Mathematical analysis of temperature accelerated dynamics

David Aristoff (joint work with T. Lelièvre)

School of Mathematics
University of Minnesota

August 12, 2013

Molecular dynamics

- We consider *overdamped Langevin dynamics*:

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

used to model the evolution of the position vector X_t of N particles in an energy landscape defined by the potential energy $V : \mathbb{R}^{3N} \to \mathbb{R}$.

The dynamics (1) are obtained as a limit as $m \to 0$ or $\gamma \to \infty$ of the *Langevin dynamics*

$$\begin{aligned} dX_t &= m^{-1} P_t\,dt \\ dP_t &= -\nabla V(X_t)\,dt - \gamma m^{-1} P_t\,dt + \sqrt{2\gamma\beta^{-1}}\,dW_t. \end{aligned} \tag{2}$$

- This energy landscape typically has many metastable states, corresponding to basins of attraction of the gradient dynamics $dy/dt = -\nabla V(y)$.
- In applications it is of interest how X_t moves between these basins – this is the so-called *metastable dynamics*.

Definition.

We write D for a generic basin of attraction of $dy/dt = -\nabla V(y)$.

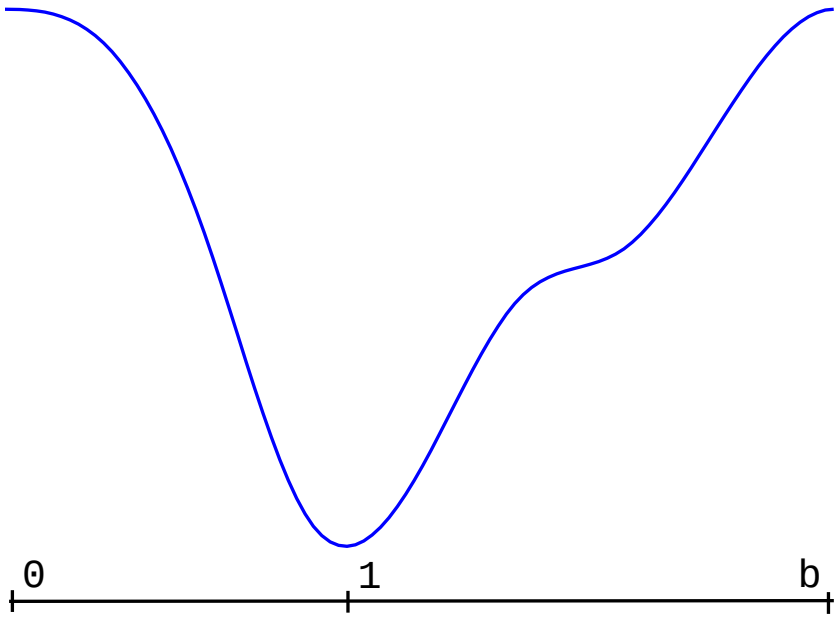

Figure : The basin $D = [0, b]$ of attraction of 1 w.r.t. $dy/dt = -\nabla V(y)$.

- Let $S : \mathbb{R}^{3N} \to \mathbb{N}$ be a function which labels the basins of attraction of $dy/dt = -\nabla V(y)$. So each basin D has the form $D = S^{-1}(i)$, $i \in \mathbb{N}$.

