A mathematical framework for Exact Milestoning

David Aristoff (joint work with Juan M. Bello-Rivas and Ron Elber)

www.math.colostate.edu/~aristoff

July 2015

- Milestoning is a technique for estimating mean first passage times (MFPTs).
- Exact Milestoning is a variant which yields exact times in a certain limit.
- Milestoning and Exact Milestoning are both practical algorithms.
- Both algorithms are appropriate for systems with rough energy landscapes.

Important application: in silico drug design.

Estimating characteristic time for a drug to dissociate from a protein target.

- Efficiency is based on the use of short trajectories simulated in parallel.
- These trajectories start on a milestone and end at a neighboring milestone.
- The milestones are usually codimension 1 hypersurfaces.

Examples:

1D milestoning: milestones = level sets of a scalar reaction coordinate
Network milestoning: milestones = faces of Voronoi cell boundaries

The problem.

Let (X_t) be a stochastic dynamics and R, P disjoint subsets of state space. We want to compute the mean first passage time of (X_t) from R to P.

Source and sink.

When (X_t) reaches P, it immediately restarts at R.

This assumption does not affect the MFPT but it is useful for theory.

Examples to keep in mind:

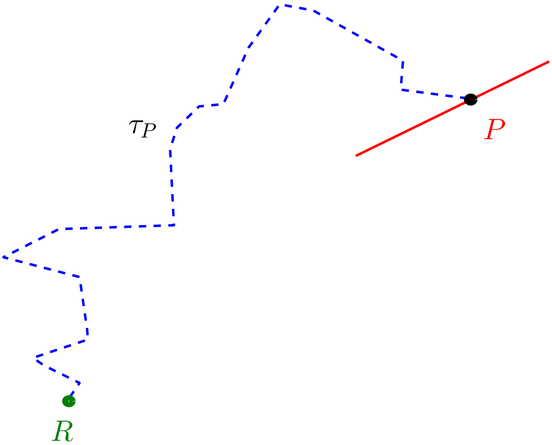
- State space is a torus, (X_t) is a diffusion;
- State space is discrete, (X_t) is a continuous time Markov process.

Below we think of R as a point, though it can be a distribution too.

Let (X_t) take values in a standard Borel space.

Our goal.

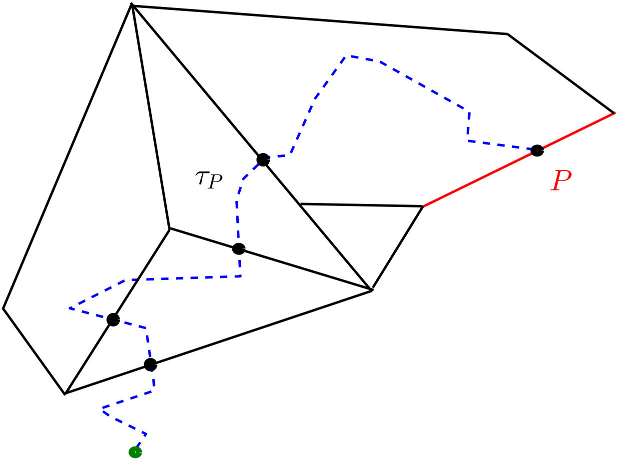
To efficiently compute $\mathbb{E}^R[\tau_P]$, where τ_P is the first time for (X_t) to hit P.



We coarse-grain (X_t) using closed sets called *milestones*.

Milestones.

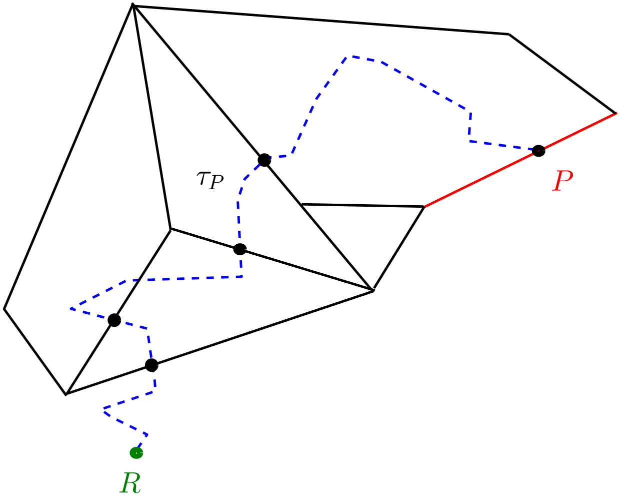
The milestones are closed sets with pairwise disjoint interiors.



