# Theoretical and computational approaches to parallel replica dynamics

Mitchell Luskin

School of Mathematics
University of Minnesota

June 7, 2012

Collaborators C. Le Bris (CERMICS), T. Lelievre (CERMICS), D. Perez (LANL), G. Simpson (UMN), A.F. Voter (LANL), A. Binder (UMN)

# Outline

# The Parallel Replica Algorithm

The Parallel Replica Algorithm proposed by A.F. Voter in 1998 is a method to accelerate a "coarse-grained projection" of a dynamics. We consider the overdamped Langevin dynamics:

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}dW_t$$

and we assume that we are given a smooth mapping

$$\mathcal{S} : \mathbb{R}^d \to \mathbb{N}$$

which to a configuration in $\mathbb{R}^d$ associates a state number (e.g., a numbering of the wells of the potential $V$).

The goal of the parallel replica dynamics is to generate very efficiently a trajectory $(S_t)_{t\geq 0}$ which has (almost) the same law as $(\mathcal{S}(X_t))_{t\geq 0}$.

# The Parallel Replica Algorithm

Initialization: Consider an initial condition $X_0^{ref}$ for a reference walker, the associated initial state $S_0 = \mathcal{S}(X_0^{ref})$, and a simulation time counter $T_{simu} = 0$.

One iteration of the algorithm goes through three steps.

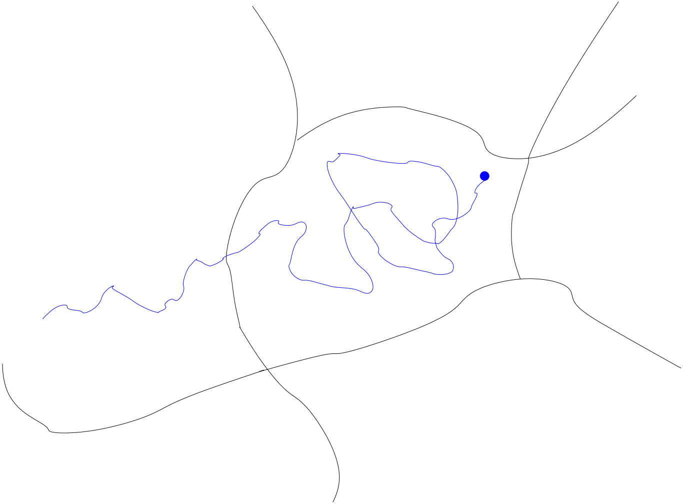
- The decorrelation step: Let the reference walker $(X_{T_{simu}+t}^{ref})_{t \geq 0}$ evolve over a time interval $t \in [0, \tau_{corr}]$. Then,
  - If the process leaves the well during the time interval (*i.e.*, $\exists t \leq \tau_{corr}$ such that $\mathcal{S}\left(X_{T_{simu}+t}^{ref}\right) \neq \mathcal{S}\left(X_{T_{simu}}^{ref}\right)$) advance the simulation clock by $\tau_{corr}$ and restart the decorrelation step ;
  - otherwise, advance the simulation clock by $\tau_{corr}$ and proceed to the dephasing step.

# The Parallel Replica Algorithm

The reference walker enter a new state

# The Parallel Replica Algorithm

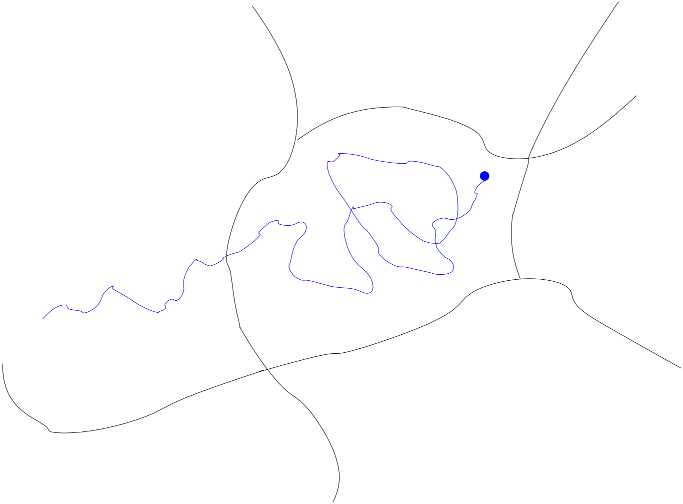
Decorrelation step: wait for a time $\tau_{corr}$.



- The dephasing step: Duplicate the walker $X^{ref}_{T_{simu}}$ into $N$ replicas. Let these replicas evolve independently and in parallel over a time interval of length $\tau_{dephase}$. If a replica leaves the well during this time interval, restart the dephasing step for this replica. Throughout this step, the simulation counter is stopped.

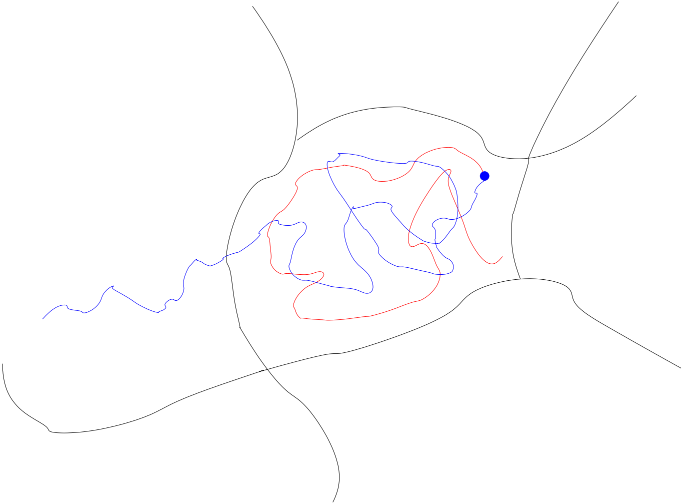
# The Parallel Replica Algorithm

Dephasing step.



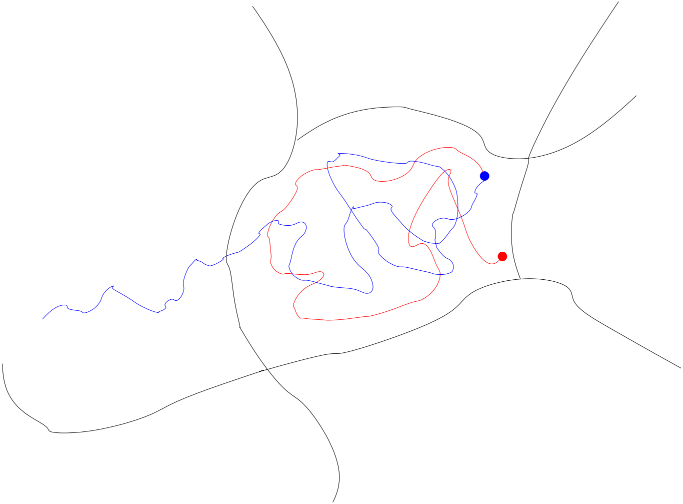
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