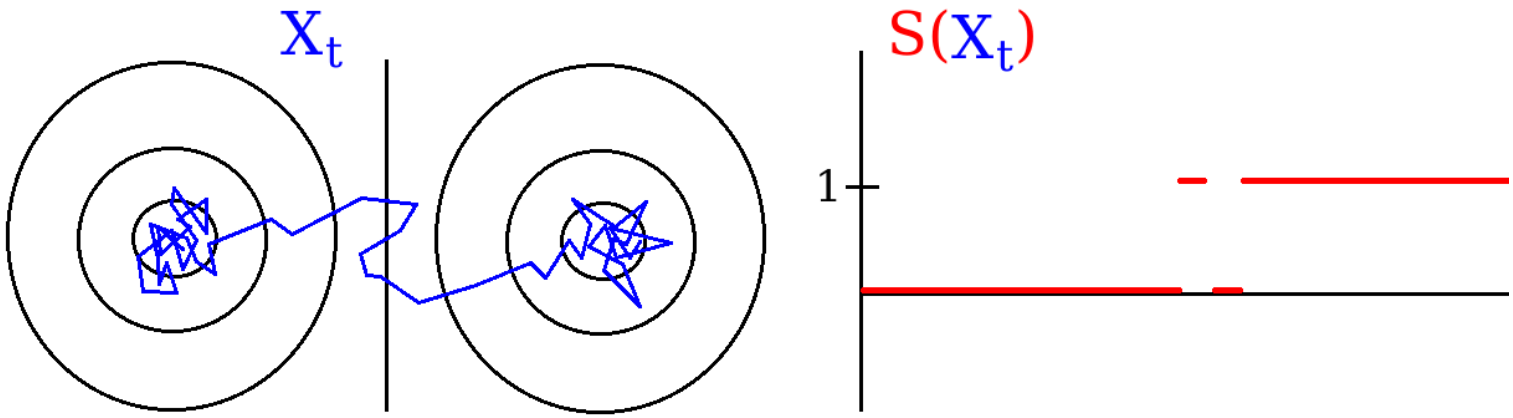

Figure : A trajectory of X_t and $S(X_t)$, with two basins labeled 0 and 1.

- The *metastable dynamics* is then $S(X_t)_{t\geq 0}$.

Problem:

Efficiently generate approximations $\hat{S}(t)_{t\geq 0}$ of $S(X_t)_{t\geq 0}$.

The quasistationary distribution

- With metastable dynamics, the time scale to reach "local equilibrium" in a basin is much smaller than the time scale to exit the basin.
- The notion of local equilibrium can be formalized using the *quasistationary distribution (QSD)*.

Definition.

The superscript in X_t^μ means X_t has initial distribution given by μ: $X_0 \sim \mu$.

Definition.

The QSD ν in D satisfies the following:

$$\nu(A) = \lim_{t\to\infty} \mathbb{P}(X_t^\mu \in A \,|\, X_s^\mu \in D \ \forall \ s \in [0,t])$$

for any probability measure μ supported in D and any measurable $A \subset D$.

- Given that X_t remains in D, the distribution of X_t converges exponentially fast to the QSD in D, no matter the initial distribution of X_0.

- If X_0 is distributed according to the QSD in D, then the first exit time of X_t from D is exponentially distributed and independent of the exit position:

Theorem.

Define $\tau = \inf\{t > 0 : X_t^\nu \notin D\}$. Then

$$\mathbb{P}(\tau > t) = e^{-\lambda t} \qquad \text{and} \qquad \tau,\ X_\tau \text{ are independent.}$$

- The theorem can be used to tackle the following:

Subproblem.

Efficiently generate an exit event of X_t^ν from a given basin D.

- The idea is to iterate the subproblem solution to generate the metastable dynamics approximation $\hat{S}(t)_{t\geq 0}$.
- The assumption $X_0 \sim \nu$ should not be drastic because the time scale for X_t to reach the QSD in D is much smaller than the time scale for X_t to exit D.

Generating exit events

Definition.

Let $\{\partial D_i\}_{i=1,2,\ldots,n}$ be a measurable partition of ∂D.

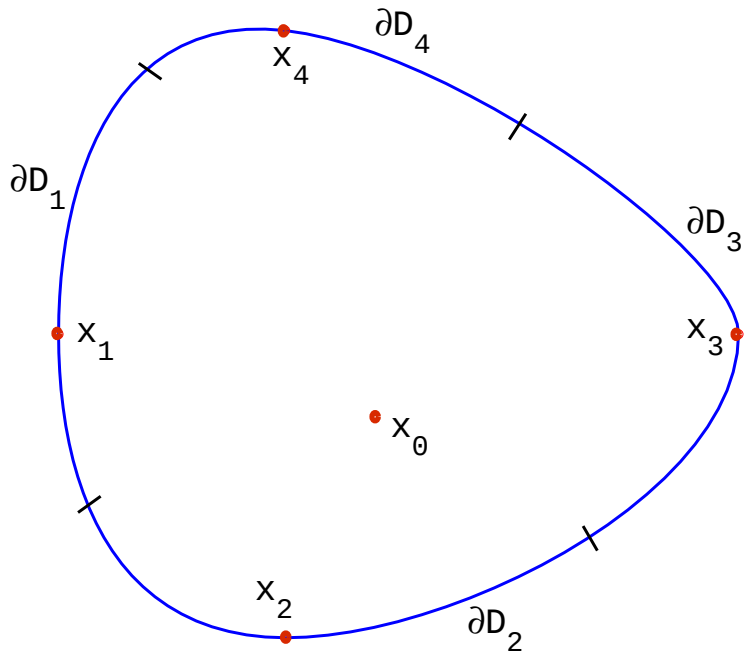

Figure : The basin D of attraction of x_0. Each ∂D_i is a neighborhood of a saddle point, x_i $(i \geq 1)$, of V in ∂D.

