

Transition Interface Sampling and Forward Flux Sampling



Sapna Sarupria

Chemical and
Biomolecular Engineering

We have postdoc openings
starting as soon as possible



Path Sampling Methods

Transition Path
Sampling



Transition
Interface
Sampling

Forward Flux
Sampling



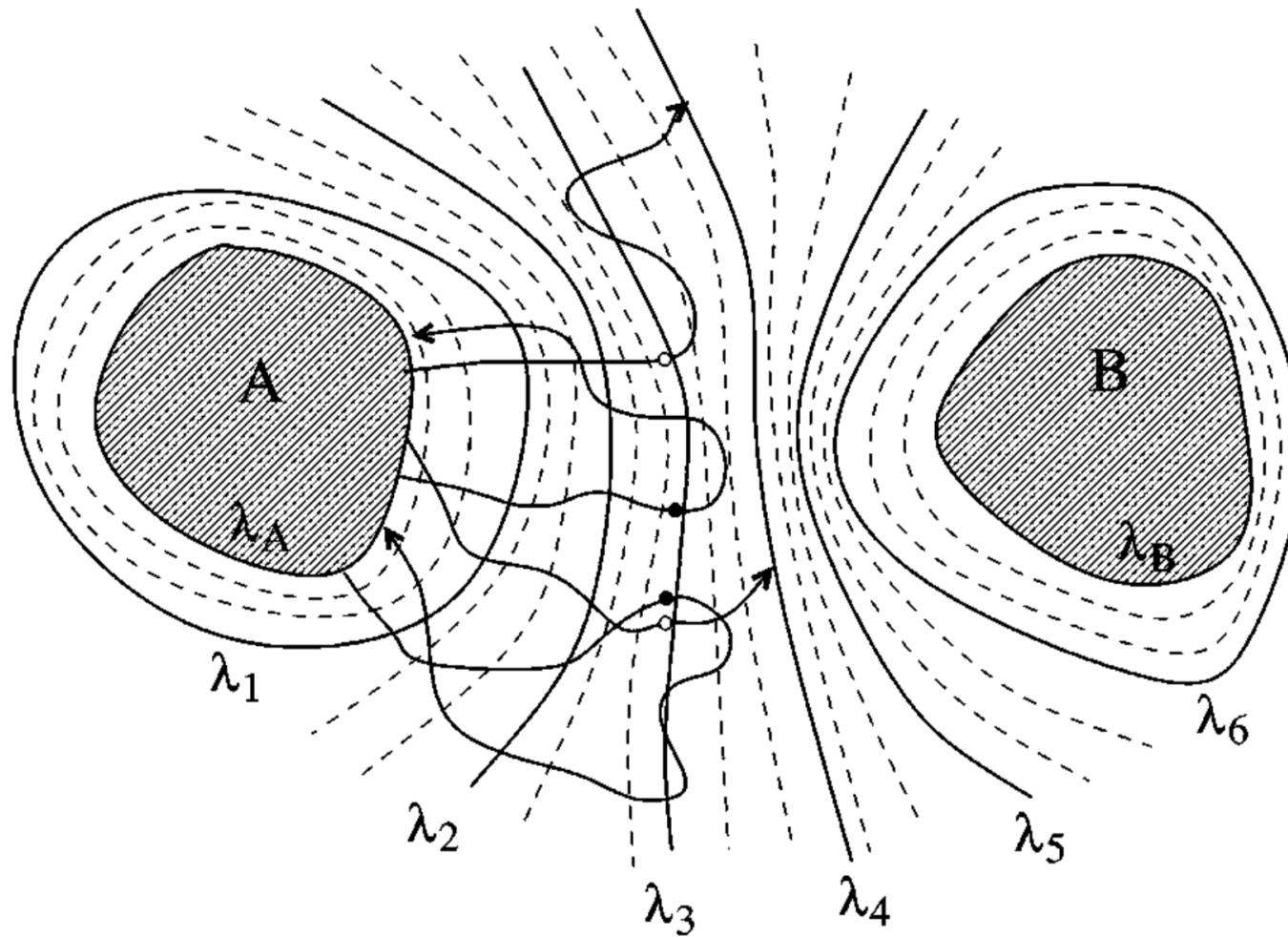
Transition Interface Sampling

A novel path sampling method for the calculation of rate constants
Titus S. van Erp, Daniele Moroni, and Peter G. Bolhuis
Citation: The Journal of Chemical Physics 118, 7762 (2003);
doi: 10.1063/1.1562614

Elaborating transition interface sampling methods
Titus S. van Erp and Peter G. Bolhuis
Journal of Computational Physics 205 (2005) 157–181

Trajectory-Based Rare Event Simulations
Peter G. Bolhuis and Christoph Dellago
Reviews in Computational Chemistry, Volume 27, Pages 111-207 (2011)

Transition Interface Sampling (TIS): Overview



TIS

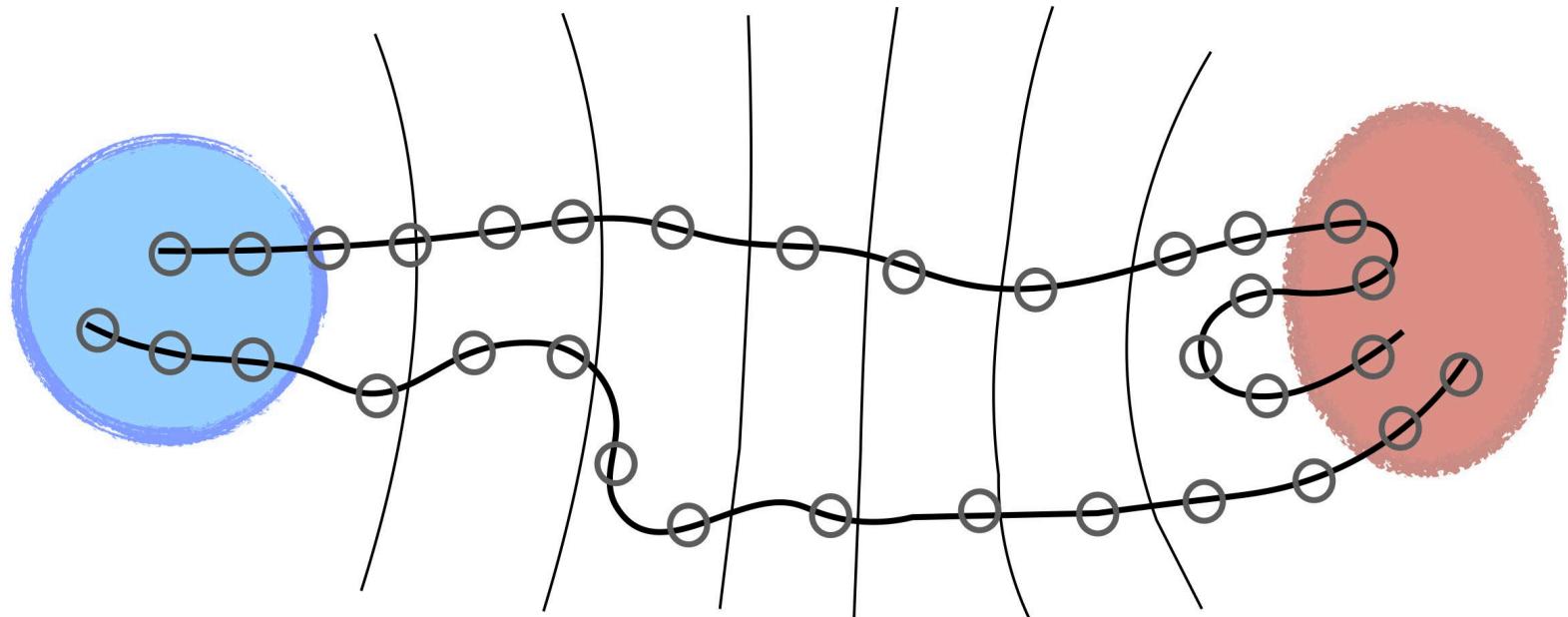
- How do we formulate the rate?
- How do we perform the sampling?

