The parallel replica method for Markov chains

David Aristoff (joint work with T. Lelièvre and G. Simpson)

Colorado State University

March 2015
Definition.
Let \((X_n)\) be a time homogeneous Markov chain:
\[
\mathbb{P}(X_{n+1} \in A \mid X_0, \ldots, X_{n-1}, \{X_n = x\}) = \int_A K(x, dy).
\]

Our goal.
To efficiently simulate \((X_n)\) when it is metastable.

- State space can be **discrete**:
  
  *random walk in random environment, Markov State Models*
  
  or **continuous**:
  
  *time discretization of molecular dynamics (MD)*

- \((X_n)\) can be **reversible**:
  
  *MCMC, time discretization of Brownian/Langevin MD*
  
  or **non-reversible**:
  
  *stochastic MD with flow/external forces*

We are interested in **both dynamics and equilibrium**!
Metastability can arise from energetic barriers or entropic barriers.

Example: Energy (left) and entropic (right) barriers for Brownian/Langevin dynamics. Here, $S$ is a metastable subset of state space.
Definition.
If $\nu$ is a probability measure supported in $S$ with

$$\nu(\cdot) = \mathbb{P}(X_n \in \cdot | X_0 \sim \nu, X_1, \ldots, X_n \in S), \quad n \geq 1,$$

then $\nu$ is a quasi-stationary distribution (QSD) in $S$.

- If for any $\mu$ supported in $S$ and any $A \subset S$,

$$\nu(A) = \lim_{n \to \infty} \mathbb{P}(X_n \in A | X_0 \sim \mu, X_1, \ldots, X_n \in S),$$

then $\nu$ is the unique QSD in $S$.

**Intuitively**: $(X_n)$ reaches $\nu$ after spending a “long” time in $S$ without leaving.

- Our setting: Time scale to reach $\nu \ll$ time scale to leave $S$.
- We need to understand exit events from $S$, starting from $\nu$. 
Let $\nu$ be the QSD in $S$ and $\tau$ the first exit time from $S$.

**Theorem.**

If $X_0 \sim \nu$, then $\tau$ is geometrically distributed, and $\tau, X_\tau$ are independent.

- The geometric distribution is

  $$\mathbb{P}(\tau > n) = (1 - p)^n, \quad p = \mathbb{P}(X_1 \notin S \mid X_0 \sim \nu).$$

  It is memoryless:

  $$\mathbb{P}(\tau > n + k \mid \tau > k) = \mathbb{P}(\tau > n).$$
Figure: Serial simulation of an exit event from $S$, starting from $\nu$. 
Figure: Parallel simulation of an exit event from $S$, starting from $\nu$. 
Figure: Parallel simulation of an exit event from $S$, with a tie.
Figure: Parallel simulation of an exit event from $S$, with a tiebreaker.
**Parallel Step.** Choose a polling time $T_{poll}$, set $M = 1$ and iterate:

1. Start $N$ i.i.d. replicas of $(X_n)$ at independent samples of $\nu$.
2. Evolve the replicas from time $(M - 1)T_{poll}$ to time $MT_{poll}$.
3. If no replica leaves $S$ during this time, update $M = M + 1$ and return to 2.
   Else, stop and compute an exit time, $\tau_{acc}$, and exit point, $X_{acc}$.
The Parallel Step

\[ (M - 1)T_{poll} \]

\[ 2T_{poll} \]

\[ \cdots \]

\[ (M - 1)T_{poll} \]

\[ MT_{poll} \]

\[ T_{poll} \]

replica 1

replica 2

\vdots

replica \( K - 1 \)

replica \( K \)

replica \( K + 1 \)

\vdots

replica \( N \)

\[ NT_{poll} \]

\[ 2NT_{poll} \]

\[ \cdots \]

\[ (M - 1)NT_{poll} \]

\[ \tau_{acc} \]
The Parallel Step

\[ T_{\text{poll}} \quad 2T_{\text{poll}} \quad \ldots \quad (M-1)T_{\text{poll}} \quad MT_{\text{poll}} \]

replica 1

replica 2

\vdots

replica \( K - 1 \)

replica \( K \)

replica \( K + 1 \)

