture (General setting)**

We should have that the Proposition holds much more generally: non-compact state space, unbounded killing rate.
Figure: Logistic regression example.
Thanks for listening!

If you are interested to learn more, come see my poster!