TIS: Rate Calculation

- Correlation function: $C(t) = \frac{\langle h_A(0)h_B(t) \rangle}{\langle h_A(0) \rangle}$
- Definitions of overall states \mathcal{A} and \mathcal{B} :

A phase point belongs to \mathcal{A} if it comes directly from A without having visited B

A phase point belongs to \mathcal{B} if it comes directly from B without having visited A



$$C(t) = \frac{\langle h_{\mathcal{A}}(0)h_{\mathcal{B}}(t) \rangle}{\langle h_{\mathcal{A}}(0) \rangle}$$

This definition maps the entire phase region.
The boundary is irregular.

TIS: Rate Calculation

$t_i^b(x_0)$ = The time it takes the backward path to reach λ_i from x_0

$t_i^f(x_0)$ = The time it takes the forward path to reach λ_i from x_0

These mark the **first crossing** with interface i on forward (backward) path

Indicator theta functions:

$$h_{i,j}^b = \theta(t_j^b(x) - t_i^b(x))$$

These measure whether backward and forward trajectories reach interface j or i.

$$h_{i,j}^f = \theta(t_j^f(x) - t_i^f(x))$$

$$h_{\mathcal{A}}(x) = h_{0,n}^b(x) = \theta(t_n^b(x) - t_0^b(x))$$

A phase point belongs to \mathcal{A} if it comes directly from A without having visited B

$$h_{\mathcal{B}}(x) = h_{n,0}^b(x) = \theta(t_0^b(x) - t_n^b(x))$$

A phase point belongs to \mathcal{B} if it comes directly from B without having visited A

TIS: Rate Calculation

$$C(t) = \frac{\langle h_{\mathcal{A}}(x_0) h_{\mathcal{B}}(x_t) \rangle}{\langle h_{\mathcal{A}}(x_0) \rangle}$$

Taking the derivative of this term at t=0

$$k_{AB} = \frac{\left\langle h_{0,n}^b(x_0) \dot{\lambda}(x_0) \delta(\lambda(x_0) - \lambda_n) \right\rangle}{\langle h_{\mathcal{A}}(x_0) \rangle}$$

Only the positive terms contribute to the rate.

$$h_{\mathcal{A}}(x) = h_{0,n}^b(x) = \theta(t_n^b(x) - t_0^b(x))$$

A phase point belongs to \mathcal{A} if it comes directly from A without having visited B

TIS: Rate Calculation

$$k_{AB} = \frac{\left\langle h_{0,n}^b(x_0) \dot{\lambda}(x_0) \delta(\lambda(x_0) - \lambda_n) \right\rangle}{\langle h_A(x_0) \rangle}$$

Numerator; in general terms;

$$\lambda_i < \lambda_j < \lambda_k$$

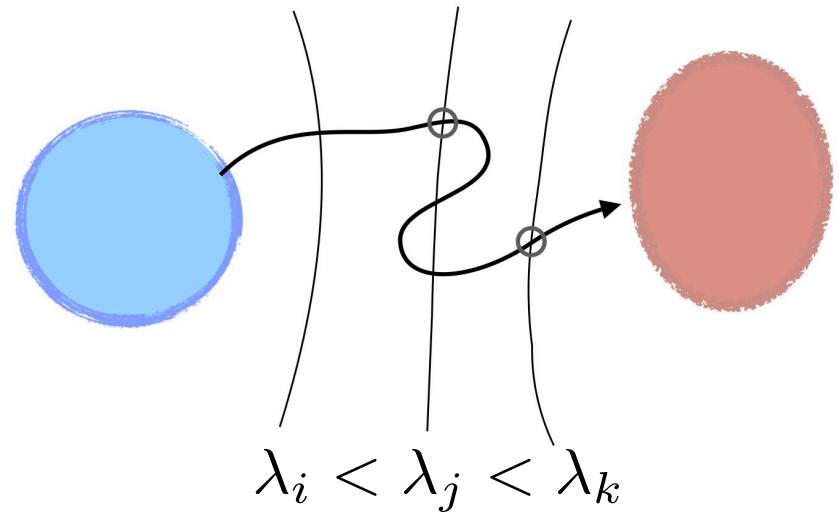
$$\left\langle h_{i,k}^b \dot{\lambda} \delta(\lambda(x) - \lambda_k) \right\rangle = \left\langle h_{i,j}^b \dot{\lambda} \delta(\lambda(x) - \lambda_j) h_{k,i}^f \right\rangle$$

flux through k

flux through j

going to k before i

The flux through k is equal to the flux through j times probability that the trajectory goes through k before going back to i for trajectories coming directly from i



TIS: Rate Calculation

Flux function

$$\phi_{j,i}(x) = h_{i,j}^b(x) |\dot{\lambda}(x)| \delta(\lambda(x) - \lambda_j) = h_{i,j}^b(x) \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \theta(\Delta t - t_j^f(x))$$

