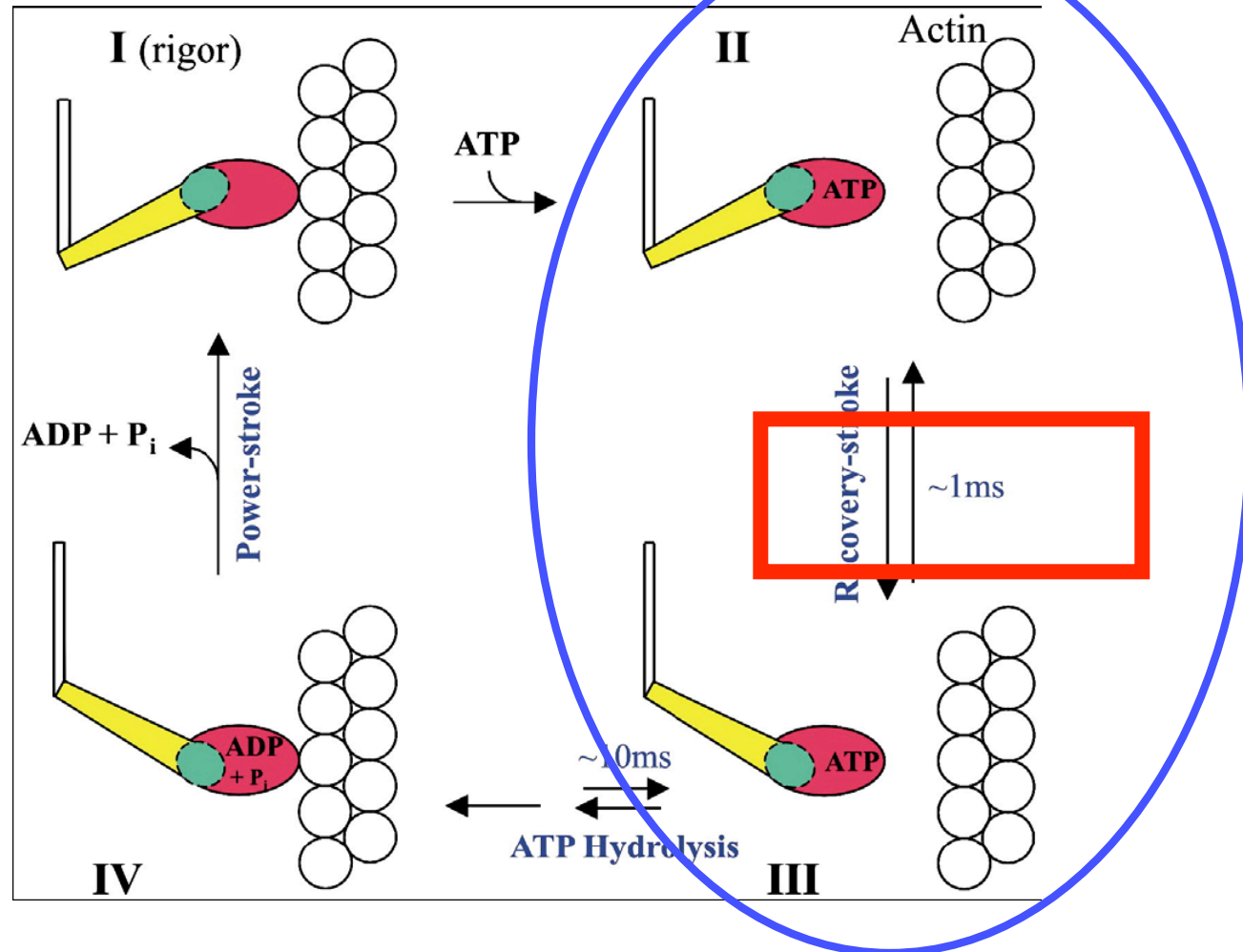# Milestoning

Ron Elber and Anthony West
Department of Chemistry and Biochemistry and
Institute of Computational Sciences and Engineering (ICES)
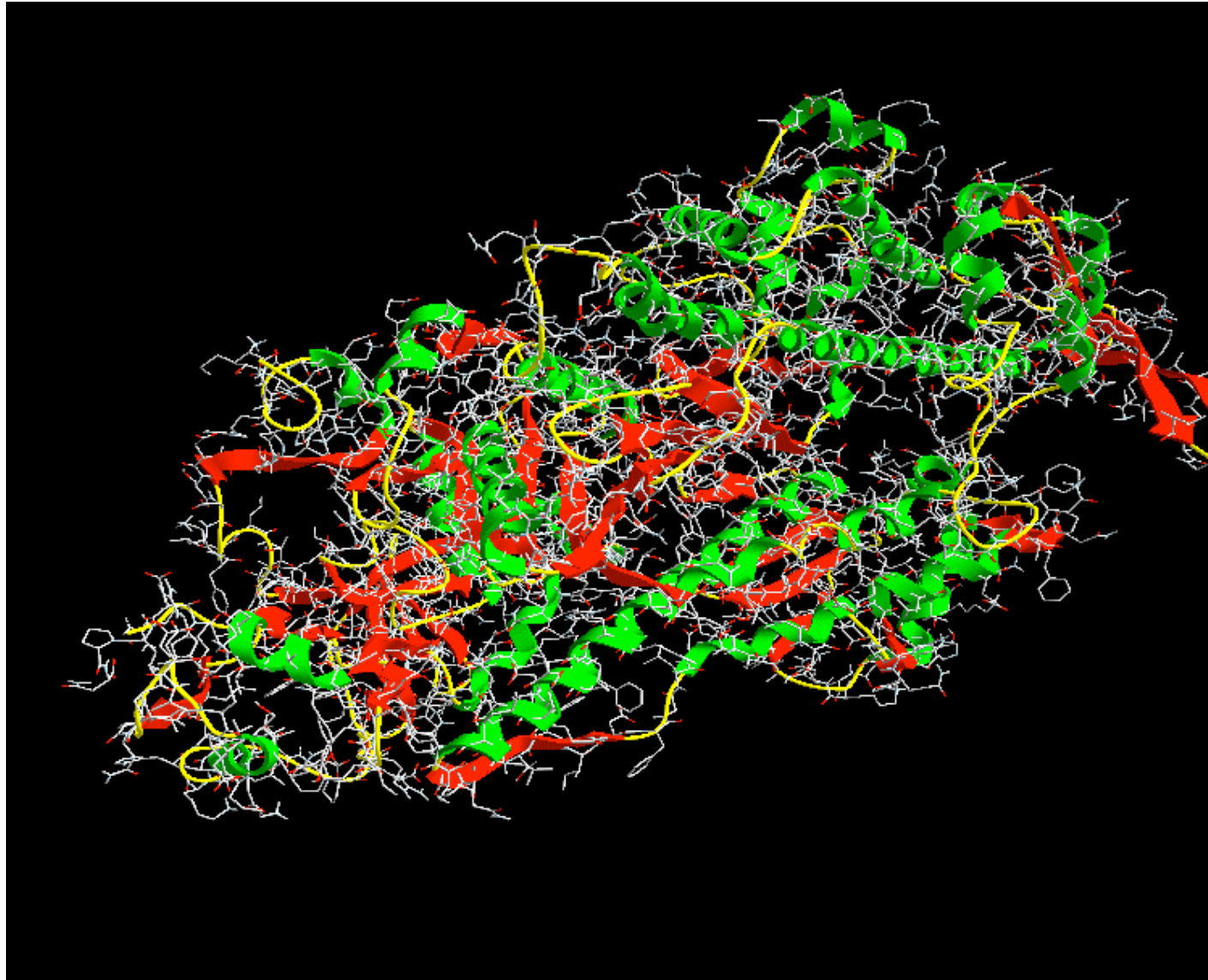University of Texas at Austin

Motivation

- ## Understanding Myosin II:
  - A protein that converts biochemical energy (ATP) to mechanical energy in the muscles

Myosin Stroke Cycle (Lymn-Taylor)

S. Mesentean et al., *JMB*, **367**, 591–602 (2007)

Movie of the reaction path for the recovery stroke in myosin

Simulation background

- Computer simulations use classical mechanics to simulate protein dynamics

$$M \frac{d^2 X}{dt^2} = -\nabla U(X)$$

$$X(t + \Delta t) = X(t) + V(t) \cdot \Delta t - \frac{\Delta t^2}{2} M^{-1} \cdot \nabla U(X(t))$$

$$V(t + \Delta t) = V(t) - \frac{\Delta t}{2} M^{-1} \cdot \left[\nabla U(X(t)) + \nabla U(X(t + \Delta t)) \right]$$

Δt about 10^{-15}s (a femtosecond) to maintain stability of the algorithm.
We are interested in time scale of 10^{-3}s (a millisecond). On a PC it takes months to compute 100 ns (10^{-7}s).

The problem

- Twelve orders of magnitude difference between the basic time steps (10^{-15}s) and the biological time scale (10^{-3}s).