We coarse-grain (X_t) using closed sets called *milestones*.

Milestones.

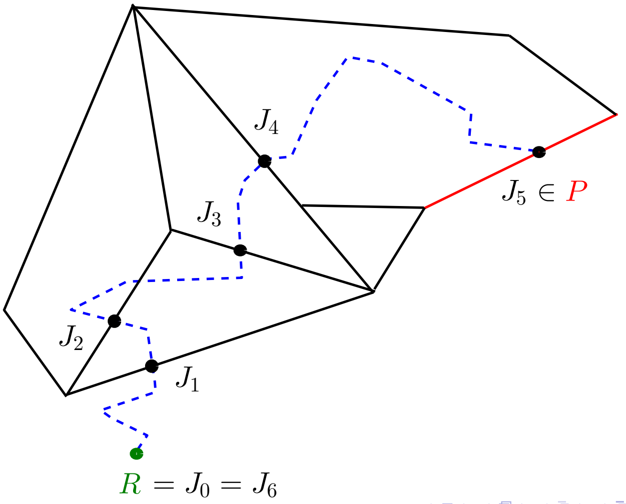
P and R are two of the milestones. M is the union of the milestones.



Assume (X_t) is strong Markov with càdlàg paths.

Jump chain.

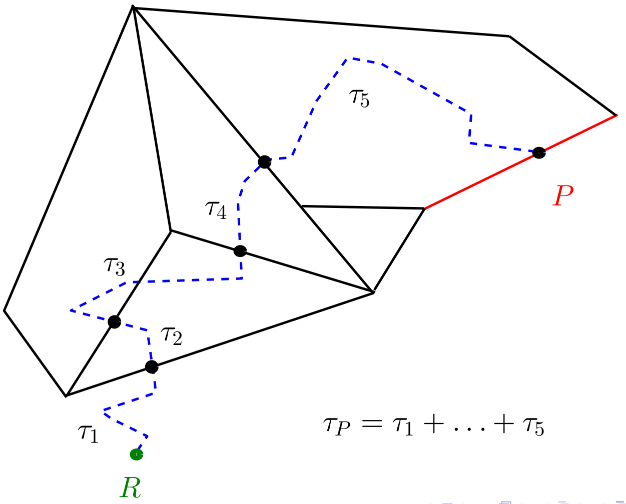
Keep the first hit points, J_n, on the milestones. (J_n) is a Markov chain on M.



Assume (X_t) is strong Markov with càdlàg paths.

Sojourn times.

Keep the times, τ_n, between the first hit points. τ_n depends only on J_{n-1} and J_n.



Define $Y_t = J_n$ for $\tau_0 + \ldots + \tau_n \le t < \tau_0 + \ldots + \tau_{n+1}$, and let (Y_t) start on M.

Theorem.

(Y_t) is a semi-Markov process on M with the same FPT to P as (X_t).

Proof: This is clear from construction...

Definition.

Let $\tau_P = \inf\{t > 0 \, : \, Y_t \in P\}$ and $\sigma_P = \min\{n \ge 0 \, : \, J_n \in P\}$.

Assumption.

$\mathbb{E}^\xi[\tau_P]$ and $\mathbb{E}^\xi[\sigma_P]$ are finite for all initial distributions ξ on M.

This ensures that (Y_t) has finite MFPTs to P and is nonexplosive.

Definition.

Let $K(x,dy)$ be the transition kernel for (J_n) and $\bar{K}(x,dy) = \begin{cases} K(x,dy), & x \notin P \\ 0, & x \in P \end{cases}$.

K and $\bar{K}$ have left/right actions on measures/bounded functions in the usual way.

Theorem.