Flux from λ_i to λ_j

for discrete trajectories

Re-writing rate equation

$$k_{AB} = \frac{\langle \phi_{n,0} \rangle}{\langle h_{\mathcal{A}} \rangle}$$

From previous slide

$$\lambda_i < \lambda_j < \lambda_k$$

$$\left\langle h_{i,k}^b \dot{\lambda} \delta(\lambda(x) - \lambda_k) \right\rangle = \left\langle h_{i,j}^b \dot{\lambda} \delta(\lambda(x) - \lambda_j) h_{k,i}^f \right\rangle$$

$$\langle \phi_{ki}(x) \rangle = \left\langle \phi_{ji}(x) h_{k,i}^f(x) \right\rangle$$

$$k_{AB} = \langle \phi_{n,0} \rangle / \langle h_{\mathcal{A}} \rangle = \left\langle \phi_{j,0}(x) h_{n,0}^f(x) \right\rangle / \langle h_{\mathcal{A}} \rangle$$

TIS: Rate Calculation

Defining a conditional probability

$$P_A (\lambda_j | \lambda_i) \equiv \left\langle \phi_{i,0}(x) h_{j,0}^f(x) \right\rangle / \langle \phi_{i,0}(x) \rangle$$

This is the probability for the system to reach interface j before interface 0 under the condition that it crosses interface i at t=0, while coming directly from interface 0 (state A) in the past.

$$\langle \phi_{n,0} \rangle = \langle \phi_{1,0} \rangle P_A (\lambda_n | \lambda_1)$$

Relates the flux through boundary of B to the flux through an interface much closer to A

TIS: Rate Calculation

$$\lambda_i < \lambda_j < \lambda_k$$

$$P_A(\lambda_k|\lambda_i) = P_A(\lambda_k|\lambda_j) P_A(\lambda_j|\lambda_i)$$

TIS: Rate Calculation

$$P_A (\lambda_k | \lambda_i) = P_A (\lambda_k | \lambda_j) P_A (\lambda_j | \lambda_i)$$

$$P_A (\lambda_n | \lambda_1) = \prod_{i=1}^{\lambda_{n-1}} P_A (\lambda_{i+1} | \lambda_i)$$

$$k_{AB} = \frac{\langle \phi_{1,0} \rangle}{\langle h_{\mathcal{A}} \rangle} P_A (\lambda_n | \lambda_1)$$

$$P_A (\lambda_n | \lambda_1) = \prod_{i=1}^{n-1} P_A (\lambda_{i+1} | \lambda_i)$$

Effective flux through λ_1 of the trajectories coming from $\lambda_0 ==$
count positive crossings

TIS: Algorithm

- This algorithm is used to sample each interface ensemble. We effectively perform transition path sampling for each interface ensemble.
- For each interface ensemble:
 - Start with an initial path that begins in A, crosses λ_i at least once, and ends in A or $\lambda > \lambda_{i+1}$.
 - If a move is rejected at any step the old path is re-counted in the interface ensemble and we return to step 1.

TIS: Algorithm

1. Select a random time slice of the current path, $x^{(o)}$
2. Change all momenta by adding displacements δp drawn from a Gaussian
3. NVE: rescale new momenta to old energy. NVT: accept the new momenta with probability $\min[1, \exp(-\beta(E(x^{(n)}) - E(x^{(o)})))]$. $E(x^{(n)})$ and $E(x^{(o)})$ are the total energy of the system at point x with the new and old momenta, respectively.
4. Integrate equations of motion backwards in time until reaching A or $\lambda > \lambda_{i+1}$. Reject if reaches $\lambda > \lambda_{i+1}$.
5. Integrate equations of motion forwards in time until reaching A or $\lambda > \lambda_{i+1}$. Reject if entire new path does not cross λ_{i+1}
6. Replace the current path with the new path with probability $\min\left(1, \frac{N^{(o)}}{N^{(n)}}\right)$, where N is the length of the path.

Features of TIS:

- Equilibrium conditions
- Takes path length into consideration
- No *a priori* knowledge of "reaction coordinate" is needed
- Relatively easy rate calculations
- Algorithm's parallelizability

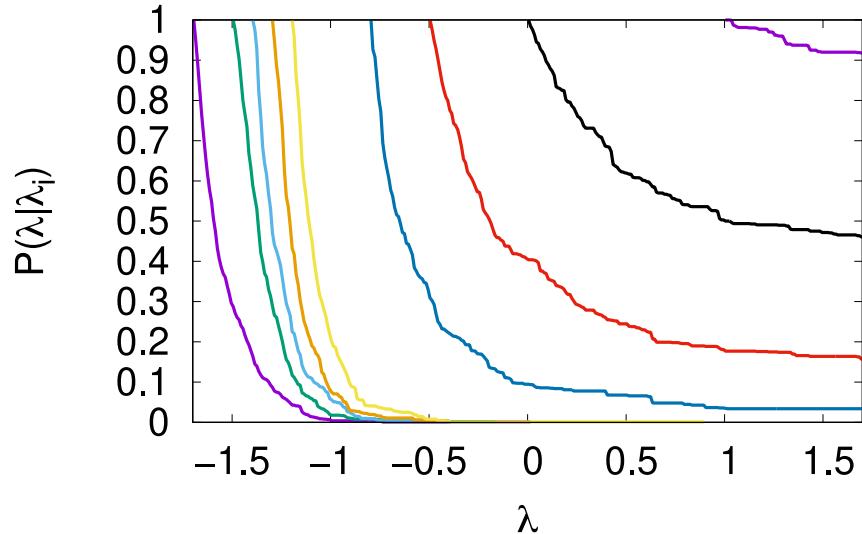
TIS Rate Calculation

Once we have a collection of paths for each interface ensemble, we can calculate $P(\lambda|\lambda_i)$ for each interface ensemble and $P(\lambda_{i+1}|\lambda_i)$.

For each interface ensemble:

1. Create a histogram, $H(\lambda|\lambda_i)$ with some bin width $\ll (\lambda_{i+1} - \lambda_i)$.
2. For each path in a given interface ensemble:
 - Calculate the max value of λ along the path, λ^{max} .
 - Increment all bins of the histogram with $\lambda < \lambda^{max}$.
3. Calculate $P(\lambda|\lambda_i)$ by dividing $H(\lambda|\lambda_i) / H(\lambda_i|\lambda_i)$. $H(\lambda_i|\lambda_i)$ should be equal to the total number of paths in the interface ensemble.