- How to bridge the time scale gap?
 - Coarsening space and time while keeping molecular details

The concept of Milestones
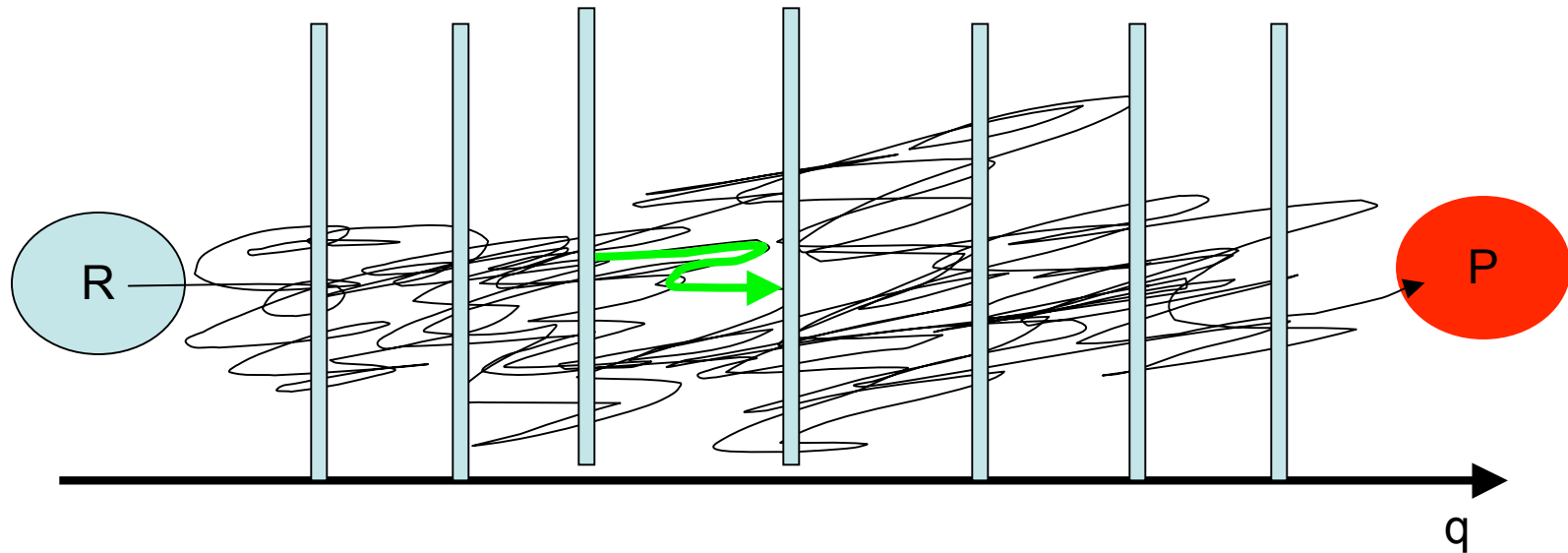
- Microscopic molecular motions tend to be diffusive at long time scales

- Following in detail individual trajectories is computationally expensive

- Can we compute the trajectories in pieces? The pieces will give us a coarse grained model.
- **Is it correct? Is it computationally efficient?**
   The interfaces between which we compute the pieces of the trajectories are called **Milestones**

Is computing trajectory pieces more efficient than computing it as a whole?

# Chopping trajectories is computationally efficient (I)

R

P

L
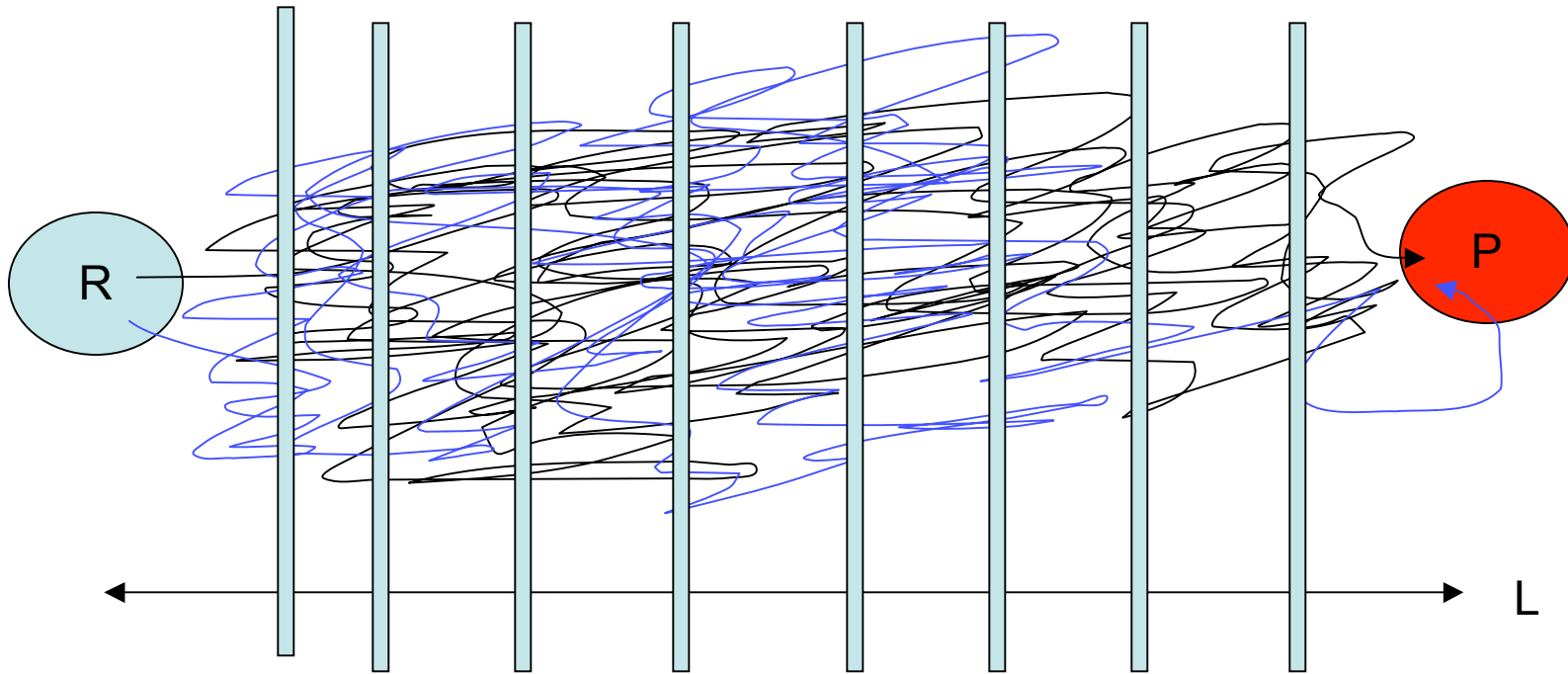
Parallelization Pieces of trajectories are trivial to parallelize.
Speed up proportional to the number of processors

Chopping trajectories is computationally efficient (II)
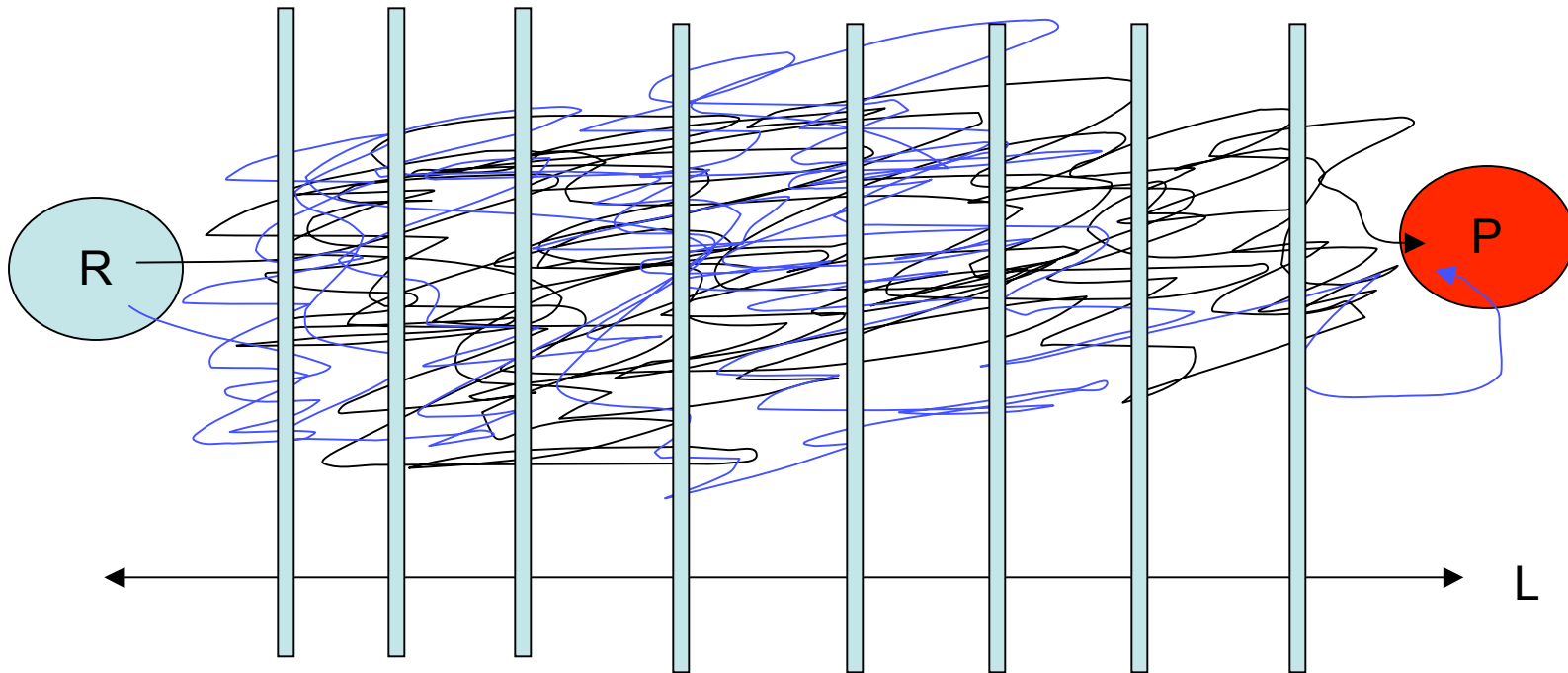
Diffusive processes The time t to diffuse a path of length L is proportional to L^2 . If we divide the path to N segments the time becomes $(L/M)^2$. There are M Milestones and therefore the time required is $M*(L/M)^2=L^2/M$. Speed up factor of M number of Milestones.