(J_n) has an invariant probability measure μ, defined by

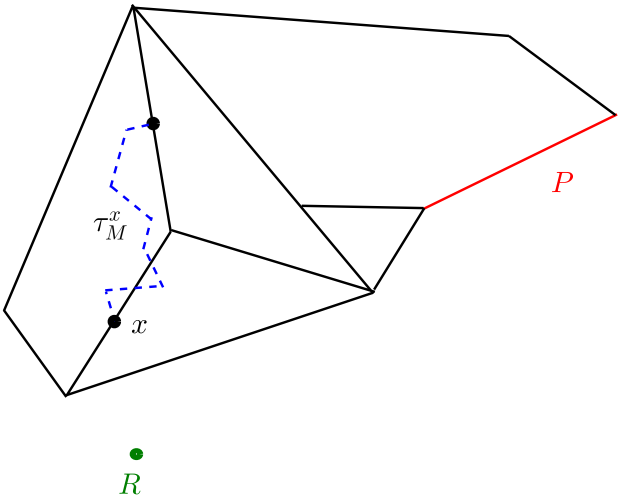
$$\mu = Z^{-1} \sum_{n=0}^{\infty} \delta_R \bar{K}^n,$$

where Z is a normalization constant. Moreover, $\mu(P) > 0$.

Proof: Show that the residence time $\mathbb{E}^R\left[\sum_{n=0}^{\sigma_P} \mathbb{1}_{J_n \in \cdot}\right]$ is invariant.

Definition.

Let M_x be the milestone containing x and $\tau_M^x = \inf\{t > 0 : Y_t \in M \setminus \mathrm{int}(M_x)\}$.



Definition.

Let M_x be the milestone containing x and $\tau_M^x = \inf\{t > 0 \,:\, Y_t \in M \setminus \mathrm{int}(M_x)\}$.

Theorem.

With μ the invariant distribution for (J_n),

$$\mu(P)\mathbb{E}^R[\tau_P] = \mathbb{E}^\mu[\tau_M] := \int_M \mu(dx)\mathbb{E}^x[\tau_M^x].$$

Proof: Write $\tau_P = (\tau_P - \tau_M^x) + \tau_M^x$, condition on $Y_{\tau_M^x} = y$, integrate w.r.t. $\mu(dy)$.

Note that $\mathbb{E}^x[\tau_M^x]$ can be obtained from short trajectories running in parallel.
So if we can sample efficiently from μ, we can efficiently estimate $\mathbb{E}^R[\tau_P]$.

Theorem.

Assume (J_n) is aperiodic in the following sense:

$$\text{g.c.d.}\{n \geq 1 \,:\, \mathbb{P}^R(\sigma_P = n-1) > 0\} = 1.$$

Then for any initial distribution ξ,

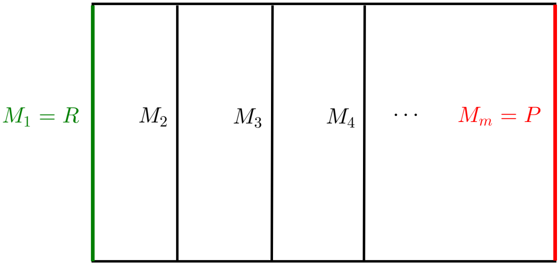
$$\lim_{n\to\infty} \|\xi K^n - \mu\|_{TV} = 0.$$

Proof: Coupling argument, using the fact that R is recurrent for (J_n).

This suggests how to estimate μ: start with a "guess" $\xi = \mu_0$, and iterate.

Aside: 1D milestoning.

The milestones are $R = M_1, M_2, \ldots, M_m = P$. If $J_n \in M_1$, then $J_{n+1} \in M_2$. And if $J_n \in M_j$ for $j = 2, \ldots, m-1$, then $J_{n+1} \in M_{j-1}$ or $J_{n+1} \in M_{j+1}$, both w.p. > 0.



Here, $\|\xi K^n - \mu\|_{TV} \to 0$ for all $\xi \iff$ the number of milestones is odd.

Why? Say there are m milestones. Then (J_n) can go from R to P in $m-1$ or $m+1$ steps, and m, $m+2$ are coprime when m is odd.

If m is even and ξ is supported on the odd indexed milestones, then after an even (resp. odd) number of steps (J_n) lies in an odd (resp. even) indexed milestone.

Error Analysis.

The error in milestoning has two sources:

- Error in the approximation $\tilde{\mu}$ of μ;
- Error due to time discretization $\tilde{X}_{n\delta t}$ of (X_t).

Define
$\tilde{\tau}_M^x = \min\{n > 0 \,:\, \text{line segment from } \tilde{X}_{(n-1)\delta t} \text{ to } \tilde{X}_{n\delta t} \text{ intersects } M \setminus \text{int}(M_x)\}$.

Theorem.