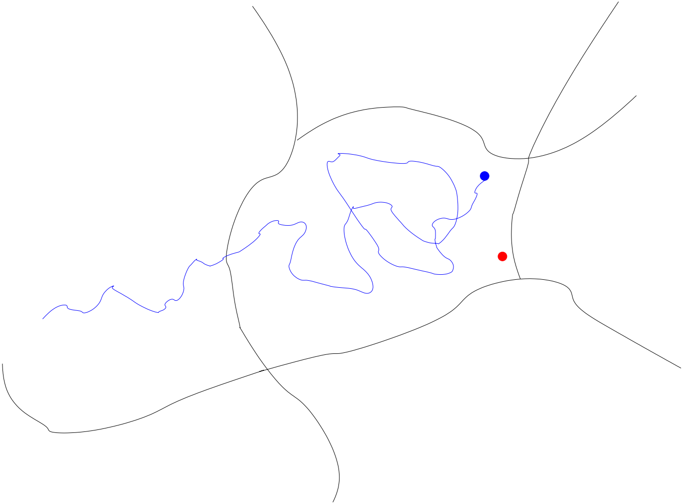
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



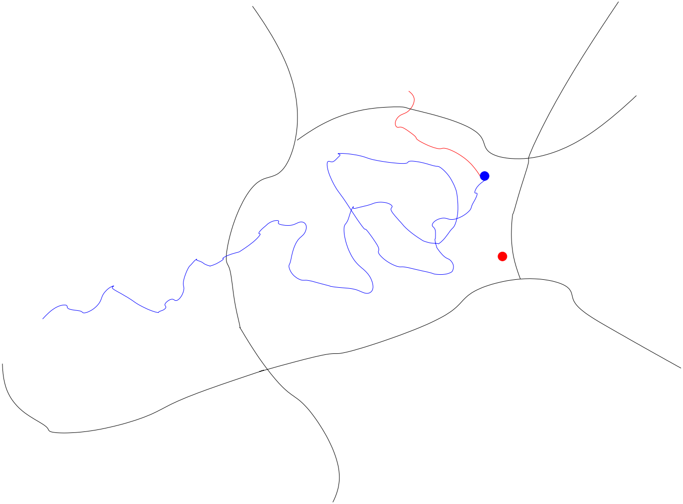
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



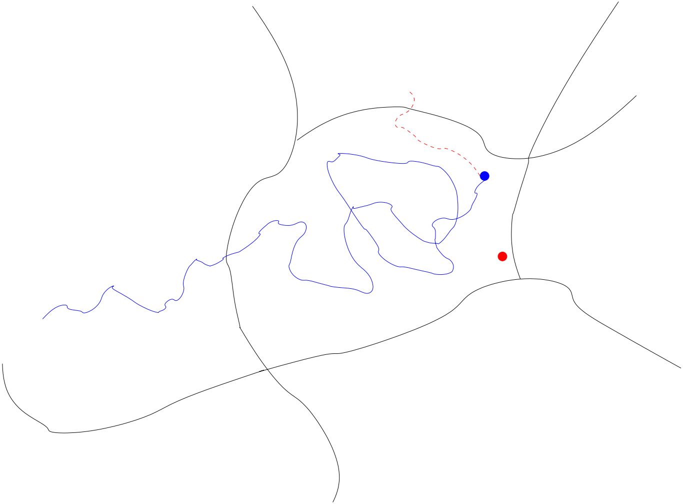
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



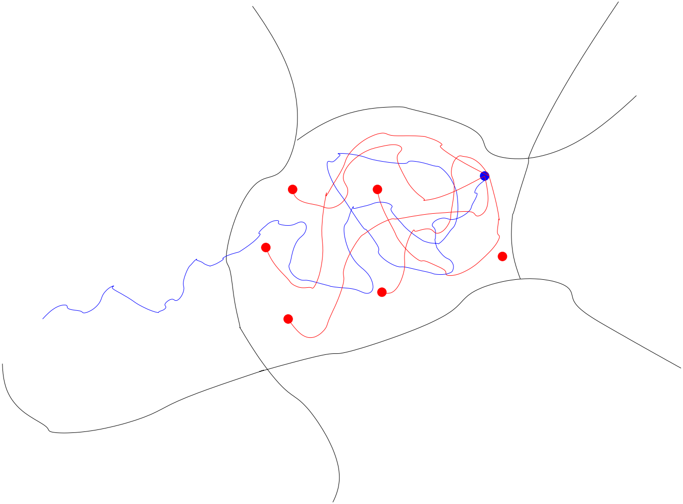
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



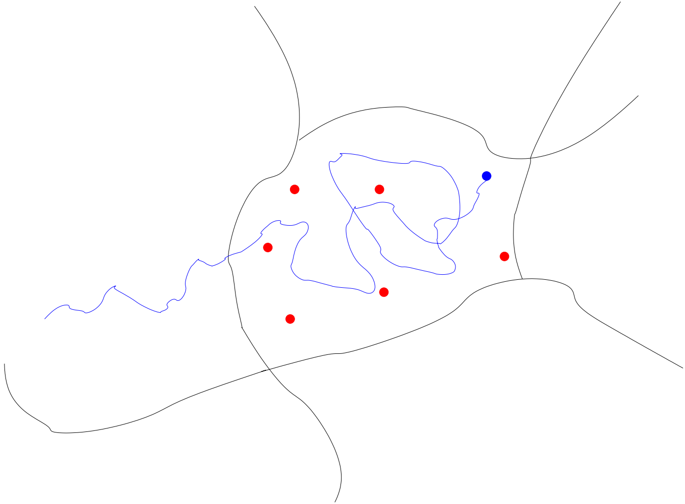
# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



# The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



- The parallel step: Let all the replicas evolve independently and track the first escape event:

$$T = \inf_k T_W^k = T_W^{K_0},$$

where $K_0 = \arg\inf_k T_W^k$ and

$$T_W^k = \inf\{t \geq 0, \mathcal{S}(X_{T_{simu}+t}^k) \neq \mathcal{S}(X_{T_{simu}}^k)\}$$

is the first time the $k$-th replica leaves the well. Then:

$$T_{simu} = T_{simu} + NT \text{ and } X_{T_{simu}+NT}^{ref} = X_{T_{simu}+T}^{K_0}.$$

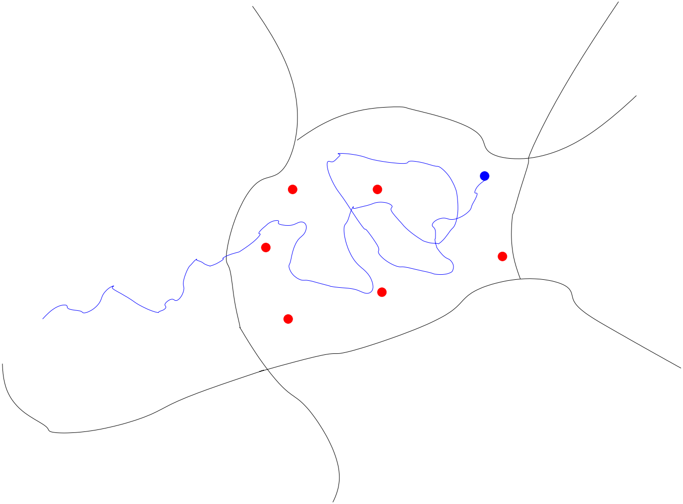
Moreover, over $[T_{simu}, T_{simu} + NT]$, the state dynamics $S_t$ is constant and defined as:

$$S_t = \mathcal{S}(X_{T_{simu}}^1).$$

Then, go back to the decorrelation step...