Chopping trajectories is computationally efficient (III)

Activated process If one of N_1 trajectories that starts in R reaches q_1, and one of N_2 trajectories that starts at q_1 reaches q_2 the total number of trajectories that we need to sample q_2 starting from R is N_1*N_2. Using Milestoning we only need N_1+N_2

Speed up exponential in the number of Milestones - M

Building a coarse grained model (which is equation-free): What do we keep from the short chopped trajectories?

$$i-1 \quad i \quad i+1$$
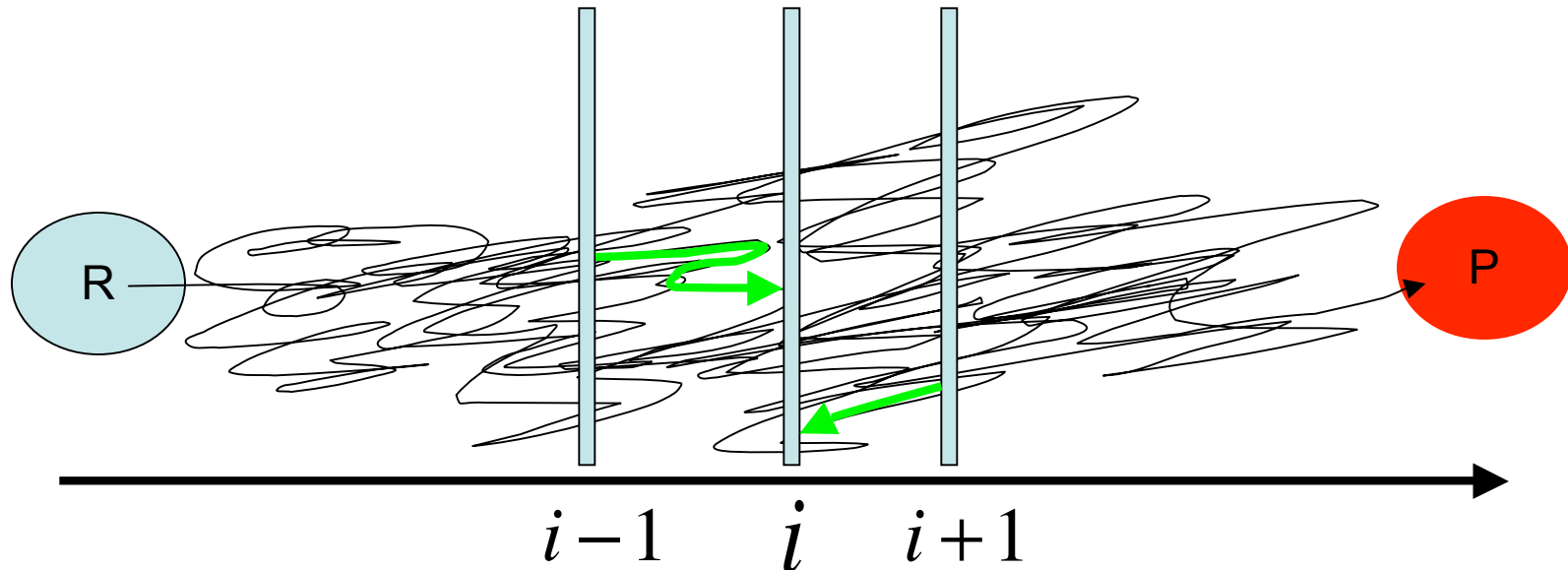
1. Initialize trajectories at the Milestones from a stationary time-independent distribution assumed known (e.g. canonical).
2. Compute trajectories between any pair of Milestones (*i,j*) with a shared volume, estimate the first passage time distribution

$$K_{ij}(\tau)$$

The local first passage time distribution

$K_{ij}(\tau)-$ The probability density that a trajectory that starts at Milestone i will terminate exactly after time τ at Milestone j

$$\int_0^\infty K_{ij}(\tau) \cdot d\tau = p_{ij} - \text{The probability that a trajectory that was}$$
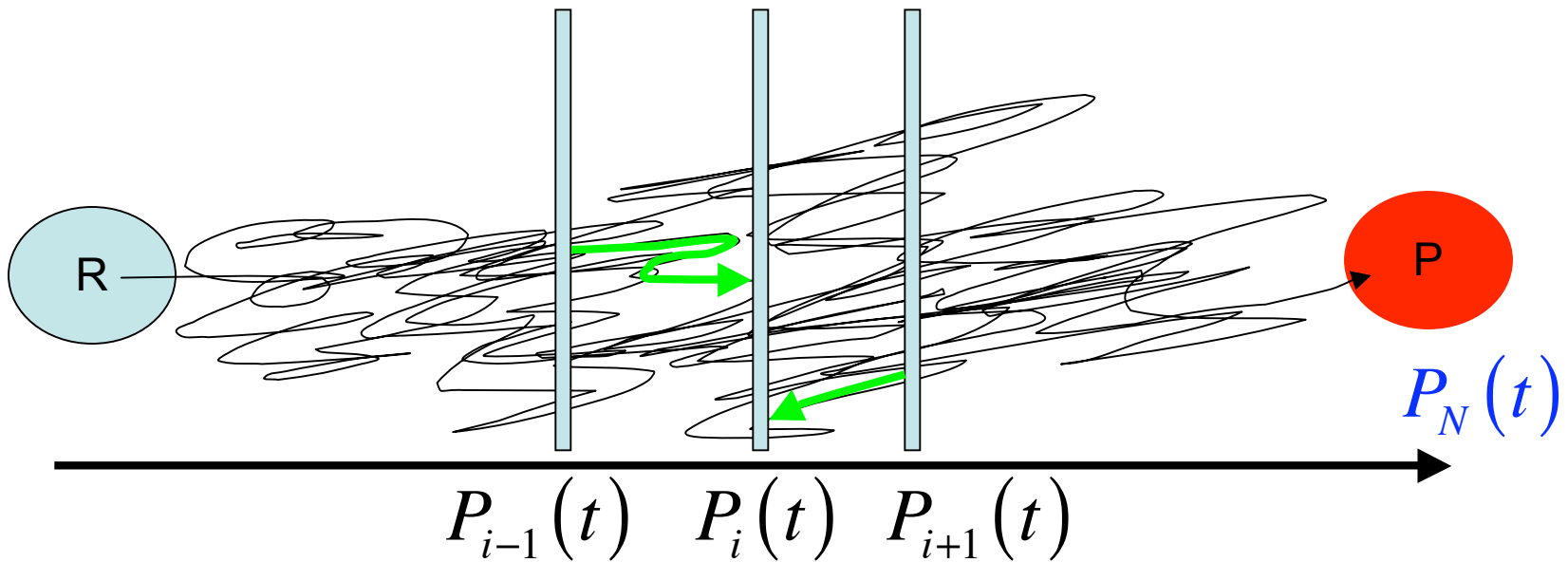
initiated at Milestone i will terminate at Milestone j

$$\sum_j p_{ij} = 1 \quad \text{-- A normalization condition on } p_{ij} \text{ (all traj terminate)}$$

This is all the microscopic information needed. No mechanical model is required.

Example for a typical $K_{ij}(\tau)$
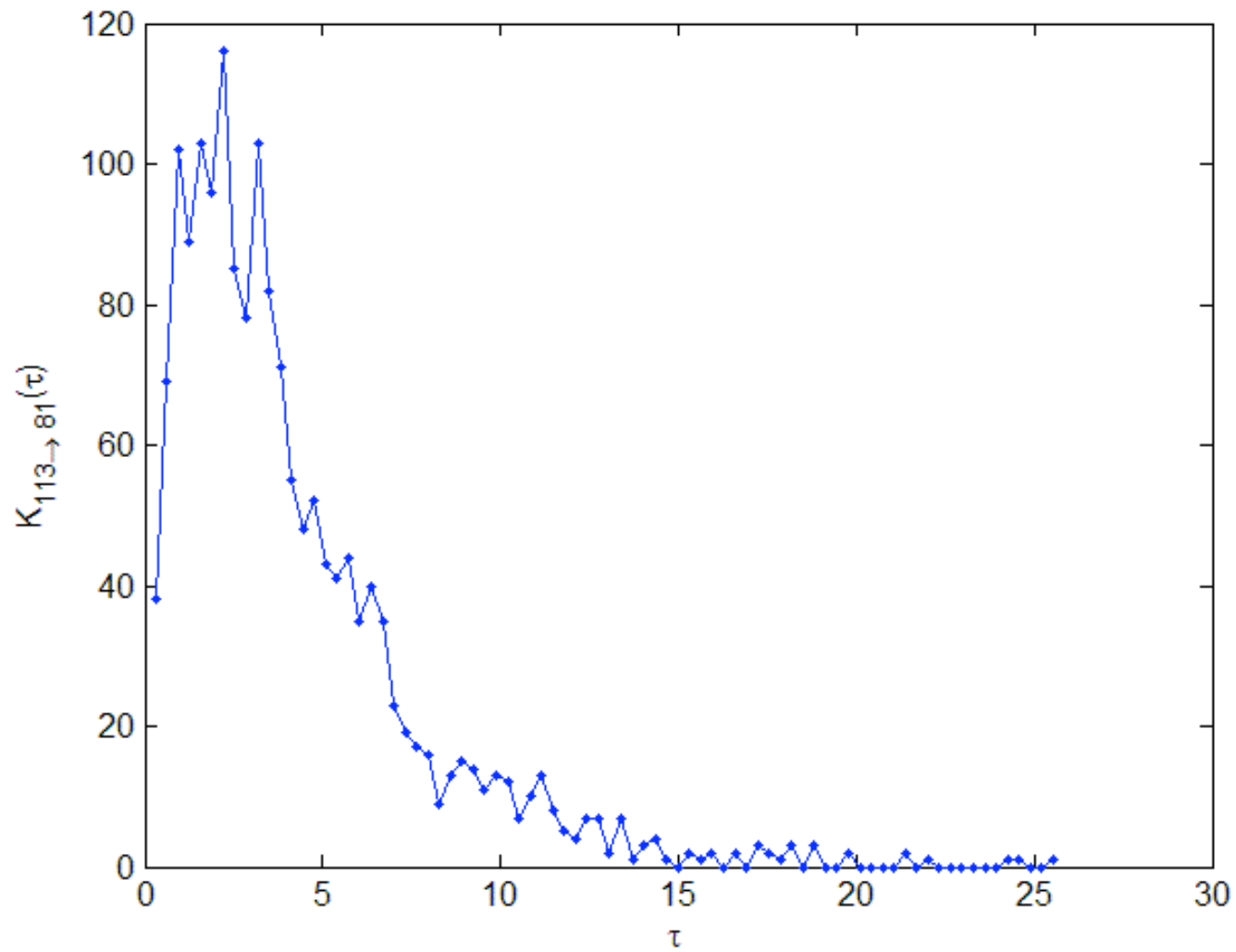
What do we want to compute from the coarse model?

$$P_N(t)$$

$$P_{i-1}(t) \quad P_i(t) \quad P_{i+1}(t)$$

$P_i(t)$ microscopically it is the probability of being somewhere between Milestones *i-1* and *i+1* at time *t* such that the last Milestone passed by the trajectory is Milestone *i*.

