Uncertainty-driven construction of Markov models from accelerated molecular dynamics

Thomas D Swinburne*, Danny Perez

T-1, Theoretical Division, Los Alamos National Laboratory

*Current address: CINaM, CNRS Aix-Marseille, France

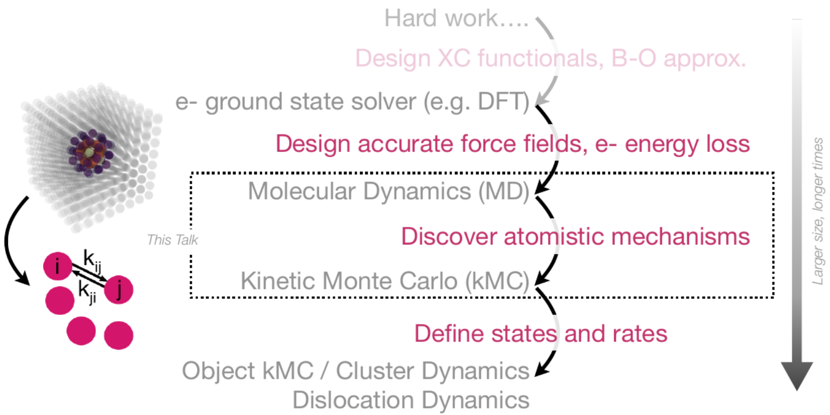
Los Alamos NATIONAL LABORATORY EST. 1943

Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

Multiscale Material Modeling

- Real life has too many atoms to simulate ⇒ **coarse graining**



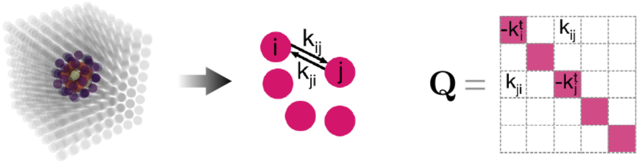
- Key challenge is to **quantify** how coarse graining degrades accuracy

Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

MD → kMC

- To coarse-grain MD→kMC we must **identify atomistic mechanisms**

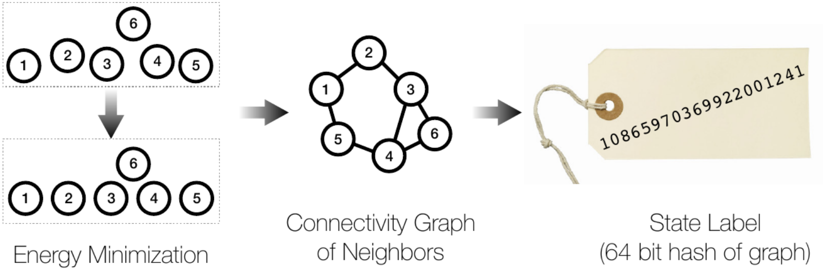


$$m\ddot{\mathsf{X}}(t) = -\nabla V(\mathsf{X}(t)) \quad \rightarrow \quad \dot{\mathsf{P}}(t) = \mathsf{P}(t) \cdot \mathsf{Q}$$

- To build a kMC model from MD we need **at least**
  - A method to define discrete states
  - A method to calculate transition rates
  - A method to find states and rates efficiently

MD → kMC: define states

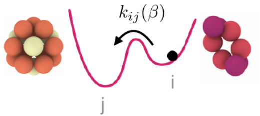
- kMC requires discrete states but MD has continuous state space
- Solids are often in a basin with unique minimum - basin ≡ state
- State identification follows ParSplice Code (Perez et al. JCTC 2015):



- Very general but is sensitive to hard neighbor cutoffs

MD → kMC: calculate rates

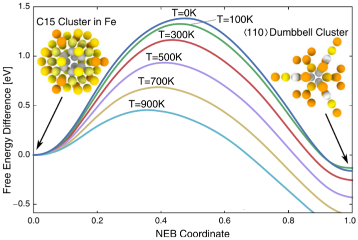
- kMC requires transition rates $k_{ij}(\beta)$:



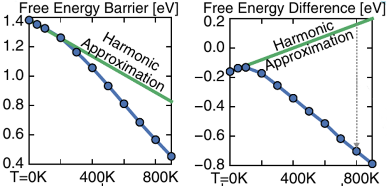
- Rate given by transition state theory (TST):

$$k_{ij}(\beta) = \omega_{ij} \exp\left[-\beta \Delta F_{ij}(\beta)\right] \simeq \nu_{ij} \exp\left[-\beta \Delta E_{ij}\right] \quad \text{(HTST)}$$