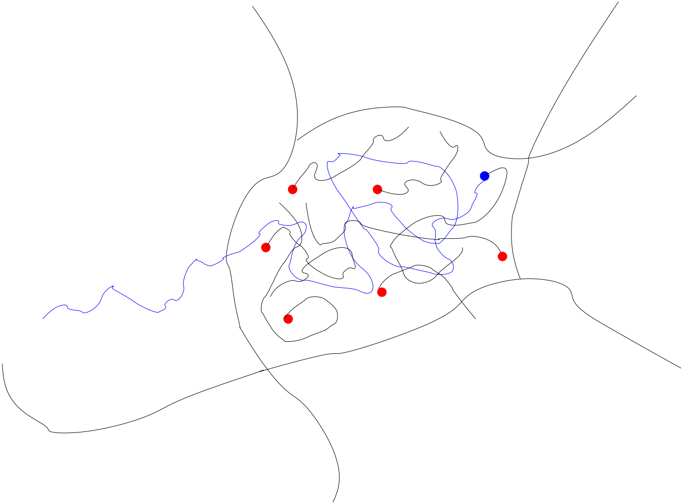
# The Parallel Replica Algorithm

Parallel step.



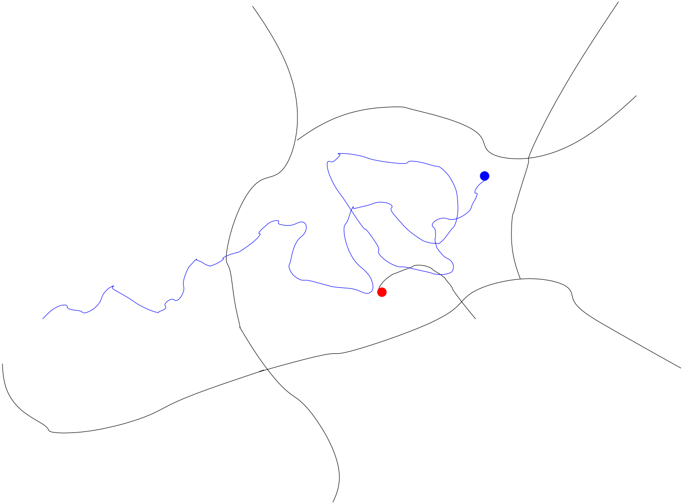
# The Parallel Replica Algorithm

Parallel step: run independent trajectories in parallel...



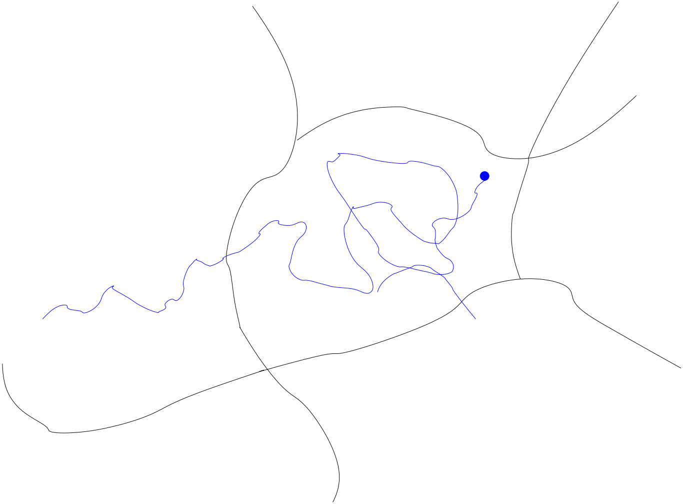
# The Parallel Replica Algorithm

Parallel step: ... and detect the first transition event.



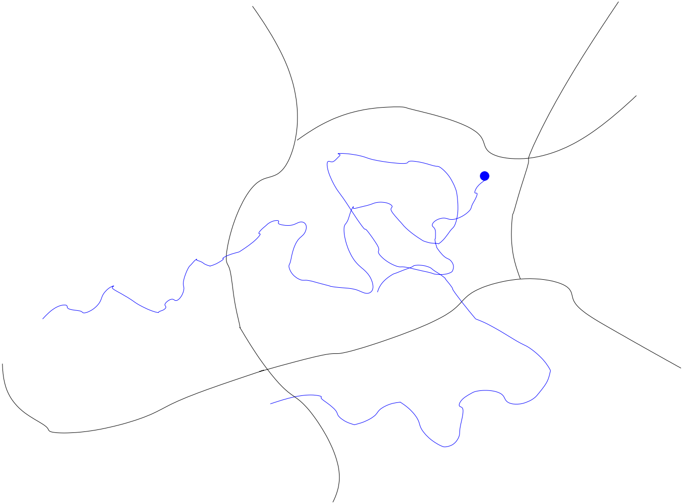
# The Parallel Replica Algorithm

Parallel step: update the time clock: $T_{simu} = T_{simu} + NT$.



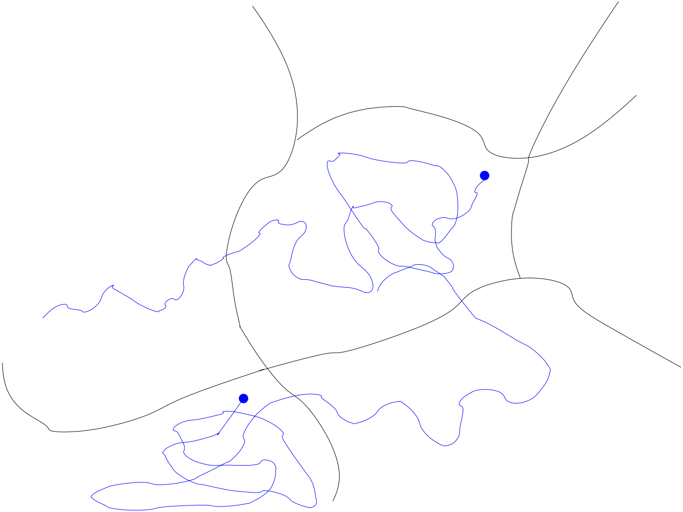
# The Parallel Replica Algorithm

A new decorrelation step starts...



# The Parallel Replica Algorithm

New decorrelation step



# Error analysis for the Parallel Replica Algorithm

The parallel step would introduce no error if

- the escape time $T_W^1$ was exponentially distributed
- and independent of the next visited state.

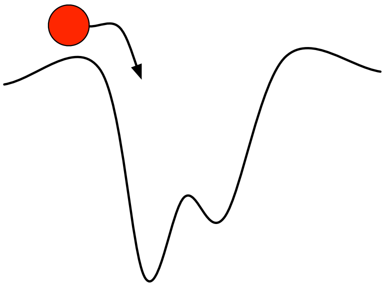
How can we analyze the error introduced by the algorithm ?