Generating exit events

Definition.

Let $\{\partial D_i\}_{i=1,2,\ldots,n}$ be a measurable partition of ∂D.

Definition.

Define r.v.'s τ and I by

$$\tau = \inf\{t > 0 : X_t^\nu \notin D\} \qquad \text{and} \qquad I = i \quad \Leftrightarrow \quad X_\tau^\nu \in \partial D_i.$$

Define λ and p_i by

$$\lambda^{-1} := \mathbb{E}[\tau] \qquad \text{and} \qquad p_i := \mathbb{P}(I = i).$$

Generating exit events

Definition.

Let $\{\partial D_i\}_{i=1,2,\ldots,n}$ be a measurable partition of ∂D.

Definition.

Define r.v.'s τ and I by

$$\tau = \inf\{t > 0 : X_t^\nu \notin D\} \qquad \text{and} \qquad I = i \quad \Leftrightarrow \quad X_\tau^\nu \in \partial D_i.$$

Define λ and p_i by

$$\lambda^{-1} := \mathbb{E}[\tau] \qquad \text{and} \qquad p_i := \mathbb{P}(I = i).$$

- An exit event from D, starting at the QSD, is represented by the pair (τ, I), with τ the exit time and I the exit pathway.

Generating exit events

Definition.

Let $\{\partial D_i\}_{i=1,2,\ldots,n}$ be a measurable partition of ∂D.

Definition.

Define r.v.'s τ and I by

$$\tau = \inf\{t > 0 : X_t^\nu \notin D\} \qquad \text{and} \qquad I = i \quad \Leftrightarrow \quad X_\tau^\nu \in \partial D_i.$$

Define λ and p_i by

$$\lambda^{-1} := \mathbb{E}[\tau] \qquad \text{and} \qquad p_i := \mathbb{P}(I = i).$$

- An exit event from D, starting at the QSD, is represented by the pair (τ, I), with τ the exit time and I the exit pathway.
- Starting at the QSD, the expected time to exit D is λ^{-1} and the probability to exit through ∂D_i is p_i.

The last theorem leads to the following:

Theorem.

Let $\{T_i\}_{i=1,2,\ldots,n}$ be independent r.v.'s with $\mathbb{P}(T_i > t) = e^{-\lambda p_i t}$. Then

$$\left(\min_{1 \le i \le n} T_i, \arg \min_{1 \le i \le n} T_i \right) \sim (\tau, I).$$

- The preceding applies to any dynamics whenever the QSD in D exists.
- The theorem can be used to sample exit events from D provided that estimates of the parameters λp_i are available.
- From now on we consider only overdamped Langevin dynamics and assume:

Assumption.

V is a Morse function, D is the basin of attraction of x_0 w.r.t. $dy/dt = -\nabla V(y)$, and each ∂D_i is a neighborhood in ∂D of a single (index one) saddle point, x_i, of V on ∂D.

The Arrhenius law

- Under the preceding assumption λp_i can be estimated by the so-called Arrhenius law:

The Arrhenius law

$$\lambda p_i \approx \eta_i e^{-\beta(V(x_i)-V(x_0))} \quad \text{for large } \beta \tag{3}$$

- Here η_i is a known function of the eigenvalues of the Hessian matrix of V at the saddle point x_i and minimum x_0.
- In particular η_i is β-independent.
- The Arrhenius law is assumed valid when

$$\beta(V(x_i) - V(x_0)) >> 1.$$

- If the locations of the saddle points are known a priori, the theorem along with equation (3) can be used to sample exit events from D.

Temperature accelerated dynamics (TAD)

Notation.

Let β^{hi} and β^{lo} be a high and low temperature. We use superscripts ^{hi} and ^{lo} to denote objects at β^{hi} and β^{lo}. (E.g. ν^{lo} is the QSD in D at temperature β^{lo}.)

We recall again:

Subproblem.