Provides a blurred (coarse) spatial description without reducing the number of degrees of freedom

The QK picture

- Define $P_i(t)$: prob of being at *i* at time *t*

- Define $Q_i(t)$: prob of transition to *i* at *t*

- Define $K_{i,j}(\tau)$: conditional probability of a transition from *i* to *j* after incubation time τ.

- Then hopping dynamics are defined by:

$$Q_i(t) = P(0)_i \, \delta\left(t - 0^+\right) + \sum_j \int_0^t \left[Q_j(t') K_{j,i}(t-t') \right] dt'$$

$$P_i(t) = \int_0^t Q_i(t') \left[1 - \int_0^{t-t'} \left[\sum_j K_{i,j}(\tau) \right] d\tau \right] dt'$$

With the matrix $K_{ij}(\tau)$ determined, compute long time kinetics

$$Q_i(t) = P_i(0)\delta(t - 0^+) + \int_0^t \left[Q_j(t')K_{ji}(t - t') \right] dt'$$

$$P_i(t) = \int_0^t Q_i(t') \left[1 - \int_0^{t-t'} \left[\sum_j K_{ij}(\tau) \right] d\tau \right] dt'$$

- by direct integration
- by Laplace transform (Shalloway)
- by trajectory statistics (Vanden Eijnden)

Computing the average first passage time from Milestone 1 to N

$$\tau = \sum_{l=1,\ldots,L} \left[\tau_{12}^{(1)} + \ldots + \tau_{ij}^{(n)} + \ldots + \tau_{kN}^{(L)} \right] \cdot \left(p_{12} \cdots p_{ij} \cdots p_{kN} \right)$$

$\tau_{ij}^{(n)}$ is a random variable sampled from $K_{ij}(\tau)$

$p_{ij} = \int_0^\infty K_{ij}(\tau) d\tau$ is the transition probability from i to j
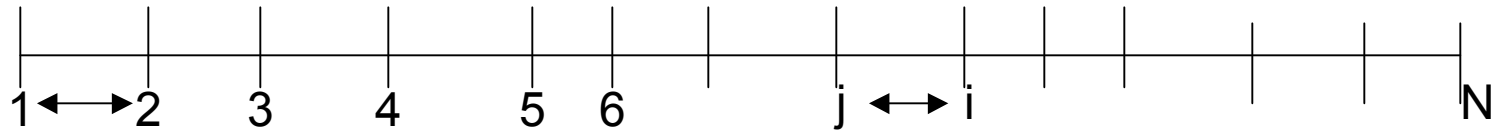
$\sum_j p_{ij} = 1$

A few points:
The limit $L \to \infty$ is considered
The N state is absorbing
$\tau_{Nk}=0$ $p_{NN}=1$

To average over τ…

Compute <τ>

Define a random matrix $(T)_{ij} \equiv p_{ij}\tau_{ij}$

the matrix $(P)_{ij} \equiv p_{ij}$ of size $N \times N$

the vector $\hat{\tau}^{(L)} \equiv \left(\tau_1^{(L)}, \tau_2^{(L)}, ..., \tau_N^{(L)}\right)^{\mathrm{T}}$ with elements $\tau_i^{(L)}$

that are the overall first passage times using L steps
to go from i to N

Then

$$\hat{\tau}^{(L)} = \sum_{l=1}^{L} P^{L-l}TP^{l-1}1 = \sum_{l=1}^{L} P^{L-l}T1$$

where $1 = (1,1,...,1)^{\mathrm{T}}$ and $P1 = 1$

The last tricks to compute <τ>

For $L \rightarrow \infty$ the trajectory is absorbed at Milestone N.

Since the time at N does not count $(\tau_{NN} = 0) \Rightarrow \hat{\tau}^{L+1} = \hat{\tau}^L$

$$\hat{\tau}^{(L)} = \sum_{l=1}^{L} P^{L-l} T1$$

$$\hat{\tau}^{(L+1)} = \sum_{l=1}^{L+1} P^{L+1-l} T1 = P \sum_{l=1}^{L} P^{L-l} T1 + T1 = P\hat{\tau}^{(L)} + T1$$

$(I - P)\hat{\tau} = T1$ (remove eigenvector 1 from the set),

$T1 = 0$ $(I - P)1 = 0$, define, $\bar{I}, \bar{P}, \bar{T}$ of size $(N-1) \times (N-1)$).

T is a random matrix and the average first passage time

is obtained by avergaing over elements of $\bar{T}$. $\left\langle \left(\bar{T}\right)_{ij} \right\rangle = p_{ij} \left\langle \tau_{ij} \right\rangle$

Only the first moments of τ_{ij} are required to compute $\left\langle \hat{\tau} \right\rangle$.

$$\left\langle \hat{\tau} \right\rangle = \left(\bar{I} - \bar{P}\right)^{-1} \left\langle \bar{T} \right\rangle 1$$

Is Milestoning correct (or what are the assumptions)?
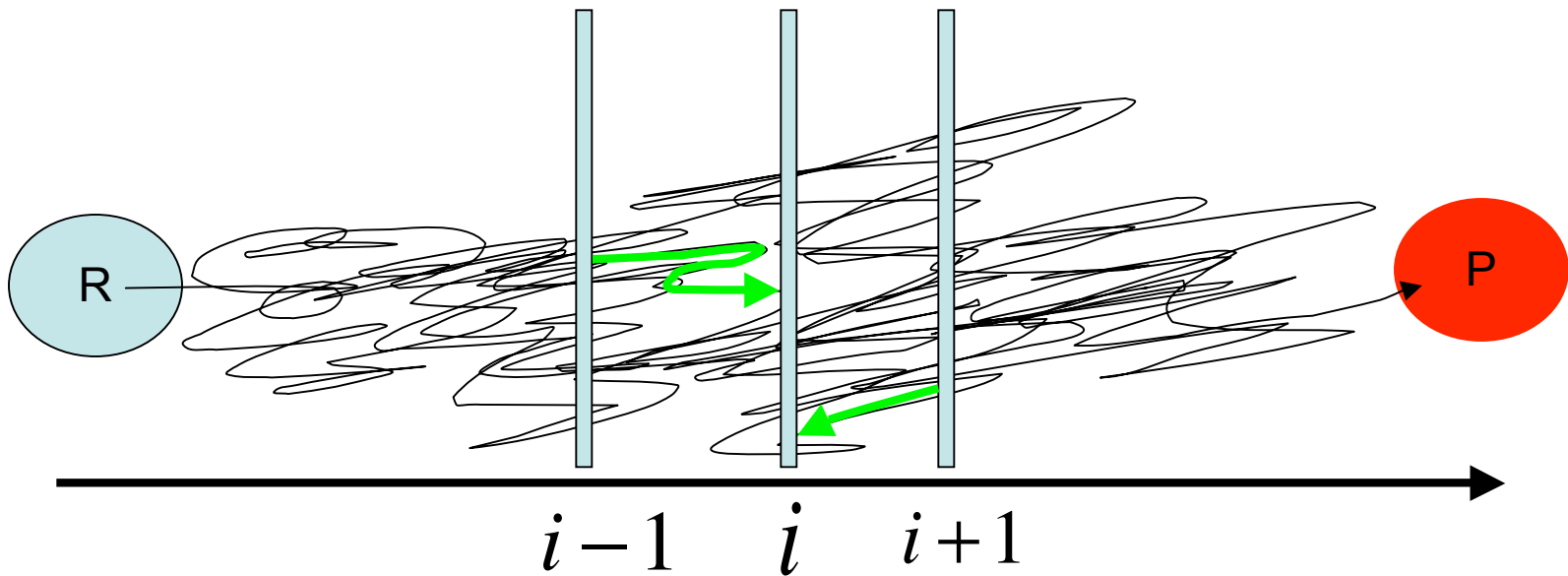
Assumption

Let S_j be the hypersurface of Milestone j.

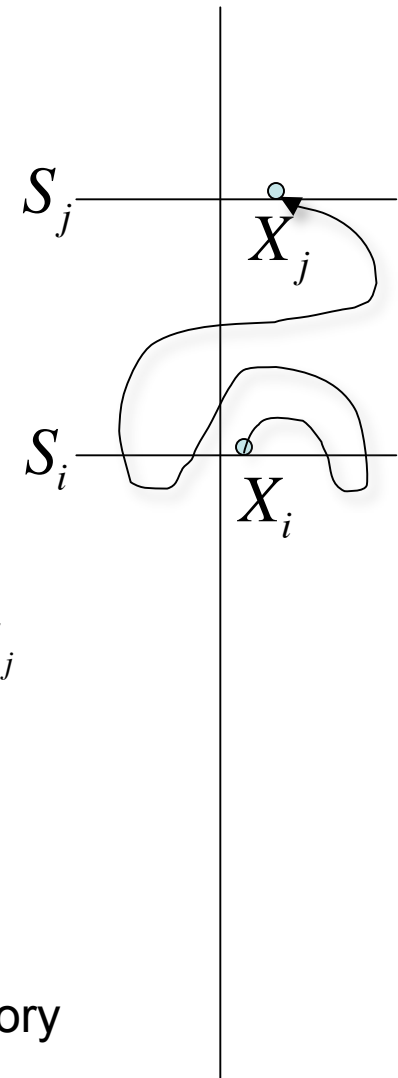Let X_j be a coordinate vector $X_j \in R^{3N}$ and $X_j \in S_j$.

$\rho\left(X_j\right)$ is the distribution at S_j initiating the short trajectories.

