LA-UR-12-24232
Approved for public release; distribution is unlimited.

Title: The Generality of Parallel Replica Dynamics

Author(s): Voter, Arthur F.

Intended for: Energy Landscapes, 2012-07-20 (Obergurgl, ---, Austria)

The Generality of Parallel Replica Dynamics

Arthur F. Voter
Theoretical Division
Los Alamos National Laboratory

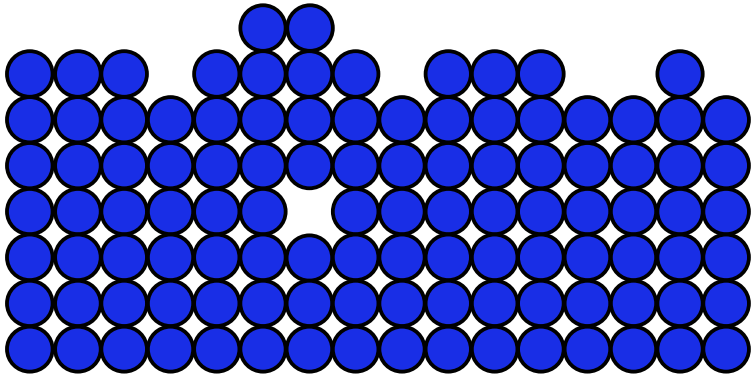
Energy Landscapes
Obergurgl, Austria
July 20, 2012

Funding: DOE Office of Basic Energy Sciences
DOE Office of Advanced Scientific Computing Research
Los Alamos LDRD

Collaborators

Danny Perez (LANL)
Tony Lelièvre (ENPC, France)
Claude Le Bris (ENPC, France)
Mitch Luskin (Minnesota)

The type of system we are typically interested in



E.g., adatoms on a surface during film growth, or bulk defects that might evolve, or radiation damage annealing, etc. ...

Infrequent atomistic jumps move the system from state to state.

The basin and pathway for each transition is typically smooth (not rugged).

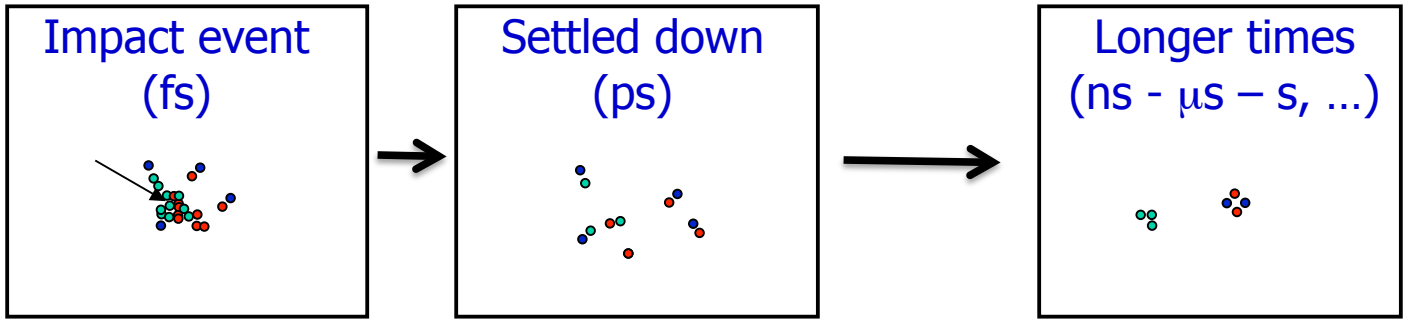
Individual transition events can be surprisingly complicated, involving many atoms.

The long-time evolution after a large number of these individual events can lead to significant complexity.

We usually don't know the final state (so path methods cannot be used).

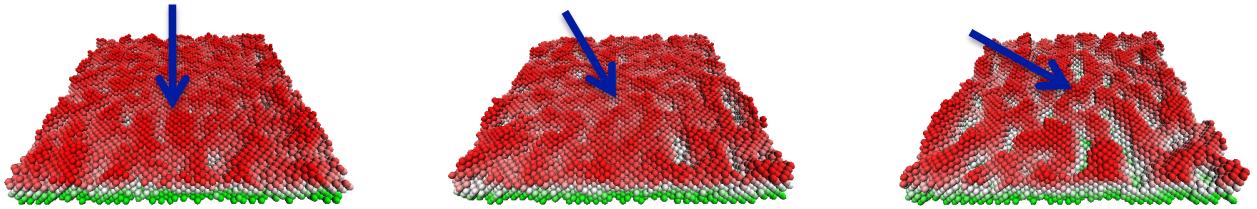
Examples

Radiation damage annealing

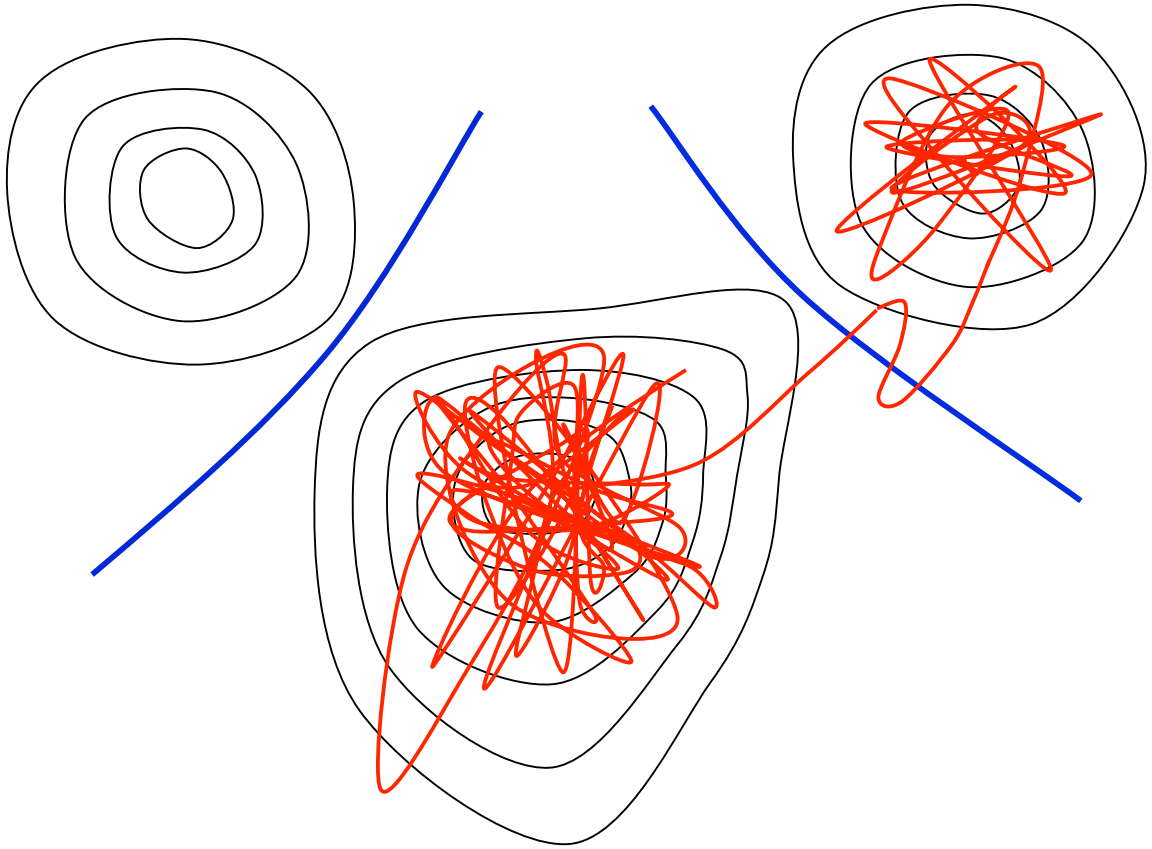


also:
voids,
bubbles,
swelling

Vapor-deposited film growth



Accelerated Molecular Dynamics

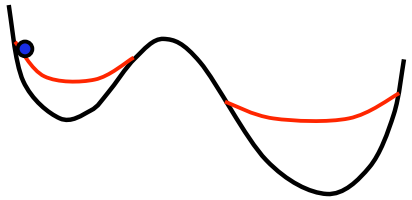


An infrequent-event system vibrates in 3N-dimensional basin many times before finding an escape path. The trajectory finds an appropriate way out (i.e., proportional to the rate constant) without knowing about any of the escape paths except the one it first sees. *In the AMD methods, we exploit this property, letting the trajectory find an appropriate escape path, but we carefully trick it into doing this more quickly.*

