

# Mathematical analysis of temperature accelerated dynamics

David Aristoff (joint work with T. Lelièvre)

[www.math.colostate.edu/~aristoff](http://www.math.colostate.edu/~aristoff)

AIMS, Madrid, July 2014

- In many scientific applications, one would like to be able to efficiently simulate molecular dynamics (MD).

### Time scale problem:

The time scales reachable by direct MD simulation are often too small.

### Example: Thermally activated reactions

- The time step in MD simulation is on the order of femtoseconds,  $10^{-14}$  s.
- Thermally activated reactions typically occur in the time scale of microseconds,  $10^{-6}$  s.
- Direct MD simulations are not fast enough to bridge this gap.

- We consider the *Langevin dynamics* in  $\mathbb{R}^d$ :

$$dQ_t = m^{-1}P_t dt$$

$$dP_t = -\nabla V(Q_t) dt - \gamma m^{-1}P_t dt + \sqrt{2\gamma\beta^{-1}} dW_t$$

- These dynamics can be understood as **Hamiltonian dynamics** plus **dissipation** and **fluctuation**. Here  $\beta = (k_B T)^{-1}$ ,  $m =$  mass, and  $\gamma =$  damping, and  $d = 6N$ , with  $N$  the number of particles.
- We also consider the *overdamped Langevin dynamics*:

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

obtained by setting  $dP_t = 0$ ,  $X_t \equiv Q_t$ , and rescaling time by  $t \rightarrow \gamma t$ .

- These dynamics can be understood as **gradient descent** plus **Brownian motion**.

- We think of  $V$  as being smooth, with several basins of attraction.

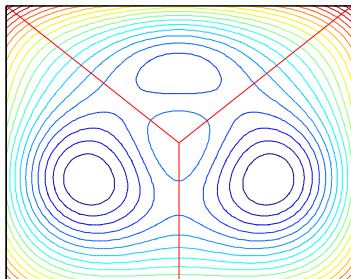


Figure: A 2D potential energy with 3 basins of attraction.

- The dynamics are *metastable* if they spend a long time each basin before hopping to the next.
- The time scale problem is often a consequence of metastability.

- For metastable dynamics, the vibrations within basins are much less interesting than the hopping between basins.

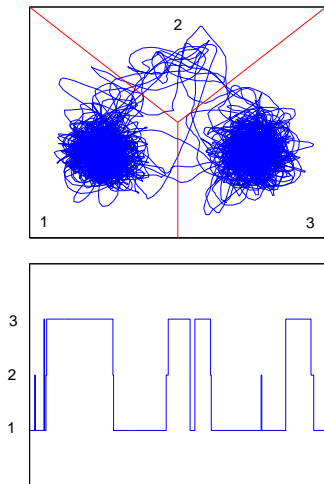


Figure: Fine and coarse scale dynamics, with basins labelled 1,2,3.

From now on we assume overdamped dynamics

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

and we assume the basins of  $V$  are bounded Lipschitz domains.

### Definition.

Let  $\Pi$  be a  $\mathbb{N}$ -valued function which labels the basins of  $V$ . The *coarse dynamics* is  $\Pi(X_t)_{t \geq 0}$ .

### Problem.

Efficiently generate approximations of the coarse dynamics.

- The trick: The dynamics reaches **local equilibrium** in a basin much faster than it hops out.

We must understand hopping events, starting at local equilibrium.

**Definition.**

A *quasistationary distribution (QSD)* in  $D$  is a probability measure  $\nu$  with support in  $D$  such that, if  $X_0 \sim \nu$  then for all  $t > 0$ ,

$$\nu(\cdot) = \mathbb{P}(X_t \in \cdot \mid X_s \in D \text{ for all } s \in [0, t]).$$

- If  $\nu$  is a probability measure with support in  $D$  such that, for any initial distribution of  $X_0$  in  $D$ ,

$$\nu(\cdot) = \lim_{t \rightarrow \infty} \mathbb{P}(X_t \in \cdot \mid X_s \in D \text{ for all } s \in [0, t]), \quad (1)$$

then  $\nu$  is the unique QSD in  $D$ . Intuitively, **we reach the QSD in a basin when the dynamics spend a sufficiently long time there without hopping out.**

**Definition.**

Let  $(u, -\lambda)$  be the principal eigenfunction/eigenvalue pair for  $L^* := \operatorname{div}(\cdot \nabla V) + \beta^{-1} \Delta$  with absorbing bc's on  $\partial D$ , and with  $\int_D u \, dx = 1$ .

