Kinetic Monte Carlo Simulation
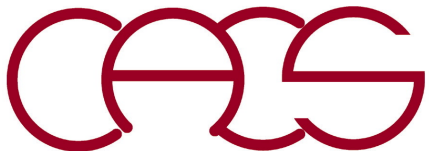
Aiichiro Nakano

Collaboratory for Advanced Computing & Simulations
Department of Computer Science
Department of Physics & Astronomy
Department of Quantitative & Computational Biology
University of Southern California

Email: anakano@usc.edu

Goal: Predicting long-time dynamics

What to Learn with Monte Carlo

- **Stochastic approaches to obtain approximate solutions to intractable computational problems**

- **Probability-distribution sampling by Markov-chain (or Metropolis) Monte Carlo (MC) algorithm**
  https://aiichironakano.github.io/cs653/06MCbasics.pdf

  1. Random state change
  2. Conditional acceptance with probability $\exp(-\Delta V/k_\mathrm{B}T)$
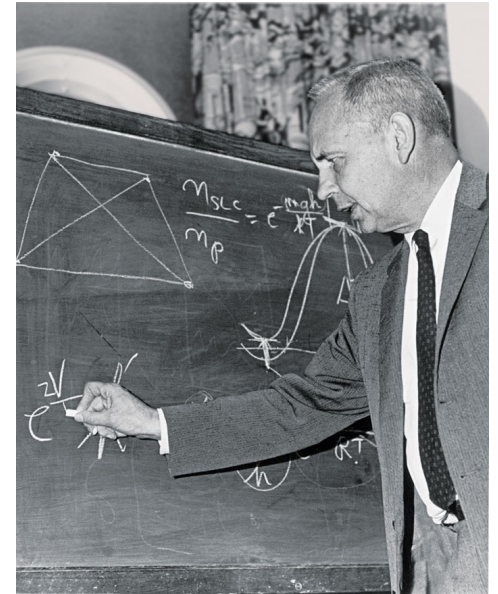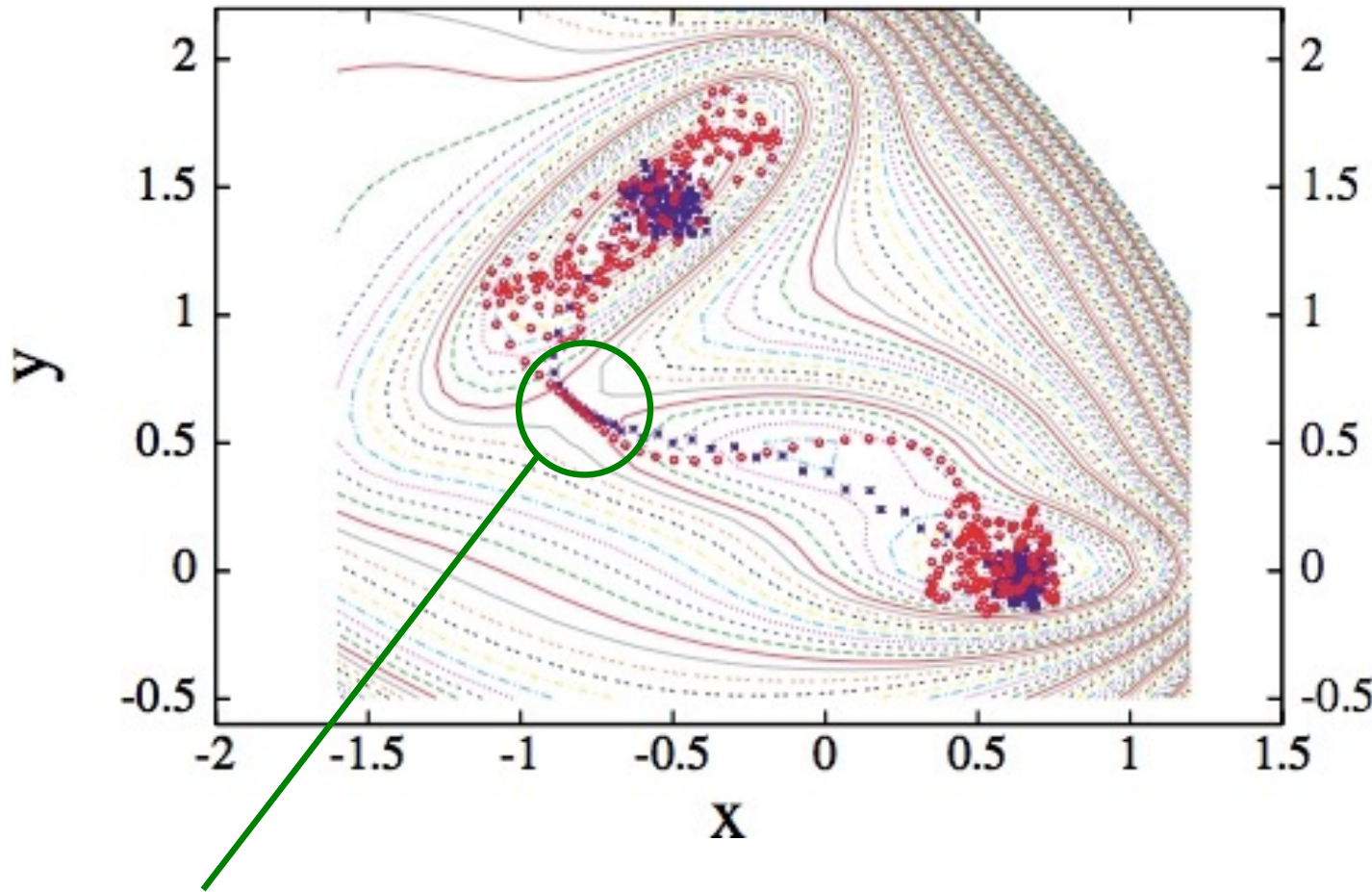
- **Applied to simulated annealing (SA) for global optimization**

- **Kinetic Monte Carlo (KMC) to predict long-time dynamics**

  1. Pick the next event i proportional to its transition rate r_i
  2. Advance the time by $-\ln(random\# \in [0,1))/\sum_k r_k$

- **Divide-&-conquer KMC algorithm for scalable parallel computing**

Rare Events

- **Infrequent transitions from a local minimum to another local minimum**

- **Transition state theory to understand chemical reaction rates: Michael Polanyi & Henry Eyring in 1920's & 1930's**

- **Renewed interests in understanding self-organization (protein folding, life, *etc.*)**
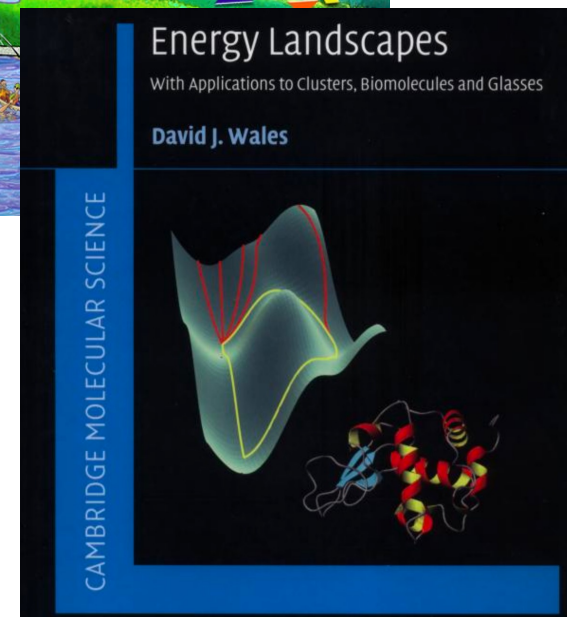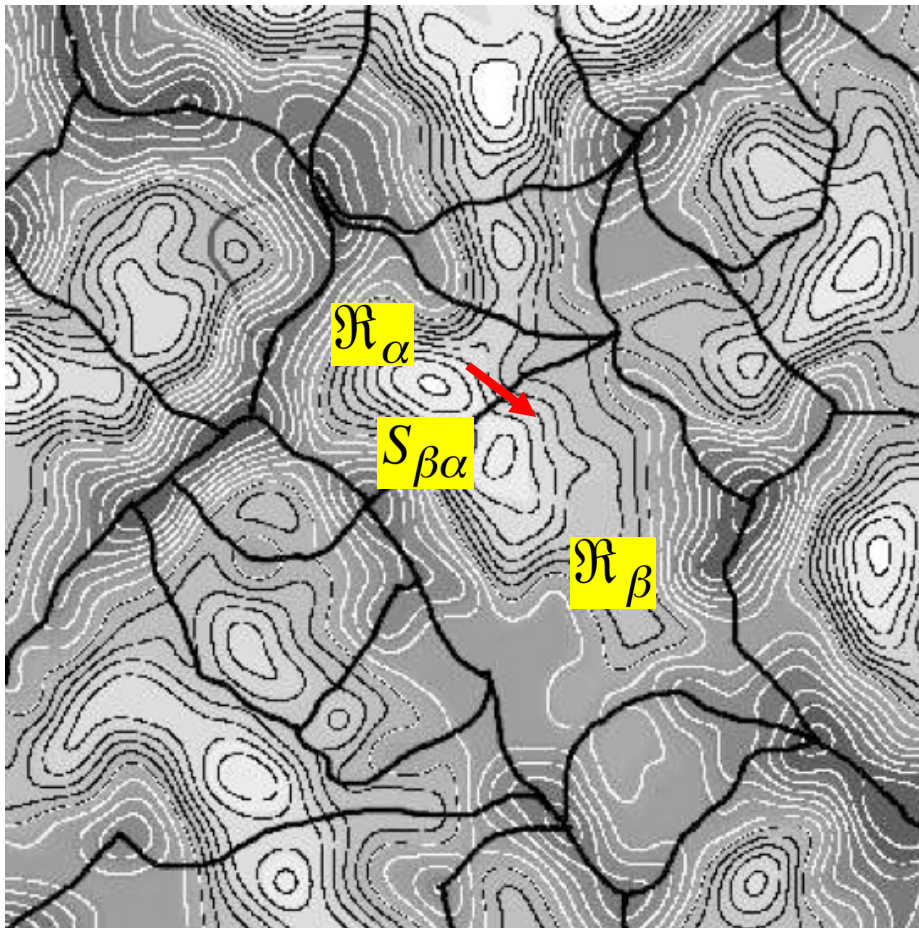
cf. H. Ostrom *et al*., *Science* **347**, 978 ('15)

Energy Landscape

- **Discrete abstraction:** Partitioned configuration space

$$\mathfrak{R}^{3N} = \bigcup_{\alpha} \mathfrak{R}_\alpha \; ; \; \mathfrak{R}_\alpha \cap \mathfrak{R}_\alpha = \emptyset$$

where a 3*N*-dimensional configuration $q \in \mathfrak{R}_\alpha$ converges to the α-th local minimum upon local minimization (*N* is the number of atoms)


Steepest descent

A Watershed

Energy Landscapes
With Applications to Clusters, Biomolecules and Glasses
David J. Wales
CAMBRIDGE MOLECULAR SCIENCE

Phase Space Distribution

- **Phase-space distribution $f(q, p, t)$: probability to find the system at $3N$-dim. position $q = (q_1, ..., q_N)$ & $3N$-dim. momentum $p = (p_1, ..., p_N)$ at time t**