Generate an exit event of $X_t^{\nu^{lo}}$ from D at temperature β^{lo}.

- In TAD[1], the exit event at β^{lo} is generated by simulating multiple exit times and pathways at β^{hi}, then extrapolating what would have happened at β^{lo}.
- In TAD, the saddle point locations are not assumed to be known a priori, and it is not necessary that all the saddle points be found. Furthermore in TAD it is not required that any of the η_i be calculated.

[1] proposed in A.F. Voter and M.R. Sørensen, *J. Chem. Phys.* **112** (2000).

Exit Algorithm (for generating an exit event of $X_t^{\nu^{lo}}$ from D).

Let $N = 1$, $T_{stop} = \infty$ and iterate the following steps:

1. Let $X_0^{(N)}$ be an sample of ν^{hi}, the QSD in D at temperature β^{hi}.
2. Evolve $X_t^{(N)}$ at temperature β^{hi} until the first time, $\tau^{(N)}$, at which it exits D.
3. Now $X_{\tau^{(N)}}^{(N)} \in \partial D_i$ for some $i \in \{1, \ldots, n\}$. If $X_{\tau^{(k)}}^{(k)} \notin \partial D_i \ \forall\ 1 \leq k < N$, let

$$T_i^{lo} = \left(\tau^{(1)} + \ldots + \tau^{(N)}\right) e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))} \quad \text{extrapolated low temp exit time}$$

$$T_{min}^{lo} = \min\{T_{min}^{lo}, T_i^{lo}\}, \quad I_{min}^{lo} = i \Leftrightarrow T_{min}^{lo} = T_i^{lo} \quad \text{update fastest low temp exit event}$$

$$T_{stop} = T_{min}^{lo} / \min_{1 \leq i \leq n} e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))} \quad \text{update stopping time}$$

The minimum above can be replaced with any a priori lower bound.

4. If $\tau^{(1)} + \ldots + \tau^{(N)} < T_{stop}$, let $N = N + 1$ and return to Step 1. Otherwise store the exit event $(T_{min}^{lo}, I_{min}^{lo})$.

Remarks:

- At low temperatures the QSD in D can be efficiently sampled[2].
- By construction, the fastest extrapolated low temperature exit event will be found by time T_{stop}.

The **Exit Algorithm** is expected to be accurate when the Arrhenius law is valid:

$$\min_{1\leq i\leq n} \beta^{hi}(V(x_i) - V(x_0)) \gg 1. \tag{4}$$

The **Exit Algorithm** will be efficient when also

$$\beta^{hi} \ll \beta^{lo}. \tag{5}$$

- To see that latter, recall the stopping time T_{stop} is updated via

$$T_{stop} = T^{lo}_{min} / \min_{1\leq i\leq n} e^{-(\beta^{hi}-\beta^{lo})(V(x_i)-V(x_0))}, \tag{6}$$

and notice from (4) and (5) the denominator in the RHS of (6) is $\gg 1$.

[2] See for example G. Simpson and M. Luskin, *M2AM* (to appear) arxiv:1204.0819.

Mathematical analysis

Assumption.

In the **Exit Algorithm**, $X_0^{(N)} \sim \nu^{hi}$ for all $N \geq 1$ and:

(A1) $e^{-(\beta^{hi}-\beta^{lo})(V(x_i)-V(x_0))}$ is everywhere replaced with $\frac{\lambda^{hi} p_i^{hi}}{\lambda^{lo} p_i^{lo}}$.

- Under the above assumption, the **Exit Algorithm** exactly replicates the low temperature exit event:

Theorem.

Under the above assumption,

$$(T_{min}^{lo}, I_{min}^{lo}) \sim (\tau^{lo}, I^{lo})$$

where we recall

$$\tau^{lo} = \inf\{t > 0 : X_t^{\nu^{lo}} \notin D\} \quad \text{and} \quad I^{lo} = i \Leftrightarrow X_{\tau^{lo}}^{\nu^{lo}} \in \partial D_i.$$

- We investigate (A1) shortly.