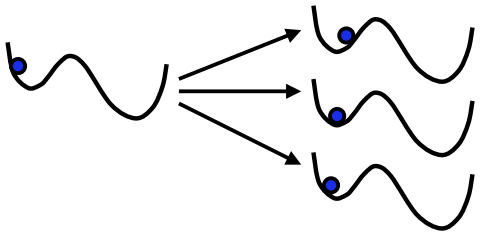
Los Alamos

Accelerated Molecular Dynamics Methods

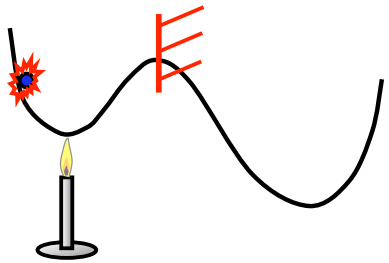
Hyperdynamics



Parallel Replica Dynamics

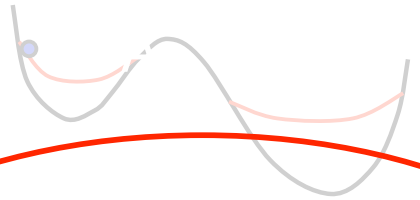


Temperature Accelerated Dynamics

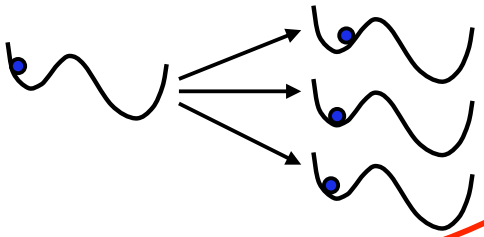


Accelerated Molecular Dynamics Methods

Hyperdynamics



Parallel Replica Dynamics



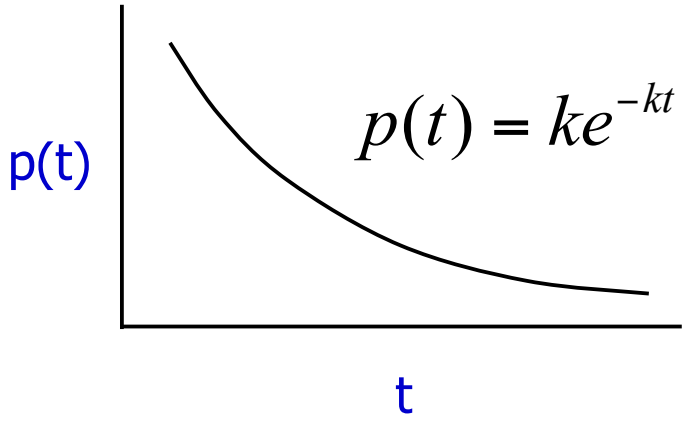
Temperature Accelerated Dynamics

Parallel Replica Dynamics

Parallelizes time evolution

Assumptions:

- infrequent events
- transitions can be detected
- exponential distribution of first-escape times

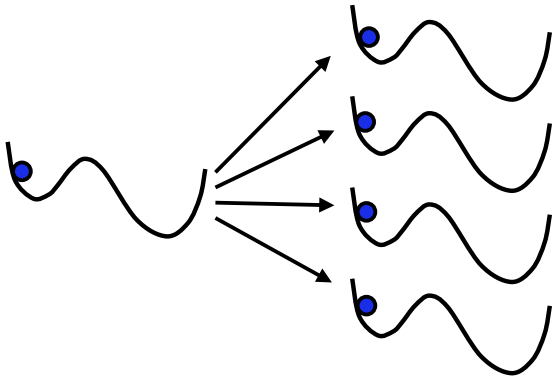


- correlation time known

AFV, Phys. Rev. B, 57, R13985 (1998)

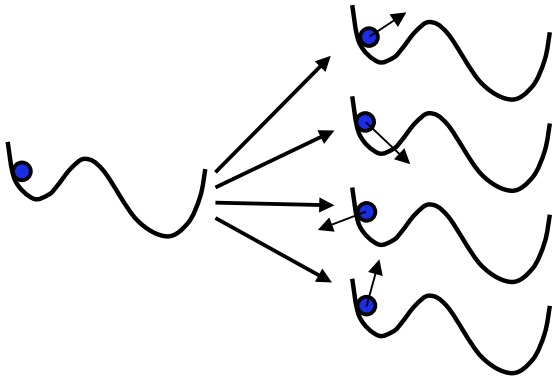
Parallel Replica Dynamics Procedure

Replicate entire system on each of M processors.



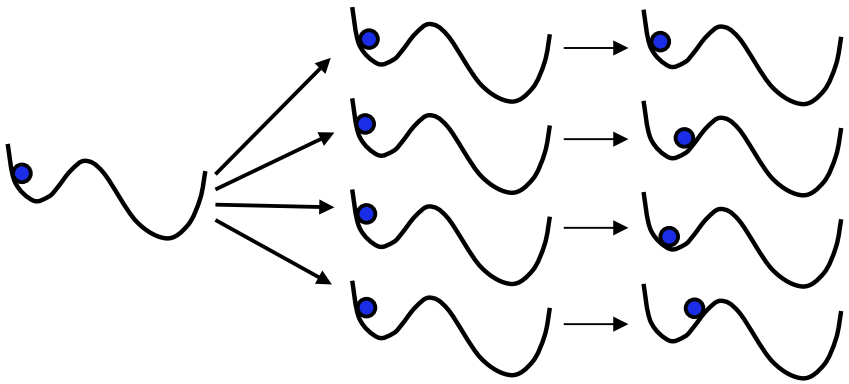
Parallel Replica Dynamics Procedure

Randomize momenta independently on each processor.



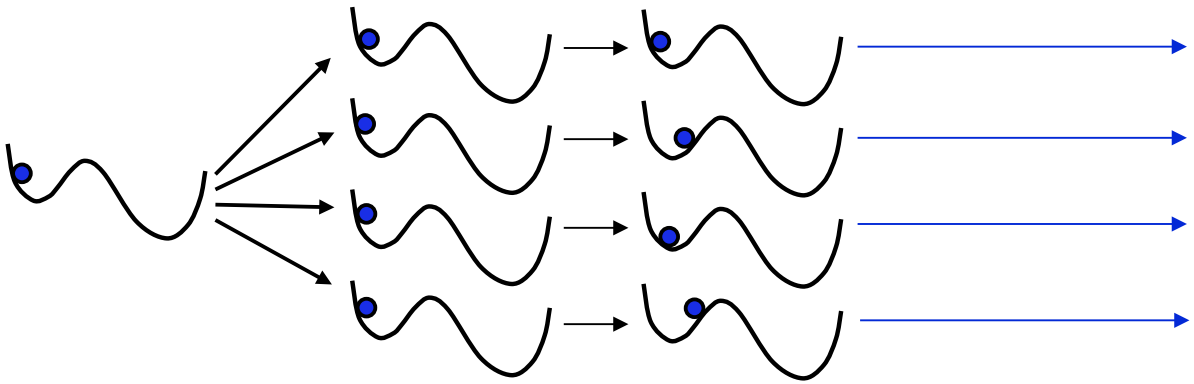
Parallel Replica Dynamics Procedure

Run MD for short time ($\tau_{dephase}$) to dephase the replicas.



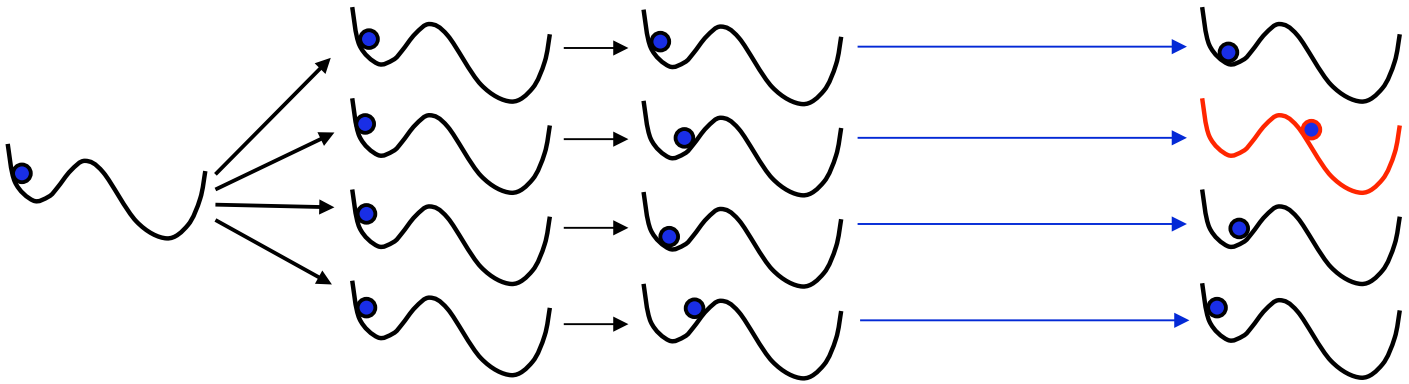
Parallel Replica Dynamics Procedure

Start clock and run thermostatted MD on each processor.
Watch for transition…



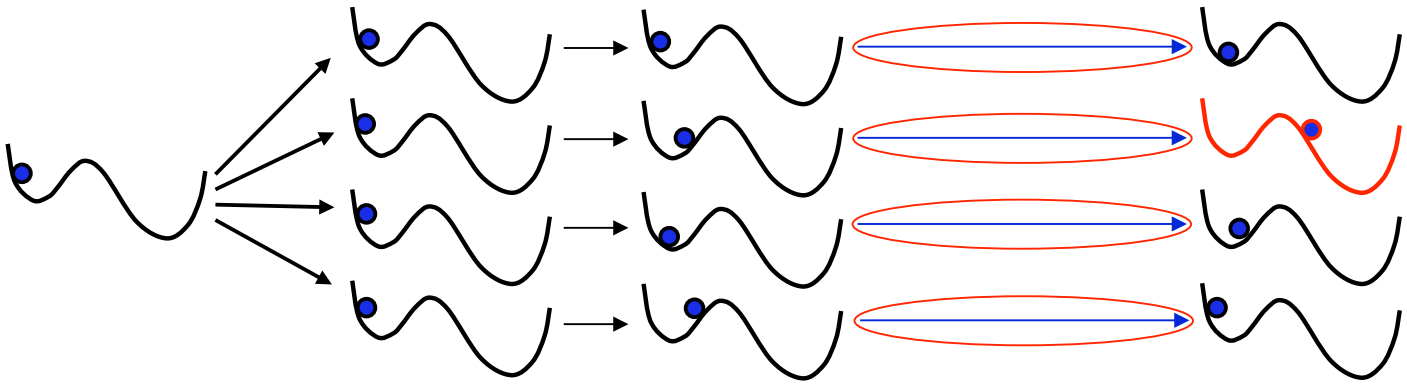
Parallel Replica Dynamics Procedure

Stop all trajectories when first transition occurs on *any* processor.