# Parallel Replica Dynamics – Steps

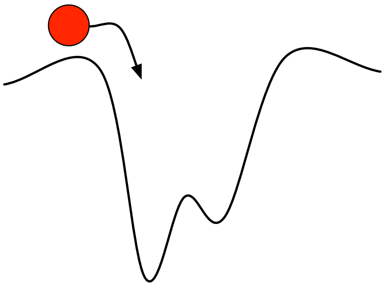
## Escaping a Single Well

1. Decorrelation Step – Let a reference process sample a well for some time
2. Dephasing Step – Simultaneously create independent replicas that further sample the well
3. Parallel Step – Run the replicas until one exits the well

# Decorrelating the Reference Process



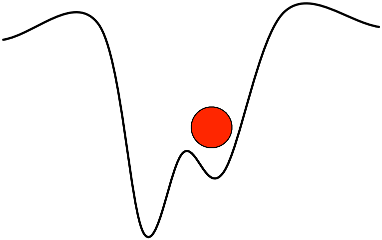
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\rm corr}$

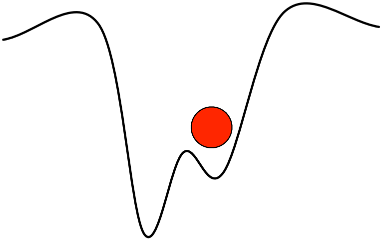
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\mathrm{corr}}$
- If $X_t$ leaves the well, begin again, in the new well

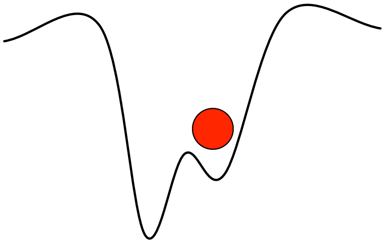
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\rm corr}$
- If $X_t$ leaves the well, begin again, in the new well
- $t_{\rm corr}$ must be long enough for it to "forget" its previous state

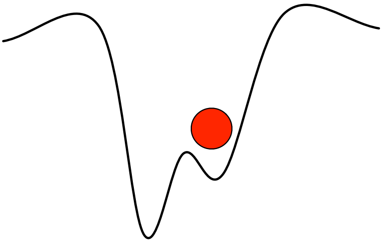
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\mathrm{corr}}$
- If $X_t$ leaves the well, begin again, in the new well
- $t_{\mathrm{corr}}$ must be long enough for it to "forget" its previous state
- If $t_{\mathrm{corr}}$ is too long, it reproduces a serial computation

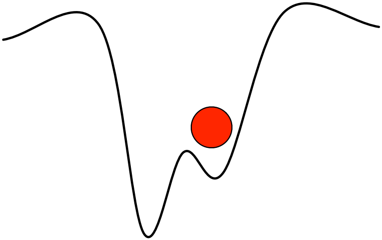
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\mathrm{corr}}$
- If $X_t$ leaves the well, begin again, in the new well
- $t_{\mathrm{corr}}$ must be long enough for it to "forget" its previous state
- If $t_{\mathrm{corr}}$ is too long, it reproduces a serial computation
- $t_{\mathrm{corr}}$ is one of the user parameters

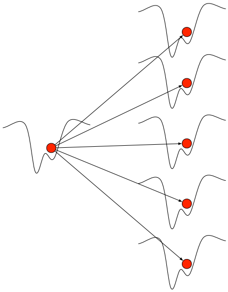
# Decorrelating the Reference Process



## Structure of the Decorrelation Step

- Run for $t \leq t_{\mathrm{corr}}$
- If $X_t$ leaves the well, begin again, in the new well
- $t_{\mathrm{corr}}$ must be long enough for it to "forget" its previous state
- If $t_{\mathrm{corr}}$ is too long, it reproduces a serial computation
- $t_{\mathrm{corr}}$ is one of the user parameters
- Simulation clock is advanced by $t_{\mathrm{corr}}$

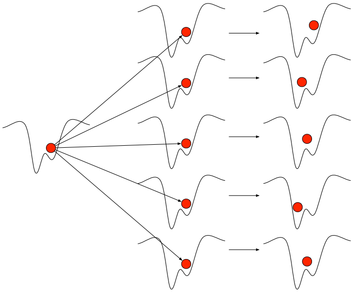
# Dephasing by Direct Simulation



## Structure of the Dephasing Step

- Replicas begin running at $t_{\mathrm{launch}} < t_{\mathrm{corr}}$ ($t_{\mathrm{launch}}$ can be zero, as in the previous Par Rep version)

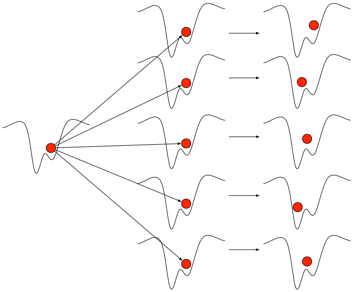
# Dephasing by Direct Simulation



## Structure of the Dephasing Step

- Replicas begin running at $t_{\mathrm{launch}} < t_{\mathrm{corr}}$ ($t_{\mathrm{launch}}$ can be zero, as in the previous Par Rep version)
- Run replicas for $t_{\mathrm{launch}} \leq t \leq t_{\mathrm{launch}} + t_{\mathrm{phase}} = t_{\mathrm{corr}}$

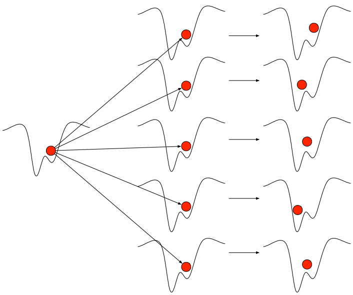
# Dephasing by Direct Simulation



## Structure of the Dephasing Step

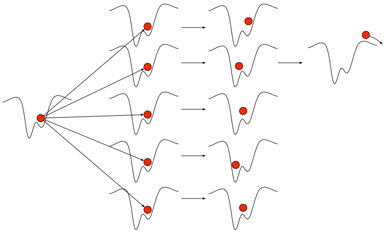
- Replicas begin running at $t_{\text{launch}} < t_{\text{corr}}$ ($t_{\text{launch}}$ can be zero, as in the previous Par Rep version)
- Run replicas for $t_{\text{launch}} \leq t \leq t_{\text{launch}} + t_{\text{phase}} = t_{\text{corr}}$
- If $X_t^k$ leaves the well, restart it

# Dephasing by Direct Simulation



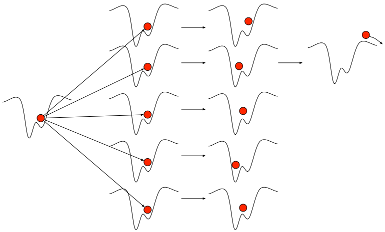
## Structure of the Dephasing Step

- Replicas begin running at $t_{\text{launch}} < t_{\text{corr}}$ ($t_{\text{launch}}$ can be zero, as in the previous Par Rep version)
- Run replicas for $t_{\text{launch}} \leq t \leq t_{\text{launch}} + t_{\text{phase}} = t_{\text{corr}}$
- If $X_t^k$ leaves the well, restart it
- $t_{\text{launch}}$ and $t_{\text{phase}}$ are other user parameters



## Structure of the Parallel Step

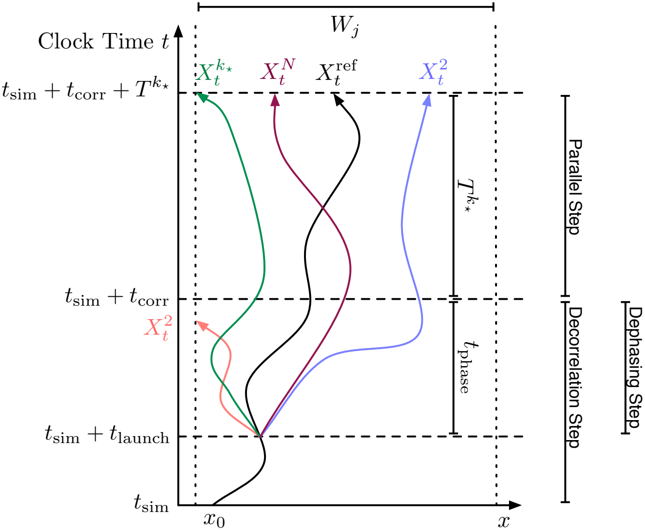
- The first process $k_*$ to leave the well, at time $T_{\rm exit} := T_{k_*}$, becomes the new reference process, and the algorithm restarts



## Structure of the Parallel Step

- The first process $k_*$to leave the well, at time $T_{\mathrm{exit}} := T_{k_*}$, becomes the new reference process, and the algorithm restarts
- The simulation clock is advanced by $NT_{\mathrm{exit}}$

# Parallel Replica Dynamics – Recap



# Goals

- Given that we begin in well $W \subset \mathbb{R}^n$, determine the properties of $T_{\mathrm{exit}}$, the first exit time from $W$

# Goals

- Given that we begin in well $W \subset \mathbb{R}^n$, determine the properties of $T_{\mathrm{exit}}$, the first exit time from $W$
- What is the distribution for $T_{\mathrm{exit}}$?