The error in the Milestoning approximation of the MFPT satisfies

$$|\mathbb{E}^R[\tau_P] - \tilde{\mu}(P)^{-1}\mathbb{E}^{\tilde{\mu}}[\tilde{\tau}_M]| \le c_1|\mu(P)^{-1} - \tilde{\mu}(P)^{-1}| + \tilde{\mu}(P)^{-1}\left(c_2\|\mu - \tilde{\mu}\|_{TV} + \phi(\delta t)\right),$$

where ϕ is a function depending only on the time step error, and

$$c_1 = \mathbb{E}^{\mu}[\tau_M], \qquad c_2 = \sup_{x \in M} \mathbb{E}^x[\tau_M^x].$$

Proof: Triangle inequalities.

Theorem.

The error in the Milestoning approximation of the MFPT satisfies

$$|\mathbb{E}^R[\tau_P]-\tilde{\mu}(P)^{-1}\mathbb{E}^{\tilde{\mu}}[\tilde{\tau}_M]| \le c_1|\mu(P)^{-1}-\tilde{\mu}(P)^{-1}|+\tilde{\mu}(P)^{-1}\left(c_2\|\mu-\tilde{\mu}\|_{TV}+\phi(\delta t)\right),$$

where ϕ is a function depending only on the time step error, and

$$c_1 = \mathbb{E}^{\mu}[\tau_M], \qquad c_2 = \sup_{x\in M} \mathbb{E}^x[\tau_M^x].$$

This holds for Exact Milestoning *as well as the original Milestoning*.

Open problem.

Suppose (X_t) is Brownian dynamics: $dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t$, and $\tilde{\mu} = \mu_0 = Z^{-1}e^{-\beta V}\,dx$. When/how much can $\|\tilde{\mu}-\mu\|_{TV}$ be controlled?

Is an iteration scheme based on powers of a numerical approximant $\tilde{K}$ consistent?

Theorem. (Ferré et. al., 2013)

Suppose (J_n) is geometrically ergodic (GE): there exists $\kappa \in (0,1)$ such that

$$\sup_{x\in M} \|\delta_x K^n - \mu\|_{TV} = O(\kappa^n).$$

Given $\epsilon > 0$, if $\tilde{K}$ is sufficiently close to K in operator norm, then

$$\sup_{x\in M} \|\delta_x \tilde{K}^n - \tilde{\mu}\|_{TV} = O(\tilde{\kappa}^n), \qquad \|\mu - \tilde{\mu}\|_{TV} < \epsilon,$$

where $\tilde{\mu}$ is some probability measure on M and $\tilde{\kappa} \in (\kappa, 1)$.

Lemma.

(J_n) is GE if the probability to reach P in n steps has a uniform lower bound.

Lemma.

(J_n) is GE if it is strong Feller and aperiodic, and state space is compact.

Example.

Let (X_t) be Brownian dynamics:

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

and for $N = 10$ let

$$V(x_1, x_2) = \sum_{k_1=-N}^{N-1} \sum_{k_2=-N}^{N-1} C_{k_1,k_2}\, f_{k_1,k_2}(x_1, x_2), \qquad (1)$$

where C_{k_1,k_2} is either 0 or sampled uniformly from $\left(-\frac{1}{2\pi}, \frac{1}{2\pi}\right)$, each w.p. $\frac{1}{2}$, and

$$f_{k_1,k_2}(x_1, x_2) = \begin{cases} \cos(2\pi k_1 x_1)\cos(2\pi k_2 x_2), & \text{w.p. } \frac{1}{3}, \\ \cos(2\pi k_1 x_1)\sin(2\pi k_2 x_2), & \text{w.p. } \frac{1}{3}, \\ \sin(2\pi k_1 x_1)\sin(2\pi k_2 x_2), & \text{w.p. } \frac{1}{3}. \end{cases}$$

V defines a rough energy landscape on the torus $\mathbb{R}^2/\mathbb{Z}^2$.

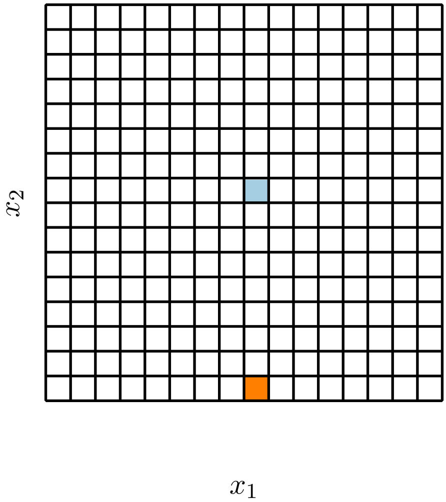


Figure: Source R (center bottom), sink P (center), and other milestones (line segments).