Interface 0 — Interface 3 — Interface 6 —
Interface 1 — Interface 4 — Interface 7 —
Interface 2 — Interface 5 — Interface 8 —



van Erp, T.S.; Moroni, D.; Bolhuis, P.G. *J. Chem. Phys.* 118(17), 2003. pp. 7762-7774

TIS Rate Calculation

The rate constant is calculated from the flux from A to λ_0 , $\Phi_{A,0}$ and $P(\lambda_B|\lambda_0)$.

$$k_{AB} = \Phi_{A,0} P(\lambda_B|\lambda_0)$$

$\Phi_{A,0}$ is calculated by running a straightforward simulation in A and counting the number of forward crossings of λ_0 .

$$\Phi_{A,0} = \frac{N_{cross}}{t_{total}}$$

$P(\lambda_B|\lambda_0)$ is estimated from the interface ensembles. Either as

$$P(\lambda_B|\lambda_0) = \prod_{i=0}^{n-1} P(\lambda_{i+1}|\lambda_i)$$

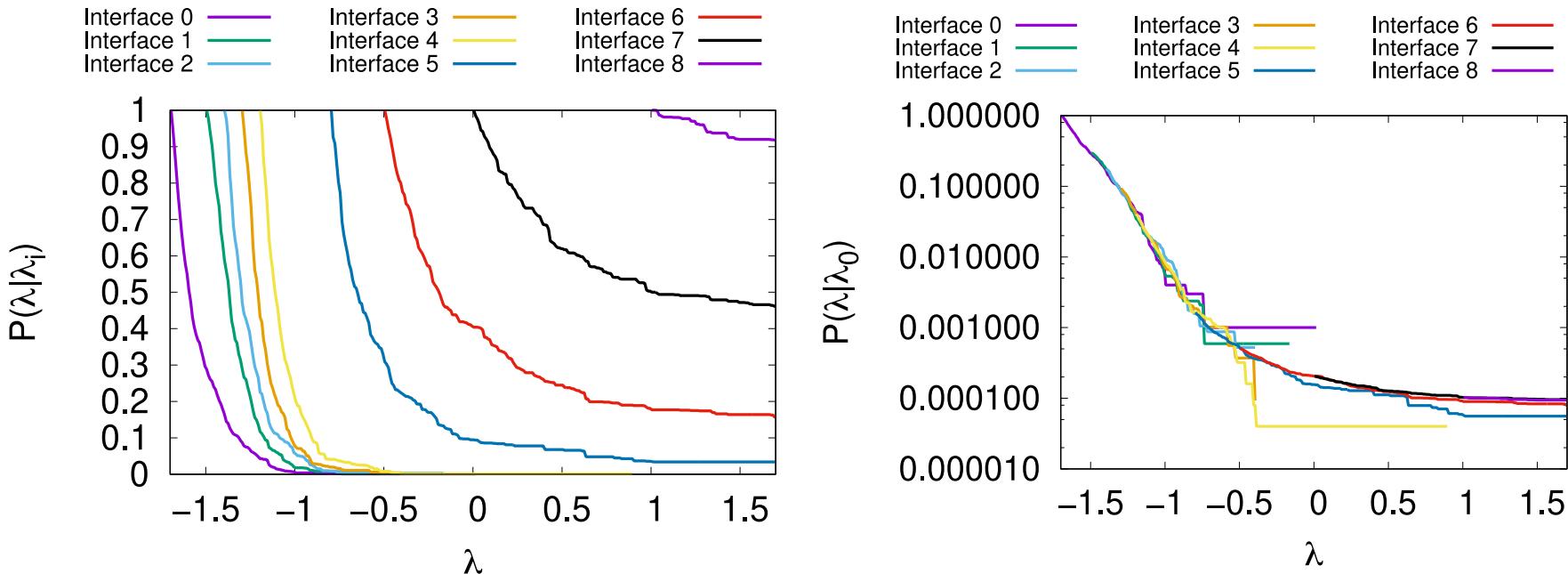
or alternatively by stitching together the $P(\lambda|\lambda_i)$ profiles from each interface ensemble and reading the value of $P(\lambda_B|\lambda_0)$.

TIS Rate Calculation

All the $P(\lambda|\lambda_i)$ can be stitched together to create $P(\lambda|\lambda_0)$.

Either:

- Multiply each $P(\lambda|\lambda_i)$ with $P(\lambda_i|\lambda_0)$, or
- Stitch all $P(\lambda|\lambda_i)$ together with the weighted histogram analysis method (WHAM). The final $P(\lambda_B|\lambda_0)$ can be extracted from the connected profile. This approach reduces the error in $P(\lambda_B|\lambda_0)$.



van Erp, T.S.; Moroni, D.; Bolhuis, P.G. *J. Chem. Phys.* 118(17), 2003. pp. 7762-7774

Constrained forward shooting TIS algorithm

Requires stochastic dynamics (for trajectory divergence) but has a simpler algorithm and 100% acceptance ratio for shooting moves.

Start with an initial path that begins in A, crosses λ_i at least once, and ends in A or B (notice we terminate in A or B rather than A or $\lambda > \lambda_{i+1}$).

Bolhuis, P.G. *J. Chem. Phys.* **129**, 2008. pp. 114108

Constrained forward shooting TIS algorithm

For each move select between a shooting move or time reversal move with 50% probability assigned to each.

Shooting move:

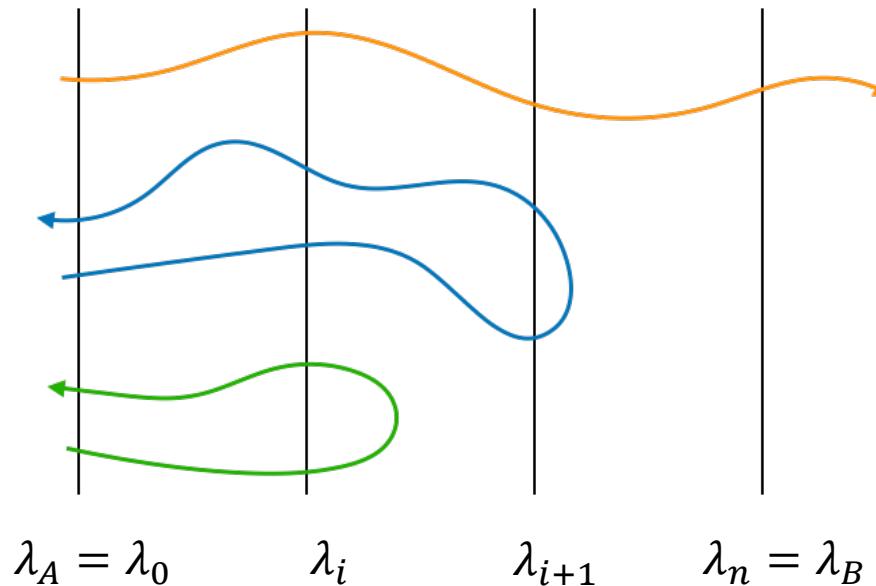
1. Select the first crossing of λ_i as the shooting point
2. Integrate equations of motion forwards until reaching A or B.
3. Connect new path section with old path section from A to λ_i .