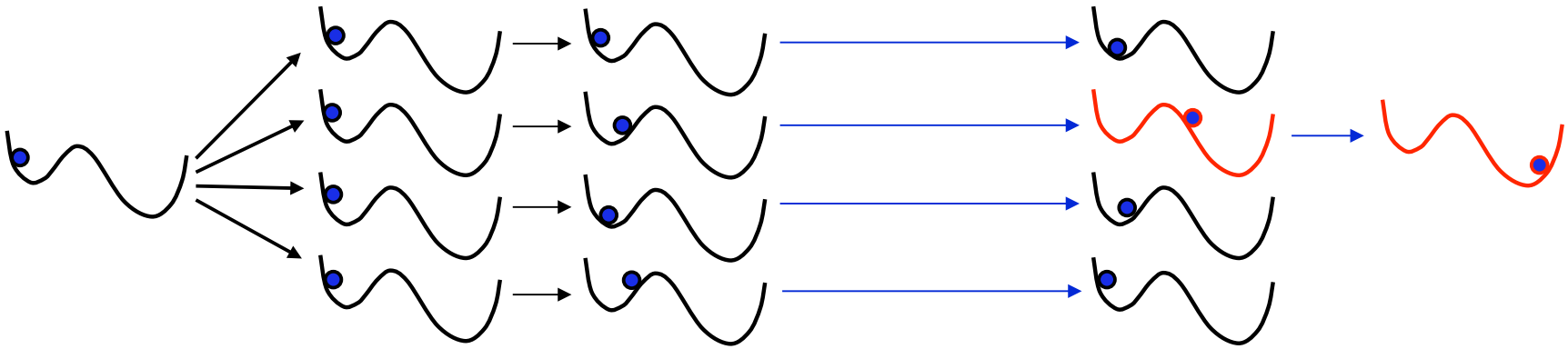
Parallel Replica Dynamics Procedure

Sum the trajectory times over all M processors. Advance simulation clock by this t_{sum}



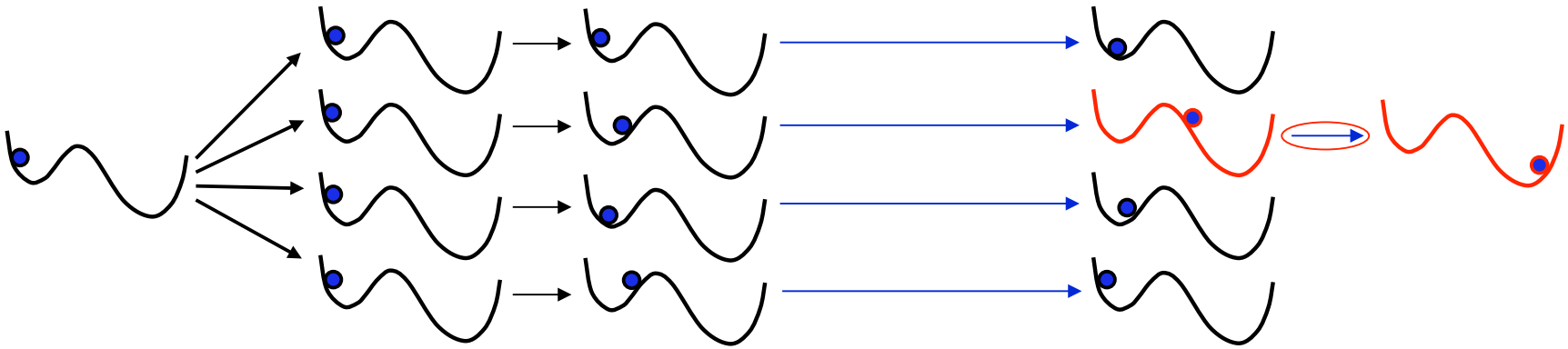
Parallel Replica Dynamics Procedure

On the processor where a transition occurred, continue trajectory for a time τ_{corr} to allow correlated dynamical events.



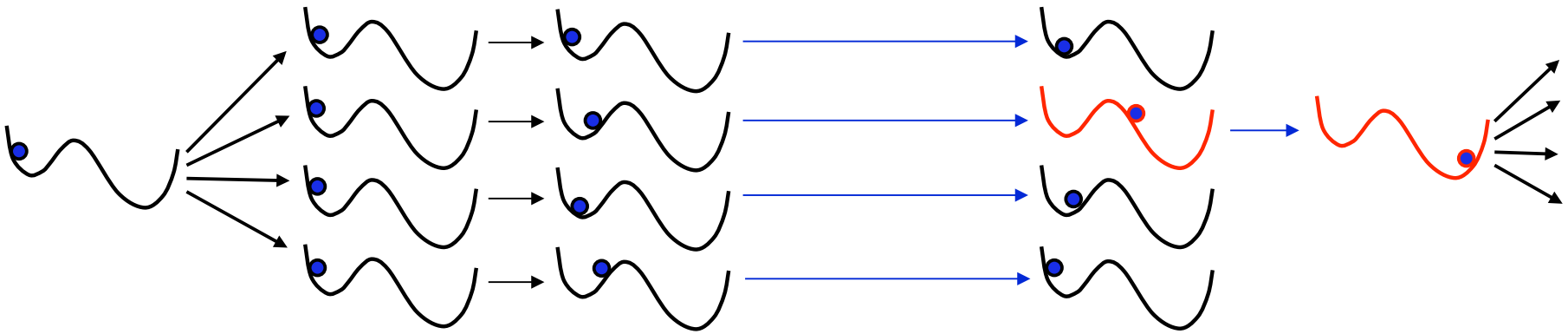
Parallel Replica Dynamics Procedure

Advance simulation clock by τ_{corr}.

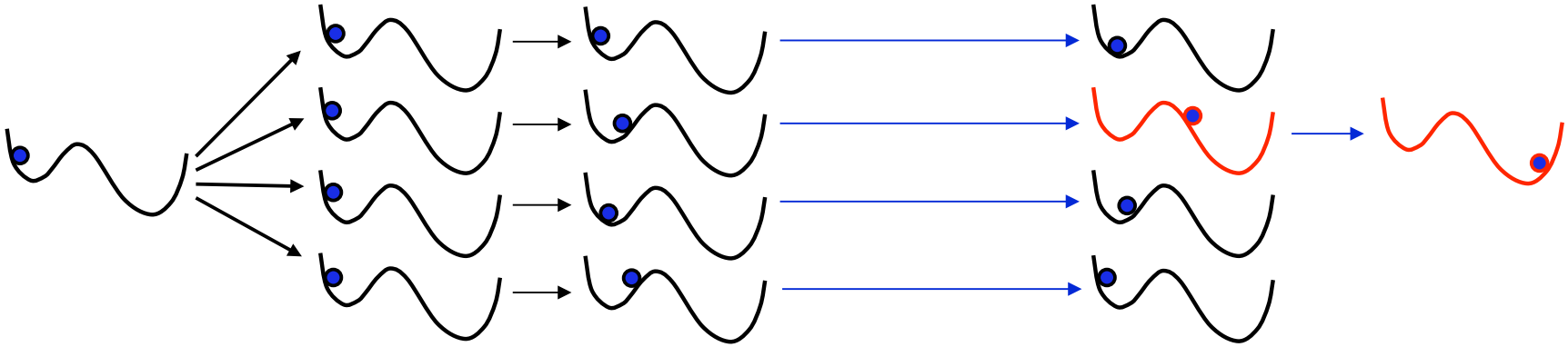


Parallel Replica Dynamics Procedure

Replicate the new state and begin procedure again.



Parallel Replica Dynamics



The summed time (t_{sum}) obeys the correct exponential distribution, and the system escapes to an appropriate state.

State-to-state dynamics are thus correct; τ_{corr} stage even releases the TST assumption [AFV, Phys. Rev. B, 57, R13985 (1998)].

Good parallel efficiency if $\tau_{rxn} / M >> \tau_{dephase} + \tau_{corr}$

Applicable to any system with exponential rare-event behavior.