- **Probability to find the system in $\Re_\alpha$ at time t (h is the Planck constant)**

$$P_\alpha(t) = \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p}\, f(q, p, t)$$

cf. Plane-wave basis

$$\sum_k = \frac{L}{2\pi} \int dk = \frac{L}{2\pi\hbar} \int \hbar dk = \frac{L}{h} \int dp \rightarrow \frac{1}{h} \int dq \int dp$$
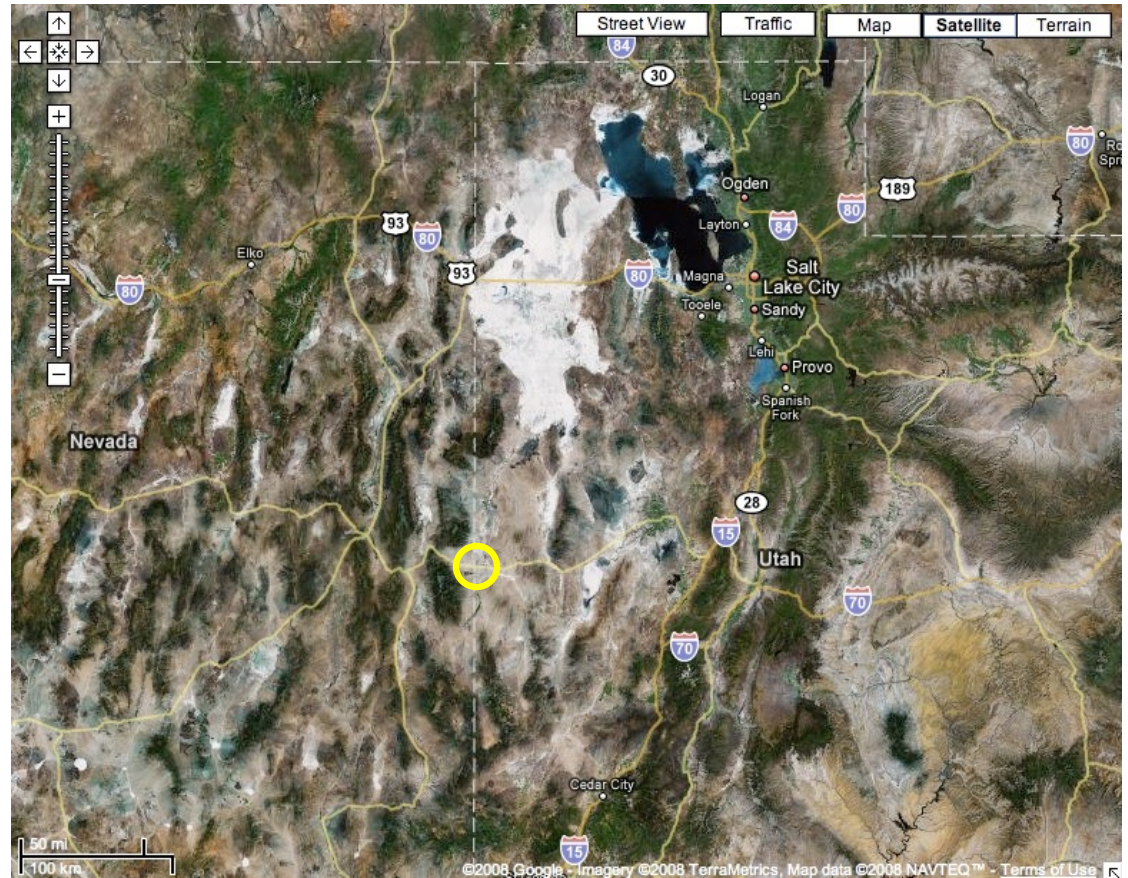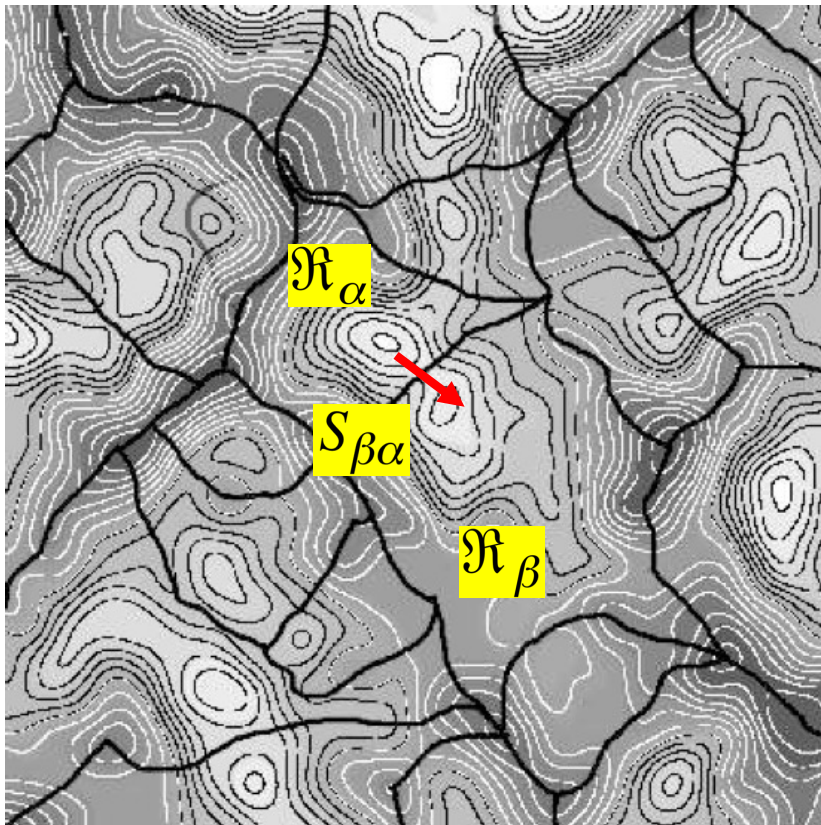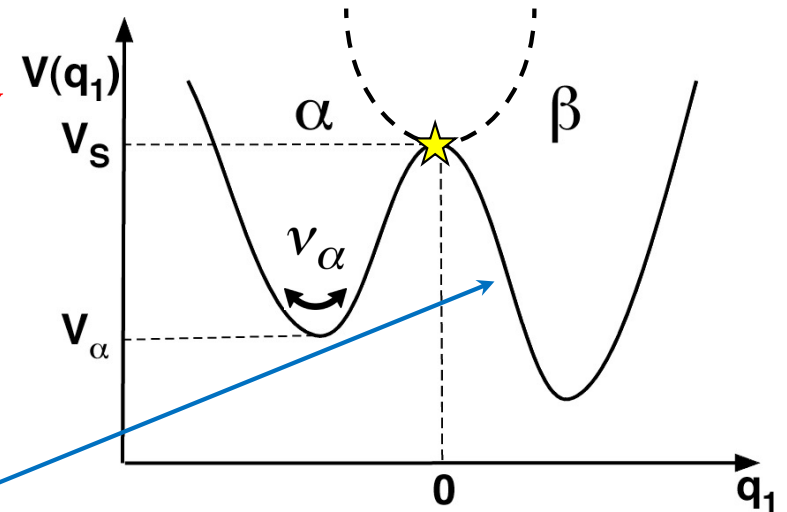
Outline

- **Master equation**

$$\frac{dP_\alpha}{dt} = -\sum_\beta W_{\beta\alpha} P_\alpha(t) + \sum_\beta W_{\alpha\beta} P_\beta(t) \quad W_{\beta\alpha}: \text{Transition rate from state } \alpha \text{ to state } \beta$$

- **How to compute $W_{\alpha\beta}$? Transition state theory**

$$W_{\beta\alpha} \cong \nu_\alpha \exp\left(-\frac{V_s - V_\alpha}{k_B T}\right)$$

Vibration frequency at the α-th local minimum

Minimum energy path along a saddle point (transition state)

- **How to simulate the master equation: Kinetic Monte Carlo simulation**

  Let $\{r_1, r_2, ...\}$ be a set of possible escape events, $r = \Sigma_i\, r_i$, and u_1 & u_2 are uniform random numbers in [0,1]:

  1. **Pick the next event i as $i = \min_j \left\{ \sum_{k=1}^{j} \frac{r_k}{r} > u_1 \right\}$**

  2. **Advance the time by $t = -\ln(u_2)/r$**

Transition state

Dynamics of Phase Space Distribution

- **Phase-space distribution $f(q, p, t)$: probability to find the system at $3N$-dim. position q = $(q_1, ..., q_N)$ & $3N$-dim. momentum p = $(p_1, ..., p_N)$ at time t**

- **Probability to find the system in $\Re_\alpha$ at time t (h is the Planck constant)**

$$P_\alpha(t) = \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} f(q,p,t)$$

- **Time derivative (L = Liouville operator; $H(q,p)$ = Hamiltonian)**

$$\frac{dP_\alpha}{dt} = \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \frac{\partial}{\partial t} f(q,p,t)$$

$$= \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \left[-Lf(q,p,t)\right]$$

$$= \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \left[-\left(\frac{\partial H}{\partial \mathbf{p}} \bullet \frac{\partial}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \bullet \frac{\partial}{\partial \mathbf{p}}\right) f(q,p,t)\right]$$
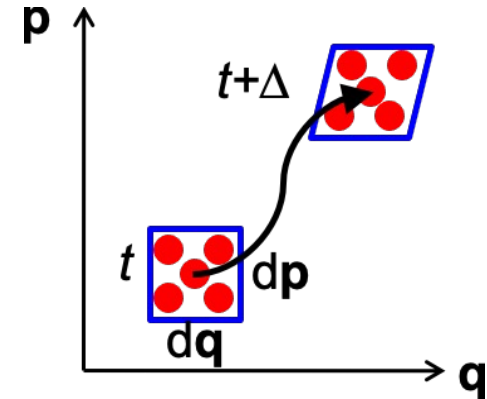
See supplementary note 1: <u>Liouville equation</u>

Sketch of Proof: Liouville Operator

$$\dot{\mathbf{X}} \bullet \frac{\partial}{\partial \mathbf{X}} = \dot{\mathbf{q}} \bullet \frac{\partial}{\partial \mathbf{q}} + \dot{\mathbf{p}} \bullet \frac{\partial}{\partial \mathbf{p}} = \frac{\partial H}{\partial \mathbf{p}} \bullet \frac{\partial}{\partial \mathbf{q}} - \frac{\partial H}{\partial \mathbf{q}} \bullet \frac{\partial}{\partial \mathbf{p}}$$

$$\frac{d}{dt} \overset{(\mathbf{q},\mathbf{p})}{\underbrace{\tilde{\mathbf{X}}}} = \underbrace{\tilde{L}} \qquad \mathbf{X}$$

Jacobian: Volume magnification factor $\left| \frac{\partial \mathbf{X}_t}{\partial \mathbf{X}_0} \right| d\mathbf{X}_0 = d\mathbf{X}_0$

Probability density as point crowd

$$N_{\text{sample}} f_0(\mathbf{X}_0) d\mathbf{X}_0 = N_{\text{sample}} f(\mathbf{X}_t, t) \underbrace{d\mathbf{X}_t}$$

Liouville's theorem

$$\left| \frac{\partial \mathbf{X}_t}{\partial \mathbf{X}_0} \right| = 1$$

Liouville lecture, p. 4

$$(1 + \delta t)^{\frac{1}{\delta t}} \xrightarrow[\delta t \to 0]{} e$$

$$(1 + \delta t^2)^{\frac{1}{\delta t}} \xrightarrow[\delta t \to 0]{} 1$$

$$f(x) = \ln(1+x) = 0 \ (x = 0); \quad \frac{df}{dx} = \frac{1}{1+x} = 1 \ (x = 0)$$

$$\therefore \ln(1+x) = f(0) + \left. \frac{df}{dx} \right|_{x=0} x + \cdots = x + \cdots$$

$$\ln(1 + \delta t)^{\frac{1}{\delta t}} = \frac{1}{\delta t} \ln(1 + \delta t) \xrightarrow[\delta t \to 0]{} \frac{1}{\delta t} \delta t = 1$$

$$\therefore (1 + \delta t)^{\frac{1}{\delta t}} \xrightarrow[\delta t \to 0]{} e$$

$$\ln(1 + \delta t^2)^{\frac{1}{\delta t}} = \frac{1}{\delta t} \ln(1 + \delta t^2) \xrightarrow[\delta t \to 0]{} \frac{1}{\delta t} \delta t^2 = 0$$

$$\therefore (1 + \delta t^2)^{\frac{1}{\delta t}} \xrightarrow[\delta t \to 0]{} 1$$

$$\overbrace{f(\mathbf{X} + \dot{\mathbf{X}} dt, dt)}^{f(\mathbf{X}, dt) + \dot{\mathbf{X}} dt \bullet \frac{\partial f}{\partial \mathbf{X}}} = f(\mathbf{X}, 0)$$

$$\therefore \frac{f(\mathbf{X}, dt) - f(\mathbf{X}, 0)}{dt} = -\dot{\mathbf{X}} \bullet \frac{\partial f}{\partial \mathbf{X}} = -Lf$$

$$\text{or} \ \left(\frac{\partial}{\partial t} + L \right) f = 0$$

population flux

Liouville lecture, p. 6

$$\frac{\partial}{\partial t} f(\mathbf{X}, t) + \frac{\partial}{\partial \mathbf{X}} \bullet (\dot{\mathbf{X}} f) = 0 \ (\mathbf{X} = (\mathbf{q}, \mathbf{p}) \in \mathfrak{R}^{6N})$$

$$cf. \ \frac{\partial}{\partial t} \rho + \nabla \bullet \vec{j} = 0 \ \text{(mass conservation)}$$