$\theta_{ij}\left(X_i\right)$ is the distribution obtained from first passage traj on S_j
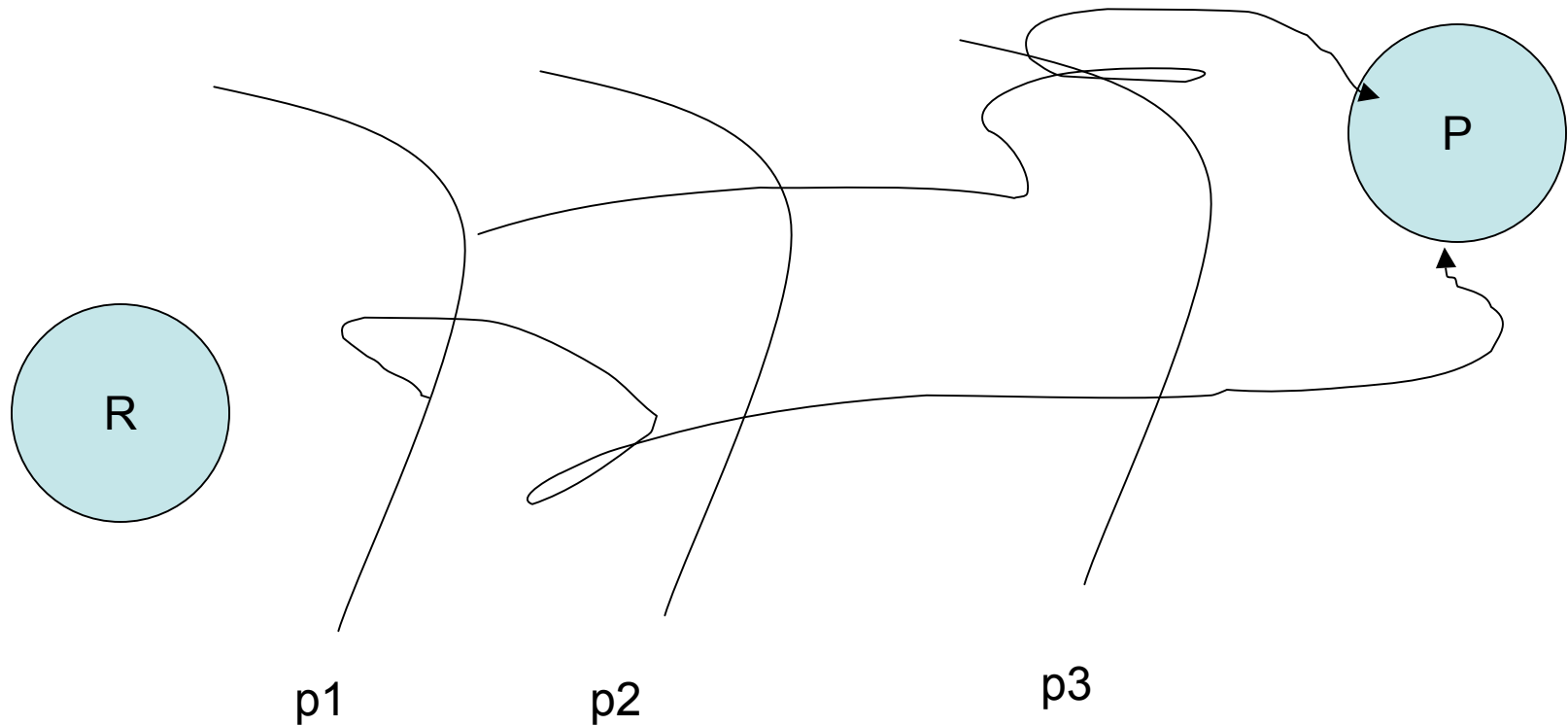
if initiated at S_i according to $\rho\left(X_i\right)$

Assumption : $\theta_{ij}\left(X_i\right) = \rho\left(X_j\right)$

Implies:
- Loss of memory
- committers

S_j

X_j

S_i

X_i

Committers are special surfaces with equal probability of reaching for the first time the product and the not the reactants
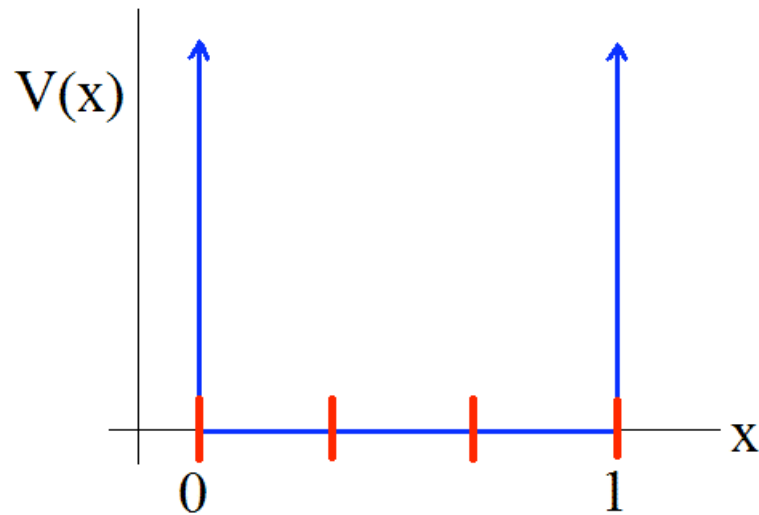
R

P

p1

p2

p3

Committer surfaces can be calculated exactly (solving partial differential equations) in 2-3 dimensions for Brownian dynamics and approximated at higher dimensions and other types of equations of motion.

Or a tunnel picture
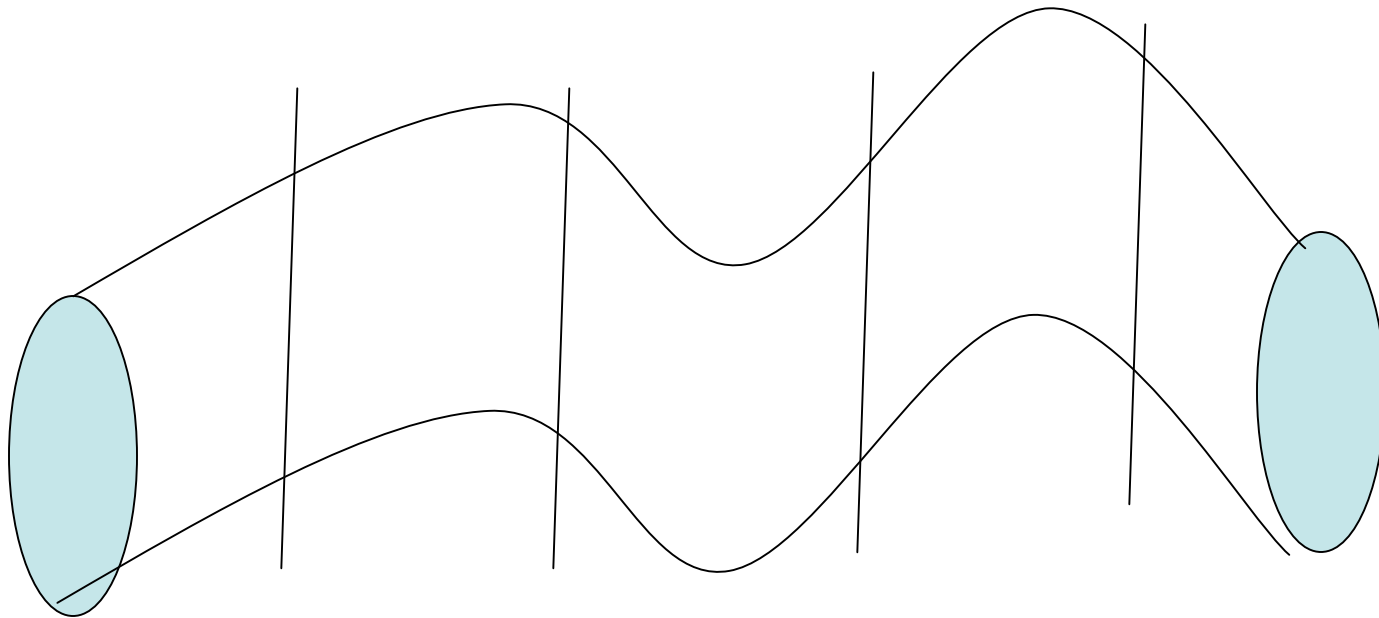
Relaxation in planes faster than transition between planes.
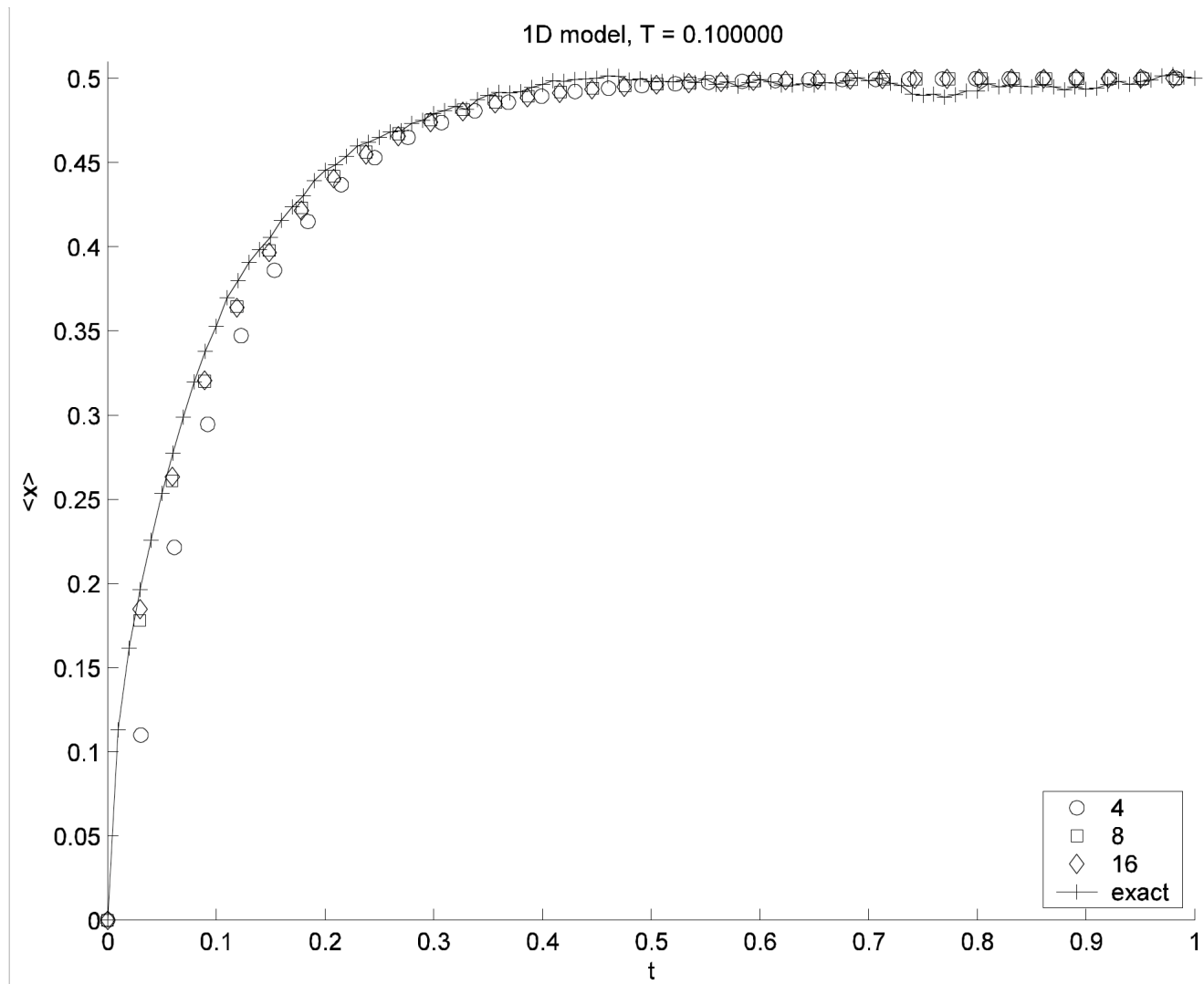General but heuristic.