\vdots

replica \( N \)

0 \quad NT_{\text{poll}} \quad 2NT_{\text{poll}} \quad \ldots \quad (M-1)NT_{\text{poll}} \quad \tau_{\text{acc}}
The Parallel Step

\[ T_{\text{poll}} \quad 2T_{\text{poll}} \quad \ldots \quad (M - 1)T_{\text{poll}} \quad MT_{\text{poll}} \]

\[
\begin{array}{c}
\text{replica 1} \\
\text{replica 2} \\
\vdots \\
\text{replica } K - 1 \\
\text{replica } K \\
\text{replica } K + 1 \\
\vdots \\
\text{replica } N
\end{array}
\]

\[
\begin{array}{c}
0 \\
NT_{\text{poll}} \\
2NT_{\text{poll}} \\
\vdots \\
(M - 1)NT_{\text{poll}} \\
\tau_{\text{acc}}
\end{array}
\]
The Parallel Step

\[ (M - 1)T_{poll} \]

\[ M T_{poll} \]

replica 1
replica 2
::
replica \( K - 1 \)
replica \( K \)
replica \( K + 1 \)
::
replica \( N \)

\[ 0 \]
\[ NT_{poll} \]
\[ 2NT_{poll} \]
\[ \cdots \]
\[ (M - 1)NT_{poll} \]
\[ \tau_{acc} \]
The Parallel Step

\[ T_{poll} \quad 2T_{poll} \quad \ldots \quad (M - 1)T_{poll} \quad MT_{poll} \]

\[ \text{replica 1} \]
\[ \text{replica 2} \]
\[ \vdots \]
\[ \text{replica } K - 1 \]
\[ \text{replica } K \]
\[ \text{replica } K + 1 \]
\[ \vdots \]
\[ \text{replica } N \]

\[ 0 \quad NT_{poll} \quad 2NT_{poll} \quad \ldots \quad (M - 1)NT_{poll} \quad \tau_{acc} \]

D. Aristoff (Colorado State University)
March 2015 15 / 29
The Parallel Step

\[ (M - 1)T_{\text{poll}} \]  \[ 2T_{\text{poll}} \]  \[ \cdots \]  \[ (M - 1)T_{\text{poll}} \]  \[ MT_{\text{poll}} \]

\[ \tau^K \]

\[ \tau_{\text{acc}} = \text{sum of lengths of blue lines} = (M - 1)NT_{\text{poll}} + (K - 1)T_{\text{poll}} + [\tau^K - (M - 1)T_{\text{poll}}] \]

\[ X_{\text{acc}} = \text{position of } K\text{-th replica at red cross (time } \tau^K) \]
Parallel Step. Choose a polling time $T_{poll}$, set $M = 1$ and iterate the following.

1. Start $N$ i.i.d. replicas of $(X_n)$ at independent samples of $\nu$.
2. Evolve the replicas from time $(M - 1)T_{poll}$ to time $MT_{poll}$.
3. If no replica leaves $S$ during this time, update $M = M + 1$ and return to 2.
   Else, record the exit event $(\tau_{acc}, X_{acc})$.

Theorem.

The Parallel Step is exact:

$$(\tau_{acc}, X_{acc}) \sim (\tau, X_\tau),$$

with $\tau$ and $X_\tau$ the first exit time and exit point from $S$, starting at $\nu$. 
Let $S$ be a collection of disjoint subsets of state space.
Let $T_{corr} = T_{corr}(S)$ be $\approx$ the time it takes to reach the QSD in $S$.

**Main Algorithm.**

1. Evolve $(X_n)$ until it spends time $T_{corr}$ in some set $S \in S$ without leaving.
2. Run the **Parallel Step** in $S$. Then, return to 1.