Detecting a transition

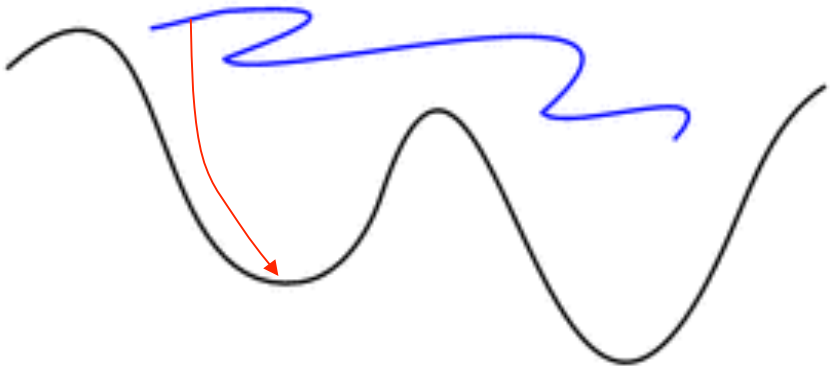
- best method depends on the system

- simple method for EAM metal systems:
  periodically perform steepest-descent quench;
  see if geometry at basin minimum has changed

Detecting a transition

- best method depends on the system

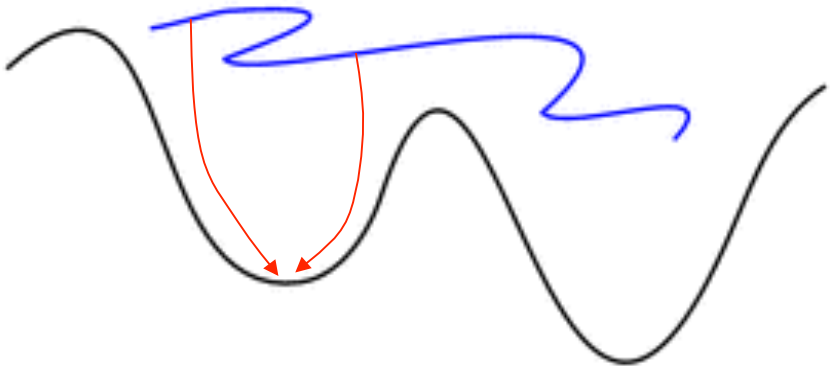
- simple method for EAM metal systems:
  periodically perform steepest-descent quench;
  see if geometry at basin minimum has changed



Detecting a transition

- best method depends on the system

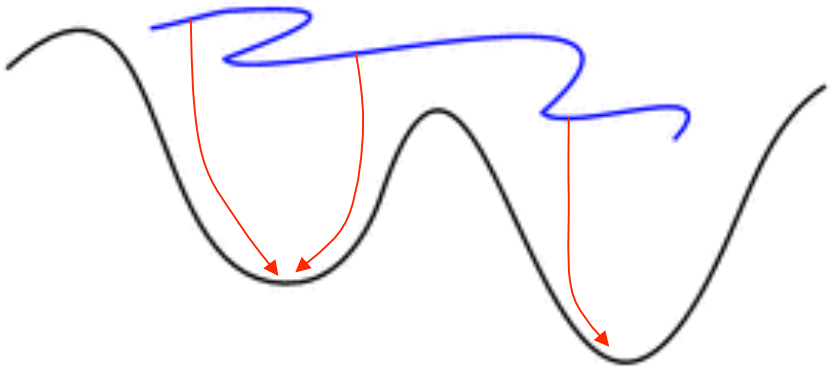
- simple method for EAM metal systems:
  periodically perform steepest-descent quench;
  see if geometry at basin minimum has changed



Detecting a transition

- best method depends on the system

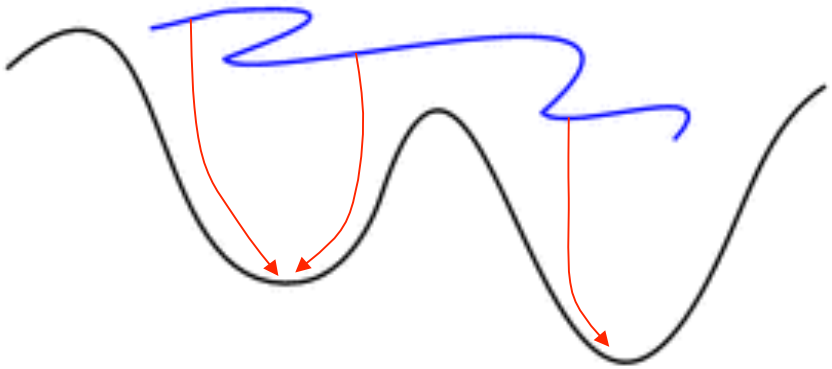
- simple method for EAM metal systems:
  periodically perform steepest-descent quench;
  see if geometry at basin minimum has changed



Detecting a transition

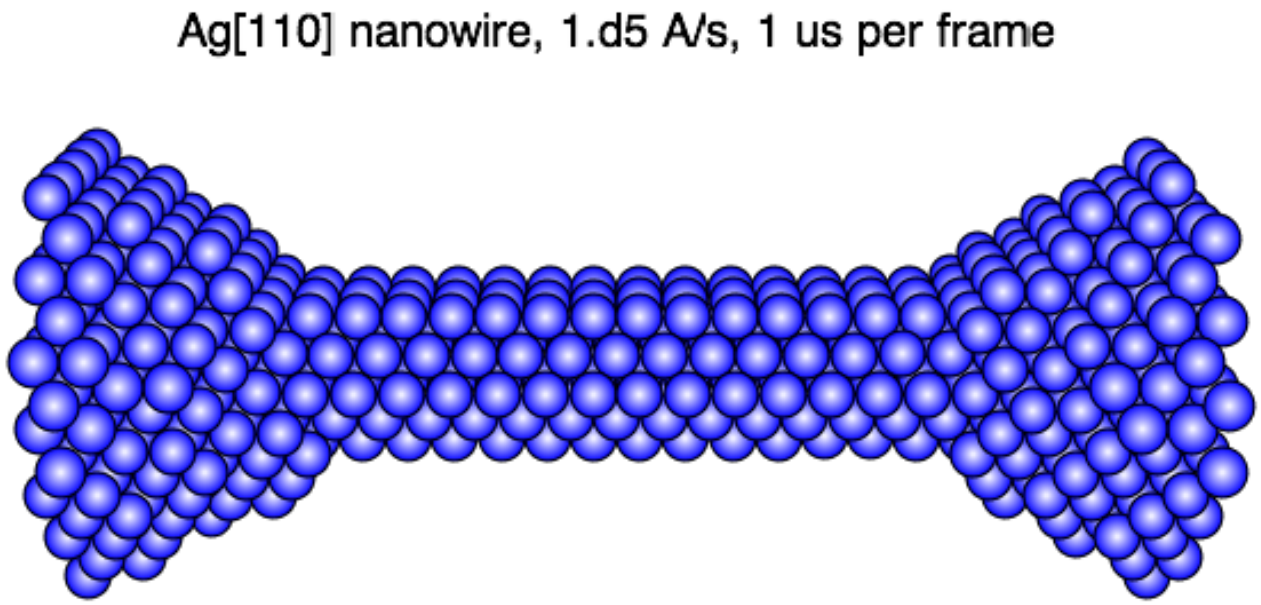
- best method depends on the system

- simple method for EAM metal systems:
  periodically perform steepest-descent quench;
  see if geometry at basin minimum has changed



- Alternatively:
  - change in bond connectivity (especially good for freely rotating clusters, for example)

Ag nanowire simulated w/ ParRep

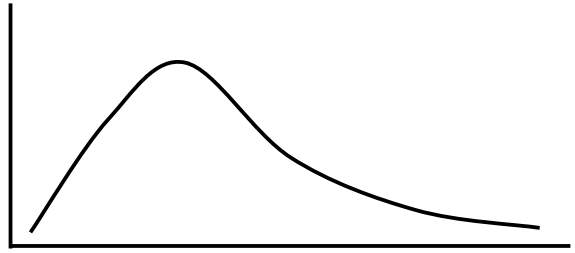


Perez, Swaminarayan, Pao and Voter, to be published (someday!).

Los Alamos

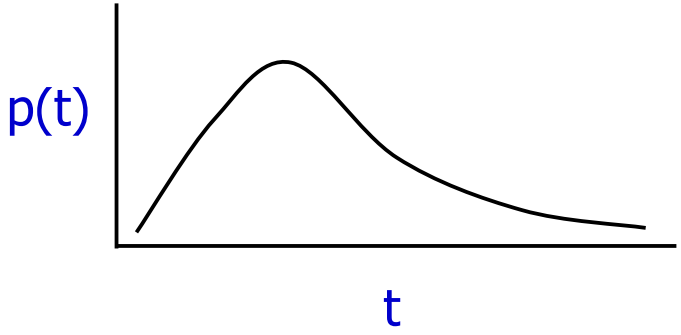
Must p(t) be exponential?

What if p(t) looks like this?