**Theorem.** (Le Bris, Lelièvre, Luskin, Perez)

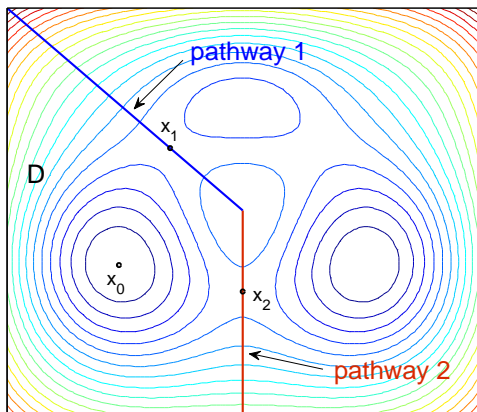
There is a unique QSD  $\nu$  in  $D$  satisfying (1), and  $d\nu = u \, dx$ .

- TAD is an algorithm for generating the coarse dynamics at low temperature, proposed by A.F. Voter and M. Sørensen in 2000.
- TAD is a popular algorithm: **the original article is cited 255 times**, according to Web of Science, and the algorithm is used by many independent research groups.
- **Our analysis focuses on a modified version of TAD which we propose**, which we show should decrease error at a small cost to efficiency.
- In TAD, temperature is increased to accelerate hopping events. When temperature is increased, the relative rates of hopping via different pathways will change. TAD accounts for the change in relative rates of hopping, so that the hopping events at the original temperature can be extrapolated.



## Definition.

Let  $x_1, \dots, x_n$  be the saddle points of  $V$  on  $\partial D$ , and  $x_0$  the minimum of  $V$  in  $D$ . Assume  $\partial D$  is partitioned into pathways  $1, \dots, n$  containing  $x_1, \dots, x_n$ .

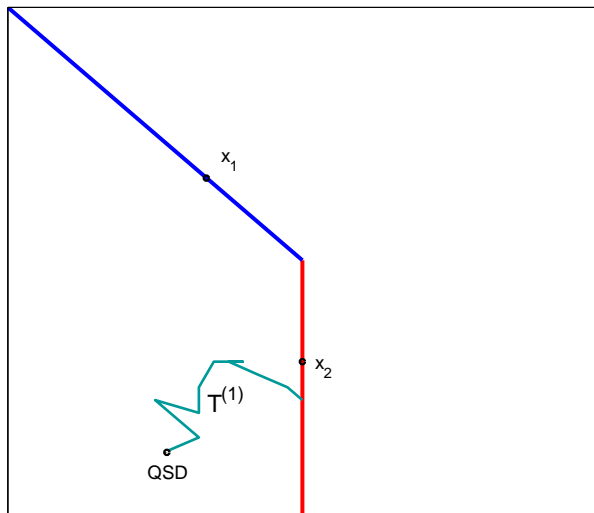


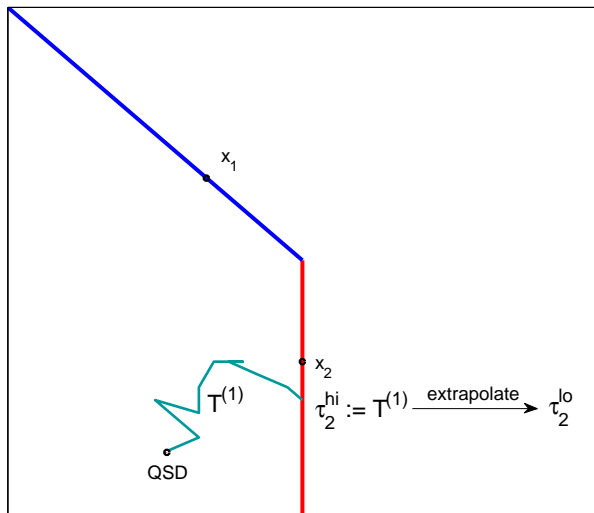
## TAD Exit Algorithm: Generating an exit (hopping) event from $D$ .