# Goals

- Given that we begin in well $W \subset \mathbb{R}^n$, determine the properties of $T_{\mathrm{exit}}$, the first exit time from $W$
- What is the distribution for $T_{\mathrm{exit}}$?
- What are the properties of $X_{T_{\mathrm{exit}}}$, the first hitting point distribution?

# Goals

- Given that we begin in well $W \subset \mathbb{R}^n$, determine the properties of $T_{\mathrm{exit}}$, the first exit time from $W$
- What is the distribution for $T_{\mathrm{exit}}$?
- What are the properties of $X_{T_{\mathrm{exit}}}$, the first hitting point distribution?
- Can we estimate the accuracy of ParRep?
- Can we optimize the efficiency of ParRep?

# Fokker-Planck Equation

The Fokker-Planck Equation for the overdamped Langevin equation $dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dB_t$ and absorbing boundary conditions:

$$\frac{\partial \rho}{\partial t} = L^* \rho := \nabla \cdot \left[ (\nabla V)\, \rho + \beta^{-1} \nabla \rho \right] \qquad \forall x \in W, \ t \geq 0,$$

$$\rho(x,t) = 0 \qquad \forall x \in \partial W, \ t \geq 0,$$

$$\rho(x,0) \geq 0 \quad \forall x \in W, \qquad \int_W \rho(x,0)\, dx = 1,$$

is given by the series expansion

$$\rho(x,t) = \sum_1^\infty a_j e^{-\lambda_j t} \psi_j(x),$$

for eigenvalues $0 < \lambda_1 < \lambda_2 \leq \cdots$ and eigenfunctions $\psi_j(x)$ of

$$L^* \psi_j = \nabla \cdot \left[ (\nabla V)\, \psi_j + \beta^{-1} \nabla \psi_j \right] = -\lambda_j \psi_j \qquad \forall x \in W,$$

$$\psi_j = 0 \qquad \forall x \in \partial W.$$

# The Exit Density

The exit density through the boundary point $x \in \partial W$ at time $t \geq 0$ is

$$\beta^{-1} \frac{\partial \rho}{\partial n}(x, t),$$

the first exit time density is

$$\int_{\partial W} \beta^{-1} \frac{\partial \rho}{\partial n}(x, t)\, dx,$$

and the first hitting point density is

$$\int_0^\infty \beta^{-1} \frac{\partial \rho}{\partial n}(x, t)\, dt.$$

# The Quasistationary Distribution (QSD)

The renormalized density $\rho(x,t)$ converges to $\psi_1(x)$ at rate $\lambda_2 - \lambda_1$ (where $\psi_1(x) > 0$ is normalized by $\int_W \psi_1(x,t)\,dx = 1$):

$$\frac{\rho(x,t)}{\int_W \rho(x,t)\,dx} = \psi_1(x) + O\left(e^{-(\lambda_2-\lambda_1)t}\right) \quad \text{as } t \to \infty.$$

The Fokker-Planck solution $\rho(x,t) = \psi_1(x)e^{-\lambda_1 t}$ has exit density

$$\beta^{-1}\frac{\partial \psi_1}{\partial n}(x)\,e^{-\lambda_1 t} \qquad \forall x \in \partial W,\ \ t \geq 0,$$

with independent exit time and hitting point.

# The Quasistationary Distribution (QSD)

The first exit time density of $\rho(x,t) = \psi_1(x) e^{-\lambda_1 t}$ is exponential:

$$\int_{\partial W} \beta^{-1} \frac{\partial \psi_1}{\partial n}(x)\, e^{-\lambda_1 t}\, dx = \lambda_1 e^{-\lambda_1 t},$$

and independent of the hitting point density:

$$\int_0^\infty \beta^{-1} \frac{\partial \psi_1}{\partial n}(x)\, e^{-\lambda_1 t}\, dt = \frac{1}{\lambda_1 \beta} \frac{\partial \psi_1}{\partial n}(x).$$

# The Quasistationary Distribution (QSD)

### Definition

On well $W$, a QSD is a distribution $\nu$ such that for all $A \subset W$ and $t \geq 0$,

$$\nu(A) = \int_W \mathbb{P}^x \left[X_t \in A \mid t < T_{\text{exit}}\right] d\nu(x). \tag{1}$$

The dephasing stage of the Par Rep Method converges to the QSD as $t_{\text{phase}} \to \infty$. (1) states that the QSD is invariant for the dephasing step.

### Theorem

*$\psi_1(x)\,dx$ is a QSD where $\psi_1(x) > 0$ is the unique ground state of the Fokker-Planck operator with eigenvalues* $0 < \lambda_1 < \lambda_2 \leq \ldots$

$$L^*\psi_j = \nabla \cdot \left[(\nabla V)\,\psi_j + \beta^{-1}\nabla\psi_j\right] = -\lambda_j\psi_j \qquad \forall x \in W,$$
$$\psi_j = 0 \qquad \forall x \in \partial W.$$

# Utility of the QSD

**Theorem**

*Let $X_t^k$ be $N$ i.i.d. processes in the well $W$, and assume:*

- $T_{\mathrm{exit}}^k$ *are exponentially distributed,*
- *Exit time is independent of hitting point.*

*If*

$$T_{\mathrm{exit}} \equiv T_{\mathrm{exit}}^{k_\star}, \quad X_{T_{\mathrm{exit}}} \equiv X_{T_{\mathrm{exit}}^{k_\star}}^{k_\star}, \quad k_\star \equiv \mathrm{argmin}_k\, T_{\mathrm{exit}}^k,$$

*then $NT_{\mathrm{exit}}$ has the same law as $T_{\mathrm{exit}}^k$, and $X_{T_{\mathrm{exit}}}$ is independent of first hitting time.*

# Utility of the QSD

## Theorem

*Let $X_t^k$ be $N$ i.i.d. processes in the well $W$, and assume:*

- *$T_{\mathrm{exit}}^k$ are exponentially distributed,*
- *Exit time is independent of hitting point.*

*If*

$$T_{\mathrm{exit}} \equiv T_{\mathrm{exit}}^{k_\star}, \quad X_{T_{\mathrm{exit}}} \equiv X_{T_{\mathrm{exit}}^{k_\star}}^{k_\star}, \quad k_\star \equiv \mathrm{argmin}_k\, T_{\mathrm{exit}}^k,$$

*then $NT_{\mathrm{exit}}$ has the same law as $T_{\mathrm{exit}}^k$, and $X_{T_{\mathrm{exit}}}$ is independent of first hitting time.*

## QSD and ParRep

Goal of the decorrelation/dephasing step: Produce $N$ processes distributed as close as possible to $\nu$.