Population Dynamics

- **Let**

$$H(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} + V(\mathbf{q})$$

then

Note $\int_{-\infty}^{\infty} dp_i \frac{\partial}{\partial p_i} f(p_i) = [f(p_i)]_{-\infty}^{\infty} = 0$ (telescoping)

$$\frac{dP_\alpha}{dt} = -\frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \frac{\partial}{\partial q_i} f(\mathbf{q},\mathbf{p},t) + \frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \sum_{i=1}^{3N} \frac{\partial V}{\partial q_i} \frac{\partial}{\partial p_i} f(\mathbf{q},\mathbf{p},t)$$

$$= -\frac{1}{h^{3N}} \int_{\Re_\alpha} d\mathbf{q} \int d\mathbf{p} \sum_{i=1}^{3N} \frac{\partial}{\partial q_i} \left(\frac{p_i}{m_i} f(\mathbf{q},\mathbf{p},t) \right)$$
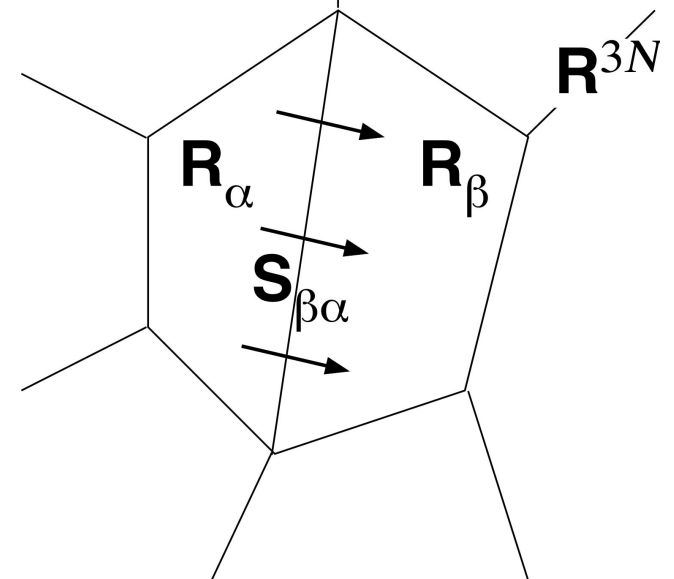
$$= -\frac{1}{h^{3N}} \int_{\partial\Re_\alpha} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} f(\mathbf{q},\mathbf{p},t)$$

Gauss' theorem ~ telescoping

$$\int_{\Re} d\mathbf{q} \nabla = \int_{\partial\Re} d\mathbf{S}$$

$\mathbf{R}^{3N}$

$\mathbf{R}_\alpha$ $\mathbf{R}_\beta$

$\mathbf{S}_{\beta\alpha}$

where dS is the surface element pointing outward normal to the surface $\partial\Re_\alpha$ that outlines $\Re_\alpha$

- **dP_α/dt is negative of the outward flux through $\partial\Re_\alpha$**

$\int dq \frac{\partial}{\partial q} f = [f]_{lower}^{upper} = f dS_+ + f dS_-$

dS_+

$\Re$

dS_-

$\partial\Re$

Population Flux

- **Partition the surface $\partial\mathfrak{R}_\alpha$ into**

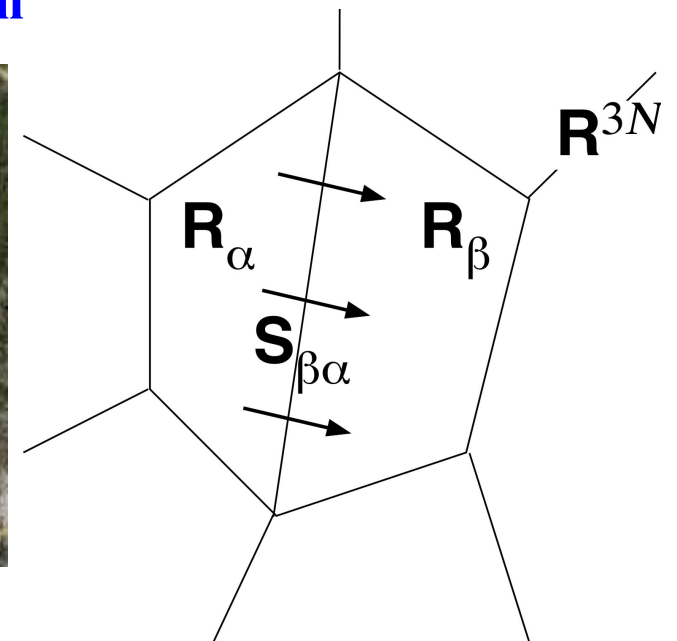$$\partial\mathfrak{R}_\alpha = \sum_\beta S_{\beta\alpha}$$

where $S_{\beta\alpha}$ is the surface splitting $\mathfrak{R}_\alpha$ & $\mathfrak{R}_\beta$ (normal pointing from α to β)

$$\frac{dP_\alpha}{dt} = -\sum_\beta \frac{1}{h^{3N}} \int\limits_{S_{\beta\alpha}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) f(\mathbf{q},\mathbf{p},t) \;\text{outgoing}$$

$$+\sum_\beta \frac{1}{h^{3N}} \int\limits_{\boxed{S_{\alpha\beta}}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) f(\mathbf{q},\mathbf{p},t) \;\text{incoming}$$

Partition of unity

$$\Theta(x) + \Theta(-x) = 1$$

$\Theta(-x)$ $\Theta(x)$

where $\Theta(x) = 1$ $(x \geq 0)$ & 0 $(x < 0)$ is the step function

flux$_{\text{UT}\leftarrow\text{NV}}$

flux$_{\text{NV}\leftarrow\text{UT}}$

$\mathbf{R}^{3N}$

$\mathbf{R}_\alpha$ $\mathbf{R}_\beta$

$\mathbf{S}_{\beta\alpha}$

Local Equilibration Approximation

- **Assume that within each $\mathfrak{R}_\alpha$, the phase space distribution is locally in thermal equilibrium, weighted to reproduce the current population (time scale of inter-state population transfer >> intra-state thermal equilibration time)**

$$f(\mathbf{q},\mathbf{p},t) \cong \frac{P_\alpha(t)}{P_\alpha(\text{eq})} f_{\text{eq}}(\mathbf{q},\mathbf{p})$$

where

$$f_{\text{eq}}(\mathbf{q},\mathbf{p}) = \frac{1}{Q} \exp\left(-H(\mathbf{q},\mathbf{p})/k_{\text{B}}T\right)$$

and the partition function is split into

$$Q = \iint \frac{d\mathbf{q}\,d\mathbf{p}}{h^{3N}} \exp\left(-H(\mathbf{q},\mathbf{p})/k_{\text{B}}T\right) = \sum_\alpha \iint_{\mathfrak{R}_\alpha} \frac{d\mathbf{q}\,d\mathbf{p}}{h^{3N}} \exp\left(-H(\mathbf{q},\mathbf{p})/k_{\text{B}}T\right) = \sum_\alpha Q_\alpha$$
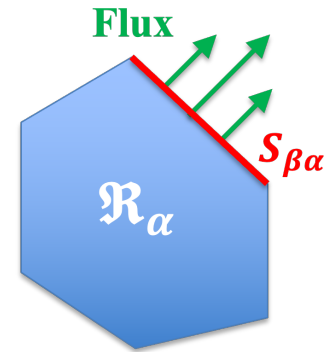
$$\therefore P_\alpha(\text{eq}) = \frac{Q_\alpha}{Q} = \frac{1}{Q} \iint_{\mathfrak{R}_\alpha} \frac{d\mathbf{q}\,d\mathbf{p}}{h^{3N}} \exp\left(-H(\mathbf{q},\mathbf{p})/k_{\text{B}}T\right)$$

Master Equation

- **Substituting the local-equilibration approximation into the population-flux equation, we obtain**

$$\frac{dP_\alpha}{dt} = -\sum_\beta \frac{1}{h^{3N}} \int_{S_{\beta\alpha}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) f_{eq}(\mathbf{q},\mathbf{p}) \frac{P_\alpha(t)}{P_\alpha(eq)}$$

$$+\sum_\beta \frac{1}{h^{3N}} \int_{S_{\alpha\beta}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) f_{eq}(\mathbf{q},\mathbf{p}) \frac{P_\beta(t)}{P_\beta(eq)}$$

Flux

$S_{\beta\alpha}$

$\Re_\alpha$

$$\therefore \frac{dP_\alpha}{dt} = -\sum_\beta W_{\beta\alpha} P_\alpha(t) + \sum_\beta W_{\alpha\beta} P_\beta(t)$$

$$W_{\alpha\beta} = \frac{1}{h^{3N}} \int_{S_{\alpha\beta}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) f_{eq}(\mathbf{q},\mathbf{p}) \bigg/ P_\beta(eq)$$

Weighted average of flux on surface

$$= \frac{1}{h^{3N}} \int_{S_{\alpha\beta}} \sum_{i=1}^{3N} dS_i \int d\mathbf{p} \sum_{i=1}^{3N} \frac{p_i}{m_i} \Theta\left(\sum_i dS_i \frac{p_i}{m_i}\right) e^{-H(\mathbf{q},\mathbf{p})/k_B T} \bigg/ \frac{1}{h^{3N}} \iint_{\Re_\beta} d\mathbf{q}d\mathbf{p}\, e^{-H(\mathbf{q},\mathbf{p})/k_B T}$$

See supplementary note 2: Master equation

A. P. J. Jansen, *Introduction to KMC Simulations of Surface Reactions* (Springer, '12)

Transition State Theory

- **Reaction coordinate** q_1 **along a** **minimum-energy path** **separates the phase space into 2 regions — A ($q_1 < 0$) & B ($q_1 > 0$); all the other coordinates & momenta are collectively denoted as X = (q, p) = (q_2, ..., q_{3N}, p_2, ..., p_{3N})**

- **Probability to find the system being in B**

$$P_B(t) = \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \Theta(q_1) f(q_1, p_1, \mathbf{X}, t)$$

S_{BA} q_1

Saddle point

Along a path!

- **Time derivative (L = Liouville operator)**

$$\frac{dP_B}{dt} = \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \Theta(q_1) \frac{\partial}{\partial t} f(q_1, p_1, \mathbf{X}, t)$$

$$= \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \Theta(q_1)(-Lf(q_1, p_1, \mathbf{X}, t))$$

Integration by part

$$= \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} (L\Theta(q_1)) f(q_1, p_1, \mathbf{X}, t)$$

$$= \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \left(\frac{p_1}{m_1} \delta(q_1) \right) f(q_1, p_1, \mathbf{X}, t)$$

Weighted average of flux on surface

$$= \iint \frac{dp_1 d\mathbf{X}}{h^{3N}} \frac{p_1}{m_1} f(0, p_1, \mathbf{X}, t)$$

$V(q_1)$ **A**   **B**

0 q_1
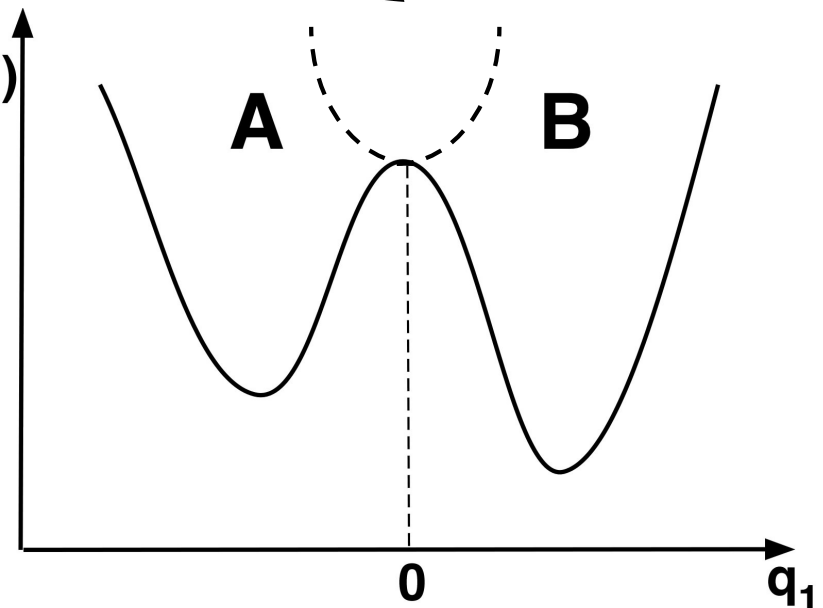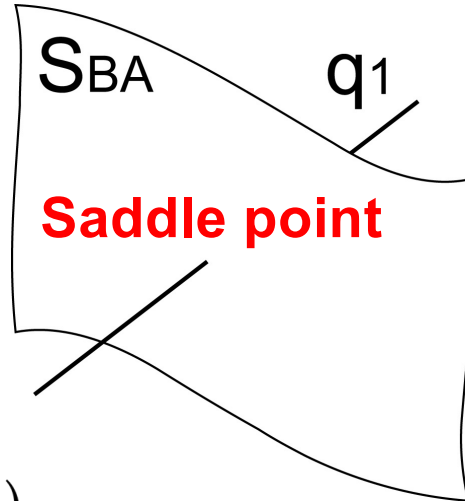
Local Equilibration Approximation

- **Split the integral into the gain (A → B) & loss (B → A) terms**

$$\frac{dP_B}{dt} = \left(\frac{dP_B}{dt}\right)_{A \to B} + \left(\frac{dP_B}{dt}\right)_{B \to A} \qquad \Theta(p_1) + \Theta(-p_1) = 1$$

$$= \int_0^\infty \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f(0, p_1, \mathbf{X}, t) + \int_{-\infty}^0 \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f(0, p_1, \mathbf{X}, t)$$

