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
Outline

- MD → 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: isomorphemic compression
To coarse-grain MD→kMC we must **identify** atomistic mechanisms.

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

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 $\rightarrow$ kMC: define states

- kMC requires discrete states but MD has continuous state space
- Solids are often in a basin with unique minimum - basin $\equiv$ state
- State identification follows ParSplice Code (Perez et al. JCTC 2015):
  - Very general but is sensitive to hard neighbor cutoffs

![Energy Minimization](attachment:energy_minimization.png)

![Connectivity Graph of Neighbors](attachment:connectivity_graph.png)

![State Label](attachment:state_label.png)
MD $\rightarrow$ kMC: calculate rates

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

  - Rate given by transition state theory (TST):

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

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

- Uses ABF result in constrained sampling along MEP to obtain MFEP+$\Delta F$
MD → kMC: find transitions

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

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

$$P(\tau_{ij} \in [t, t + dt]) = k_{ij}dt \exp(-k_{ij}t) \quad \leftarrow \text{Poisson distribution}$$

$$\langle \tau_{ij}(\beta) \rangle \simeq \nu_{ij}^{-1} \exp(\beta \Delta E_{ij}) \quad \leftarrow \text{Exponentially large!}$$

- One approach is **temperature accelerated dynamics (TAD)**:

  - run MD at high temperature $\beta_H^{-1}$
  - 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 $\rightarrow$ 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 $\tau_\delta(\beta)$

$$\tau_\delta(\beta) \equiv \log(1/\delta)\nu^{-1}_{min} \exp(\beta E_{min})$$

- 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 → 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 $\subseteq$ 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 $S_i$ with a total rate $k^t_i(\beta) = \sum_{i \in S_i} k_{ij}(\beta)$

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

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

• 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
• **Temperature Accelerated construction of Markov Models with Bayesian Estimation of Rates**

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

• We maximize the **global UQ** \( \langle \tau_{res} \rangle \)

\( \tau_{res} \): *residence time in found states*

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

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: isomorphich 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 \( \sim \text{Poisson} \) (allows TAD UQ)

- We can thus derive the likelihood for Bayes:

\[
\text{Posterior(Params.}\mid\text{Data}) = \text{Prior(Params.)} \times \text{Likelihood(Data}\mid\text{Params.})
\]

- Poisson likelihood of waiting \( \delta t \) for a new event, given \( k^t \):

\[
\pi(\delta t\mid k^t) = [k^t - k^0(\tau)] \exp\left( -[k^t - k^0(\tau)]\delta t \right)
\]

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

\[
\pi(k^t\mid\{\tau_{ij}, k_{ij}\}) = \frac{1}{k^t} e^{-[k^t - k^0(\tau)]\delta t_w} \prod_j [k^t - k^0(\tau_{ij})] e^{-[k^t - k^0(\tau_{ij})]\delta \tau_{ij}} \Theta(k^t - k^0) = (\text{Jefferies}) \times L(\text{last wait}) \times L(\text{waits}) \times \text{Step function}
\]
Estimating the unknown rate

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

$$\pi(k^t|\{\tau_{ij}, k_{ij}\}) = \frac{1}{k^t} e^{-[k^t-k^0(\tau)]\delta t_W} \prod_{j} [k^t - k^0(\tau_{ij})] e^{-[k^t-k^0(\tau_{ij})]\delta \tau_{ij}} \Theta(k^t - k^0)$$

- $k^u_i$ can then be estimated as

$$\langle k^u_i \rangle = \int k \pi(k|\{\tau_{ij}, k_{ij}\}) dk - k^0_i$$

Simply HTST rescale to change T
Optimal TAD temperature

- Sampling \( i \) at \( \beta \) requires work \( c_i(\beta) \) for MD, state ID and NEB:
  \[
  \frac{d c_i(\beta_H)}{d \tau_i(\beta_H)} = \dot{c}_{\text{MD}} + c_{\text{ST}} k_i^0(\beta_H) + c_{\text{NEB}} k_i^u(\beta_H)
  \]

- TAD temperature for each state should reduce \( k^u \) as fast as possible:
  \[
  \beta_i^\text{TAD} = \arg \max_i \left[ -\frac{d \langle k_i^u(\beta_L) \rangle}{d c_i(\beta_H)} \right]
  \]