# Decorrelation Result

### Theorem

*Let $X_0$ be distributed by $\mu_0$ on $W$, then for any observable $f$*

$$\left|\mathbb{E}^{\mu_t}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\, \|f\|_{L^\infty}\, e^{-(\lambda_2-\lambda_1)t}.$$

*where*

$$d\mu_t(x) := \frac{\rho(x,t)\,dx}{\int_W \rho(x,t)\,dx}.$$

- $d(\mu_t, \nu)$ measures the difference between $\mu_t$ and $\nu$; vanishes as $t \to \infty$.

# Decorrelation Result

**Theorem**

*Let $X_0$ be distributed by $\mu_0$ on $W$, then for any observable $f$*

$$\left|\mathbb{E}^{\mu_t}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\, \|f\|_{L^\infty}\, e^{-(\lambda_2-\lambda_1)t}.$$

*where*

$$d\mu_t(x) := \frac{\rho(x,t)\,dx}{\int_W \rho(x,t)\,dx}.$$

- $d(\mu_t, \nu)$ measures the difference between $\mu_t$ and $\nu$; vanishes as $t \to \infty$.
- Exponential convergence with decorrelation time scale is

$$\frac{1}{\lambda_2 - \lambda_1}$$

# Decorrelation Result

### Theorem

*Let $X_0$ be distributed by $\mu_0$ on $W$, then for any observable $f$*

$$\left|\mathbb{E}^{\mu_t}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\left\|f\right\|_{L^\infty} e^{-(\lambda_2-\lambda_1)t}.$$

*where*

$$d\mu_t(x) := \frac{\rho(x,t)\,dx}{\int_W \rho(x,t)\,dx}.$$

- $d(\mu_t, \nu)$ measures the difference between $\mu_t$ and $\nu$; vanishes as $t \to \infty$.
- Exponential convergence with decorrelation time scale is

$$\frac{1}{\lambda_2 - \lambda_1}$$

- ParRep is efficient when the decorrrelation time scale is much less than the mean first exit time

# Decorrelation Example

- We have

$$\left|\mathbb{E}^{\mu_{t_{\mathrm{corr}}}}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\left\|f\right\|_{L^\infty} e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

# Decorrelation Example

- We have

$$\left|\mathbb{E}^{\mu_{t_{\mathrm{corr}}}}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\left\|f\right\|_{L^\infty} e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

- For any $t \geq 0$, to obtain an error estimate for the first exit time let $f(\tau, \xi) = \chi_{\tau > t}$; then

$$\left|\mathbb{P}^{\mu_{t_{\mathrm{corr}}}}\left[T > t\right] - e^{-\lambda_1 t}\right| \lesssim d(\mu_0, \nu) e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

# Decorrelation Example

- We have

$$\left|\mathbb{E}^{\mu_{t_{\mathrm{corr}}}}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \lesssim d(\mu_0, \nu)\, \|f\|_{L^\infty}\, e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

- For any $t \geq 0$, to obtain an error estimate for the first exit time let $f(\tau, \xi) = \chi_{\tau > t}$; then

$$\left|\mathbb{P}^{\mu_{t_{\mathrm{corr}}}}\left[T > t\right] - e^{-\lambda_1 t}\right| \lesssim d(\mu_0, \nu) e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

- For any $t \geq 0$, to obtain an error estimate for the exit point distribution, let $f(\tau, \xi) = \phi(\xi)$; then

$$\left|\mathbb{E}^{\mu_{\mathrm{corr}}}\left[\phi(X_T) \mid T > t\right] - \int_{\partial W} \phi d\rho\right| \lesssim d(\mu_0, \nu) e^{-(\lambda_2-\lambda_1)t_{\mathrm{corr}}}$$

# Parallel Step Error

## Theorem

*Assume at time $t_{\mathrm{corr}}$, there are $N$ processes $X^k_{t_{\mathrm{corr}}}$ distributed according to $\mu_{\mathrm{corr}}$ and such that*

$$\left|\mathbb{E}^{\mu_{\mathrm{corr}}}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \leq \epsilon_{\mathrm{corr}} \left\|f\right\|_{L^\infty}.$$

*Then for any $\phi : \partial W \to \mathbb{R}$, smooth,*

$$\left|\mathbb{P}^{\mu_{\mathrm{corr}}}\left[T^{k_*} > t\right] - e^{-N\lambda_1 t}\right| \lesssim N\epsilon_{\mathrm{corr}},$$

$$\left|\mathbb{E}^{\mu_{\mathrm{corr}}}\left[\phi(X_{T^{k_*}}) \mid T^{k_*} > t\right] - \int_{\partial W} \phi \, d\rho\right| \lesssim N \left\|\phi\right\|_{L^\infty} \epsilon_{\mathrm{corr}} e^{N\lambda_1 t}.$$

# Parallel Step Error

## Theorem

*Assume at time $t_{\mathrm{corr}}$, there are $N$ processes $X^k_{t_{\mathrm{corr}}}$ distributed according to $\mu_{\mathrm{corr}}$ and such that*

$$\left|\mathbb{E}^{\mu_{\mathrm{corr}}}\left[f(T, X_T)\right] - \mathbb{E}^{\nu}\left[f(T, X_T)\right]\right| \leq \epsilon_{\mathrm{corr}} \left\|f\right\|_{L^\infty}.$$

*Then for any $\phi : \partial W \to \mathbb{R}$, smooth,*

$$\left|\mathbb{P}^{\mu_{\mathrm{corr}}}\left[T^{k_*} > t\right] - e^{-N\lambda_1 t}\right| \lesssim N\epsilon_{\mathrm{corr}},$$

$$\left|\mathbb{E}^{\mu_{\mathrm{corr}}}\left[\phi(X_{T^{k_*}}) \mid T^{k_*} > t\right] - \int_{\partial W}\phi d\rho\right| \lesssim N \left\|\phi\right\|_{L^\infty} \epsilon_{\mathrm{corr}} e^{N\lambda_1 t}.$$

- Factor of $N$ speedup

# Remarks on ParRep

- ParRep converges as $t_{\mathrm{corr}} \to \infty$, shrinking $\epsilon_{\mathrm{corr}} \to 0$ over a single well,

# Remarks on ParRep

- ParRep converges as $t_{\mathrm{corr}} \to \infty$, shrinking $\epsilon_{\mathrm{corr}} \to 0$ over a single well,
- Computing $\lambda_2$ and $\lambda_1$ directly from the parabolic problem is non-trivial/intractable in high dimensional systems,

$$L^* \psi_j = \nabla \cdot \left[ (\nabla V)\, \psi_j + \beta^{-1} \nabla \psi_j \right] = -\lambda_j \psi_j \qquad \forall x \in W,$$

$$\psi_j = 0 \qquad \forall x \in \partial W.$$

# Remarks on ParRep

- ParRep converges as $t_{\mathrm{corr}} \to \infty$, shrinking $\epsilon_{\mathrm{corr}} \to 0$ over a single well,
- Computing $\lambda_2$ and $\lambda_1$ directly from the parabolic problem is non-trivial/intractable in high dimensional systems,

$$L^* \psi_j = \nabla \cdot \left[ (\nabla V)\, \psi_j + \beta^{-1} \nabla \psi_j \right] = -\lambda_j \psi_j \qquad \forall x \in W,$$
$$\psi_j = 0 \qquad \forall x \in \partial W.$$

- Currently investigating ways of approximating $\lambda_2 - \lambda_1$ on-the-fly.