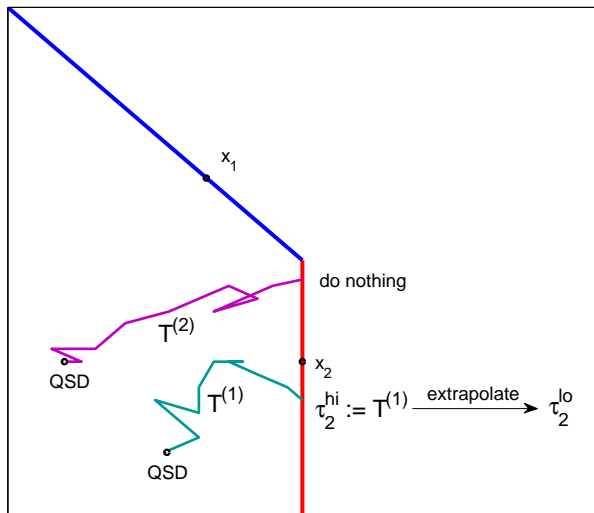
Fix a high and low temperature. Set  $N = 1$ ,  $T_{stop} = \infty$  and iterate:

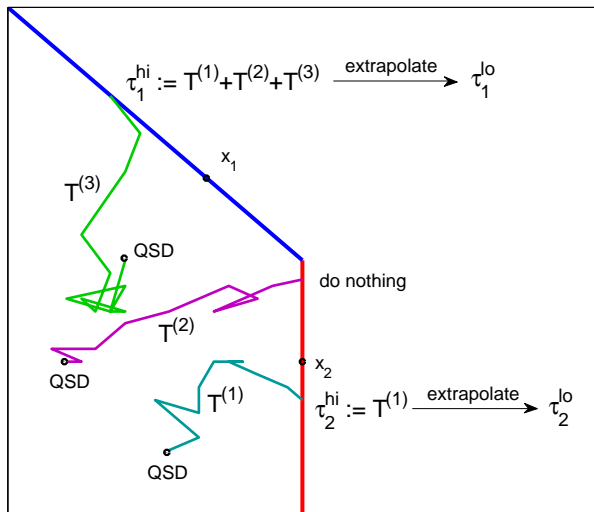
1. Starting at the QSD at high temperature and time 0, evolve the dynamics at high temperature.
2. Stop when the dynamics leave  $D$ , say at time  $T^{(N)}$  and through pathway  $i$ .
3. If this is the first time we leave through pathway  $i$ , use  $\tau_i^{hi} \equiv T^{(1)} + \dots + T^{(N)}$  and  $x_i$  to extrapolate a low temperature exit time, call it  $\tau_i^{lo}$ .
- 4-6. Update the current smallest low temperature extrapolated exit time and corresponding pathway. Update  $N = N + 1$ , and if a certain total simulation time has not been reached\*, return to Step 1.

\* To be described shortly.









## TAD Exit Algorithm: Generating an exit event from $D$ .

Fix a high and low temperature. Set  $N = 1$ ,  $T_{stop} = \infty$  and iterate:

1. Starting at the QSD at high temperature and time 0, evolve the dynamics at high temperature.
2. Stop when the dynamics leave  $D$ , say at time  $T^{(N)}$  and through pathway  $i$ .
3. If this is the first time we leave through pathway  $i$ , use  $\tau_i^{hi} \equiv T^{(1)} + \dots + T^{(N)}$  and  $x_i$  to extrapolate a low temperature exit time, call it  $\tau_i^{lo}$ .
4. Update  $\tau_{min}^{lo}$ , the current smallest extrapolated low temperature exit time, and  $I_{min}^{lo}$ , the corresponding pathway.
5. Update the simulation stopping time,  $T_{stop}$ .
6. If  $T^{(1)} + \dots + T^{(N)} < T_{stop}$ , update  $N = N + 1$  and return to Step 1. Otherwise, return  $(\tau_{min}^{lo}, I_{min}^{lo})$ , an approximation of the exit time and pathway at low temperature.

- How do we extrapolate  $\tau_i^{lo}$ ? How do we update  $T_{stop}$ ?

Fix an (inverse) temperature  $\beta$ .