- Using perturbation theory to consider \( \delta t \) more MD, we derived that
  \[
  -\frac{d \langle k_i^u(\beta_L) \rangle}{d c_i(\beta_H)} = \left( \frac{d c_i(\beta_H)}{d \tau_i(\beta_H)} \right)^{-1} \left( \min \left[ \langle k_i^u(\beta_L) \rangle \cup \{ k_{ij}(\beta_L) \} \right] + \left( \frac{\tau_i(\beta_L)}{\tau_i(\beta_H)} - \frac{\langle k_i^u(\beta_H) \rangle}{\langle k_i^u(\beta_L) \rangle} \right) \left( \langle (k_i^u(\beta_L))^2 \rangle - \langle k_i^u(\beta_L) \rangle^2 \right) \right)
  \]
Absorbing Markov chain

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

\[
P(t) = P_K(t) \oplus P_\Delta(t)
\]

\[
Q = \begin{bmatrix}
Q_K & k^u \\
0^T & 0
\end{bmatrix}
\]

\[
\dot{P}(t) = P(t) \cdot Q
\]

• Unknown are to an absorbing state \( \delta \), representing unseen states/transition

• We take averages over the known state probabilites at a time \( t \)

\[
P(t) = P(0) \cdot \exp(Qt) \Rightarrow P_K(t) = P_K(0) \cdot \exp(Q_Kt)
\]
Absorbing Markov chain

- Single solve gives the expected residence time

\[
\langle \tau_{res} \rangle = \int_0^\infty t P_K(t) \cdot k^u dt = -P_K(0) \cdot Q_K^{-1} \cdot 1_K
\]

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

\[
\frac{\delta \langle \tau_{res} \rangle}{\delta c} = \sum_{i \in K} S_i \frac{\delta \langle \tau_{res} \rangle}{\delta c_i} \Rightarrow S_i \propto \frac{\delta \langle \tau_{res} \rangle}{\delta c_i}
\]

\[
S_i = \eta [P_K \cdot Q_K^{-1}]_i \times [Q_K^{-1} \cdot 1_K]_i \times \frac{\delta k^u_i(\beta_L)}{\delta c_i}
\]

\[
= \langle \tau_{res} \mid \text{start in } i \rangle \times \langle \text{time in } i \rangle \times \text{(Change in unknown rate)}
\]

- Globally sensitive scheme avoids sampling states of little influence
Outline

- MD → 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 $\tau_{res}$ from B: 80s
  total 300K time: $7.4 \times 10^6$ s
- Wasted sampling due to lack of symmetry awareness

TD Swinburne, tomswinburne@gmail.com, tiny.cc/tds110  CIRM, 20/09/2018
Outline

• MD → 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 A to breakup B
Mesoscale uncertainty quantification

- Markov Model can probe breakup of SIA+C15
- Use MM to estimate times between $A, B, \Delta$
- Absorbing MC gives UQ on observable:
  \[ P(A \rightarrow B|!\Delta) = P_{B<\Delta} = P_A(0) \cdot Q_A^{-1} \cdot k_B^u \]
- When $P_{B<\Delta} \ll 1$ results are uncertain
- More certain when target $T \rightarrow$ simulation $T$
- Lack of symmetry awareness leads to wasted sampling effort

Approximate unbiasing:

\[
\left\langle \tau_{A\rightarrow B}^{\text{abs}} \right\rangle_\infty = \left( P_{B<\Delta}^{-1} - 1 \right) \left\langle \tau_{\text{res}} \right\rangle \\
\left\langle \tau_{A\rightarrow B}^{\text{abs}} \right|_{k_{\Delta}^u = 0} = \lim_{k_{\Delta}^u \rightarrow 0} \left\langle \tau_{A\rightarrow B}^{\text{abs}} \right\rangle
\]
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

A₁:

```
1 2 3 4 5
```

A₂:

```
1 2 3 4 5
```

A₁ and A₂ are isomorphic.
In development: Canonical-TAMMBBER

- Symmetry vastly simplifies the energy landscape in crystals

- 12 hours on 2160 cores

- 49 states, 101 NEBs (450 barriers)

- Residence time from B: $4.6 \times 10^6 \text{s}$
  Total 300K time: $5.1 \times 10^6 \text{s}$
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 observables
- Symmetry is hugely beneficial
- Code in active dev; email if interested!
tiny.cc/tds110
tomswinburne@gmail.com