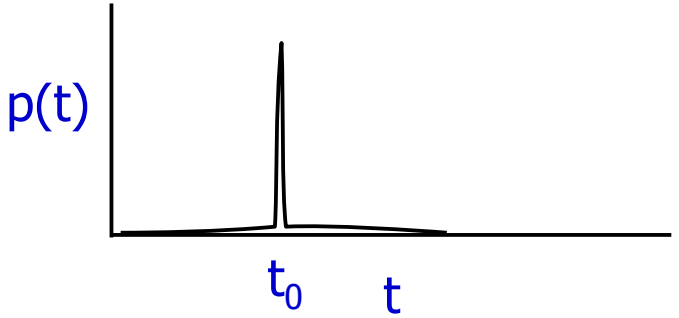
Must p(t) be exponential?

What if p(t) looks like this?



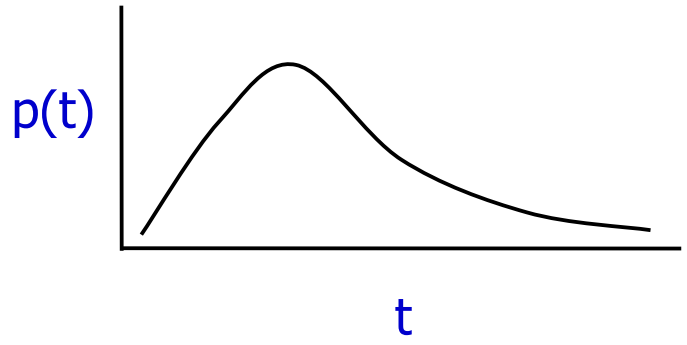
This is a problem.

Imagine the limiting case:



If $p(t)=\delta(t-t_0)$, parallel replica simulation will detect an event on all M processors at the same time (t_0), giving no boost. Worse, the summed time will predict a transition time too large by exactly a factor of M.

General comments about non-exponential p(t)



A non-exponential p(t) is almost certainly caused by "hidden" transitions, or, similarly, a diffusive character to some coordinate(s).

Thus, applying parallel-replica dynamics when p(t) is non-exponential will give incorrect times, and incorrect dynamics.

If every transition is detected, and the correlated events are allowed to finish, each p(t) will be exponential for a (non-diffusive) infrequent-event system, and parallel-replica dynamics will be valid.

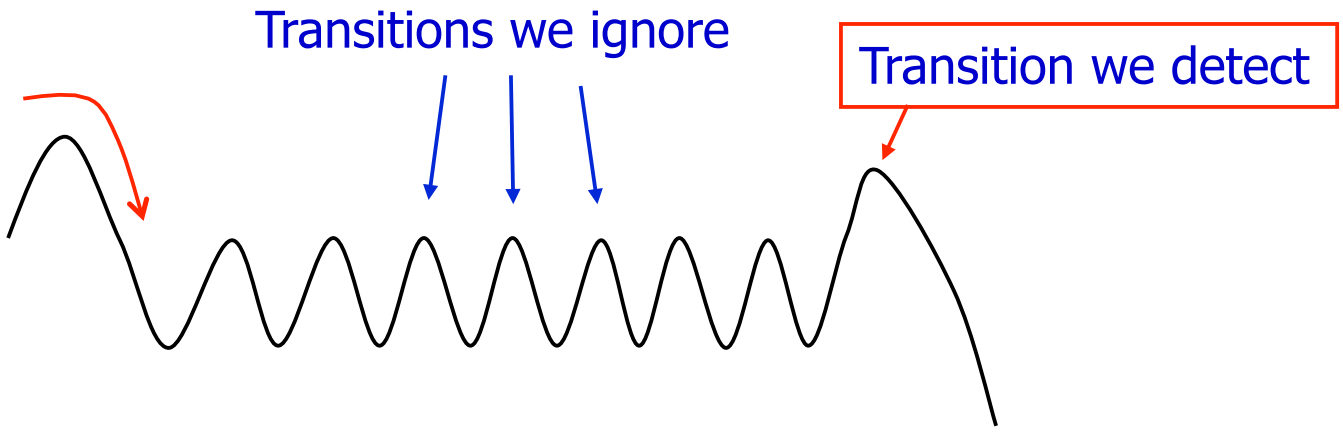
Parallel Replica Dynamics
Can we overlook fast transitions?

The maximum boost is limited by the fastest processes in the system.

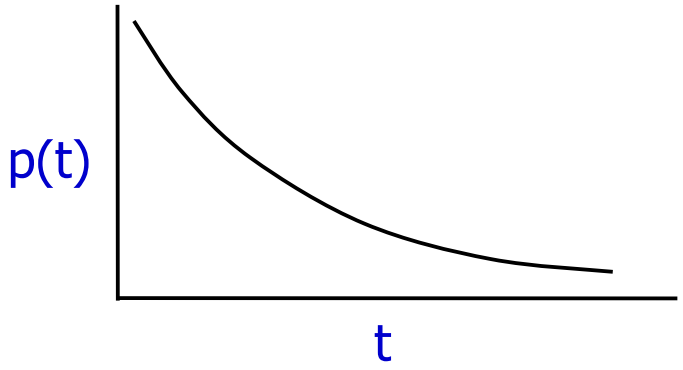
If the rapid transitions seem to be unimportant, can we ignore them?

Yes - but be careful…

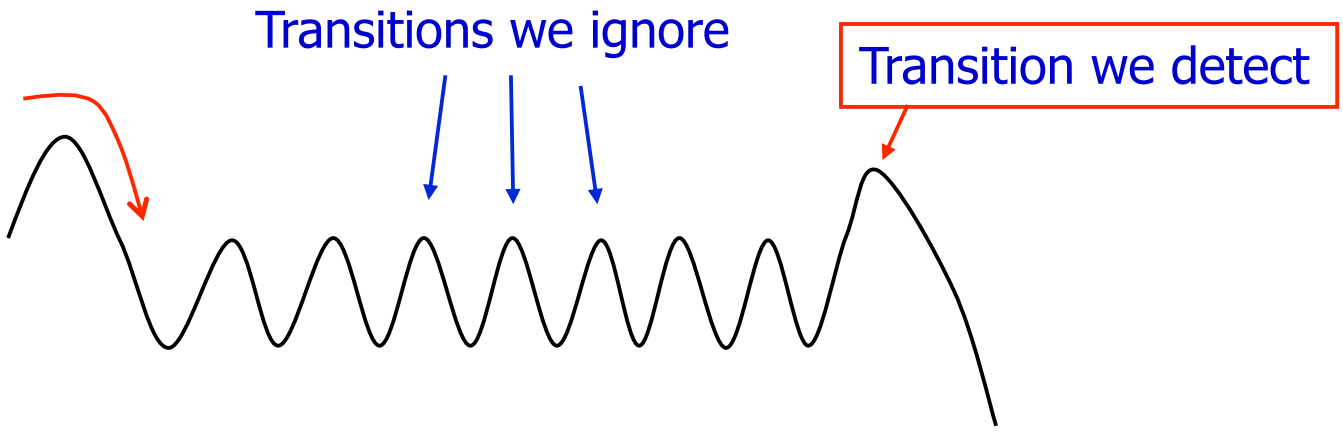
Safe case



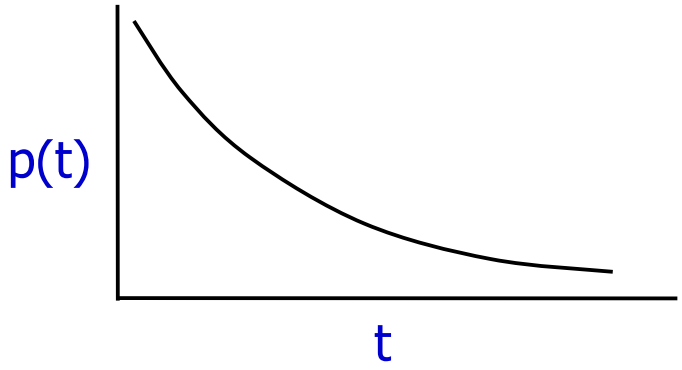
Trajectory will go back and forth many times before escaping