- Can treat anharmonicity in TST (e.g. TDS and Marinica, PRL 2018)



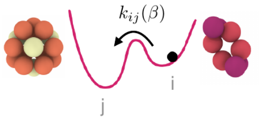
$$\partial_\xi \mathcal{F}(\xi;T) = \left\langle \frac{\mathbf{w}\cdot\nabla V}{\mathbf{w}\cdot\nabla\tilde{\xi}} + \beta^{-1}\nabla\cdot\frac{\mathbf{w}}{\mathbf{w}\cdot\nabla\tilde{\xi}} \right\rangle_\xi ,$$



- Uses ABF result in constrained sampling along MEP to obtain MFEP+ΔF

MD → kMC: find transitions

- kMC requires transition rates $k_{ij}(\beta)$:



- Problem: transitions are often rare ⇒ expensive to find in MD

$$P(\tau_{ij} \in [t, t+\mathrm{d}t]) = k_{ij}\mathrm{d}t \exp\left(-k_{ij}t\right) \quad \leftarrow \text{Poisson distribution}$$
$$\langle \tau_{ij}(\beta) \rangle \simeq \nu_{ij}^{-1} \exp(\beta \Delta E_{ij}) \quad \leftarrow \text{Exponentially large!}$$

- One approach is **temperature accelerated dynamics (TAD)**:
  - run MD at high temperature β_H^{-1}
  - Record transition times $\{\tau_{ij}(\beta_H)\}$
  - $X \sim \text{Poisson}(k) \Rightarrow (k/k')X \sim \text{Poisson}(k')$
  - Simple to infer the low temperature transition times if HTST holds:

$$\tau_{ij}(\beta_L)/\tau_{ij}(\beta_H) = \exp(\beta_L \Delta F_{ij}(\beta_L) - \beta_H \Delta F_{ij}(\beta_H)) \simeq \exp([\beta_L - \beta_H]\Delta E_{ij})$$

MD → kMC: find transitions

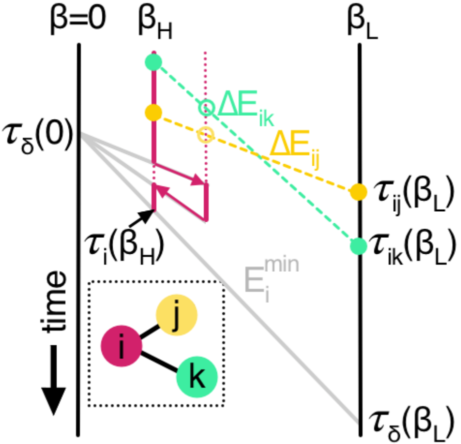
- Real TAD is more subtle!
  Sørensen and Voter JCP 2000

- Assumes all $\nu_{ij} \geq \nu_{min}$ to infer **effective sampling time $\boldsymbol{\tau_\delta(\beta)}$**

  $$\tau_\delta(\beta) \equiv \log(1/\delta)\nu_{\mathrm{min}}^{-1} \exp(\beta E_{\mathrm{min}})$$
  $$\tau_\delta(\beta_H) \equiv \tau_i(\beta_H) \Rightarrow E_{\mathrm{min}}$$

- Original TAD finds $\min_j\{\tau_{ij}(\beta_L)\}$

- We use TAD to give $\forall\beta \in [\beta_H, \beta_L]$
  - Escape data $\{k_{ij}(\beta), \tau_{ij}(\beta)\}$
  - Total sampling time $\tau_\delta(\beta)$



- Can now find states, rates and transitions - **but this is not enough!**

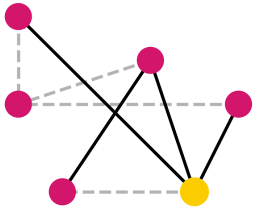
Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

The incompleteness problem

- Explored state space $\subsetneq$ full state space
- All kMC models are limited by **network incompleteness**

- kMC on subset will diverge from MD on set
- $\langle$Time to unknown$\rangle$ $\sim$Prediction timescale
- Uncertainty quantification (UQ) essential but often ignored