Time reversal move:

1. Reverse the existing path.
2. Reject move if reversed path ends in A. Accept otherwise.

Replica Exchange TIS (Parallel path swapping) (RETIS)

Define interface ensembles $[i^+]$ as the ensemble of paths that begins in A, crosses λ_i at least once, and ends in A or B. Thus any path in $[(i + 1)^+]$ ensemble is a valid path in the $[i^+]$ ensemble.



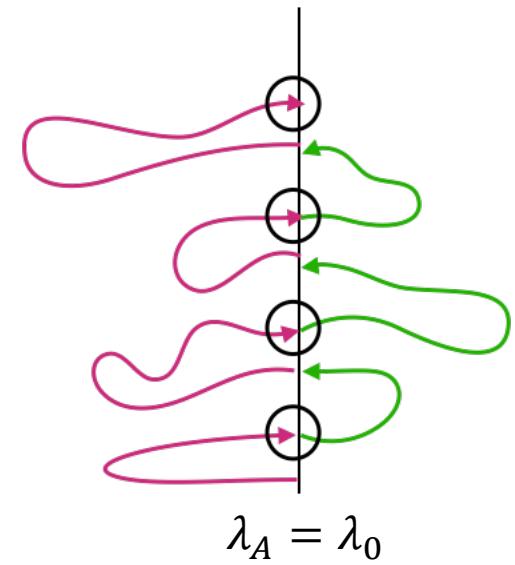
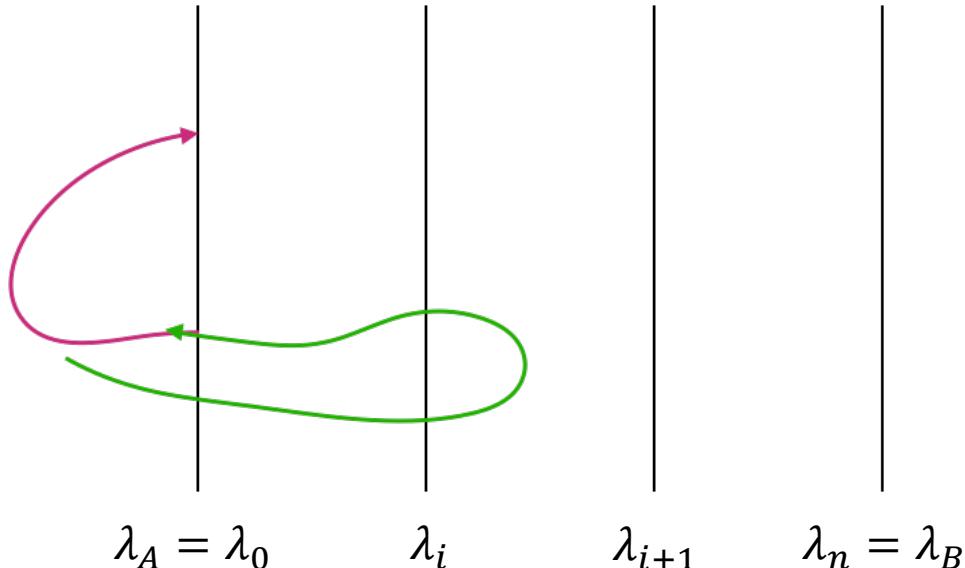
Green → valid path in $[i^+]$ ensemble

Blue → valid path in $[i^+]$ ensemble, in $[(i + 1)^+]$ ensemble

Orange → valid path in all path $[i^+]$ ensemble for all $i = 0 \dots n$

Replica Exchange TIS (RETIS)

The flux, $\Phi_{A,0}$ is going to be calculated with a new $[0^-]$ ensemble. These paths start at the end of a $[0^+]$ path, and end when $\lambda > \lambda_0$. These are generated as part of the $[0^-]$ move.



One valid first crossing of λ_0 for each $[0^-]$, $[0^+]$ sequence

$$\Phi_{A,0} = \frac{1}{\langle t_{\text{path}}^{[0^-]} \rangle + \langle t_{\text{path}}^{[0^+]} \rangle}$$

Green → valid path in $[0^+]$ ensemble
Magenta → valid path in $[0^-]$ ensemble

Replica Exchange TIS (RETIS)

Each interface ensemble will be run in parallel. Start with valid paths in each interface ensemble.

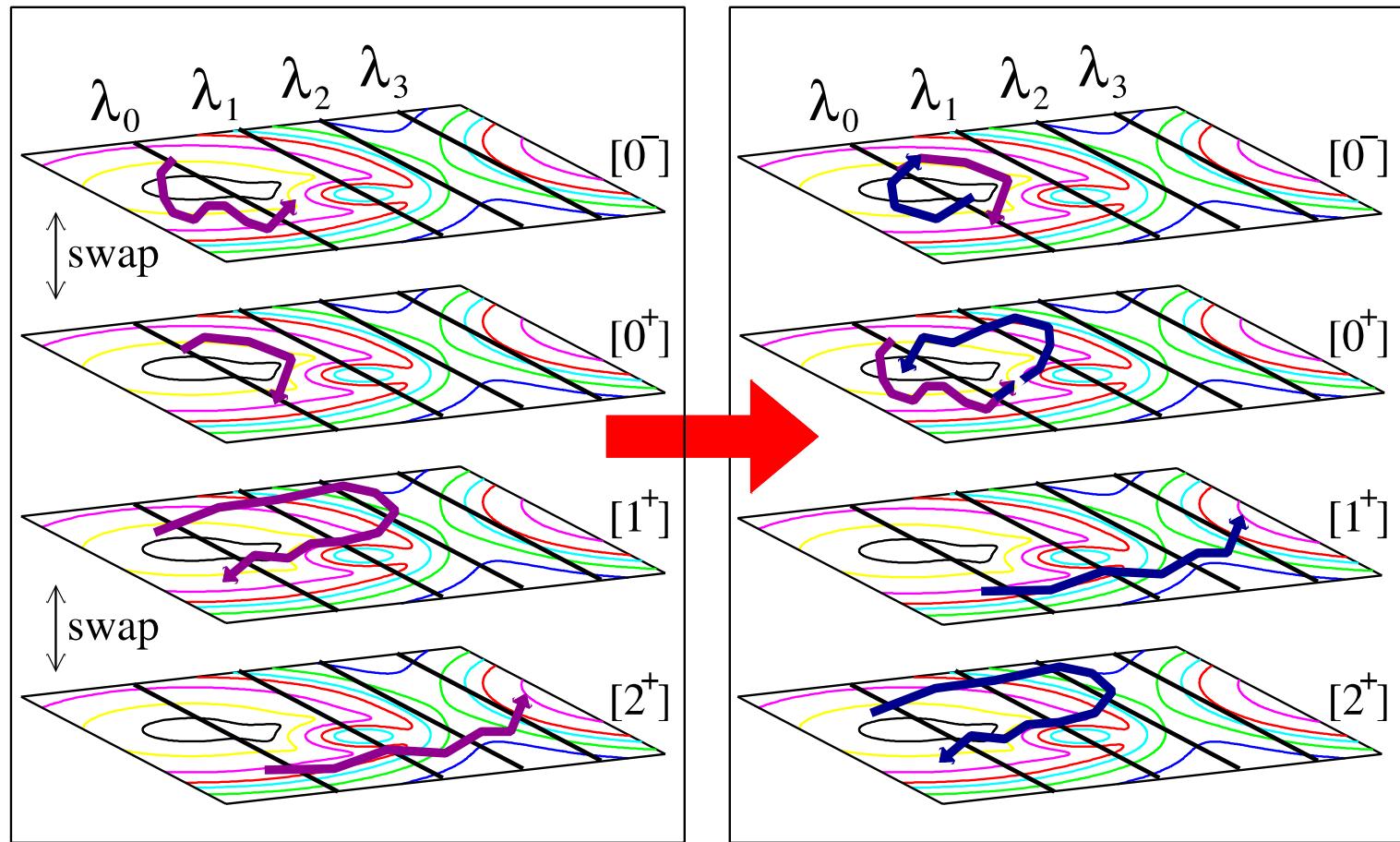
For each ‘global move’, choose a swapping move or a update move with 50% probability assigned to each.