### Definition.

Let  $X_0 \sim \nu$ . Define  $\tau = \inf\{t > 0 : X_t \notin D\}$  and  $I = i \Leftrightarrow X_\tau \in \text{pathway } i$ .

- Starting at the QSD,  $\tau$  is the first exit time and  $I$  is the exit pathway.

### Theorem. (Le Bris, Lelièvre, Luskin, Perez)

$\tau$  and  $I$  are independent. Moreover,

$$\mathbb{P}(\tau \leq t) = 1 - e^{-\lambda t} \quad \text{and} \quad \mathbb{P}(I = i) = -\frac{1}{\beta\lambda} \int_i \partial_n u \, d\sigma_{\partial D}$$

where  $\partial_n$  is the normal to  $\partial D$  and  $\sigma_{\partial D}$  is Lebesgue measure on  $\partial D$ .

- Starting at the QSD, the first exit time is exponentially distributed and independent of the exit position.



- Recall: Starting at the QSD,  $\tau$  is the first exit time and  $I$  is the exit pathway.

### Theorem.

The last theorem implies that  $\lambda^{-1} = \mathbb{E}[\tau]$ .

### Definition.

Let  $p_i \equiv \mathbb{P}(I = i)$ .

- Starting at the QSD,  $\lambda^{-1}$  is the expected exit time and  $p_i$  is the probability to exit through pathway  $i$ .
- We want to be able to sample  $(\tau, I)$ . We will see that knowing  $\lambda p_i$  is sufficient to do this...

## Theorem.

Let  $\tau_1, \dots, \tau_n$  be independent r.v.'s with  $\mathbb{P}(\tau_i \leq t) = 1 - e^{-\lambda p_i t}$ . Then

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

- At low temperature  $\lambda p_i$  is estimated by:

## The Arrhenius law

$$\lambda p_i \approx \eta_i e^{-\beta(V(x_i) - V(x_0))} \quad \text{when} \quad \beta^{-1} \ll V(x_i) - V(x_0).$$

- $\eta_i$  is a known function of the eigenvalues of the Hessian of  $V$  at the saddle point  $x_i$  and minimum  $x_0$ .  $\eta_i$  does not depend on  $\beta$ .
- If the locations of the saddle points are known a priori, the theorem and Arrhenius law can be used to sample exit events.

## TAD Exit Algorithm: Generating an exit event from $D$ .

Fix a high and low (inverse) temperature  $\beta^{hi}$  and  $\beta^{lo}$  (so  $\beta^{hi} < \beta^{lo}$ ). Set  $N = 1$ ,  $T_{stop} = \infty$  and iterate:

1. Starting at the QSD at high temperature and time 0, evolve the dynamics at high temperature. Stop when the dynamics leave  $D$ , say at time  $T^{(N)}$  and through pathway  $i$ . Define  $\tau_i^{hi} = T^{(1)} + \dots + T^{(N)}$ .
3. If this is the first time we leave through pathway  $i$ , extrapolate a low temperature exit time  $\tau_i^{lo}$  by:

$$\tau_i^{lo} = \tau_i^{hi} e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))}.$$

4. Update  $\tau_{min}^{lo}$ , the current smallest extrapolated low temperature exit time, and  $I_{min}^{lo}$ , the corresponding pathway.
5. Update the simulation stopping time,  $T_{stop}$ , by:  $T_{stop} = \tau_{min}^{lo} / C$ , where  $C \leq \min_{1 \leq i \leq n} e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))}$ .
6. If  $T^{(1)} + \dots + T^{(N)} < T_{stop}$ , update  $N = N + 1$  and return to Step 1. Otherwise, return  $(\tau_{min}^{lo}, I_{min}^{lo})$ .

**Remarks:**

- At relatively low temperatures the QSD in  $D$  can be efficiently sampled.
- Some a priori knowledge about minimum energy barriers is needed to define  $T_{stop}$ .
- If the simulation continues after time  $T_{stop}$ , the value of  $\tau_{min}^{lo}$  will no longer change (which is why we stop at  $T_{stop}$ ).
- The algorithm **does not require that all the saddle points be found**. Also, **the prefactors  $\eta_i$  are not needed**.

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

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

The **Exit Algorithm** will be **efficient** when also  $\beta^{hi} \ll \beta^{lo}$ .