- Sampling incompleteness ideas previously investigated by
  - Chill & Henkelman JCP 2014
  - Aristoff, Chill & Simpson CAMCS 2016
  - Bhute & Chatterjee JCP 2013, Bhoutekar *et al.* JCP 2017

The incompleteness problem

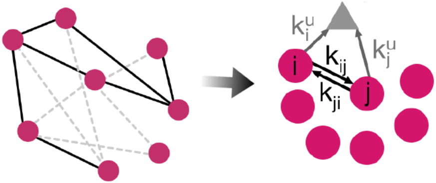
- A state i has connections $\mathcal{S}_i$ with a **total rate** $k^t_i(\beta) = \sum_{i \in \mathcal{S}_i} k_{ij}(\beta)$

- We only find a subset $\mathcal{K}_i \subsetneq \mathcal{S}_i$ with an **observed rate** $k^o_i(\beta) = \sum_{i \in \mathcal{K}_i} k_{ij}(\beta)$

- $\Rightarrow$ **unknown rate** $k^u_i(\beta) \equiv k^t_i(\beta) - k^o_i(\beta) = \sum_{i \in \mathcal{S}_i \backslash \mathcal{K}_i} k_{ij}(\beta)$

- **With exact $\{k^u_i\}$, incomplete network gives statistically exact dynamics until exit to unknown**

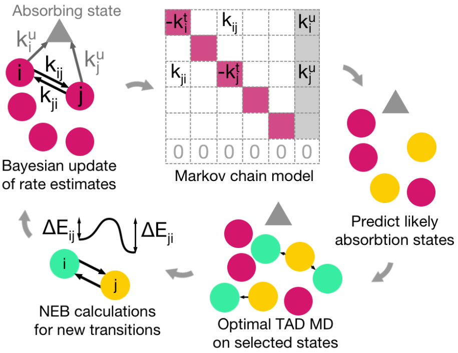
- We estimate $\{k^u_i\}$ for use as **local** UQ to aid global sampling strategy



TAMMBER

- **T**emperature **A**ccelerated construction of **M**arkov **M**odels with **B**ayesian **E**stimation of **R**ates

- Autonomous, computationally optimal MD→kMC method using TAD+Bayes/Markov analysis



- We maximize the **global UQ** $\langle \tau_{res} \rangle$
  τ_{res}: *residence time in found states*

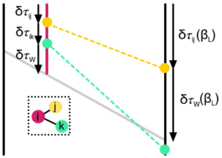
- Built on ParSplice/LAMMPS: O(N) shown up to N=11,000

- Details: Swinburne and Perez, Phys. Rev. Materials, 2, 053802 (2018)

Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

Estimating the unknown rate



- TAD gives passage times+rates $\{\tau_{ij}, k_{ij}\}$ for a given sampling time and target temperature
- **Rare MD escape times $\sim$ Poisson** (allows TAD UQ)
- We can thus **derive** the likelihood for Bayes:

  Posterior(Params.|Data) = Prior(Params.)$\times$Likelihood(Data|Params.)

- Poisson likelihood of waiting δt for a new event, given k^t:

$$\pi(\delta t|k^t) = [k^t - k^o(\tau)] \exp\left(-[k^t - k^o(\tau)]\delta t\right)$$

- For the entire TAD escape trajectory the posterior for k^t reads

$$\pi(k^t|\{\tau_{ij}, k_{ij}\}) = \frac{1}{k^t} e^{-[k^t - k^o(\tau)]\delta t_W} \prod_j [k^t - k^o(\tau_{ij})] e^{-[k^t - k^o(\tau_{ij})]\delta\tau_{ij}} \Theta(k^t - k^o)$$

$$= (Jefferies) \times L(last\ wait) \times L(waits) \times Step\ function$$

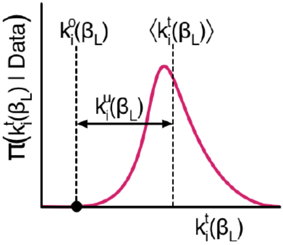
Estimating the unknown rate

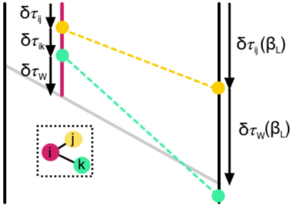
- For the entire TAD escape trajectory the posterior for k^t reads

$$\pi(k^t|\{\tau_{ij},k_{ij}\}) = \frac{1}{k^t}e^{-[k^t-k^o(\tau)]\delta t_W}\prod_j[k^t-k^o(\tau_{ij})]e^{-[k^t-k^o(\tau_{ij})]\delta\tau_{ij}}\Theta(k^t-k^o)$$