- **Regions A & B locally (*i.e.*, within the region) maintain the equilibrium distribution weighted to reproduce the current population**

$$f_{\alpha,\text{local}} \cong \frac{P_\alpha(t)}{P_\alpha(\text{eq})} f_{\text{eq}} \qquad (\alpha = A, B)$$

$$f_{\text{eq}} = \frac{1}{Q} \exp(-H/k_B T)$$

$$Q = \iiint \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \exp(-H/k_B T) = Q_A + Q_B$$

$$Q_A = \iiint_{q_1 < 0} \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \exp(-H/k_B T) \qquad Q_B = \iiint_{q_1 > 0} \frac{dq_1 dp_1 d\mathbf{X}}{h^{3N}} \exp(-H/k_B T)$$

$$P_\alpha(\text{eq}) = \frac{Q_\alpha}{Q} \qquad (\alpha = A, B)$$
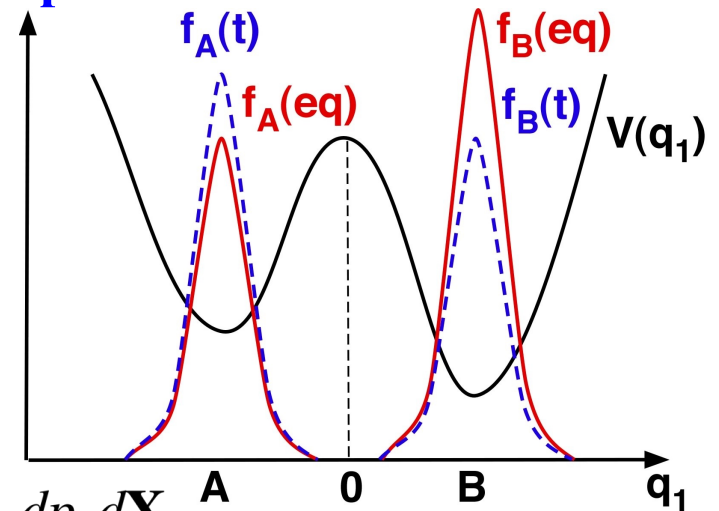
Transition State Theory

- **Substituting the local equilibration approximation to the flux equation**

$$\left(\frac{dP_B}{dt}\right)_{A\to B} = \int_0^\infty \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f_{eq}(0,p_1,\mathbf{X}) \frac{P_A(t)}{P_A(eq)}$$

$$\left(\frac{dP_B}{dt}\right)_{B\to A} = -\int_0^\infty \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f_{eq}(0,p_1,\mathbf{X}) \frac{P_B(t)}{P_B(eq)}$$

$$\therefore \frac{dP_B(t)}{dt} = k_{BA} P_A(t) - k_{AB} P_B(t)$$

$$k_{BA} = \int_0^\infty \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f_{eq}(0,p_1,\mathbf{X}) \frac{1}{P_A(eq)}$$

$$k_{AB} = \int_0^\infty \frac{dp_1}{h} \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{p_1}{m_1} f_{eq}(0,p_1,\mathbf{X}) \frac{1}{P_B(eq)}$$

- **Analytical integration over p_1**

$$k_{BA} = \int_0^\infty \frac{dp_1}{h} \frac{p_1}{m_1} \exp\left(-p_1^2/2m_1 k_B T\right) \int \frac{d\mathbf{X}}{h^{3N-1}} \frac{\exp(-H/k_B T)_{q_1=p_1=0}}{Q} \frac{Q}{Q_A} = \frac{k_B T}{h} \frac{Q^*}{Q_A}$$

$$Q^* = \int \frac{d\mathbf{X}}{h^{3N-1}} \exp(-H/k_B T)_{q_1=p_1=0}$$

$k_B T/h \sim 10^{13}\ \text{s}^{-1}$
at room temperature

Harmonic Transition State Theory

- **In region A, we assume**

$$V(q_1, \ldots, q_{3N}) = V_A + \frac{1}{2}\sum_j m_j\left(\omega_j^A\right)^2\left(q_j - b_j\right)^2$$

$$\therefore Q_A = \left(\frac{2\pi k_B T}{h}\right)^{3N} \frac{\exp(-V_A / k_B T)}{\displaystyle\prod_{j=1}^{3N} \omega_j^A}$$

- **At the dividing surface, we assume**

$$V(q_1, \ldots, q_{3N}) = V_S - \frac{1}{2}a_{11}q_1^2 + \frac{1}{2}\sum_{j=2}^{3N} m_j\left(\omega_j^*\right)^2 q_j^2$$

$$Q^* = \int \frac{d\mathbf{X}}{h^{3N-1}}\exp(-H / k_B T)_{q_1 = p_1 = 0} = \left(\frac{2\pi k_B T}{h}\right)^{3N-1}\frac{\exp(-V_s / k_B T)}{\displaystyle\prod_{j=2}^{3N}\omega_j^*}$$

attempt frequency

acceptance probability

$$\therefore k_{BA} = \frac{1}{2\pi}\exp\left(-\frac{V_s - V_A}{k_B T}\right)\frac{\displaystyle\prod_{j=1}^{3N}\omega_j^A}{\displaystyle\prod_{j=2}^{3N}\omega_j^*} \cong \boxed{\frac{\omega_1^A}{2\pi}}\boxed{\exp\left(-\frac{V_s - V_A}{k_B T}\right)}$$

$V(q_1)$
V_s

A   B

V_A

0   q_1
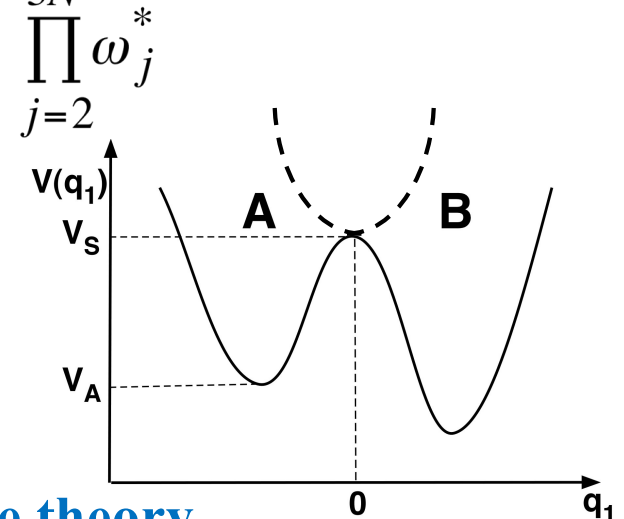
See supplementary note 3: <u>Transition state theory</u>

Digression: Save the World?

- **Solar land-area requirement (with 10% energy conversion efficiency) to supply the global energy** [Nathan Lewis, *Caltech*]

- **Need better catalyst for splitting water** [Lewis & Nocera, *PNAS* **103**, 15729 ('06)]

$$2H_2O \xrightarrow{h\nu} 2H_2 + O_2$$
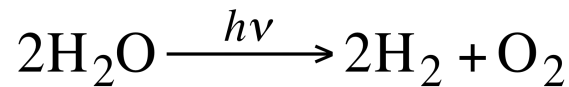
COMPUTATIONAL METHODS

A search engine for catalysts

Trial and error has been the traditional method of finding the best catalyst for a reaction. A computational approach can reduce the lab work required.

Just bring down the transition-state energy!

Nature Mater. **5**, 847 ('06)

MANOS MAVRIKAKIS
is in the Department of Chemical & Biological Engineering, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.

Need > 2,000°C
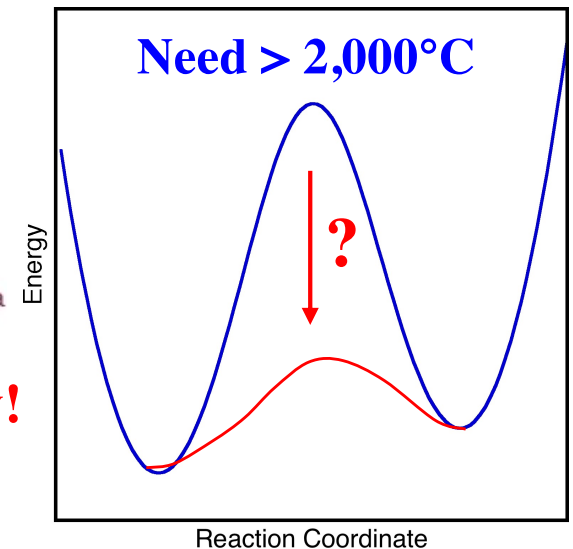
Hydrogen Production@Home

A Job for Superatom

Molecular Dynamics Simulations of Rapid Hydrogen Production from Water Using Aluminum Clusters as Catalyzers

Fuyuki Shimojo,[1,2] Satoshi Ohmura,[1,2] Rajiv K. Kalia,[1] Aiichiro Nakano,[1] and Priya Vashishta[1]

[1]Collaboratory for Advanced Computing and Simulations, Department of Computer Science, Department of Physics & Astronomy, Department of Chemical Engineering & Materials Science, University of Southern California, Los Angeles, California 90089-0242, USA
[2]Department of Physics, Kumamoto University, Kumamoto 860-8555, Japan
(Received 1 December 2009; published 26 March 2010)

$$k_B T_{room}/h \sim 10^{13} \ (\text{s}^{-1})$$

$$1 \ \text{eV}/k_B \sim 10{,}000 \ \text{K}$$

$$e^{-x} \sim 10^{-x/2}$$
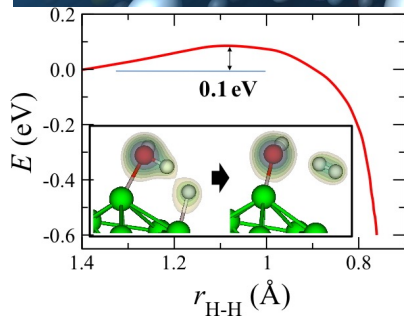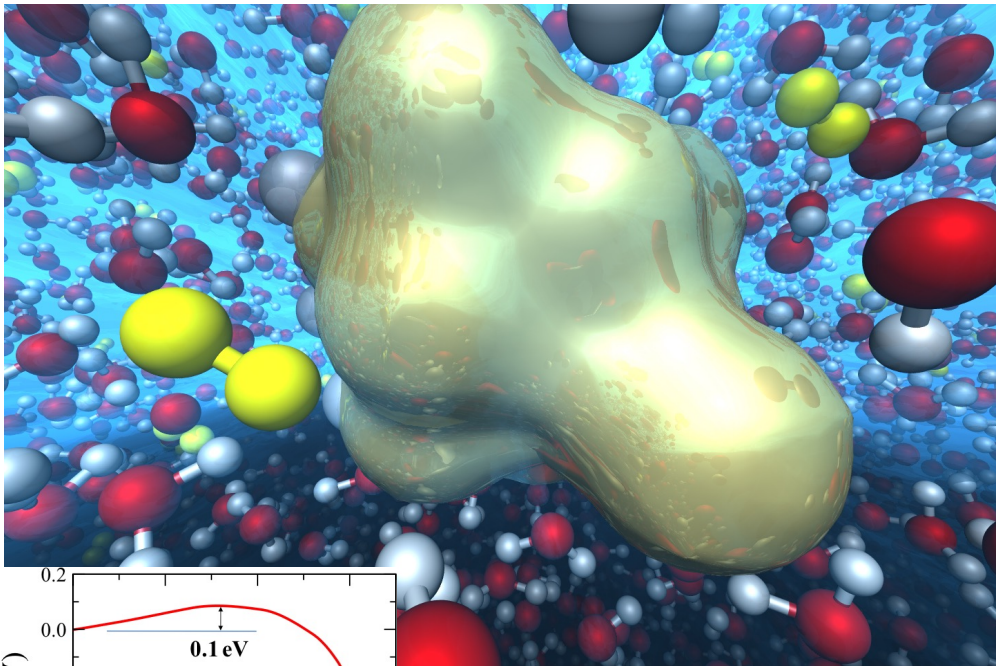
$$k_{H_2} = \frac{k_B T}{h} \exp\left(-\frac{\Delta}{k_B T_{room}}\right)$$
$$= 10^{11} (\text{s}^{-1})$$
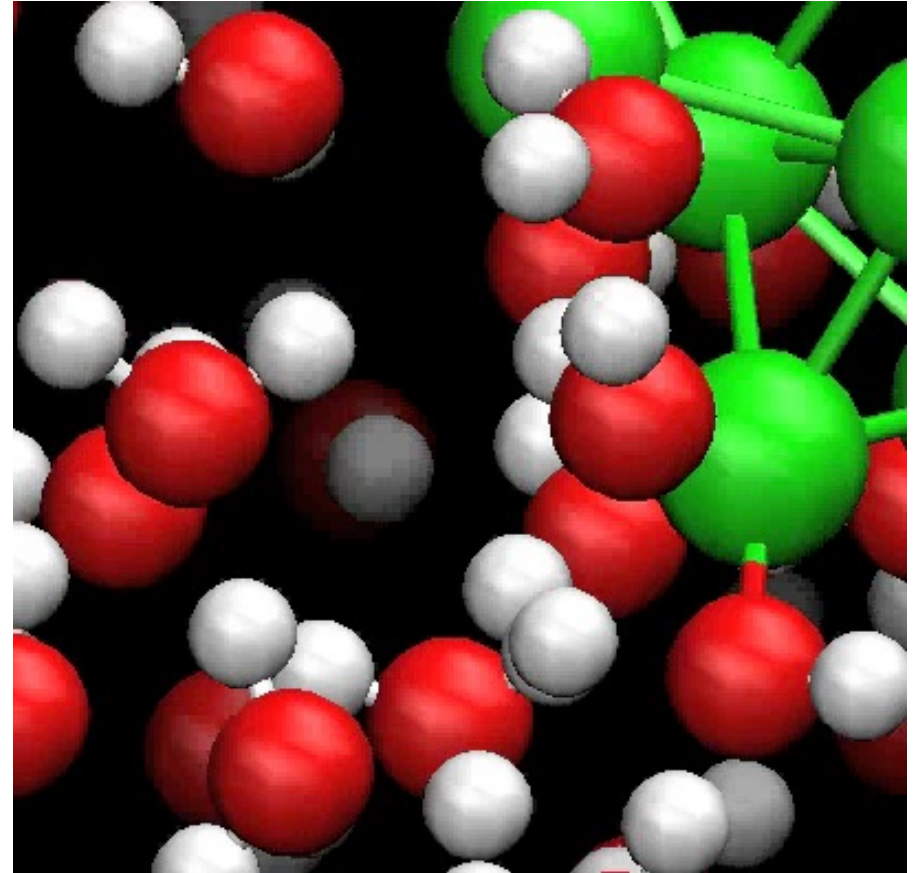
H$_2$ Production from Water Using LiAl Particles