# Remarks on ParRep

- ParRep converges as $t_{\mathrm{corr}} \to \infty$, shrinking $\epsilon_{\mathrm{corr}} \to 0$ over a single well,
- Computing $\lambda_2$ and $\lambda_1$ directly from the parabolic problem is non-trivial/intractable in high dimensional systems,

$$L^* \psi_j = \nabla \cdot \left[ (\nabla V) \psi_j + \beta^{-1} \nabla \psi_j \right] = -\lambda_j \psi_j \qquad \forall x \in W,$$
$$\psi_j = 0 \qquad \forall x \in \partial W.$$

- Currently investigating ways of approximating $\lambda_2 - \lambda_1$ on-the-fly.
- Study of error over many cycles is underway.

# Remarks on ParRep

- ParRep converges as $t_{\mathrm{corr}} \to \infty$, shrinking $\epsilon_{\mathrm{corr}} \to 0$ over a single well,
- Computing $\lambda_2$ and $\lambda_1$ directly from the parabolic problem is non-trivial/intractable in high dimensional systems,

$$
\begin{aligned}
L^* \psi_j = \nabla \cdot \left[ (\nabla V)\, \psi_j + \beta^{-1} \nabla \psi_j \right] = -\lambda_j \psi_j \qquad & \forall x \in W, \\
\psi_j = 0 \qquad \forall x \in \partial W. &
\end{aligned}
$$

- Currently investigating ways of approximating $\lambda_2 - \lambda_1$ on-the-fly.
- Study of error over many cycles is underway.
- Study of Langevin and other generalizations is underway.

# Toy Problem

## Set Up

- $$V(x) = -k\cos(\pi x).$$
- Wells boundaries at odd integers, centered at even integers.
- $\beta = 1$.
- $\mu_0 = \delta_0$.
- After decorrelating a single trajectory, the QSD is sampled exactly.

# Toy Problem

## Set Up

- $$V(x) = -k\cos(\pi x).$$
- Wells boundaries at odd integers, centered at even integers.
- $\beta = 1$.
- $\mu_0 = \delta_0$.
- After decorrelating a single trajectory, the QSD is sampled exactly.

## Questions

- For what values of $k$ will there be a spectral gap?

# Toy Problem

## Set Up

- $$V(x) = -k\cos(\pi x).$$
- Wells boundaries at odd integers, centered at even integers.
- $\beta = 1$.
- $\mu_0 = \delta_0$.
- After decorrelating a single trajectory, the QSD is sampled exactly.

## Questions

- For what values of $k$ will there be a spectral gap?
- How does $t_{\mathrm{corr}}$ alter the hitting time, $X_T$, in well at $\pm 10$.

# Toy Problem

## Set Up

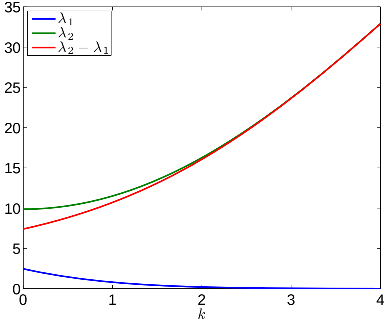
- $$V(x) = -k\cos(\pi x).$$
- Wells boundaries at odd integers, centered at even integers.
- $\beta = 1$.
- $\mu_0 = \delta_0$.
- After decorrelating a single trajectory, the QSD is sampled exactly.

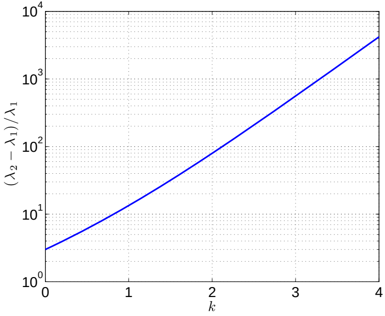
## Questions

- For what values of $k$ will there be a spectral gap?
- How does $t_{\mathrm{corr}}$ alter the hitting time, $X_T$, in well at $\pm 10$.
- How well does ParRep perform?

# Time Scale Separation

$V(x) = -k\cos(\pi x)$



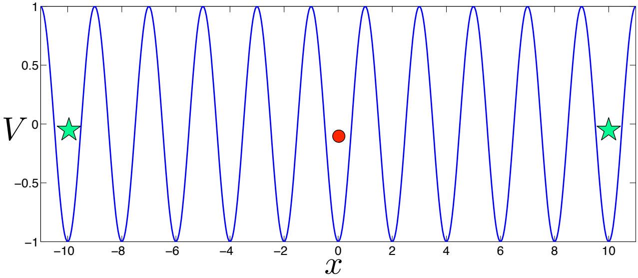

- Scale separations exist

# Rapid Convergence to the QSD

(Loading...)

- $V(x) = -2\cos(\pi x)$, $\beta = 1$.
- $W = (-1, 1)$.
- Initial distribution is $\delta_0(x)$.

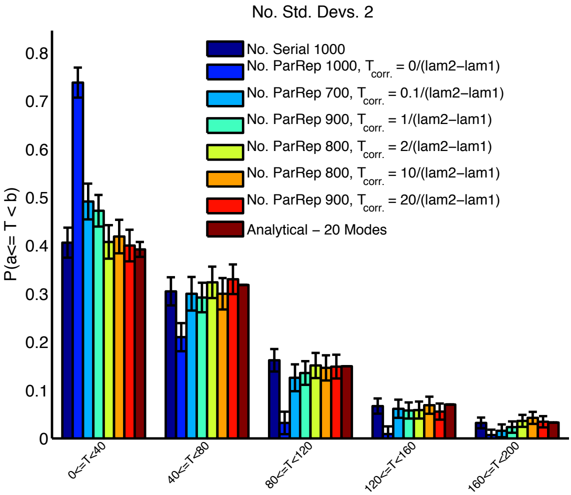
# First Exit Problem – Many Wells




- Process ends if $X_t$ enters either well at $\pm 10$.
- Run a full step of ParRep (Decorrelation, Dephasing, Parallel) every time a new well is entered.
- Dephasing is conducted "analytically" from the QSD.

# Hitting Time Distribution, $k = 1$

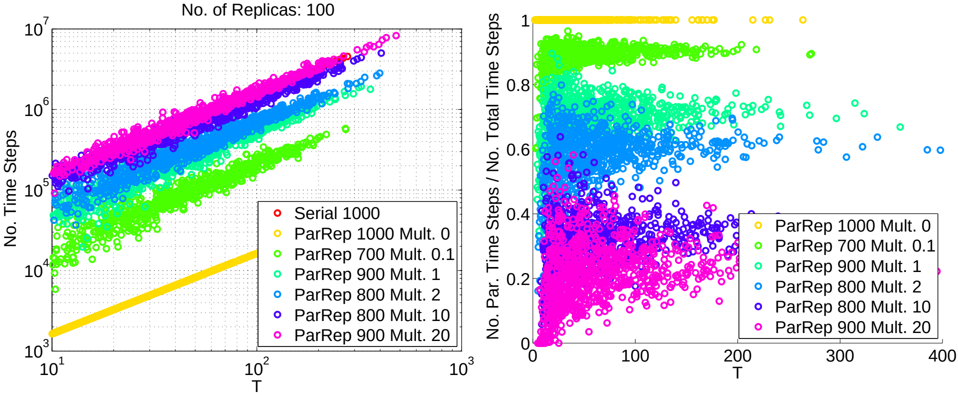
## $V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$