- k_i^u can then be estimated as

$$\langle k_i^u\rangle = \int k\pi(k|\{\tau_{ij},k_{ij}\})\mathrm{d}k - k_i^o$$





$$\langle(k_i^u)^n\rangle = \frac{\sum_{r=0}^{N-2}(r+n)!A_r\tau_i^{-r}}{\tau_i^n\sum_{r=0}^{N-2}r!A_r\tau_i^{-r}}$$

Simply HTST rescale to change T

Optimal TAD temperature

- Sampling i at β requires work $c_i(\beta)$ for MD, state ID and NEB:

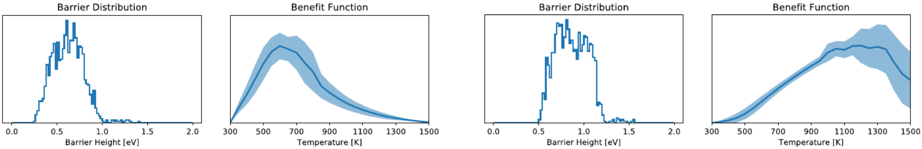
$$\frac{\mathrm{d}c_i(\beta_H)}{\mathrm{d}\tau_i(\beta_H)} = \dot{c}_{\mathrm{MD}} + c_{\mathrm{ST}} k^o_i(\beta_H) + c_{\mathrm{NEB}} k^u_i(\beta_H)$$

- TAD temperature for each state should reduce k^u as fast as possible:

$$\beta^{TAD}_i = \arg\max_i \left[-\frac{\mathrm{d}\langle k^u_i(\beta_L)\rangle}{\mathrm{d}c_i(\beta_H)} \right]$$

- Using pertubation theory to consider δt more MD, we derived that

$$-\frac{\mathrm{d}\langle k^u_i(\beta_L)\rangle}{\mathrm{d}c_i(\beta_H)} = \left(\frac{\mathrm{d}c_i(\beta_H)}{\mathrm{d}\tau_i(\beta_H)}\right)^{-1} \left(\min\left[\langle k^u_i(\beta_L)\rangle \cup \{k_{ij}(\beta_L)\}\right] + \left(\frac{\tau_i(\beta_L)}{\tau_i(\beta_H)} - \frac{\langle k^u_i(\beta_H)\rangle}{\langle k^u_i(\beta_L)\rangle}\right) \left(\langle (k^u_i(\beta_L))^2\rangle - \langle k^u_i(\beta_L)\rangle^2\right)\right)$$



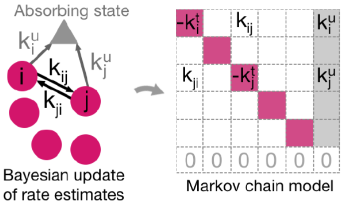
Absorbing Markov chain

- With known and unknown rates estimated we can build a Markov model

$$\mathsf{P}(t) = \mathsf{P}_{\mathcal{K}}(t) \oplus \mathsf{P}_{\triangle}(t)$$

$$\mathsf{Q} = \begin{bmatrix} (\mathsf{Q})_{\mathcal{K}} & \mathsf{k}^{u} \\ \mathsf{0}^{\top} & 0 \end{bmatrix}$$

$$\dot{\mathsf{P}}(t) = \mathsf{P}(t) \cdot \mathsf{Q}$$



- Unknown are to an absorbing state δ, representing unseen states/transitions

- We take averages over the known state probabilites at a time t

$$\mathsf{P}(t) = \mathsf{P}(0) \cdot \exp\left(\mathsf{Q}t\right) \Rightarrow \mathsf{P}_{\mathcal{K}}(t) = \mathsf{P}_{\mathcal{K}}(0) \cdot \exp\left(\mathsf{Q}_{\mathcal{K}}t\right)$$

Absorbing Markov chain

- Single solve gives the expected residence time

$$\langle \tau_{res} \rangle = \int_0^\infty t\mathbf{P}_{\mathcal{K}}(t) \cdot \mathbf{k}^u \mathrm{d}t = -\mathbf{P}_{\mathcal{K}}(0) \cdot \mathbf{Q}_{\mathcal{K}}^{-1} \cdot \mathbf{1}_{\mathcal{K}}$$