Safe case

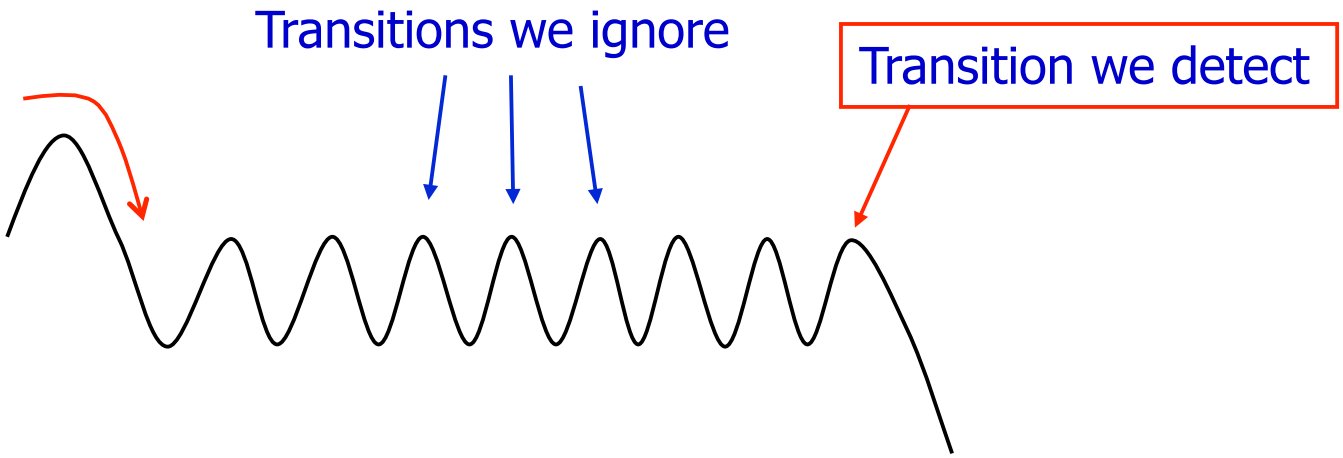


Trajectory will go back and forth many times before escaping

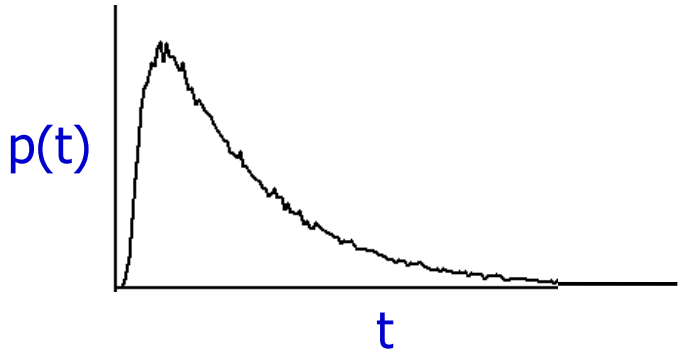


But note that we have to be careful to lengthen the dephasing time to equilibrate in the superbasin

Dangerous case



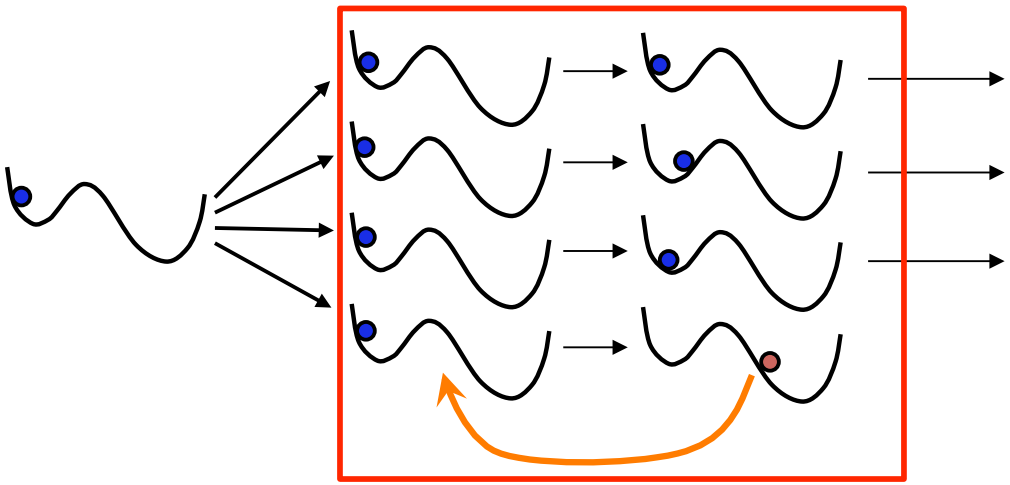
Trajectory may fall off right-hand end on first pass - acts like an exponential preceded by a lag time while system diffuses from left to right



The new understanding

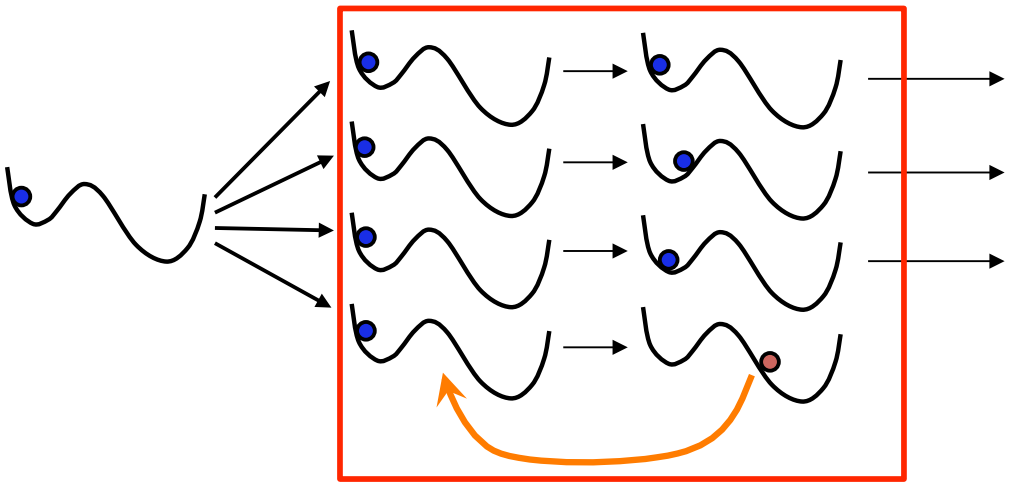
The quasi-stationary distribution

During the dephasing step, we remove (and perhaps restart) any trajectories that escape from the state.



The quasi-stationary distribution

During the dephasing step, we remove (and perhaps restart) any trajectories that escape from the state.

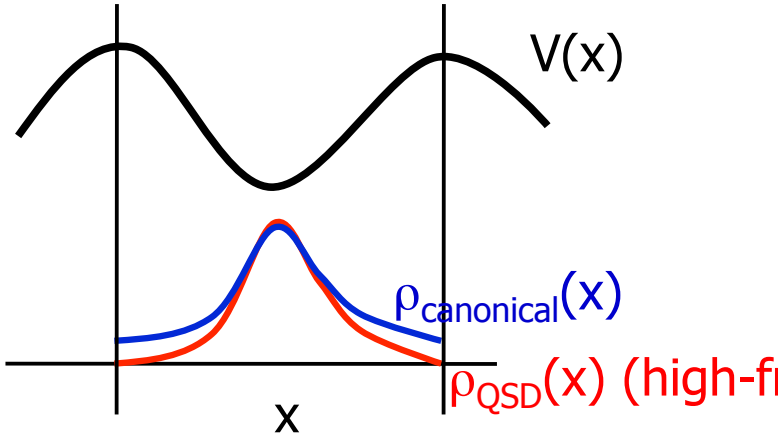


This ParRep dephasing procedure prepares a "quasi-stationary distribution" (QSD).

The quasi-stationary distribution (QSD)

The QSD is the distribution that results in the long time limit of dynamics in a potential with absorbing boundaries.

Note that it is not the same as the canonical ensemble:



(although for a high barrier, it is virtually the same)

C. Le Bris, T. Lelièvre, M. Luskin, D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications (in press); available as arXiv:1105.4636.

Properties of the QSD

The probability distribution for the first escape from the QSD is an exponential, and the escape hitting points are independent of time.

Thus, it is appropriate for ParRep.

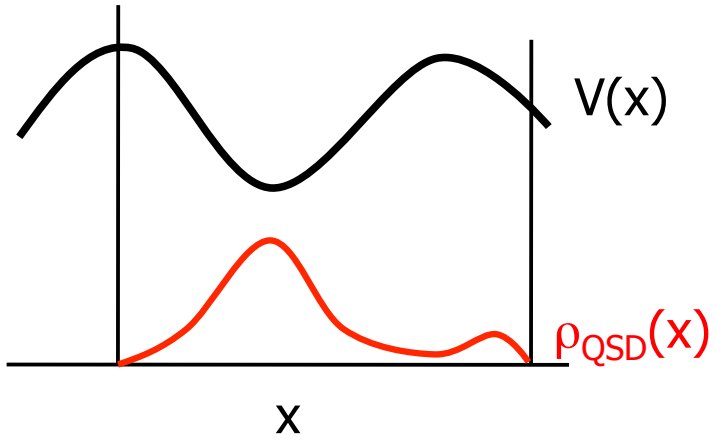
C. Le Bris, T. Lelièvre, M. Luskin, D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications (in press); available as arXiv:1105.4636.

Properties of the QSD

The probability distribution for the first escape from the QSD is an exponential, and the escape hitting points are independent of time.

Thus, it is appropriate for ParRep.