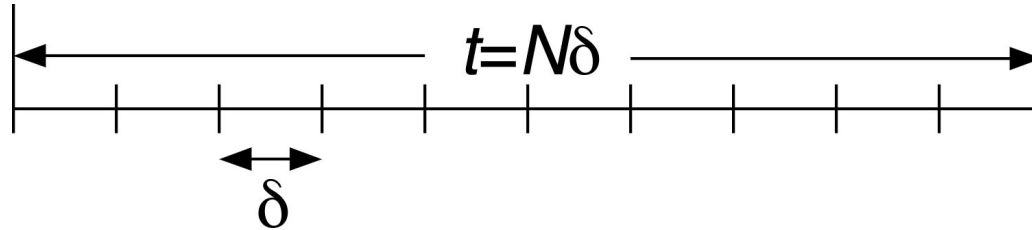**16,661-atom QMD simulation of Li$_{441}$Al$_{441}$ in water on 786,432 IBM Blue Gene/Q cores**
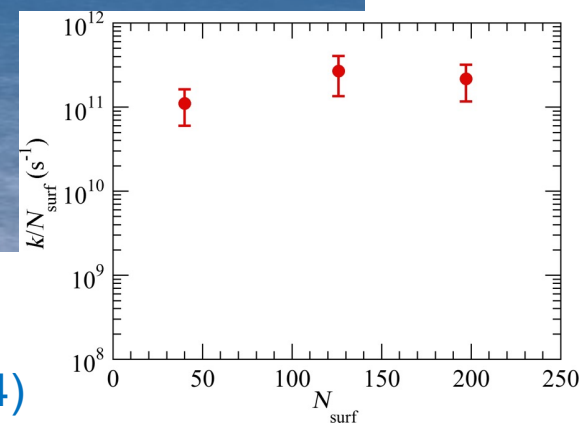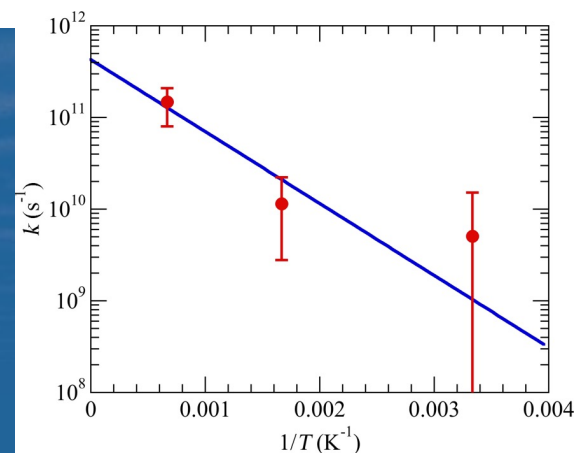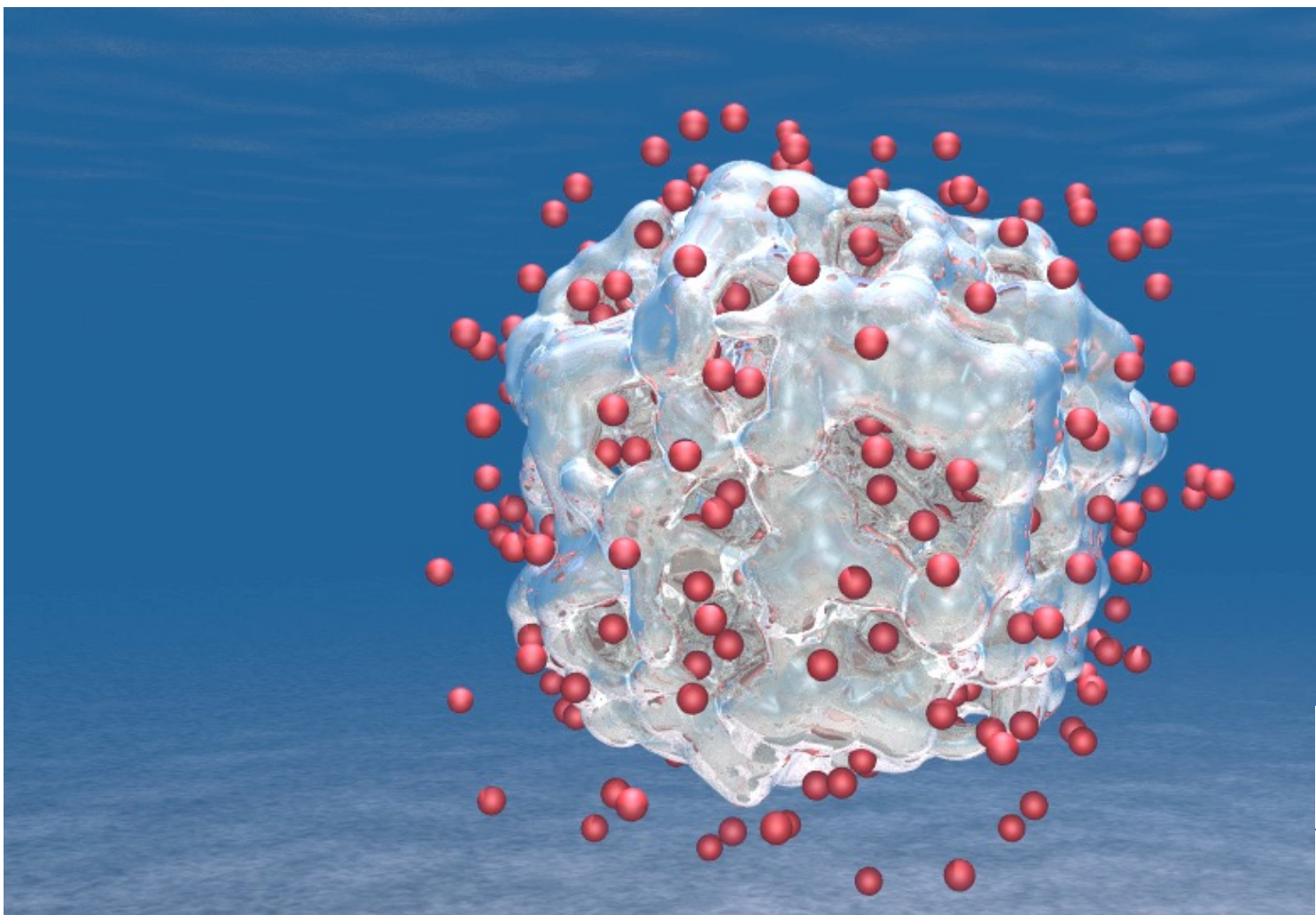
- **Scalable to industrially relevant particle sizes**

K. Shimamura *et al.*, *Nano Lett.* **14**, 4090 ('14)

Poisson Process

- **Poisson process = sequence of events, in which the probability of an event to occur in time $[t, t+\delta]$ is $r\delta$ (r is the rate) independent of history as $\delta \to 0$**
- **Probability $P(n,t)$ that n events occur in time interval $t=N\delta$**

$$\overset{\longleftarrow \quad t=N\delta \quad \longrightarrow}{\mid \quad \mid \quad \mid \quad \mid \quad \mid \quad \mid \quad \mid \quad \mid \quad \mid \quad \mid}$$

$$\overset{\longleftrightarrow}{\delta}$$

$$P(n,t) = C(N,n)(r\delta)^n (1-r\delta)^{N-n} = \frac{N!}{n!(N-n)!}(r\delta)^n (1-r\delta)^{N-n}$$

$$\xrightarrow[N\to\infty]{} \frac{(rt)^n}{n!} e^{-rt}$$

$$\because \begin{cases} (1-rt/N)^{N/rt} \xrightarrow[N\to\infty]{} e^{-1} \\ N! \xrightarrow[N\to\infty]{} \sqrt{2\pi N} N^N e^{-N} \end{cases}$$

$$\lim_{N\to\infty}\left(1 \pm \frac{1}{N}\right)^N = e^{\pm 1}$$

$$\because \ln\left(1 \pm \frac{1}{N}\right)^N = N\ln\left(1 \pm \frac{1}{N}\right)$$

$$\xrightarrow[N\to\infty]{} N\left(\pm\frac{1}{N}\right) = \pm 1$$

- **Sum rule:**

$$\sum_{n=0}^{\infty} P(n,t) = \sum_{n=0}^{\infty} \frac{(rt)^n}{n!} e^{-rt} = 1$$

Kinetic Monte Carlo Simulation

- **Probability density $P(t)$ of time t between successive events**

$P(t)dt = \text{probability}(\text{no event in } [0,t] \wedge 1 \text{ event in } [t, t+dt])$

$= P(0,t) \times rdt = e^{-rt} \times rdt$

$\therefore P(t) = re^{-rt}$

$t = Ndt$

$dt \left(1 - r\frac{t}{N}\right)^N \times rdt = \left(\left(1 - \frac{rt}{N}\right)^{\frac{N}{rt}}\right)^{rt} rdt \xrightarrow[N \to \infty]{} e^{-rt}rdt$
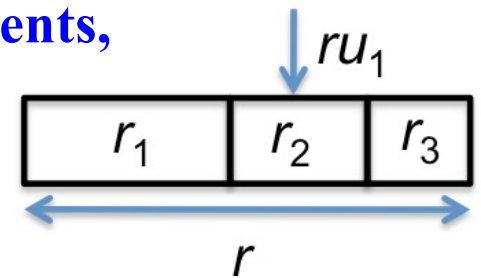
- **Random time-interval generation: Let u be a uniform random number in $[0,1]$ & generate $t = -\ln(u)/r \in [0,\infty]$**

$u = e^{-rt}$

$\because P(t) = P(u)\left|\frac{du}{dt}\right| = 1 \times re^{-rt} = re^{-rt}$

- **Kinetic MC algorithm: Let $\{r_1, r_2, ...\}$ be a set of possible events, $r = \Sigma_i\, r_i$, and u_1 & u_2 are uniform random numbers in $[0,1]$:**

ru_1

$r_1 \quad r_2 \quad r_3$

r

1. **Pick the next event i as** $i = \min_j \left\{ \sum_{k=1}^{j} \frac{r_k}{r} > u_1 \right\}$

   Pick i-th event with probability $p_i = r_i/r$

2. **Advance the time by** $t = -\ln(u_2)/r$

   $P(t)dt = \overbrace{e^{-rt}}^{\text{no evt in }[0,t]} \overbrace{(r_1 + r_2 + \cdots)dt}^{\text{evt 1 or 2 or ... in }[t,t+dt]}$

K. A. Fichthorn & W. H. Weinberg, *J. Chem. Phys.* **95**, 1090 ('91)

See supplementary note 4: Kinetic Monte Carlo simulation

Divide-&-Conquer KMC Algorithm

- **Domain decomposition:** Concurrent events among multiple domains, *d*

$$\Delta t = -\ln(rnd) / \sum_d r_d = O(N^{-1}) \quad \Rightarrow \quad -\ln(rnd) / \max_d (r_d) = O(1)$$