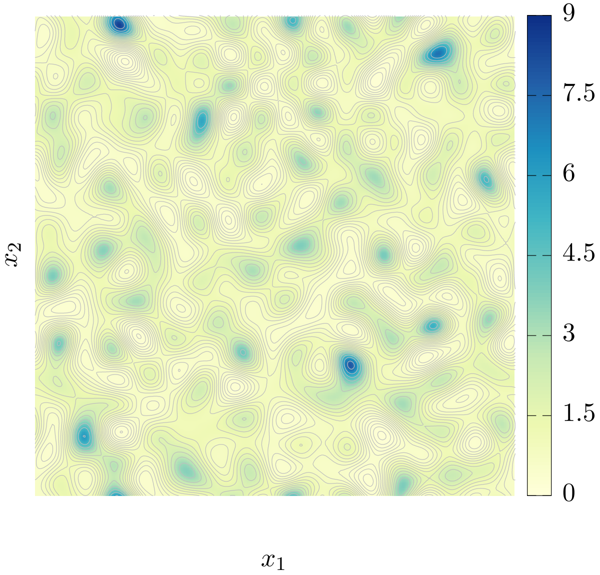


Figure: Contour map of the canonical Gibbs density $Z^{-1}e^{-\beta V}$.

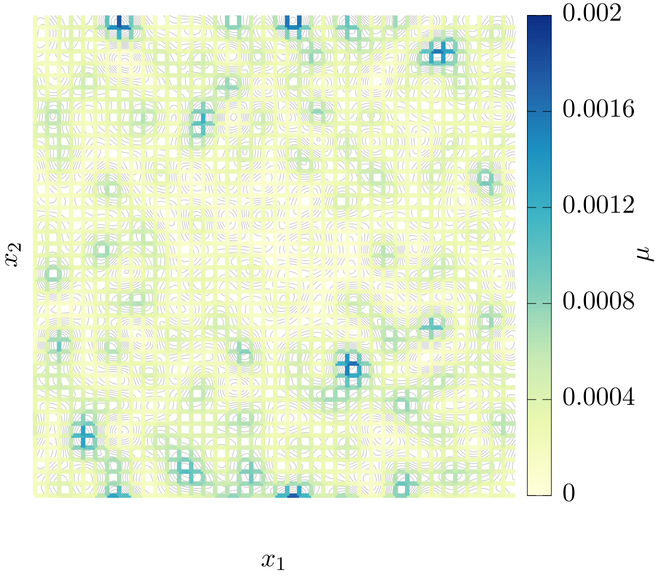


Figure: The stationary measure μ, superimposed on contour lines of V.

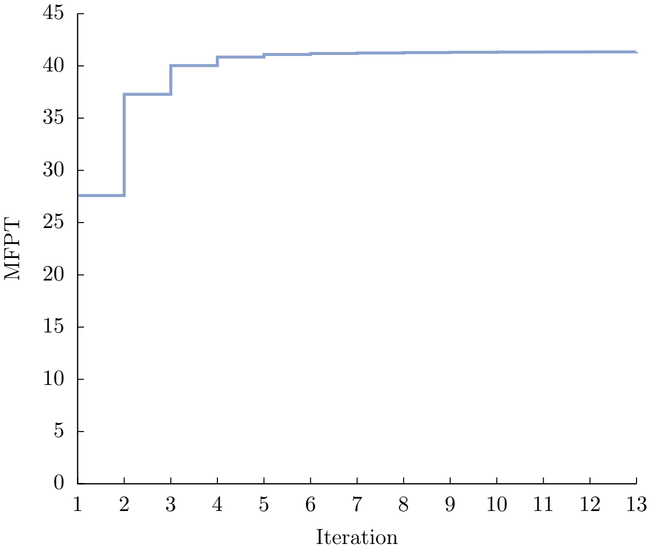


Figure: The MFPT vs. the number of iterations, starting at $\mu_0(dx) = Z^{-1}e^{-\beta V(x)}dx$.