Toy model: 1D box simulation

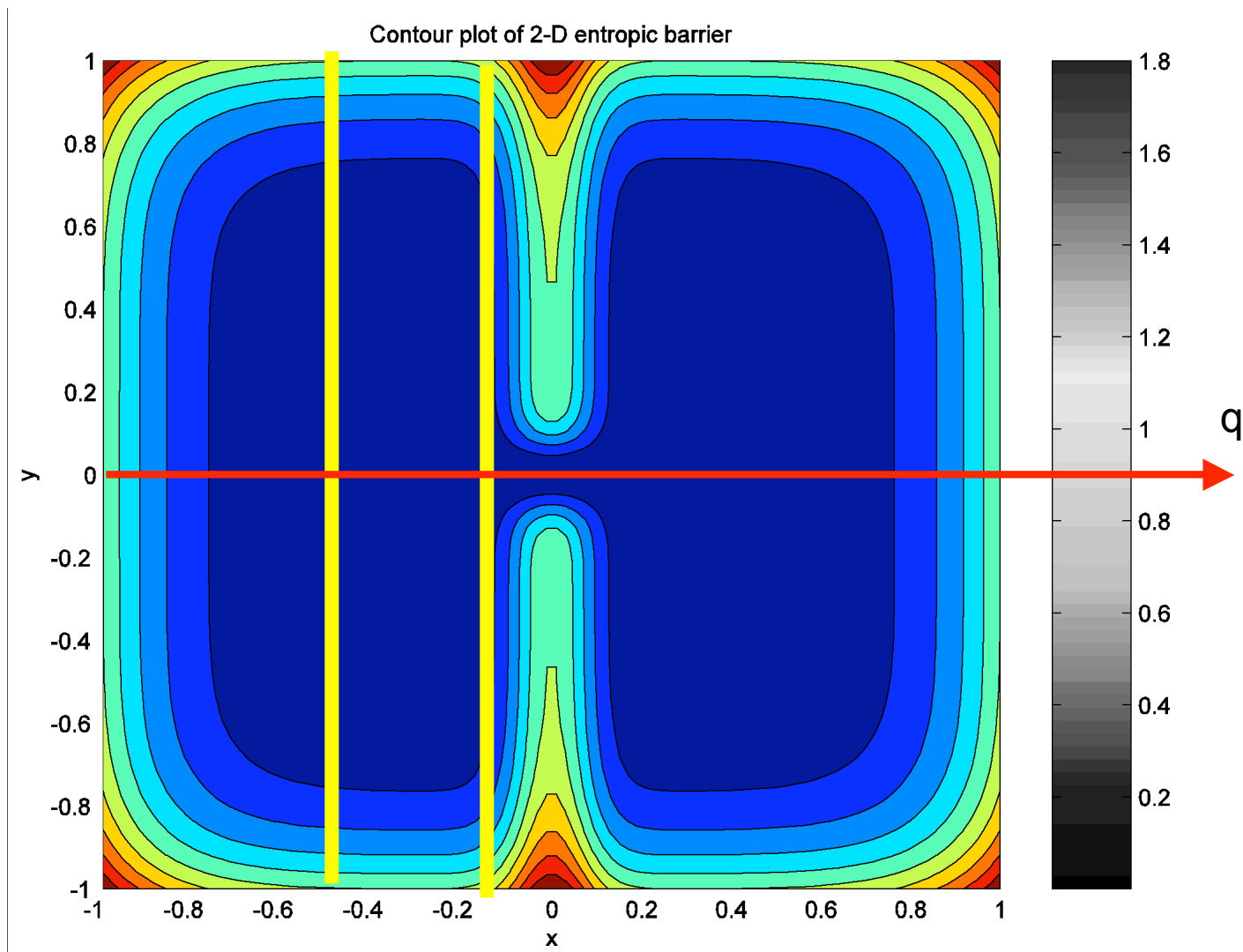$$\gamma \frac{dX}{dt} = -\nabla U + R$$

$$\langle R \rangle = 0 \quad \langle R(t)R(t') \rangle = C\delta(t - t')$$

- Microscopic dynamics are Brownian
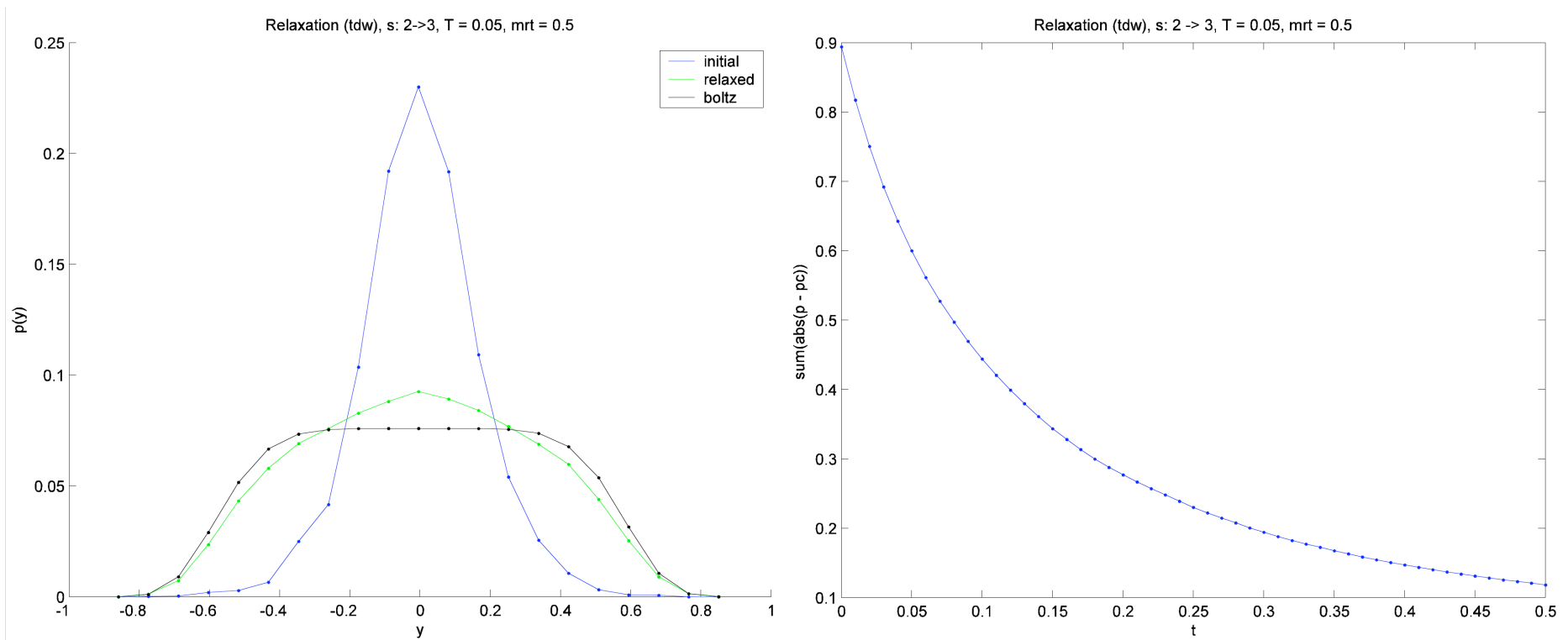- Simulations run at various temperatures and for 4, 8, and 16 milestones
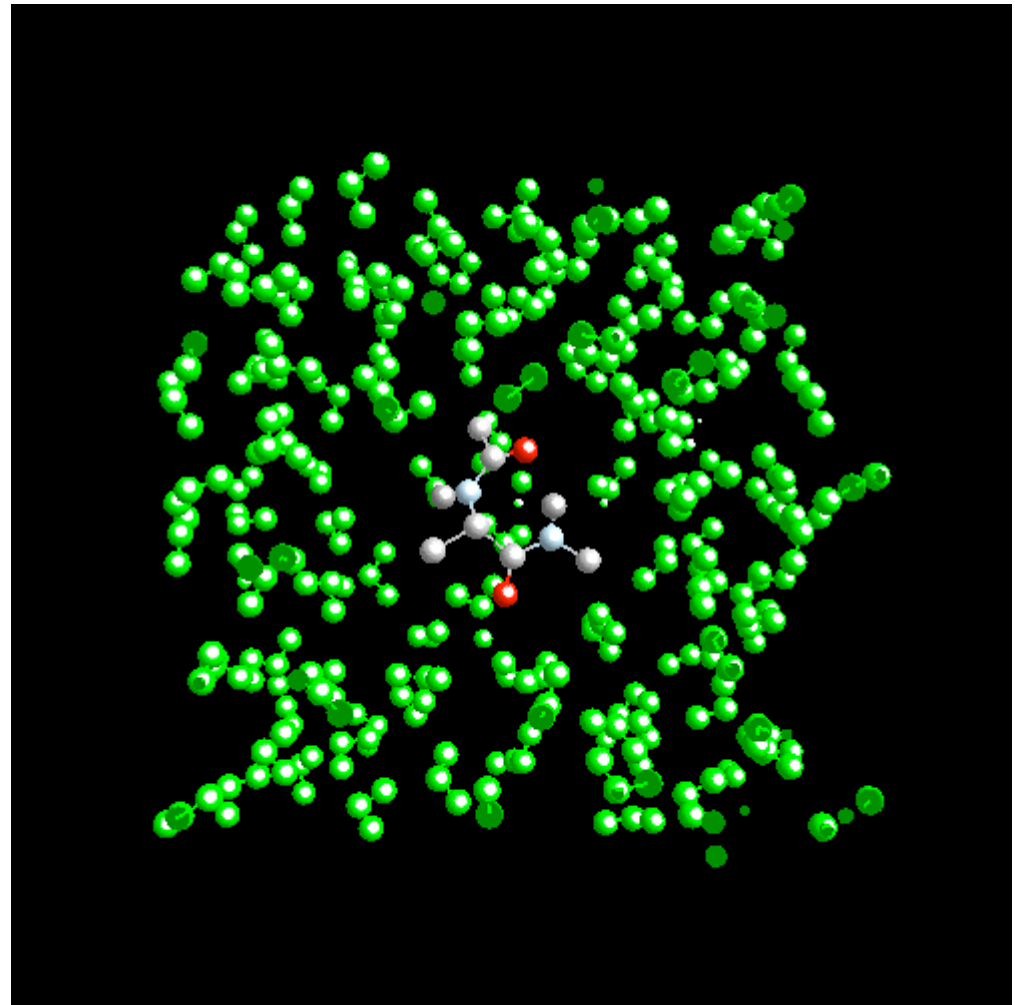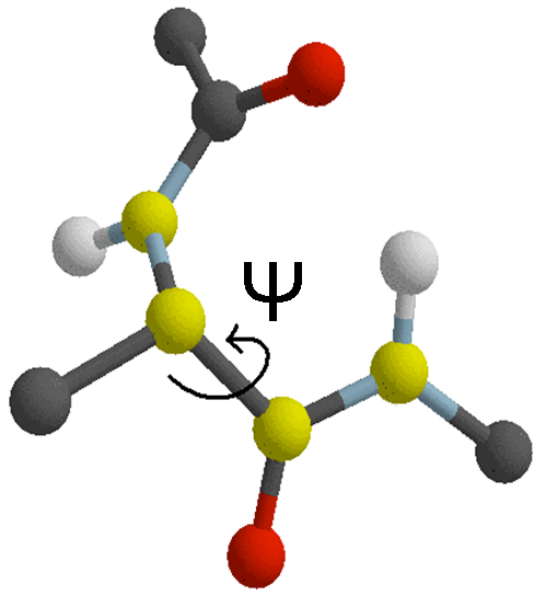
1D reaction curves (5000 trajs/MLST)