- To see that latter, from the formula defining  $T_{stop}$  we have:

$$T_{stop} \leq \tau_{min}^{lo} / \min_{1 \leq i \leq n} e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))}, \quad (2)$$

and notice  $(V(x_i) - V(x_0))^{-1} \ll \beta^{hi} \ll \beta^{lo}$  implies that  $T_{stop} \ll \tau_{min}^{lo}$ .

## Idealization.

In the **Exit Algorithm**, the QSD at high temperature is sampled **exactly**, and  $e^{-(\beta^{hi} - \beta^{lo})(V(x_i) - V(x_0))}$  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 idealization,

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

where  $(\tau_{min}^{lo}, I_{min}^{lo})$  is computed via the **Exit Algorithm**, and  $(\tau^{lo}, I^{lo})$  is the true exit time/pathway at temperature  $\beta^{lo}$ , starting at the QSD.

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

$$\tau_i^{hi} = T^{(1)} + \dots + T^{(N_i)}$$

where  $N_i$  is the first loop of the algorithm in which we exit through pathway  $i$ . One can show that the r.v.'s  $\tau_i^{hi}$  are independent with

$$\mathbb{P}(\tau_i^{hi} \leq t) = 1 - e^{-\lambda^{hi} p_i^{hi} t}.$$

So since

$$\tau_i^{lo} := \tau_i^{hi} \frac{\lambda^{hi} p_i^{hi}}{\lambda^{lo} p_i^{lo}},$$

the r.v.'s  $\tau_i^{lo}$  are independent with

$$\mathbb{P}(\tau_i^{lo} \leq t) = 1 - e^{-\lambda^{lo} p_i^{lo} t}.$$

By construction, after  $T_{stop}$  the value of the smallest  $\tau_i^{lo}$  will not change. Appealing to our earlier theorem, we are done.

## Theorem.

Let  $D = [x_1, x_2]$  be a basin of attraction for a Morse potential  $V$ . 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  $\beta^{hi} \rightarrow \infty$ ,  $\beta^{hi}/\beta^{lo} \rightarrow \text{positive const.}$

- This shows that **the Arrhenius law extrapolation becomes exact in the small temperature limit**, at least in 1D.

## Open Problem.

Prove a version of the above theorem in dimension  $\geq 2$ .

**Remark:** Recall that  $(u, -\lambda)$  is the principal eigenvector/value pair of  $L^* \equiv \text{div}(\cdot \nabla V) + \beta^{-1} \Delta$  with absorbing boundary conditions on  $\partial D$ :

$$\begin{cases} L^* u = -\lambda u & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases}$$

A PDE-based attack on the open problem would require an asymptotic analysis of  $\lambda$  and  $\partial_n u$  as  $\beta \rightarrow \infty$ .

- So far we have discussed how to use the TAD **Exit Algorithm** to efficiently generate a hopping event out of a basin.
- Now we will show how to use this repeatedly to obtain metastable dynamics.
- Our analysis shows that to accurately simulate a hopping event in the **Exit Algorithm**, we should begin at the QSD at low temperature.
- Thus we need the following definition:

### Decorrelation parameter.

Assume that at temperature  $\beta^{lo}$ , the dynamics approximately reaches the QSD in a given basin after spending time  $T_{corr}$  there without leaving.

- In applications, the practitioner will choose  $T_{corr}$ . It is problem-specific, but in metastable situations,  $T_{corr}$  is much smaller than the time to leave a basin.
- In general  $T_{corr}$  is basin-dependent.



## TAD: Main Algorithm.

Fix a low temperature, choose parameters  $T_{corr}$ , and iterate:

1. Evolve the dynamics at low temperature.
  2. Stop when the dynamics spends  $T_{corr}$  time in some basin without leaving.
  3. Use the **Exit Algorithm** to simulate an exit event from this basin. Then, return to Step 1.
- Since  $T_{corr}$  and the simulation time of the **Exit Algorithm** will be on average much smaller than the time to exit a basin at low temperature, the **Main Algorithm** will be efficient compared to direct sampling.
  - There is an additional error in the **Main Algorithm**, associated with not exactly reaching the QSD in Step 2 above. A larger  $T_{corr}$  leads to less error, but decreased efficiency.

# 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/>