- We allocate new resources $\delta c_i \equiv s_i \delta c$ to maximize $\langle \tau_{res} \rangle$

$$\frac{\delta\langle \tau_{res} \rangle}{\delta c} = \sum_{i \in \mathcal{K}} s_i \frac{\delta\langle \tau_{res} \rangle}{\delta c_i} \qquad \Rightarrow s_i \propto \frac{\delta\langle \tau_{res} \rangle}{\delta c_i}$$

$$s_i = \eta \left[\mathbf{P}_{\mathcal{K}} \cdot \mathbf{Q}_{\mathcal{K}}^{-1}\right]_i \times \left[\mathbf{Q}_{\mathcal{K}}^{-1} \cdot \mathbf{1}_{\mathcal{K}}\right]_i \times \frac{\delta k_i^u(\beta_L)}{\delta c_i}$$

$$= \langle \tau_{res} | \text{start in i} \rangle \times \langle \text{time in i} \rangle \times (\text{Change in unknown rate})$$

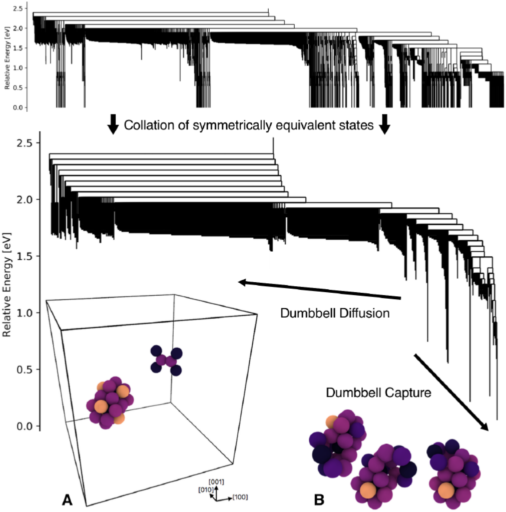
- **Globally sensitive** scheme avoids sampling states of little influence

Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

TAMMBER: Interstitial clusters in Fe

- Tetra-C15 + dumbell in ”Fe”
- Target T=300K, TAD T $\leq$900K
- 16 hours on 2160 cores
- 2664 states, 7676 barriers
- Naïve symmetric compression
  τ_{res} from B: 80s
  total 300K time: 7.4×10^6s
- **Wasted sampling due to lack of symmetry awareness**

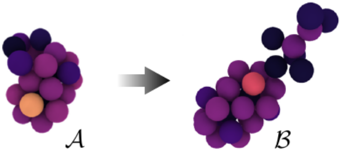


Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

Mesoscale uncertainty quantification

- MD→kMC requires **mesoscopic observables**
- In this example, we have SIA capture / breakup of C15 clusters
- We look at time from capture $\mathcal{A}$ to breakup $\mathcal{B}$

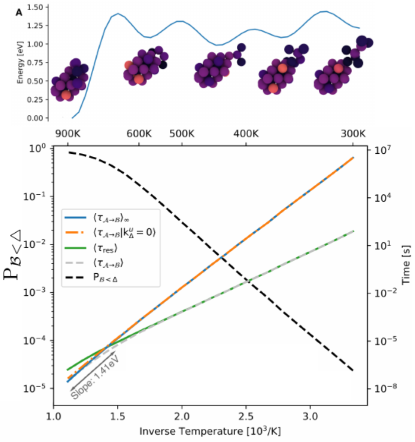


Mesoscale uncertainty quantification

- Markov Model can probe breakup of SIA+C15
- Use MM to estimate times between $\mathcal{A}, \mathcal{B}, \Delta$
- Absorbing MC gives UQ on observable:

$$P(\mathcal{A} \to \mathcal{B}|!\Delta) = P_{\mathcal{B}<\Delta} = P_{\mathcal{A}}(0) \cdot \mathbf{Q}_{\mathcal{A}}^{-1} \cdot k^{u}_{\mathcal{B}}$$

- When $P_{\mathcal{B}<\Delta} \ll 1$ results are uncertain
- More certain when target T $\to$ simulation T
- **Lack of symmetry awareness leads to wasted sampling effort**



Approximate unbiasing:

$$\langle \tau^{abs}_{\mathcal{A}\to\mathcal{B}} \rangle_\infty = \left(\mathrm{P}^{-1}_{\mathcal{B}<\Delta} - 1\right) \langle \tau_{res} \rangle$$