- **Colored domain blocks: Avoids conflicting events by allowing concurrent events only with domains of the same color, which are well-separated**
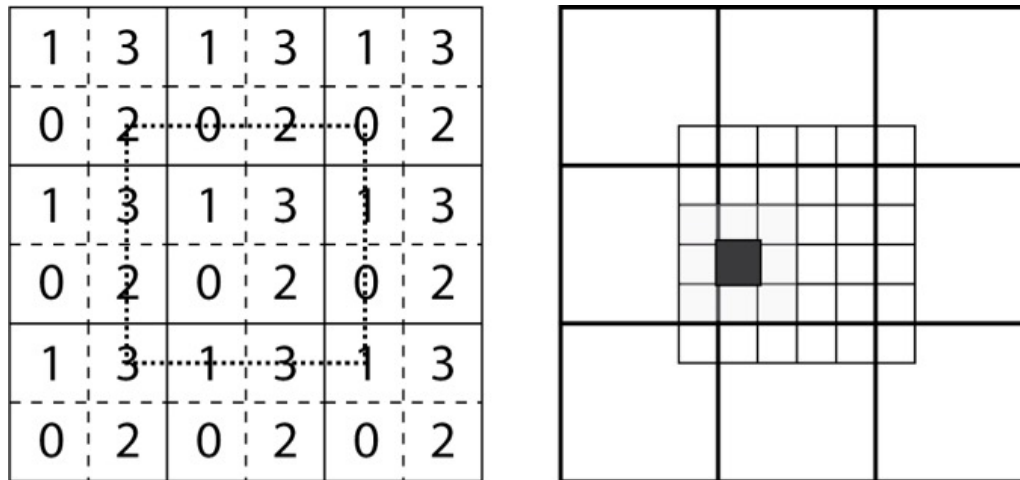

E. Martinez *et al., J. Comp. Phys.* **230**, 1359 ('11)

- **Neighbor-domain caching for spatial decomposition via message-passing**
- **Dual linked-list cell method: (1) small cells for constructing neighbor lists for nearest-neighbor hopping events; (2) large cells for domain-block coloring**

H. Byun *et al., Comput. Phys. Commun.* **219**, 246 ('17)

Scalable Parallel KMC

- **Benchmark tests on electron transfer in heme aggregates**

- **Better weak-scaling for coarser granularity (N hemes on P procesors)**
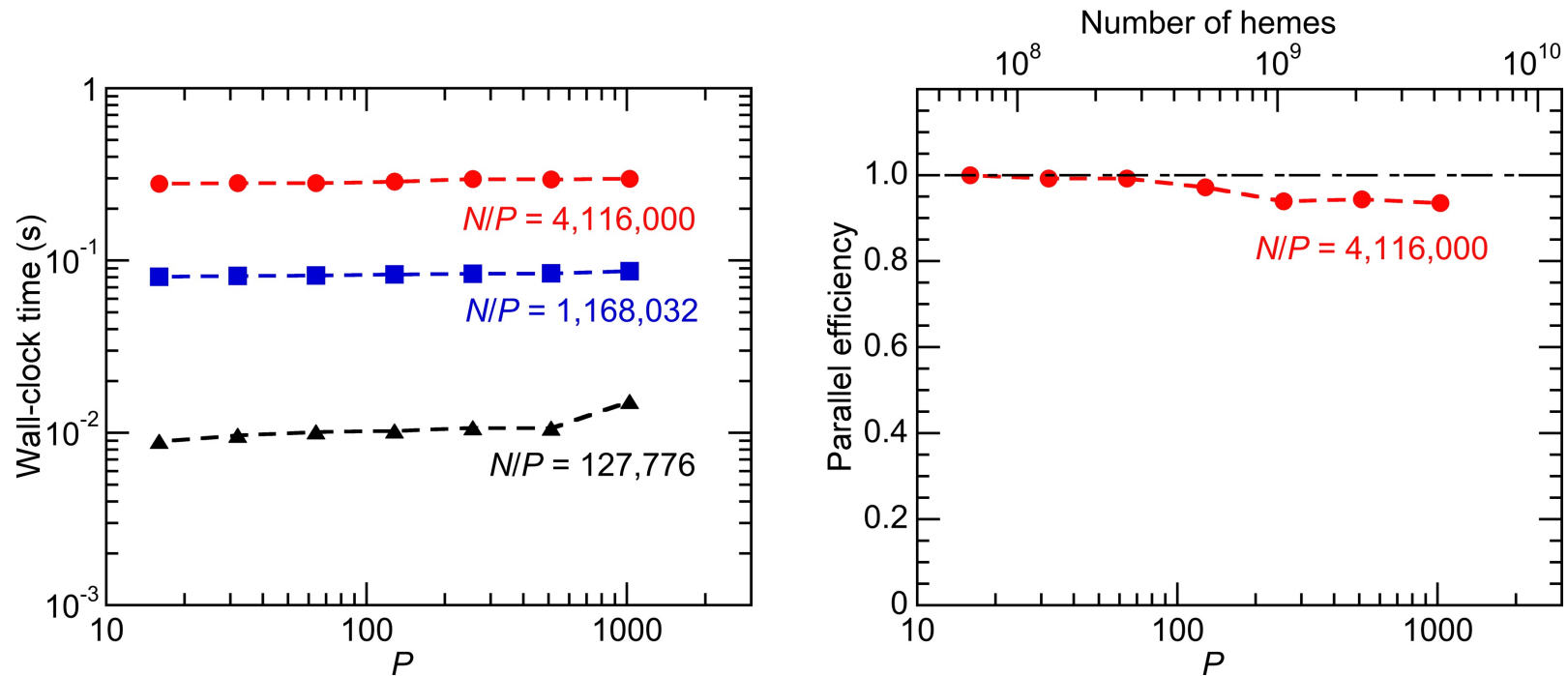
- **Weak-scaling parallel efficiency 0.935 for a 4.2 billion-heme system on 1,024 Intel Xeon processors**
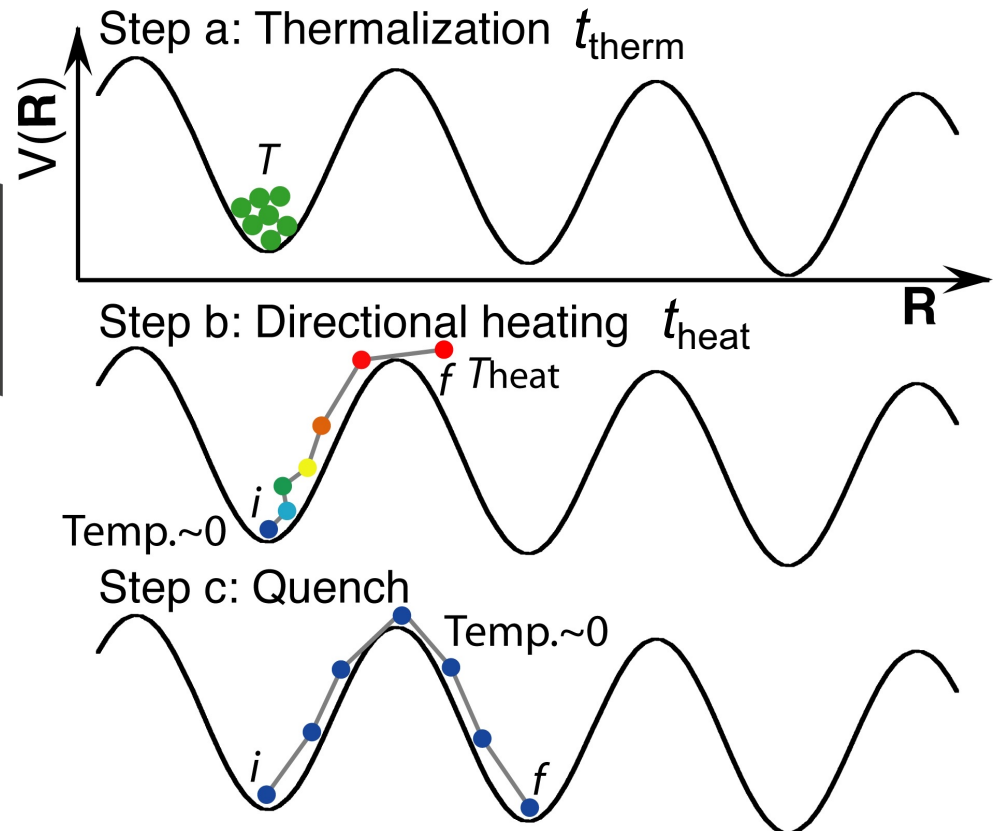
H. Byun *et al.*, *Comput. Phys. Commun.* **219**, 246 ('17)

Temporal Locality in Long-Time Dynamics

- **Temporal locality:** Rare transitions between local minimum-energy states
- **Transition state theory:** Reformulate *sequential* long-time dynamics as *parallel* search for low activation-barrier transition events
- **Discrete graph abstraction:** Linear combinations of atomistic events (LCAE)

  A. Nakano, *Comput. Phys. Commun.* **176**, 292 ('07)

- **Directionally heated nudged elastic band (NEB) method:** Search for thermally activated events without the knowledge of final states

$$\mathbf{M}\ddot{\mathbf{R}}_s = \mathbf{F}_s - M\gamma_s\dot{\mathbf{R}}_s \ (s = 0, \ldots, S-1)$$

$i = 1$
2
N
$s = 0 \quad 1 \quad \cdots \quad S-2 \quad S-1$

$0 \quad 1 \quad \cdots \quad S-2 \quad S-1$
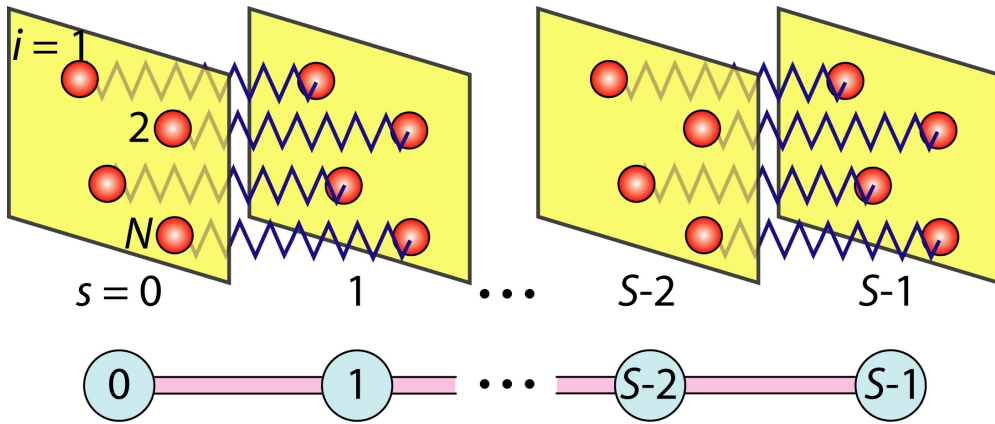
$$\mathbf{F}_s = \begin{cases} -\dfrac{\partial V}{\partial \mathbf{R}_s}\bigg|_\perp + \mathbf{F}_s^{\mathrm{spr}}\bigg|_\parallel & (1 \leq s \leq S-2) \\ -\dfrac{\partial V}{\partial \mathbf{R}_s} & (s = 0, S-1) \end{cases}$$