Idea of proof: Consider **Exit Algorithm** with no stopping criterion. Define

$$N_i = \min\{N : X^{(N)}_{\tau^{(N)}} \in \partial D_i\} \quad \text{first trial to exit thru } i\text{th pathway}$$

$$T_i^{hi} = \tau^{(1)} + \ldots + \tau^{(N_i)} \quad \text{cumulative time to first exit thru } i\text{th pathway}$$

One can show that the r.v.'s T_i^{hi} are independent and

$$\mathbb{P}(T_i^{hi} > t) = e^{-\lambda^{hi} p_i^{hi} t}.$$

So since

$$T_i^{lo} \equiv T_i^{hi} \frac{\lambda^{hi} p_i^{hi}}{\lambda^{lo} p_i^{lo}},$$

we see that the r.v.'s T_i^{lo} are independent and

$$\mathbb{P}(T_i^{lo} > t) = e^{-\lambda^{lo} p_i^{lo} t}.$$

The stopping time is chosen so that by construction, the value of the smallest T_i^{lo} will not change after T_{stop}. Appealing to our earlier theorem we are done.

Theorem. (Justifying (A1)).

Let $D = [x_1, x_2]$. Under the preceding assumptions, for $i = 1, 2$:

$$\frac{\lambda^{hi} p_i^{hi}}{\lambda^{lo} p_i^{lo}} = \left(1 + O\left(\frac{1}{\beta^{hi}} - \frac{1}{\beta^{lo}}\right)\right) e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))}$$

as $\beta^{hi} \to \infty$, $\beta^{hi}/\beta^{lo} \to$ positive const.

- This shows that the Arrhenius law extrapolation becomes exact in the small temperature limit, at least in 1D.
- We hope to prove an analogue of the theorem in any dimension.

Side note: We also have the following formal statement of the Arrhenius law:

Theorem.

Let $D = [x_1, x_2]$. Under the preceding assumptions, for $i = 1, 2$:

$$\lambda p_i = \left(1 + O\left(\beta^{-1}\right)\right) \eta_i e^{-\beta(V(x_i) - V(x_0))} \quad \text{as } \beta \to \infty.$$

Main Algorithm (for generating metastable dynamics $\hat{S}(t)_{t\geq 0}$).

Let $T_{sim} = 0$, choose a basin-dependent *decorrelation time* T_{corr}, and:

1. Starting at $t = T_{sim}$, evolve X_t^{lo} at temperature β^{lo} in the current basin D.
2. If X_t^{lo} exits D at a time $T_{sim} + \tau < T_{sim} + T_{corr}$, set

$$\hat{S}(t) = S(D), \quad t \in [T_{sim}, T_{sim} + \tau],$$

advance the clock by $T_{sim} = T_{sim} + \tau$ and go back to Step 1, with D now the new basin. Otherwise, set

$$\hat{S}(t) = S(D), \quad t \in [T_{sim}, T_{sim} + T_{corr}],$$

advance the clock by $T_{sim} = T_{sim} + T_{corr}$, and proceed to Step 3.

3. Do the **Exit Algorithm** in the current basin D. Then set

$$\hat{S}(t) = S(D), \quad t \in [T_{sim}, T_{sim} + T_{min}^{lo}],$$

advance the clock by $T_{sim} = T_{sim} + T_{min}^{lo}$, and return to Step 1, with D the new basin obtained by exiting through $\partial D_{I_{min}^{lo}}$.

Exact metastable dynamics

Theorem.

The **Main Algorithm** for generating metastable dynamics $\hat{S}(t)$ is exact in the limit $T_{corr} \to \infty$ and $\beta^{hi} \to \infty$.

- In Steps 1–2 of the **Main Algorithm**, the dynamics are simulated exactly and so no error is induced.
- We want the dynamics to be distributed according to ν^{lo} in some basin before the **Exit Algorithm** begins.
- This is why steps 1–2 of the **Main Algorithm** are included: after Step 2, we have (approximately) $X^{lo}_{T_{sim}} \sim \nu^{lo}$. Indeed, X^{lo}_t converges to ν^{lo} exponentially fast in T_{corr} (in total variation norm).
- Since T_{corr} and the simulation time of the **Exit Algorithm** will be (on average) much smaller than the time scale to exit a basin at temperature β^{lo}, the **Main Algorithm** will be efficient compared to direct sampling of trajectories at low temperature.

Acknowledgements

Collaborators T. Lelièvre (École des Ponts ParisTech)

Funding DOE Award de-sc0002085

Part of this work was completed while T. Lelièvre was an Ordway visiting professor at the University of Minnesota.

Thanks Thanks to Gideon Simpson, Art Voter and Danny Perez for fruitful discussions.

`http://www.math.umn.edu/~daristof/`