Let $\Pi$ be the quotient map which mods out state space by $S$.

**Coarse dynamics.**

The **Main Algorithm** produces an approximation $(\tilde{X}_n)$ of $(\Pi(X_n))$ via

\[
\tilde{X}_n = \Pi(X_n) \text{ in Step 1,} \quad \tilde{X}_n = \Pi(S) \text{ in Step 2.}
\]
Main Algorithm.

1. Evolve \((X_n)\) until it spends time \(T_{\text{corr}}\) in some set \(S \in S\) without leaving.
2. Run the Parallel Step in \(S\). Then, return to 1.

- Let \(f\) be a function on state space and \(\pi\) the equilibrium distribution of \((X_n)\).

Equilibrium averages.

The Main Algorithm produces an approximation \(\frac{f_{\text{sim}}}{T_{\text{sim}}}\) of \(\int f \, d\pi\) via

\[
f_{\text{sim}} = f_{\text{sim}} + f(X_n) \text{ in Step 1, } \quad f_{\text{sim}} = f_{\text{sim}} + f_{\text{acc}} \text{ in Step 2},
\]

where \(f_{\text{acc}}\) is the sum of values of \(f\) over replicas in the Parallel Step. In Step 1, \(T_{\text{sim}}\) increases by 1 for each time step; in Step 2, \(T_{\text{sim}}\) is increased by \(\tau_{\text{acc}}\).
The Main Algorithm

\( \tau^K \) = \text{sum of lengths of blue lines}

\( f_{acc} = \text{sum of values of } f \text{ over positions of replicas along blue lines} \)
Idealization.
After spending $T_{corr}$ consecutive time steps in $S$, $(X_n)$ exactly reaches the QSD.

Assumption.
$(X_n)$ is uniformly ergodic:

$$\lim_{n \to \infty} \sup_x \| \mathbb{P}(X_n \in \cdot | X_0 = x) - \pi \|_{TV} = 0.$$

Theorem.
Under the Idealization, the **Main Algorithm** generates exact coarse dynamics:

$$(\tilde{X}_n) \sim (\Pi(X_n)).$$

Under the Idealization and Assumption, the **Main Algorithm** generates exact equilibrium averages:

$$\lim_{T_{sim} \to \infty} \frac{f_{sim}}{T_{sim}} \overset{a.s.}{=} \int f \, d\pi.$$
Figure: A random walk with entropic barriers. At each step it moves up, down, left or right at random, unless that results in crossing a barrier.
Figure: Equilibrium averages vs. $T_{corr}$ when $N = 100$. Here $f = 1_{(x,y) \in \text{upper half of right box}}$. 
Figure: Parallel step stats when $T_{corr}$ = largest value from last slide and $T_{sim} = 5 \times 10^9$. 

D. Aristoff (Colorado State University)
Figure: A random walk with energetic barriers. At each step, it moves to the left with probability $1/2 + \text{slope}$, and to the right with probability $1/2 - \text{slope}$. 
\textbf{Figure:} Equilibrium averages vs. $T_{corr}$ when $N = 100$. Here, $f = 1_{y \in \text{right half of state space}}$. 

D. Aristoff (Colorado State University)
Figure: Parallel step stats when $T_{corr} =$ largest value from last slide and $T_{sim} = 10^9$. 
Remarks.

- Our algorithm applies to any Markov chain.
- It is efficient when for $S \in \mathcal{S}$,

  \[
  \text{Time scale to reach QSD in } S \ll \text{Time scale to leave } S.
  \]

- $S$ need not be known a priori. Sets in $S$ can be identified on the fly.

Future work.

- $(X_n)$ only approximately reaches $\nu$. How does this error propagate?
- When can $S$ be efficiently identified on the fly?
Acknowledgements

Collaborators  T. Lelièvre (École des Ponts ParisTech), G. Simpson (Drexel University)

Funding  DOE Award de-sc0002085 (Thanks to Mitch Luskin for the support)

http://www.math.colostate.edu/~aristoff/