Step a: Thermalization t_{therm}

$V(\mathbf{R})$

T

$\mathbf{R}$

Step b: Directional heating t_{heat}

$f \ T_{\mathrm{heat}}$

i

Temp. ~0

Step c: Quench

Temp. ~0

i

f

Space-Time-Ensemble Parallel (STEP) NEB

- **Path ensemble method (PEM): Long-time simulation in the framework of kinetic Monte Carlo—*molecular kinetics simulation***

$$r_b = \left\{ t_{\text{therm}} + t_{\text{heat}} \exp\left[\frac{\Delta_b}{k_{\text{B}}} \left(\frac{1}{T} - \frac{1}{T_{\text{heat}}} \right) \right] \right\}^{-1}$$

$$P_b = \frac{r_b}{r} = \frac{r_b}{\sum_{b=0}^{B-1} r_b}$$
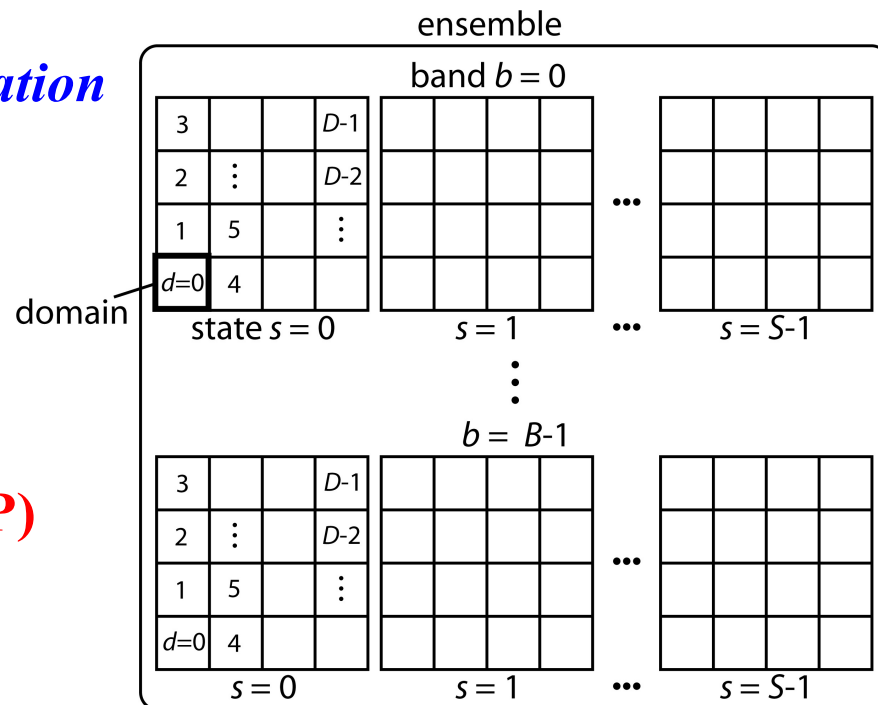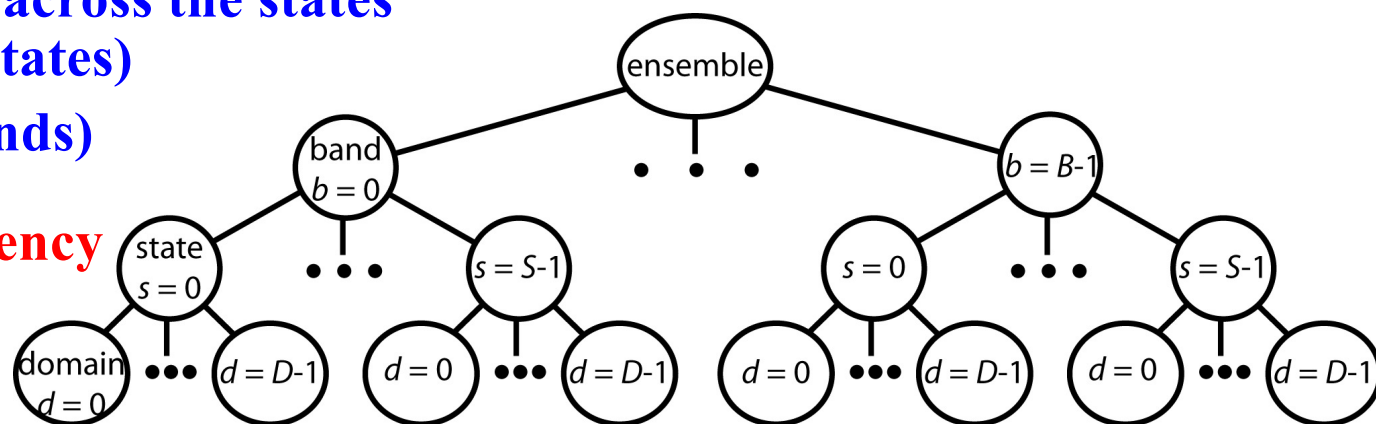
- **Space-time-ensemble parallelism (STEP)**
- = **spatial decomposition within each state (D domains)**
- + **temporal parallelism across the states within each band (S states)**
- + **band ensemble (B bands)**

- **Hierarchical concurrency**
$P = BSD$

A. Nakano, *Comput. Phys. Commun.* **178**, 280 ('08)

Divide-&-Conquer Protein Folding

- **Levinthal paradox (1968):** How the Nature folds an amino-acid sequence into a global energy minimum 3D structure (which is known to be NP complete) within microseconds (~ billion molecular-dynamics steps).

- **Sequential KMC not good enough.**

Zip-&-assembly algorithm (Ken Dill at UCSF)

1. **(Divide)** Chop the amino-acid sequence into ~10 residue fragments.

2. **(Conquer)** For each fragment, perform replica-exchange (~ temperature accelerated) molecular dynamics simulation & detect the formation of any stable hydrophobic contacts.

3. **(Combine)** Grow the stable fragments by adding surrounding residues while freezing (~ constraint) the found stable contacts.
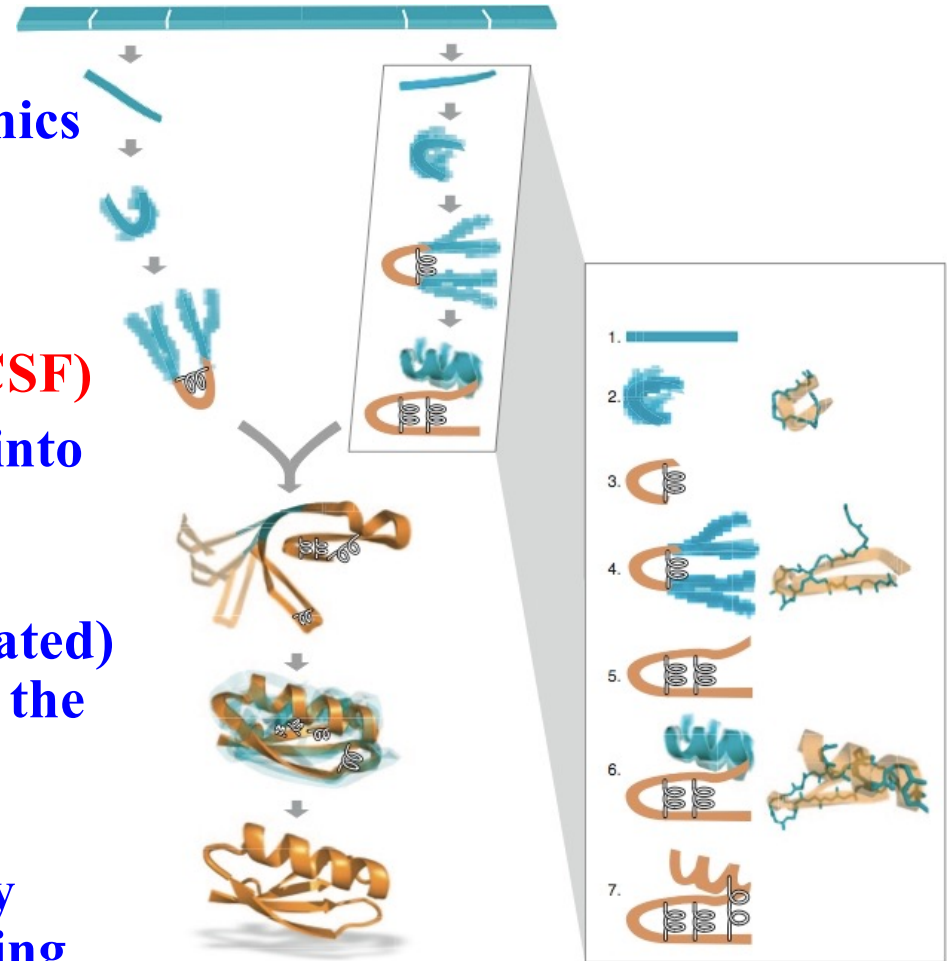
S. B. Ozkan *et al.*, *PNAS* **104**, 11987 ('07)

Parsing Protein-Folding Routes

Computational linguistics

(1) Formal grammar to describe protein-folding routes

(2) Dynamic programming for an efficient algorithm for the folding routes
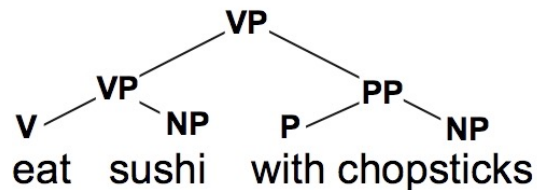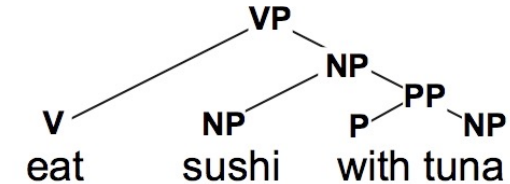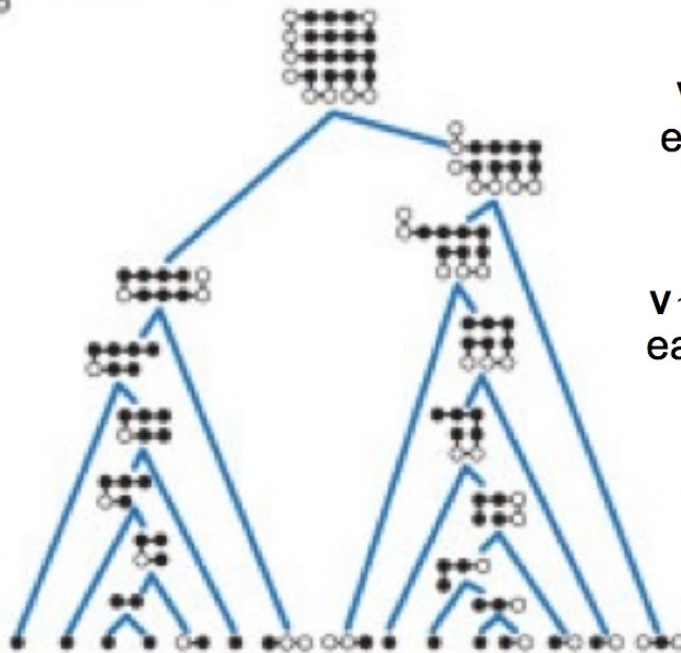
K. A. Dill *et al.*, *Polymer* **48**, 4289 ('07)
W. Dyrka *et al.*, *Alg. Mol. Biol.* **8**, 31 ('13)
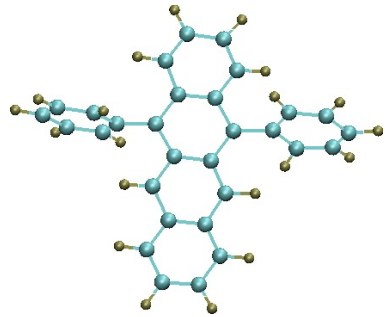
Singlet Fission in Amorphous DPT

- **Photo-current doubling by splitting a singlet exciton into 2 triplet excitons**

- **Singlet fission (SF) in mass-produced disordered organic solid**
  → efficient low-cost solar cells

- **Exp'l breakthrough: SF found in amorphous diphenyl tetracene (DPT)**

- **Ultrafast transient absorption measurements identified *two time-scales* (1 & 100 ps) for exciton population dynamics**