1D model, T = 0.100000

Legend:
- ○ 4
- □ 8
- ◇ 16
- + exact

2D simulation

Contour plot of 2-D entropic barrier
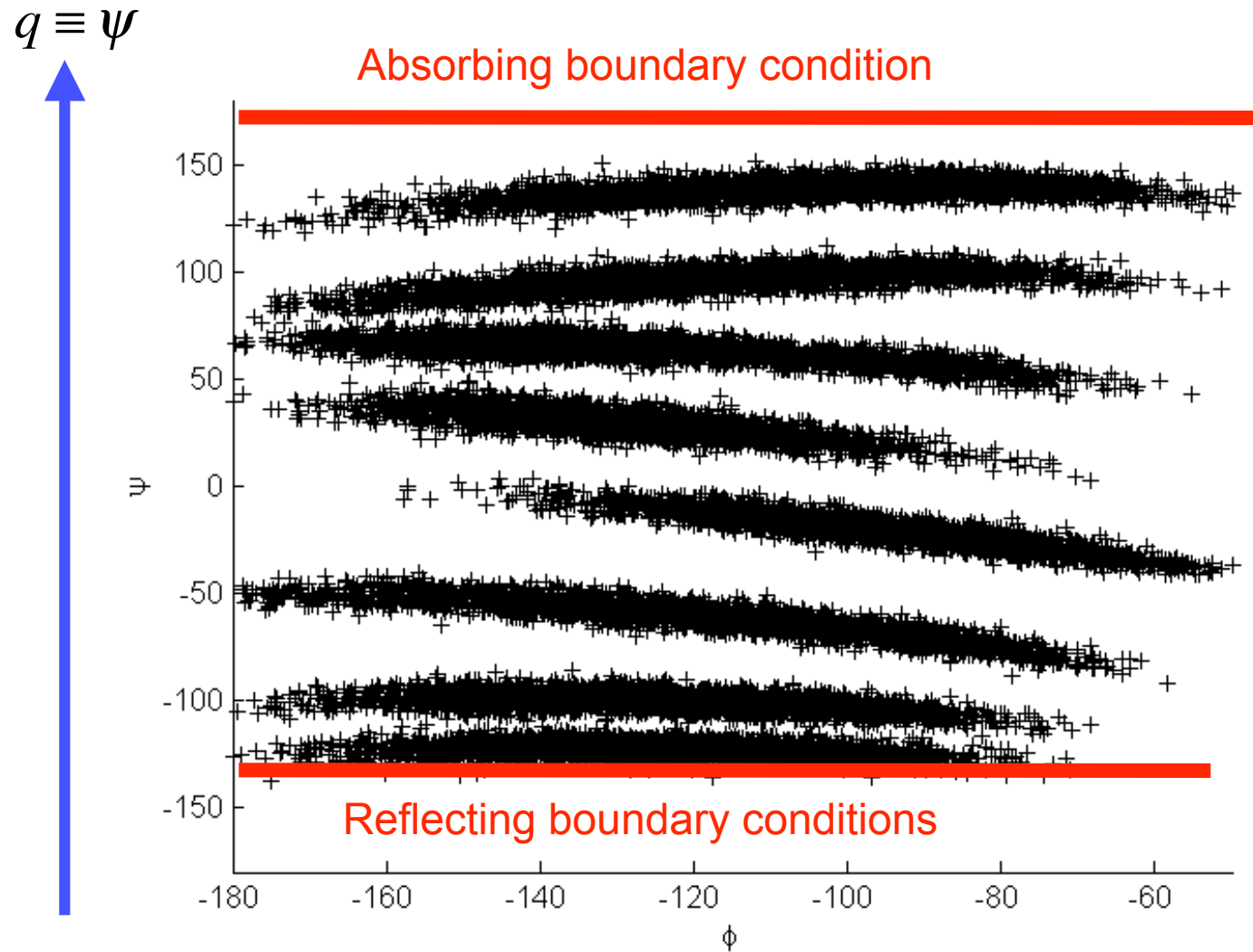
Memory loss demonstration

$$\tau_\perp \sim 0.15$$

$$\tau_\parallel \sim 6.34, \quad \left(\int\limits_0^\infty \tau K_2(\tau)\, d\tau \right)$$
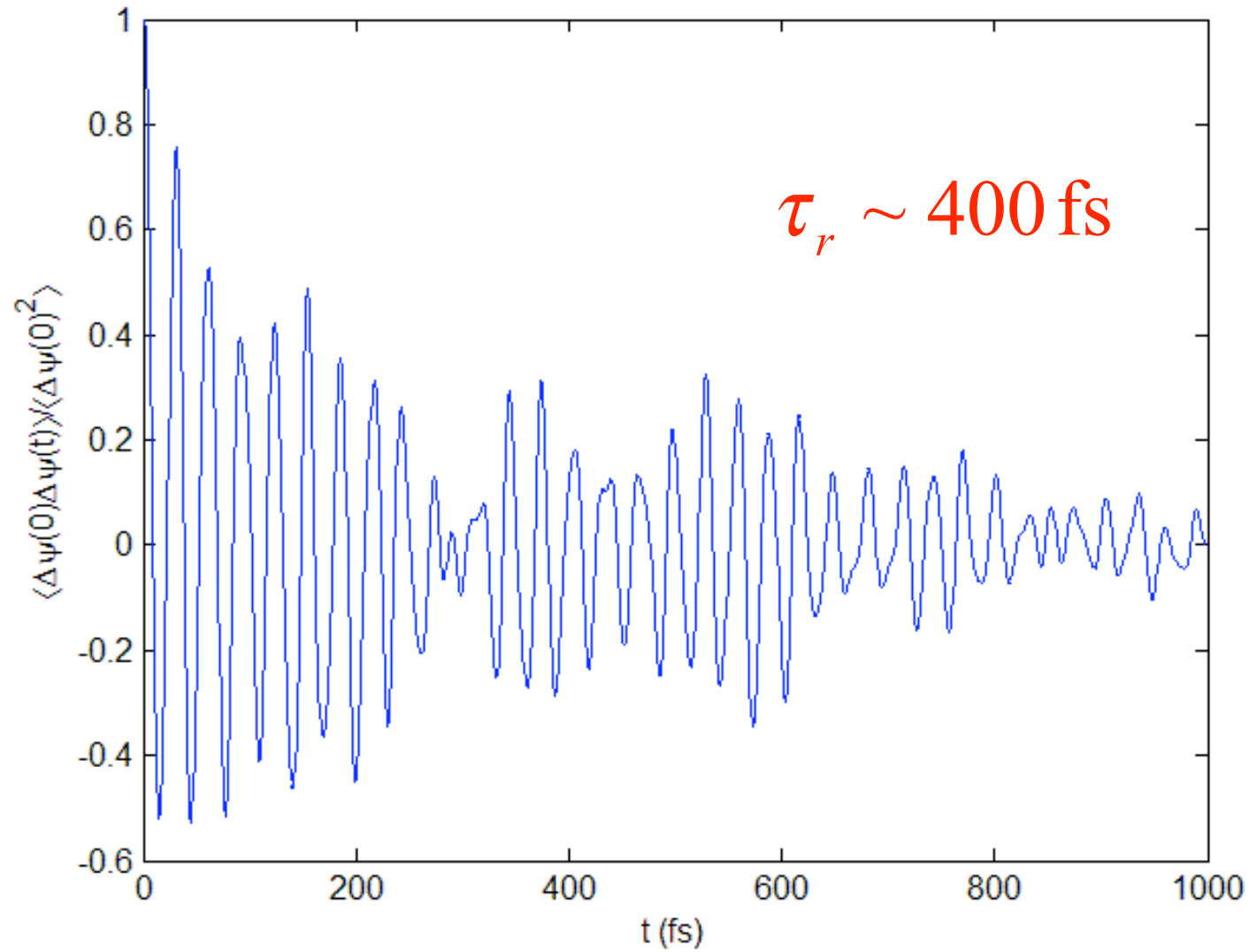
Alanine Dipeptide