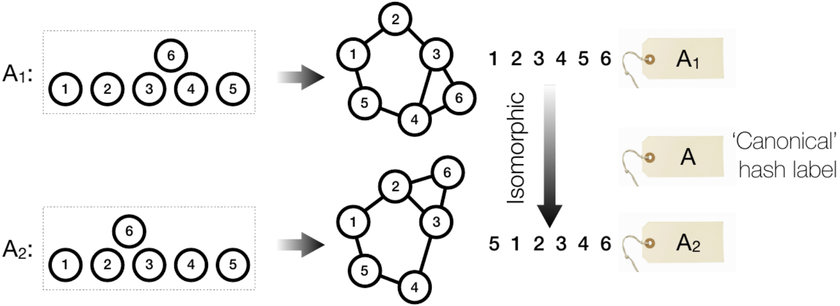
$$\langle \tau^{abs}_{\mathcal{A}\to\mathcal{B}} | k^{u}_{\Delta} = 0 \rangle = \lim_{k^{u}_{\Delta} \to 0} \langle \tau^{abs}_{\mathcal{A}\to\mathcal{B}} \rangle$$

Outline

- MD $\rightarrow$ kMC: states, rates and transitions
- The incompleteness problem
- Estimating the unknown
- Application to interstitial clusters in Fe
- Mesoscale uncertainty quantification
- In development: isomorphic compression

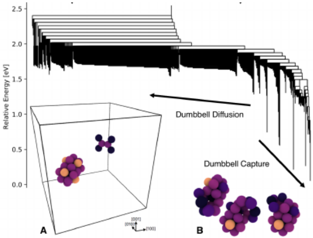
In development: Canonical-TAMMBER

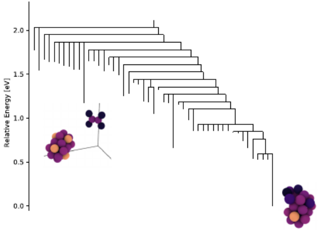
- Canonical-TAMMBER uses graph isomorphisms to capture symmetries

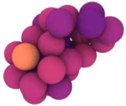


In development: Canonical-TAMMBER

- Symmetry vastly simplifies the energy landscape in crystals

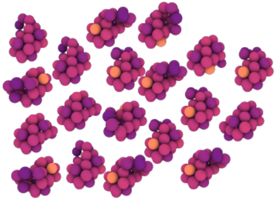






Lowest energy capture state

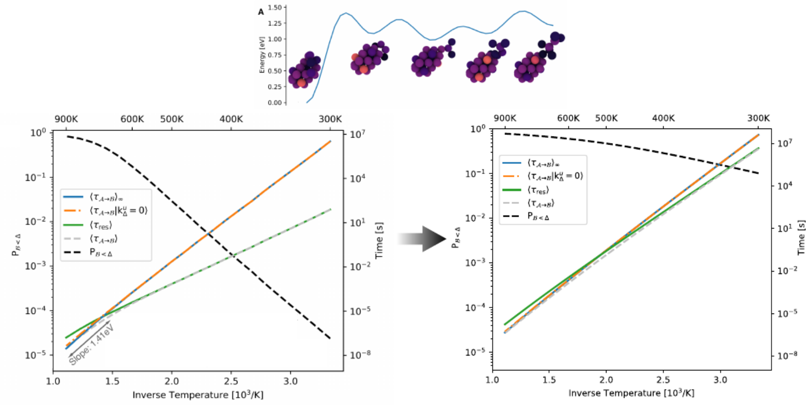
≡



- 12 hours on 2160 cores
- 49 states, 101 NEBs (450 barriers)
- Residence time from B: 4.6×10^6s
  Total 300K time: 5.1×10^6s

In development: Canonical-TAMMBER

- The useful sampling time is significantly increased, giving much better UQ



- Slightly different escape times due to slightly different definitions of sets $\mathcal{A}, \mathcal{B}$

Thank you for listening

- TAMMBER is a self-optimizing network building scheme driven by global UQ
- Can give UQ on mesoscopic observales
- Symmetry is hugley beneficial
- Code in active dev; email if interested!
- Details: TDS and Perez, Physical Review Materials 2, 053802 (2018)
  `tiny.cc/tds110`
  `tomswinburne@gmail.com`