- **Hypothesis: Existence of *SF hot spots* [S. T. Roberts *et al.*, *JACS* 134, 6388 ('12)]**
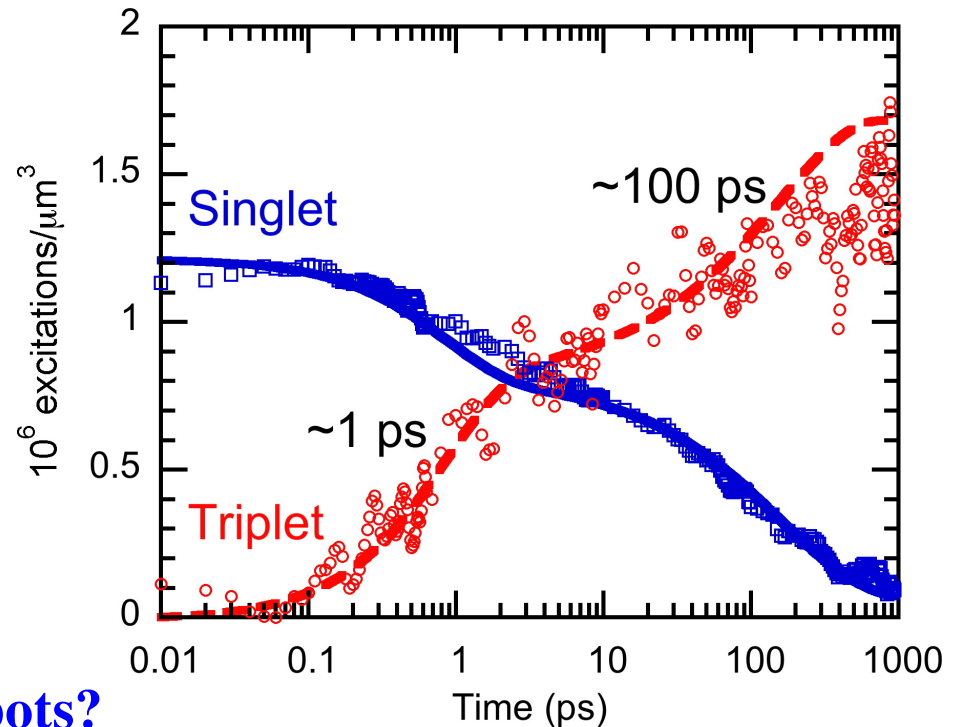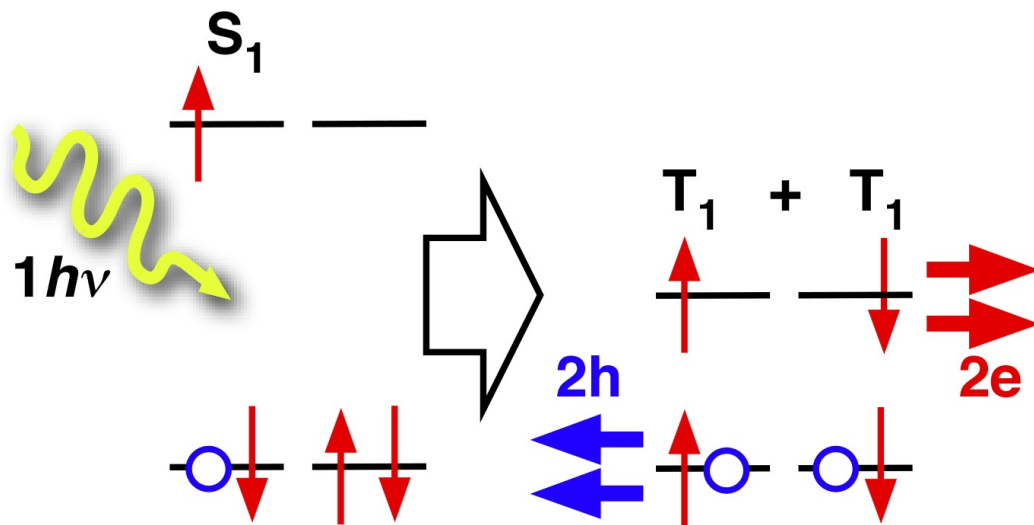
- **Problem: *Molecular origin* of SF hot spots?**
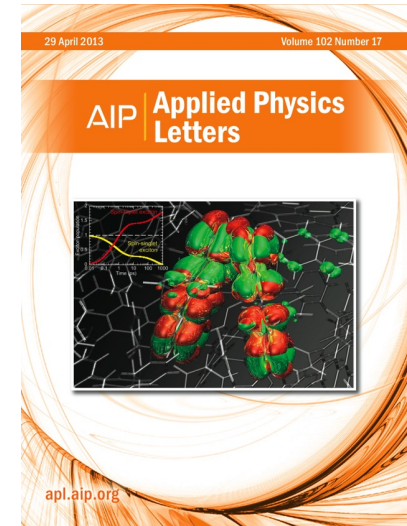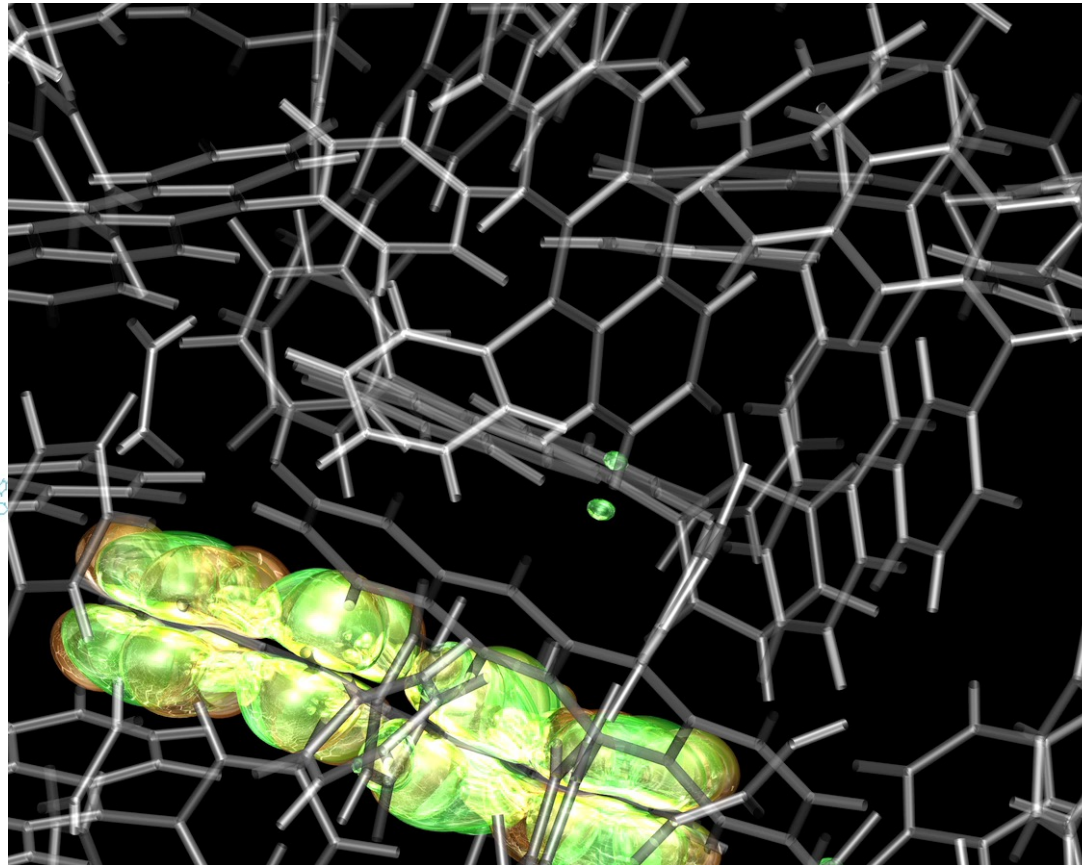
Divide-Conquer-Recombine KMC

- **Move up from molecules to microstructures**
- **Challenge:** Unprecedented 10^4-atom NAQMD simulation
- **Computational approach:** Divide-conquer-recombine (DCR) NAQMD


DPT molecule

Amorphous DPT

domain *i*

domain *j*
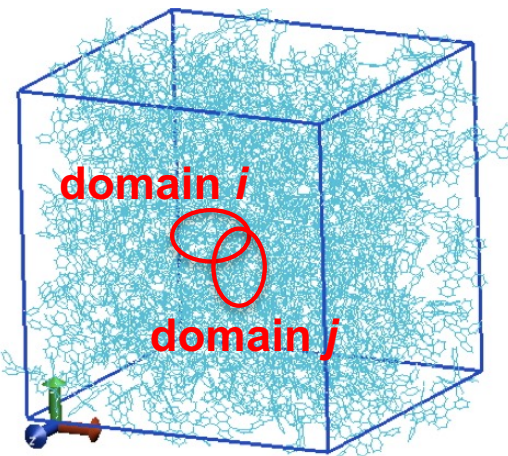
Quasi-electron

Quasi-hole

- **DCR-NAQMD (phonon-assisted exciton dynamics) + time-dependent perturbation theory (singlet-fission rate) + kinetic Monte Carlo calculations of exciton population dynamics in 6,400-atom amorphous DPT**
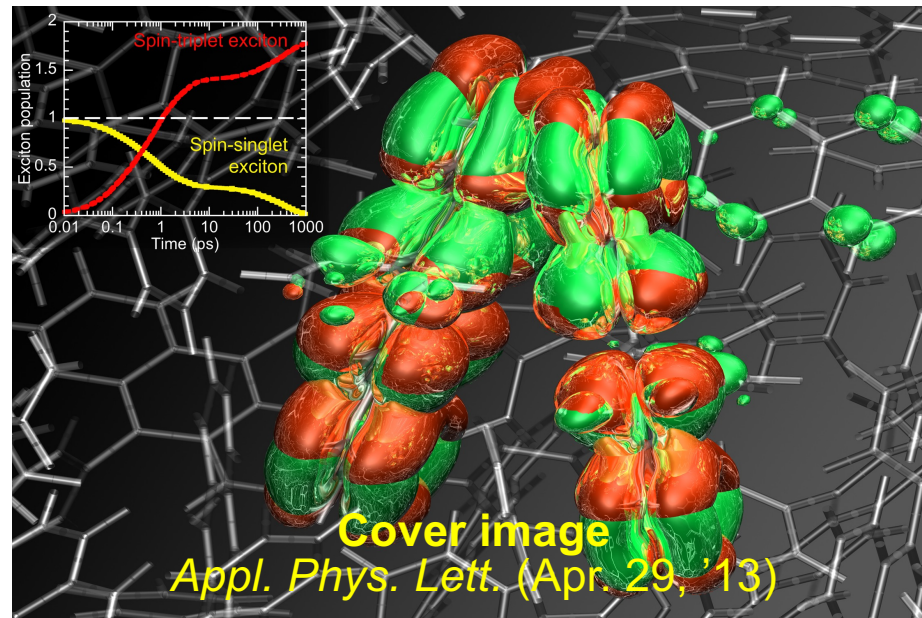
F. Shimojo *et al.*, *J. Chem. Phys.* **140**, 18A529 ('14)

NAQMD-informed Kinetic Monte Carlo

- **NAQMD-KMC exciton population dynamics reproduces the experimentally observed two time scales (~ 1 & 100 ps) in amorphous DPT**
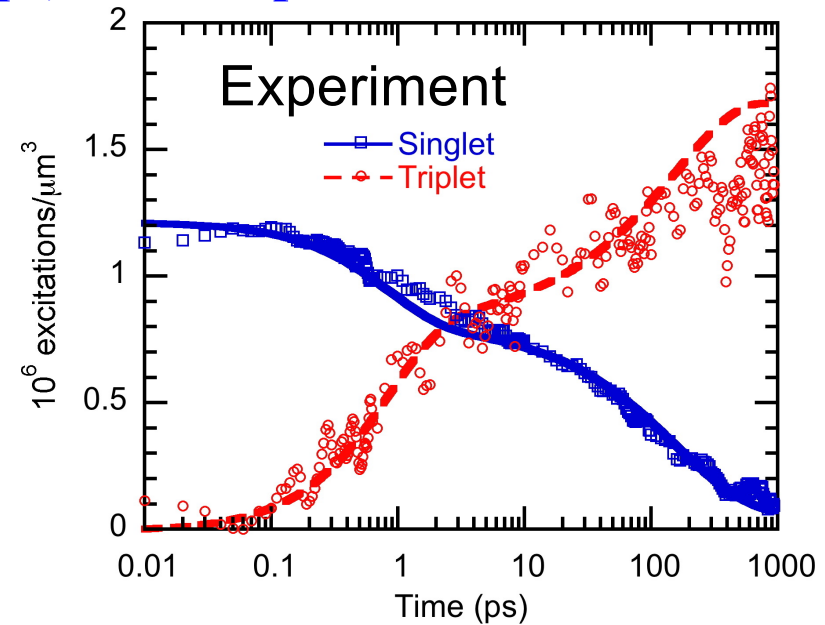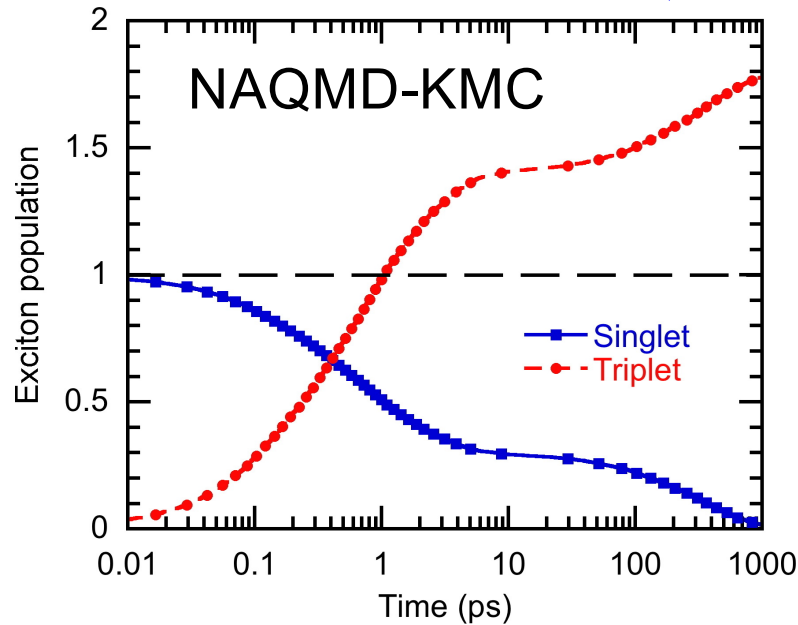
W. Mou *et al.*,
APL **102**, 173301 ('13)

S. T. Roberts *et al.*,
JACS **134**, 6388 ('12)

Cover image
Appl. Phys. Lett. (Apr. 29, '13)