If swapping:

1. Select swaps $[0^-] \leftrightarrow [0^+]$, $[1^+] \leftrightarrow [2^+]$, $[3^+] \leftrightarrow [4^+]$, ... or $[0^+] \leftrightarrow [1^+]$, $[2^+] \leftrightarrow [3^+]$, $[4^+] \leftrightarrow [5^+]$ with 50% probability assigned to each of the two options.
2. Attempt each swap.
 - a) Accept if path from $[i^+]$ is valid path in $[(i + 1)^+]$.
 - b) Reject otherwise and re-count paths in original ensembles.
3. The $[0^-] \leftrightarrow [0^+]$ move is always accepted. Starting from the end of the path in the $[0^-]$ ensemble integrate EOMs to generate new $[0^+]$ path. Starting from the end of the path in the $[0^+]$ integrate EOMs generate new $[0^-]$ path.

Replica Exchange TIS (RETIS)

Examples of swapping moves:



Replica Exchange TIS (RETIS)

If update:

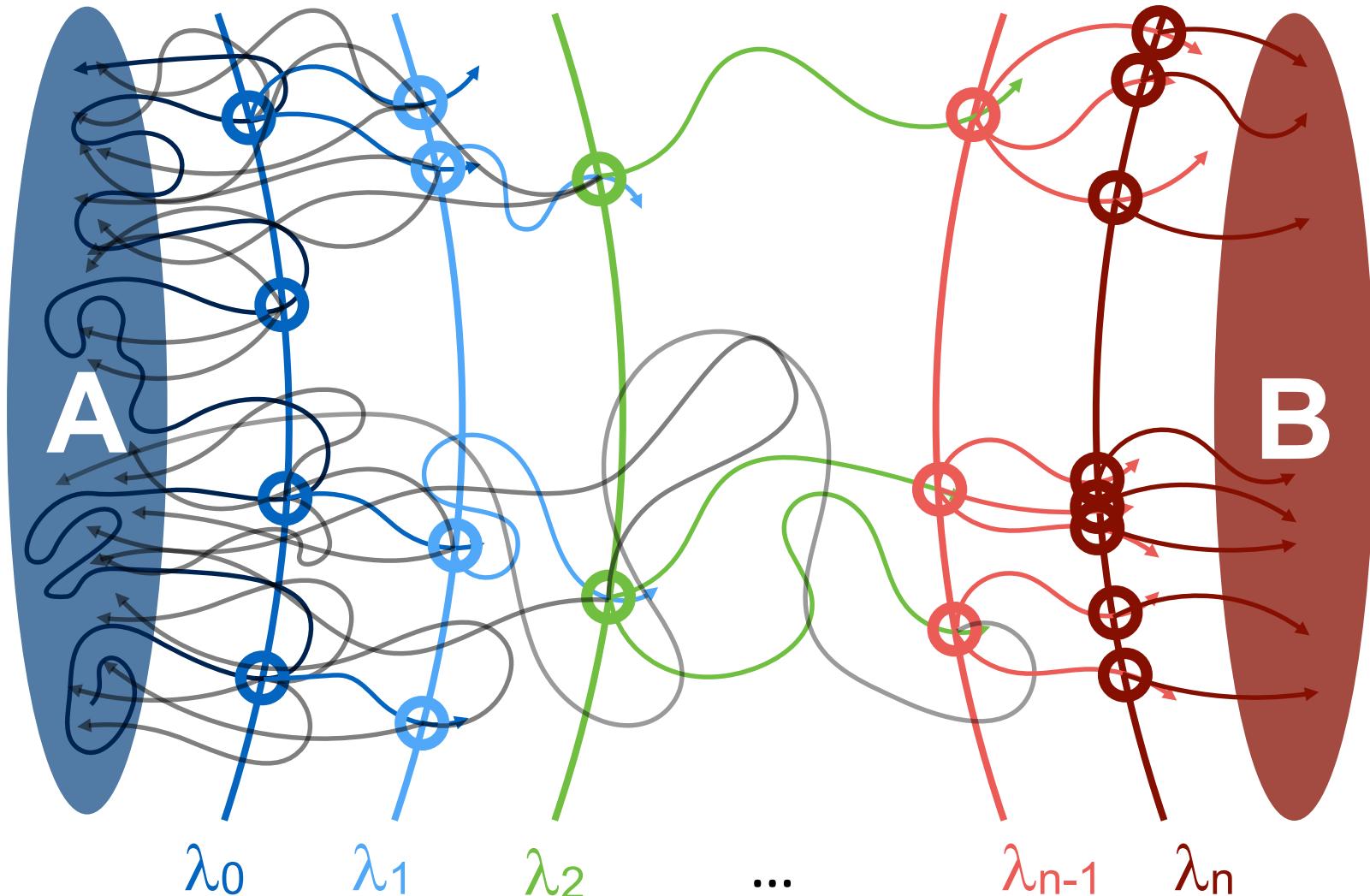
Independently attempt a update move in each interface ensemble.