- Time scale separation $\sim 10$
- Cases with $t_{\mathrm{corr}} < 2/(\lambda_2 - \lambda_1)$ give poor results

# Performance, $k = 1$

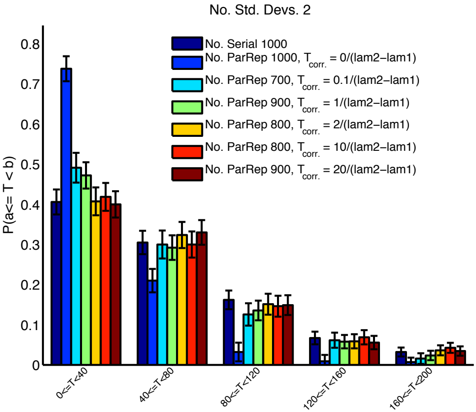
## $V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$

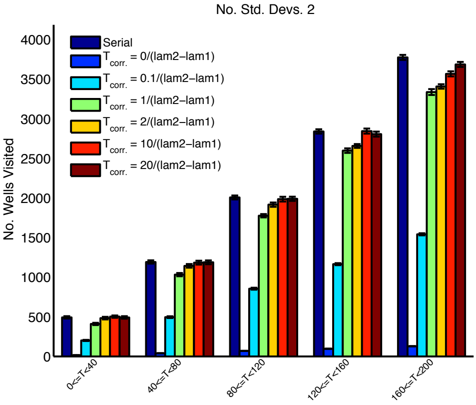



- For small separation of time scales, $\sim 10$, minimal speedup

# Number of Wells Visited, $k = 1$

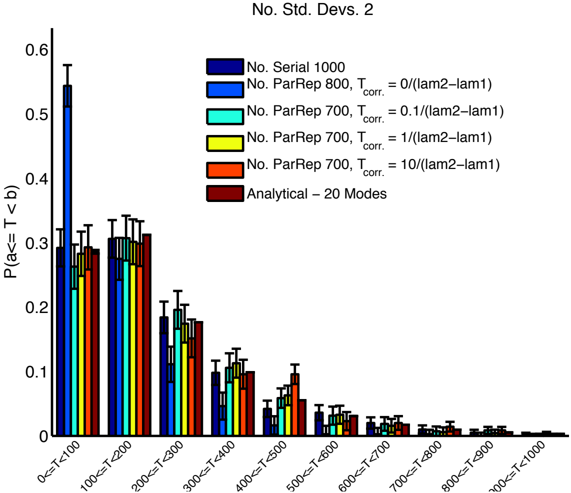
## $V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$



# Hitting Time Distribution, $k = 2$

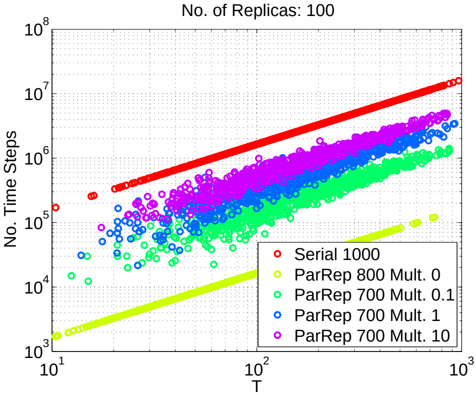
## $V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$

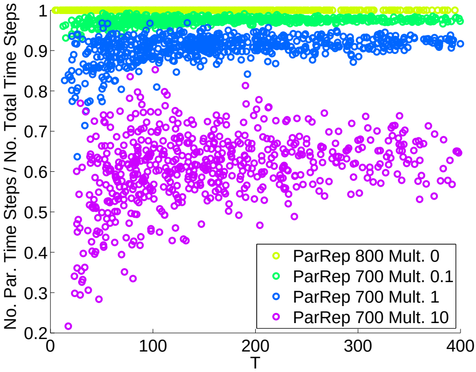



- Time scale separation $\sim 80$
- Only $T_{\mathrm{corr}} = 0$ gives poor results

# Performance, $k = 2$

$V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$

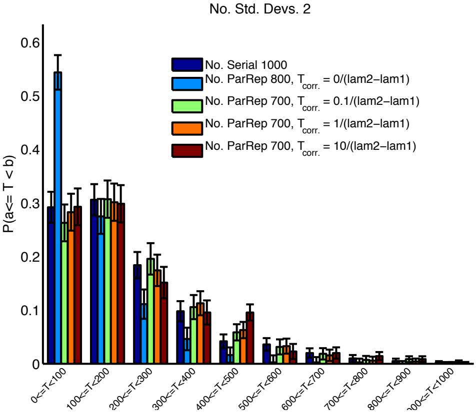


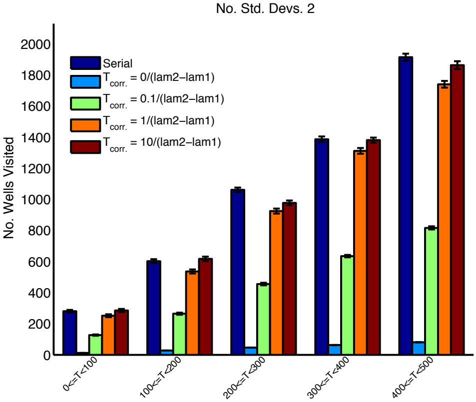




- For larger separation of time scales, $\sim 80$, speedup approaches theoretical factor of $N = 100$.

# Number of Wells Visited, $k = 2$

## $V(x) = -k\cos(\pi x)$, Target Wells $\pm 10$




- Despite agreement in the exit time distributions, there may be disagreements in the distribution in the number of wells visited

# Remarks on Results

- ParRep appears to converge as $t_{\mathrm{corr}} \to \infty$ over many wells

# Remarks on Results

- ParRep appears to converge as $t_{\mathrm{corr}} \to \infty$ over many wells
- ParRep is more efficient when the scale separation is large

# Remarks on Results

- ParRep appears to converge as $t_{\mathrm{corr}} \to \infty$ over many wells
- ParRep is more efficient when the scale separation is large
- Convergence of the hitting time distribution may not imply convergence of the state to state dynamics

# Remarks on Results

- ParRep appears to converge as $t_{\mathrm{corr}} \to \infty$ over many wells
- ParRep is more efficient when the scale separation is large
- Convergence of the hitting time distribution may not imply convergence of the state to state dynamics
- Study of Langevin and other generalizations is underway

# References


Arthur F. Voter.
Parallel replica method for dynamics of infrequent events.
*Phys. Rev. B*, 57(22):13985–13988, Jan 1998.

D. Perez, B.P. Uberuaga, Y. Shim, J.G. Amar, and A.F. Voter.
Accelerated molecular dynamics methods: introduction and recent developments.
*Annual Reports in Computational Chemistry*, 5:79–98, 2009.

Claude Le Bris, Tony Lelièvre, Mitchell Luskin, and Danny Perez.
A mathematical formalization of the parallel replica dynamics.
*Monte Carlo Methods Appl.*, to appear.
arXiv:1105.4636.

Gideon Simpson and Mitchell Luskin.
Numerical Analysis Of Parallel Replica Dynamics.
arXiv:1204.0819v2.