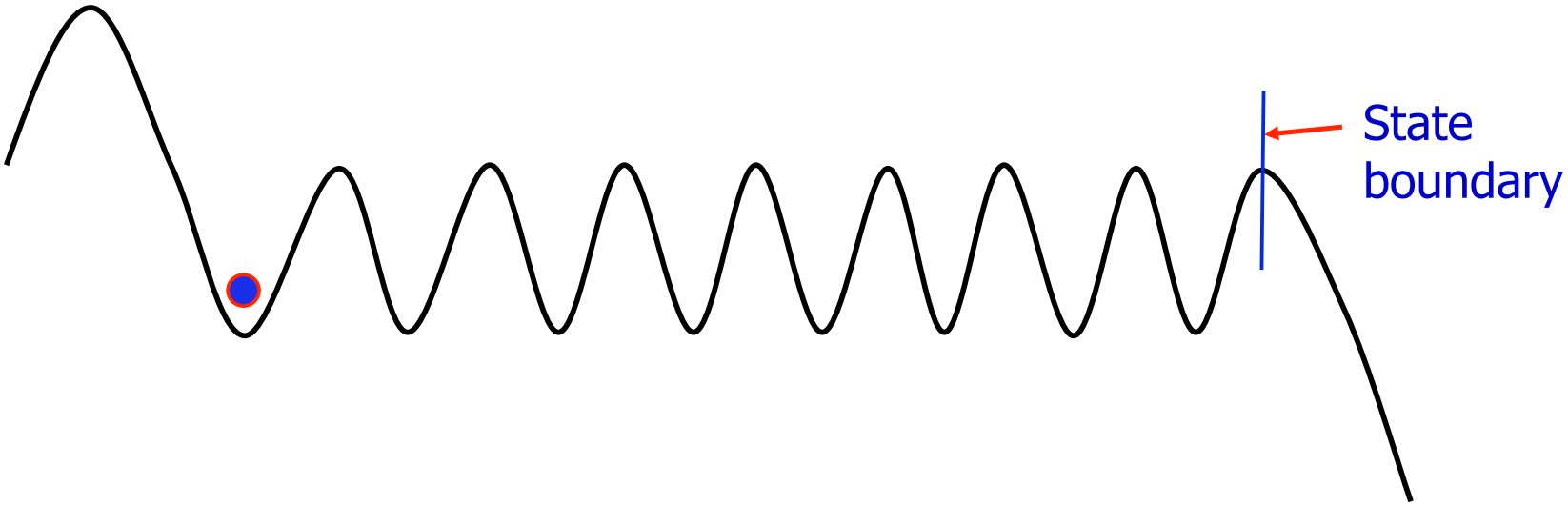
Moreover, it has these properties regardless of where the boundaries are positioned (!). E.g.



C. Le Bris, T. Lelièvre, M. Luskin, D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications (in press); available as arXiv:1105.4636.

Exploiting the QSD concept

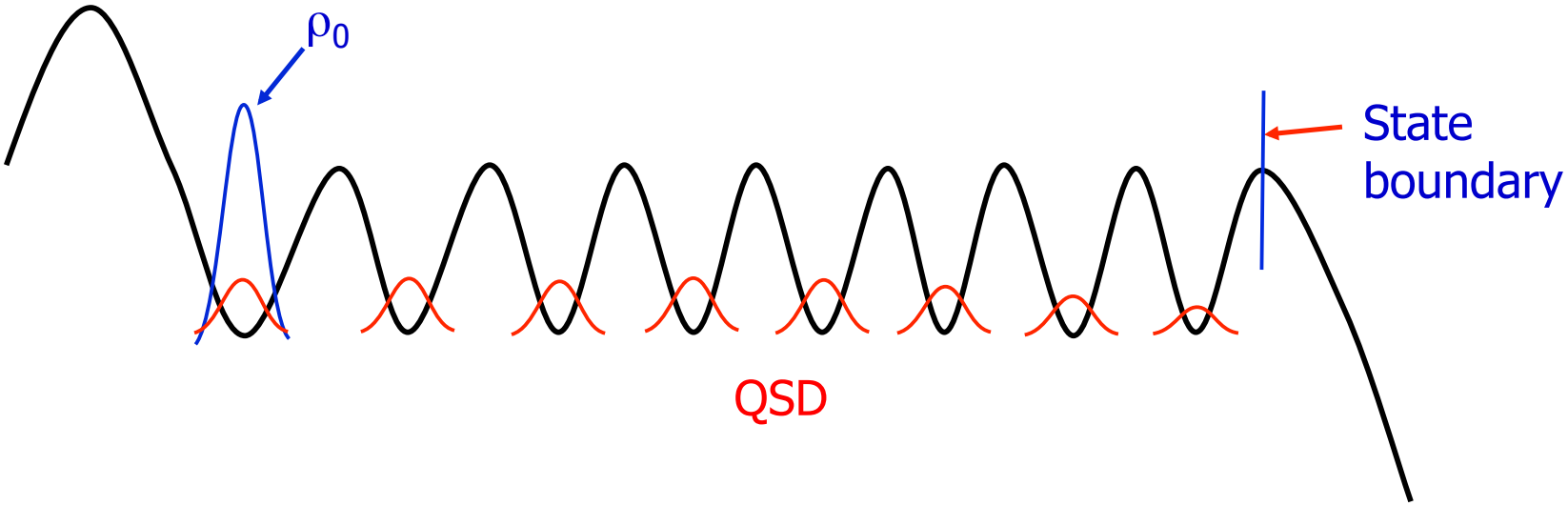
This means we can even use ParRep on a "dangerous" case, provided we dephase long enough.



The required dephasing time may be so long that it is less efficient than simply continuing the ParRep in and out of each individual basin, or simply doing direct MD, *but the ParRep will be correct*.

Exploiting the QSD concept

This means we can even use ParRep on a "dangerous" case, provided we dephase long enough.



The required dephasing time may be so long that it is less efficient than simply continuing the ParRep in and out of each individual basin, or simply doing direct MD, *but the ParRep will be correct*.

The QSD as Fokker-Planck eigenvector

For simplicity, consider overdamped Langevin dynamics.

Solving the Fokker-Planck equation with absorbing boundary conditions leads to the time dependent density

$$\rho(x,t) = \sum_k \exp(-\lambda_k t)\, c_k(0)\, u_k(x) ,$$

where λ_k and $u_k(x)$ are the k^{th} eigenvalue and eigenvector of the Fokker-Planck equation and $c_k(0)$ is set by the initial conditions.

At very long times, the higher eigenvector components decay away, leaving

$$\rho(x,t) = \exp(-\lambda_1 t)\, c_1(0)\, u_1(x) .$$

This continues to decay as $\exp(-\lambda_1 t)$, but if renormalized, it is the QSD.

C. Le Bris, T. Lelièvre, M. Luskin, D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications (in press); available as arXiv:1105.4636.

Error in preparation of the QSD

For this overdamped Langevin case, LeBris et al* showed that the ParRep dephasing step prepares the QSD exactly as $\tau_{corr} \rightarrow$ infinity, with error at finite time τ_{corr} proportional to

$$\exp[-(\lambda_2-\lambda_1)\tau_{corr}]$$

Regular Langevin also gives a QSD, but it is harder to derive the error bound.

*C. Le Bris, T. Lelièvre, M. Luskin, D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications (in press); available as arXiv:1105.4636.

The ParRep procedure for a complex superbasin

The procedure is exactly as above, but paying particular attention that τ_{corr} is large enough to obtain the QSD accurately.

This τ_{corr} is used for both the dephasing stage and the correlation stage.

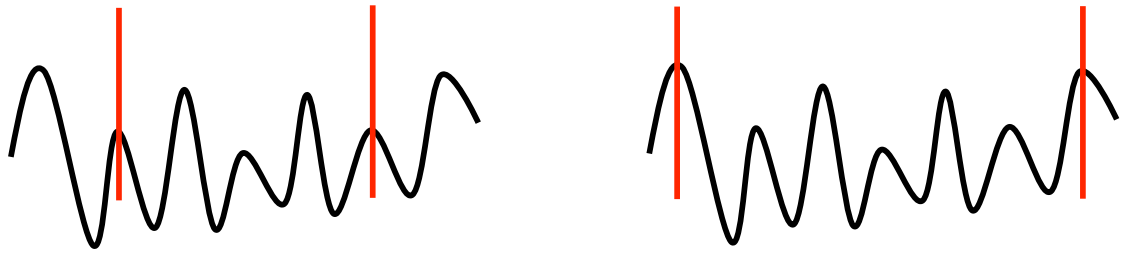
If the transitioning trajectory escapes from the superbasin before τ_{corr}, then there is no parallel step -- we just continue the MD into the next superbasin.

Determining a safe value for τ_{corr} is important, and this may be the hard part. (still learning about the best way to do this)

Choosing the superbasin definition

We are free to choose any state definition we want.