JCP, **126**, 145104 (2007)

29

Preparing initial conditions by sampling

$q \equiv \psi$

Absorbing boundary condition

Reflecting boundary conditions

Torsion velocity auto-correlation indicates when milestoning assumption is being violated
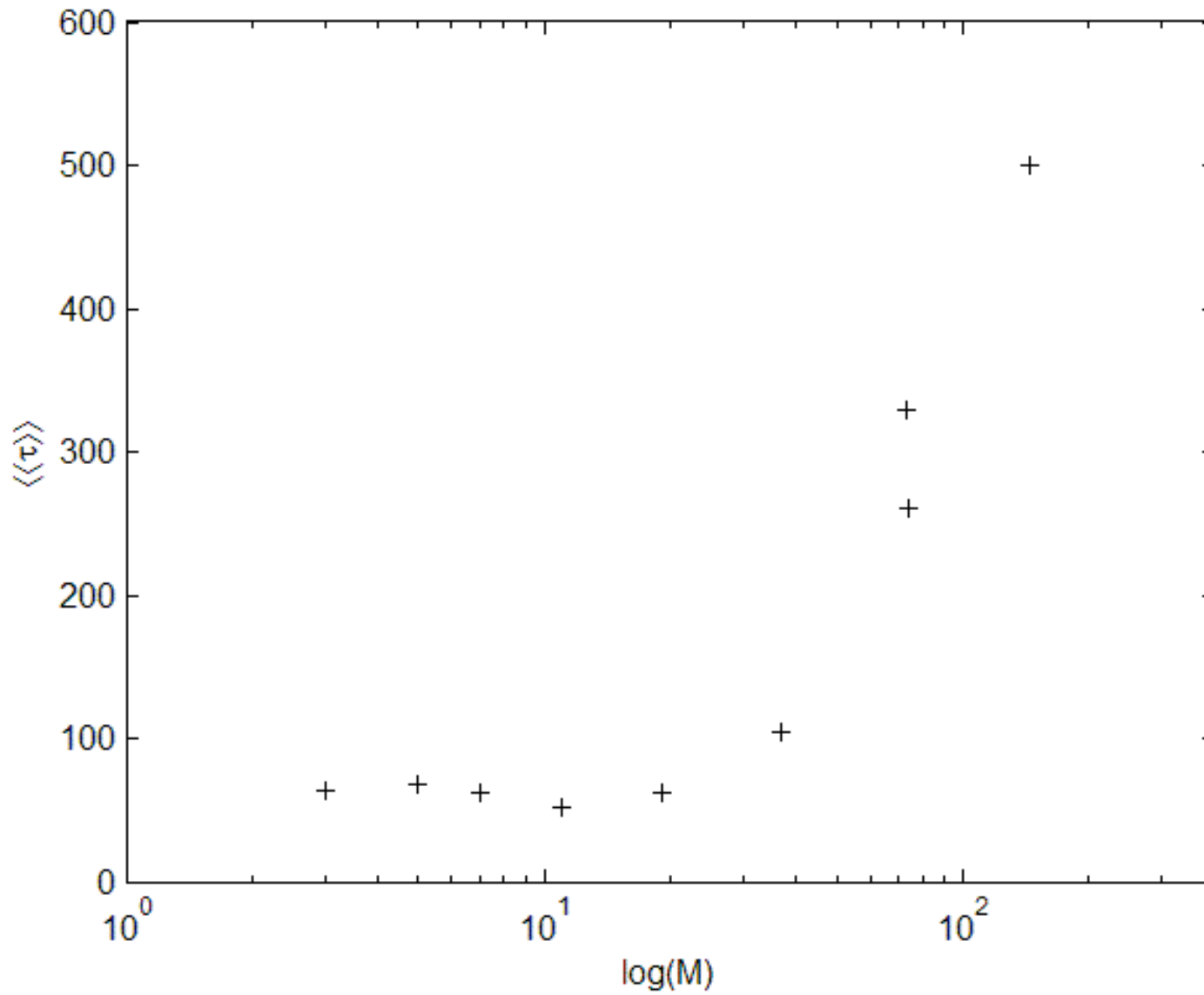

$$\tau_r \sim 400\,\text{fs}$$

Average Incubation Times vs. Velocity Relaxation Time

$$\langle \bar{\tau} \rangle \equiv \frac{1}{M} \sum_{i=1}^{M} \int_{0}^{\infty} \tau K_{s_i}(\tau) d\tau$$
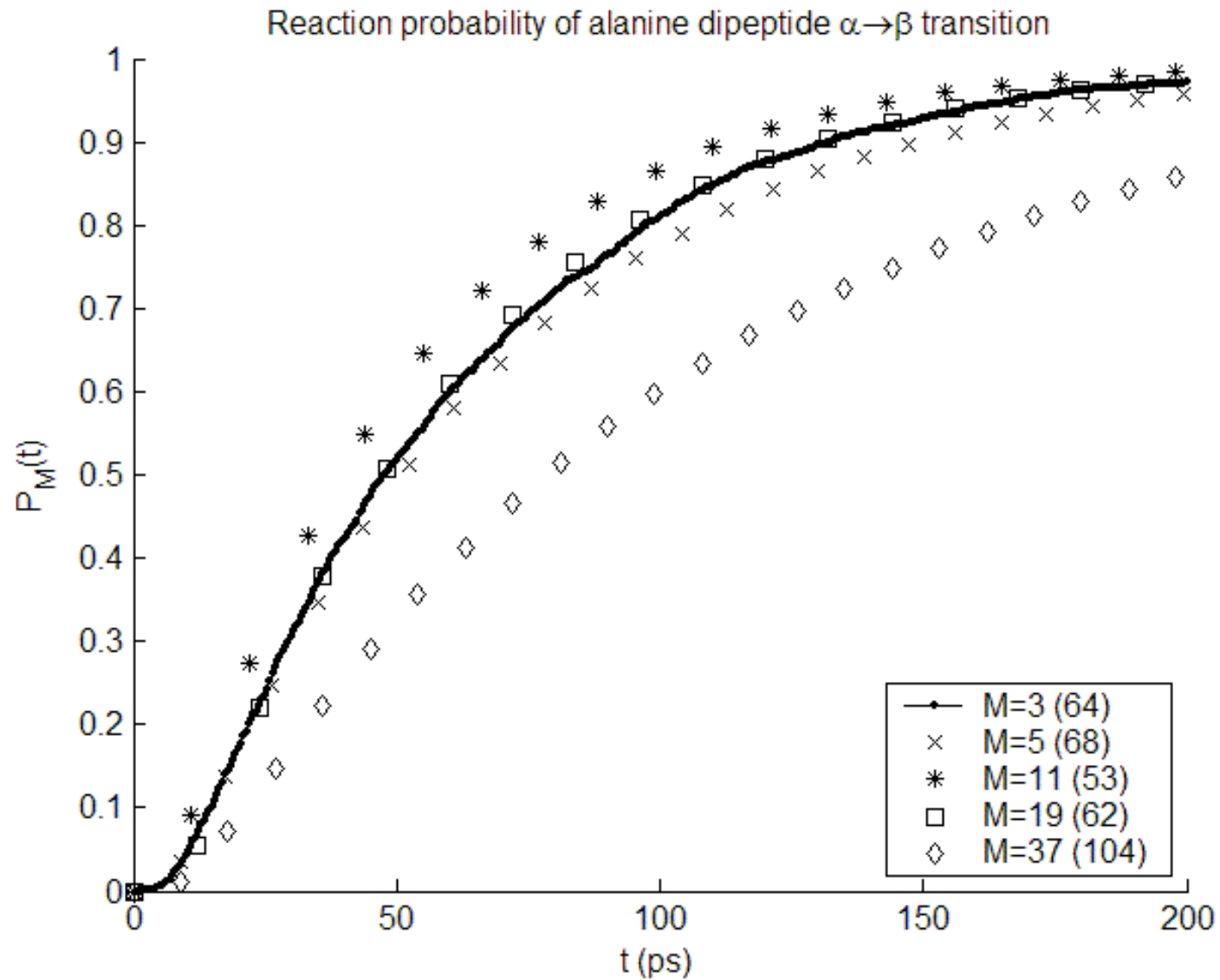
M	$\langle \bar{\tau} \rangle$ (fs)
144	31
74	58
73	58
37	129
19	373
11	1305
7	3581
5	10902

$\uparrow \tau_r > \langle \bar{\tau} \rangle$

Rate Results

Reaction curves

Reaction probability of alanine dipeptide $\alpha \rightarrow \beta$ transition

Summary

- Milestoning divides RC into fragments whose kinetics can be computed independently then "glued" together
- Provides factor of M improvement in computational efficiency on serial machines, plus exp bootstrapping
- Uses LFPTDs from microscopic dynamics: $K_s^{\pm}(\tau)$
- System distribution $P_s(t)$ given by simple integral equations that can be easily solved numerically
- Correct kinetics for solvated alanine dipeptide (x 9 speedup)
- Predicts microsecond *Scapharca* rate with ~10 ns total serial time
- Sub-millisecond rate for myosin recovery stroke with total run time (serial) - speedup : more than 10,000

$$200\,sample \times 241\,mlst \times 0.01ns + 0.1ns \times 241 = 501.6ns$$

Milestoning papers

- Ron Elber, "A milestoning study of the kinetics of an allosteric transition: Atomically detailed simulations of deoxy Scapharca hemoglobin", Biophysical J. ,2007 92: L85-L87
- Anthony M.A. West, Ron Elber, and David Shalloway, "Extending molecular dynamics timescales with milestoning: Example of complex kinetics in a solvated peptide", J. Chem. Phys. 126,145104(2007)
- Anton K. Faradjian and Ron Elber, "Computing time scales from reaction coordinates by milestoning", J. Chem. Phys. 120:10880-10889(2004)
- Code (moil + zmoil) available from https://wiki.ices.utexas.edu/clsb/wiki