1. Select time reversal move or shooting move with 50% probability assigned to each.
2. If time reversal:
 - a) Attempt time reversal.
 - b) Accept if reversed path ends in A.
 - c) Reject otherwise and recount original path
3. If shooting:
 - a) Attempt shooting move
 - b) Accept if new path is a valid member of the relevant interface ensemble
 - c) Reject otherwise and recount original path

Final rate calculation is once again: $k_{AB} = \Phi_{A,0} P(\lambda_B | \lambda_0)$



Forward flux sampling



Valeriani, C.; Allen, R. J.; Morelli, M. J.;
Frenkel, D.; Rein ten Wolde, J. Chem.
Phys. 2007, 127 (11), 114109.

$$k_{AB} = \Phi_{A\lambda_0} \prod_{i=0}^n P(\lambda_{i+1} | \lambda_i)$$

Forward flux sampling

- Equilibrium and non-equilibrium
- “Easy implementation”
- Quite popular

References and uses of FFS

Simulating rare events in equilibrium or nonequilibrium stochastic systems
Rosalind J. Allen, Daan Frenkel, and Pieter Rein ten Wolde
The Journal of Chemical Physics 124, 024102 (2006); doi: 10.1063/1.2140273

Computing stationary distributions in equilibrium and nonequilibrium systems with forward flux sampling
Chantal Valeriani, Rosalind J. Allen, Marco J. Morelli, Daan Frenkel, and Pieter Rein ten Wolde
The Journal of Chemical Physics 127, 114109 (2007); doi: 10.1063/1.2767625

Review: Forward flux sampling for rare event simulations
Rosalind J Allen, Chantal Valeriani and Pieter Rein ten Wolde
J. Phys.: Condens. Matter 21 (2009) 463102 (21pp)

References and uses of FFS

Reaction coordinates and transition pathways of rare events via forward flux sampling

Ernesto E. Borrero, and **Fernando A. Escobedo**

The Journal of Chemical Physics 127, 164101 (2007); doi: 10.1063/1.2776270

Rate constants for proteins binding to substrates with multiple binding sites using a generalized forward flux sampling expression

Adithya Vijaykumar, Pieter Rein ten Wolde, and Peter G. Bolhuis

The Journal Of Chemical Physics 148, 124109 (2018)