If we can optimize the definition to maximize the separation of time scales, we will get more boost. (may or may not be difficult, depending on system)

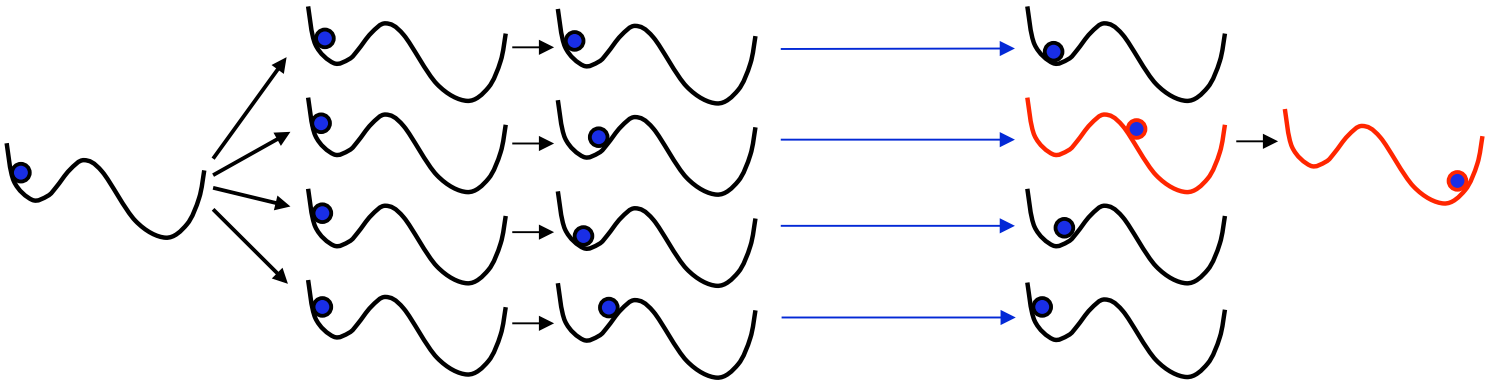


bad

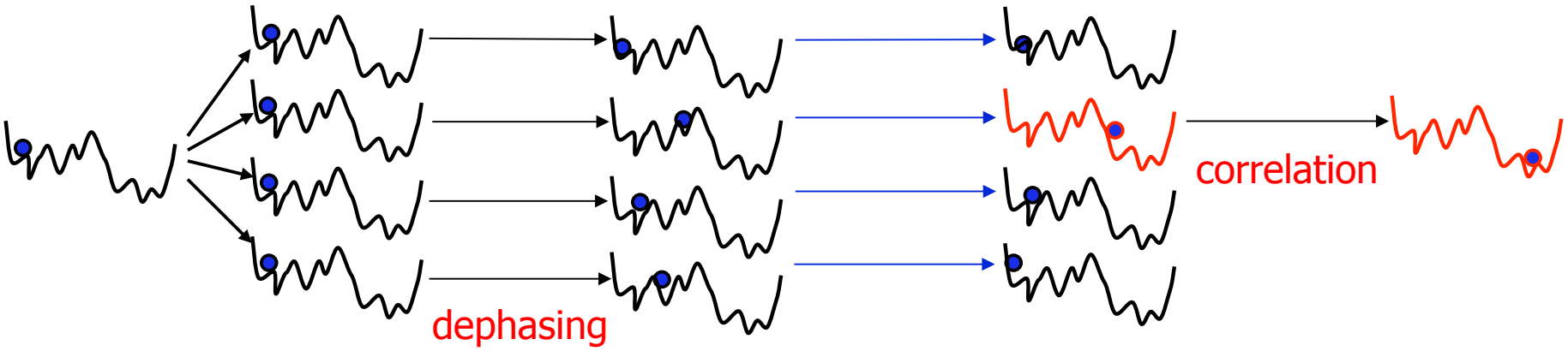
better

Increase in overhead due to ruggedness

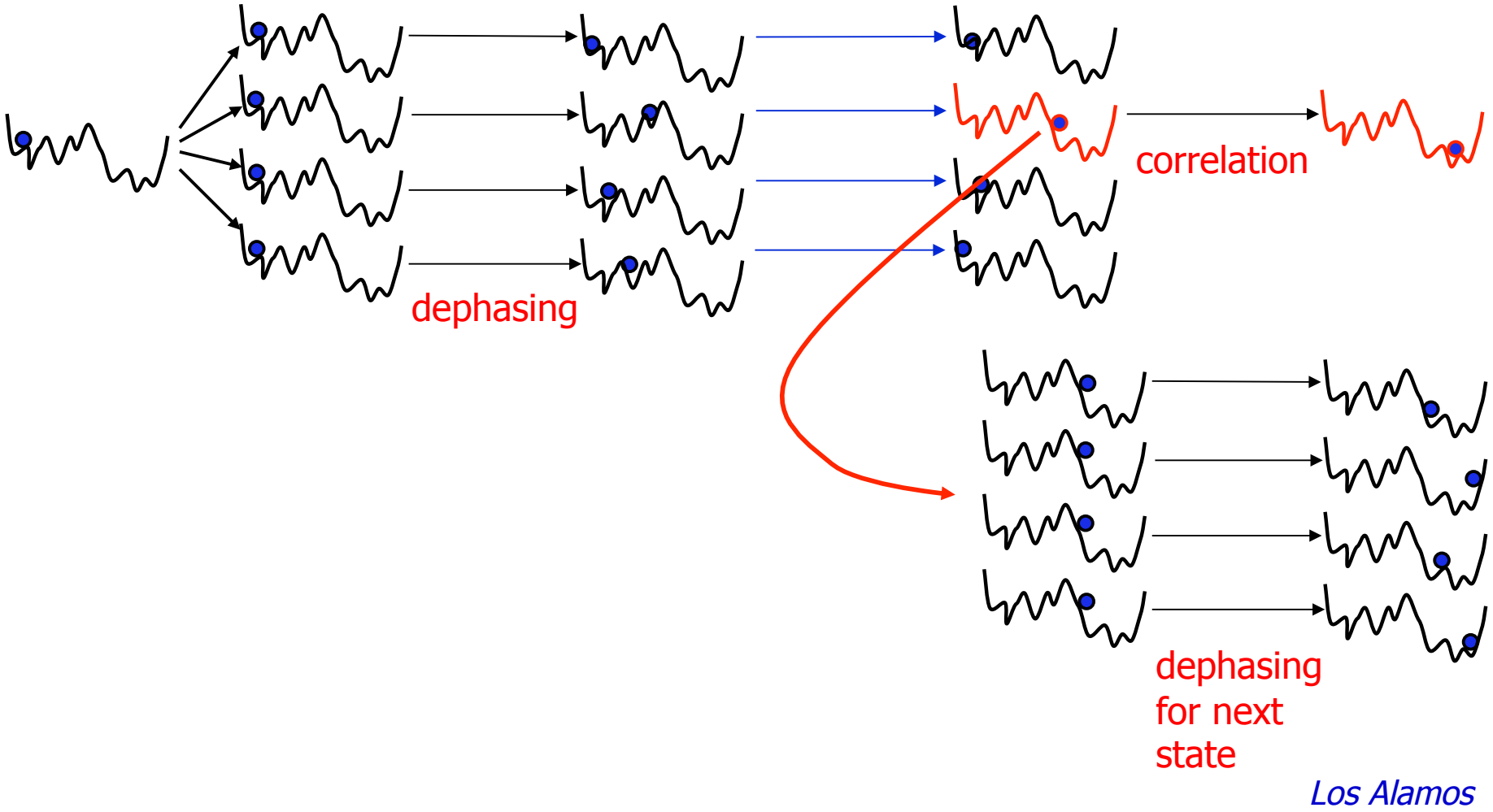
Simple basin



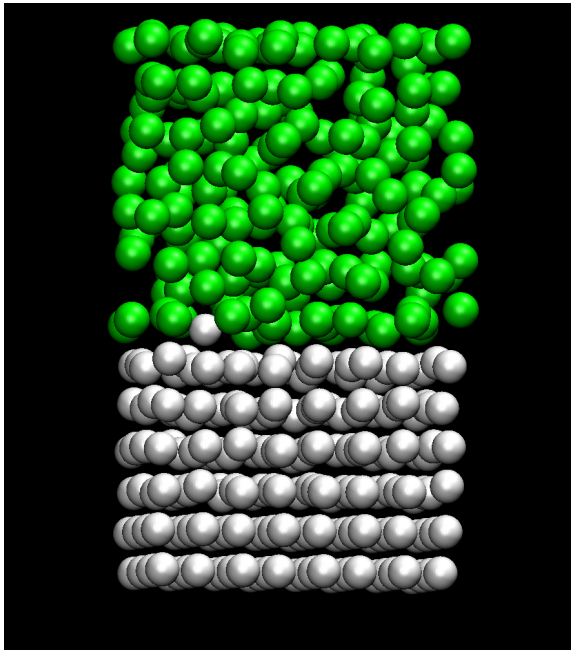
Rugged basin – longer τ_{corr} and thus greater overhead and less boost.



Overlapping dephasing and correlation stages



Examples where general state definitions will help



Adatom dynamics at liquid-solid interface.

Define state by ignoring all liquid degrees of freedom. New QSD understanding is crucial if liquid diffusion is sluggish, clouding time-scale separation issue.

Also:

Messy low-barrier situations (e.g., surface clusters, interstitials, grain boundaries, …)

Diffusion of solute in a polymer

Soft matter dynamics

Protein folding

Glass dynamics?

…

Conclusions

Parallel Replica dynamics parallelizes time.

Implemented carefully, it gives *exact* state-to-state dynamics.

Recognizing that the dephasing stage in ParRep automatically prepares a quasi-stationary distribution (QSD) means that ParRep can be applied much more generally than we previously realized.

Any state definition can be employed. The QSD enforces the exponentiality needed to make ParRep valid.

Choosing a state definition that gives a good separation of time scales will make it more efficient at parallelizing time because the necessary correlation time (τ_{corr}) will be shorter relative to the escape time.

Some review articles on AMD methods

Perez et al, Ann. Rep. Comp. Chem. **5**, 79 (2009).

Voter, Montalenti, and Germann, Ann. Rev. Mater. Res. 32, 321 (2002).