Probing Methane Hydrate Nucleation through the Forward Flux Sampling Method

Yuanfei Bi and Tianshu Li

dx.doi.org/10.1021/jp503000u | J. Phys. Chem. B 2014, 118, 13324 – 13332

Forward-flux sampling with jumpy order parameters

Amir Haji-Akbari

J. Chem. Phys. 149, 072303 (2018); doi: 10.1063/1.5018303

Quoted from Direct calculation of ice nucleation rate

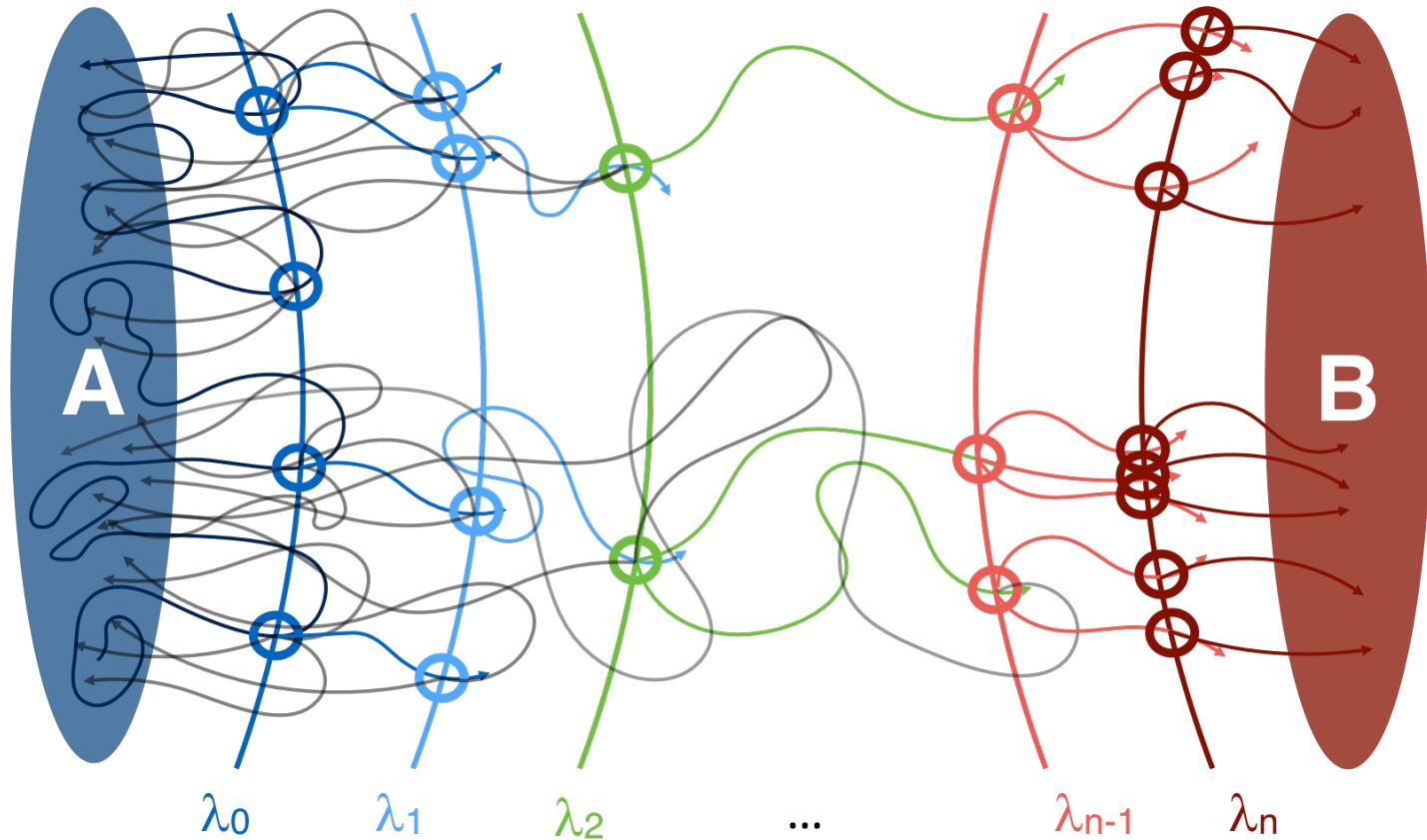
Amir Haji-Akbari, Pablo G. Debenedetti

Proceedings of the National Academy of Sciences Aug

2015, 112 (34) 10582-10588; DOI:10.1073/pnas.1509267112

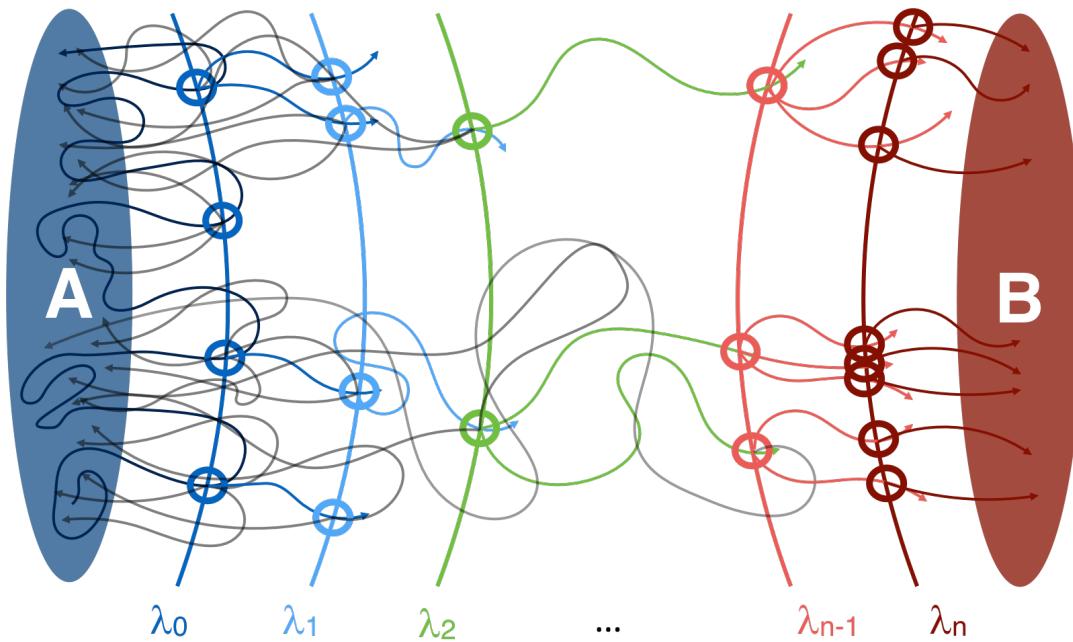
- As a result, we carried out a total of $608 \mu\text{s}$ of MD trajectories. This amounts to a total of **21,452,433 central processing unit (CPU) hours** on the Texas-based Stampede supercomputer. Due to the embarrassingly parallel nature of the FFS algorithm, we were able to distribute this very costly calculation across the following supercomputers: the Princeton-based Della and Tiger supercomputers, the Texas Advanced Computing Center-based Stampede supercomputer, the San Diego Supercomputer Center-based Gordon supercomputer, and the Rensselaer Polytechnic Institute-based Blue Gene/Q supercomputer

FFS: Implementation Challenge



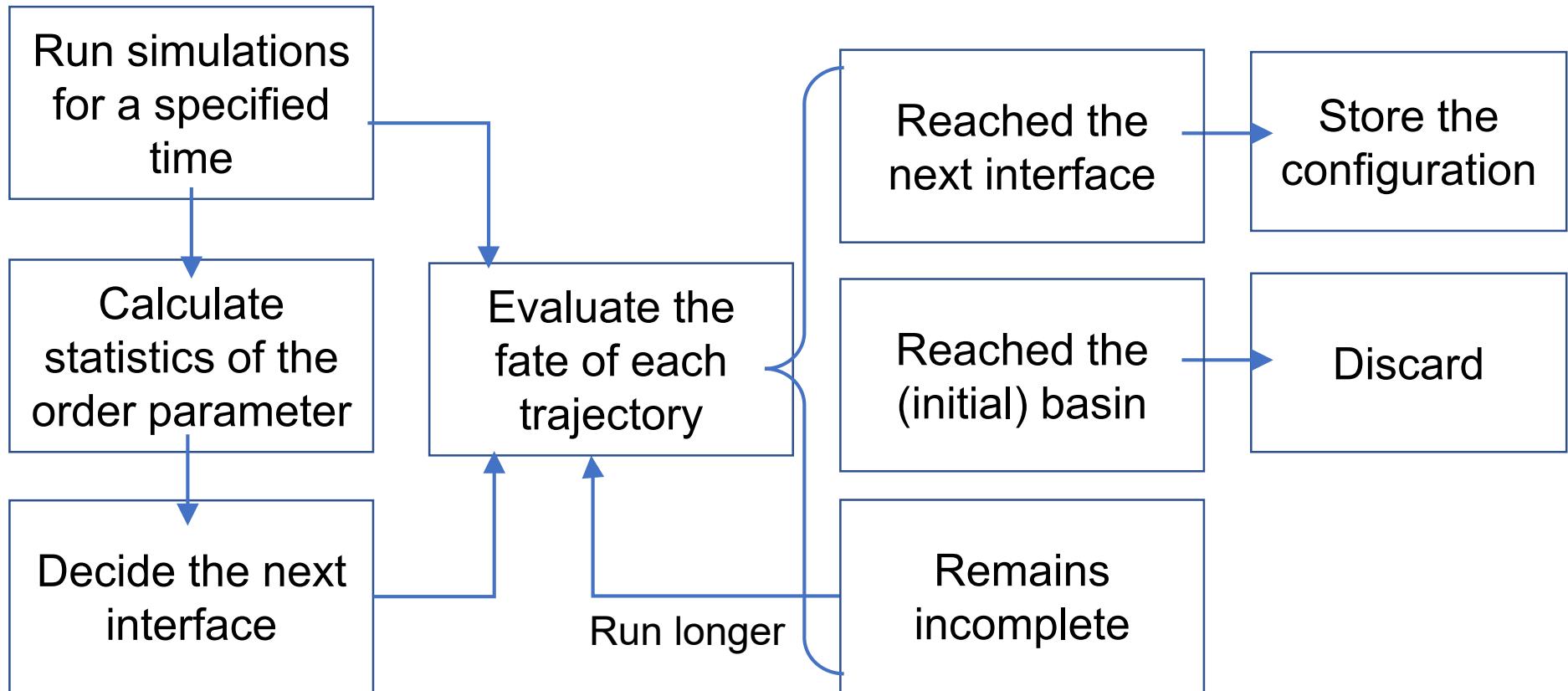
- Modification of the source code
- Maintaining a lot of data
- Very large number of files
- Debugging is a nightmare

FFS: Implementation Challenge



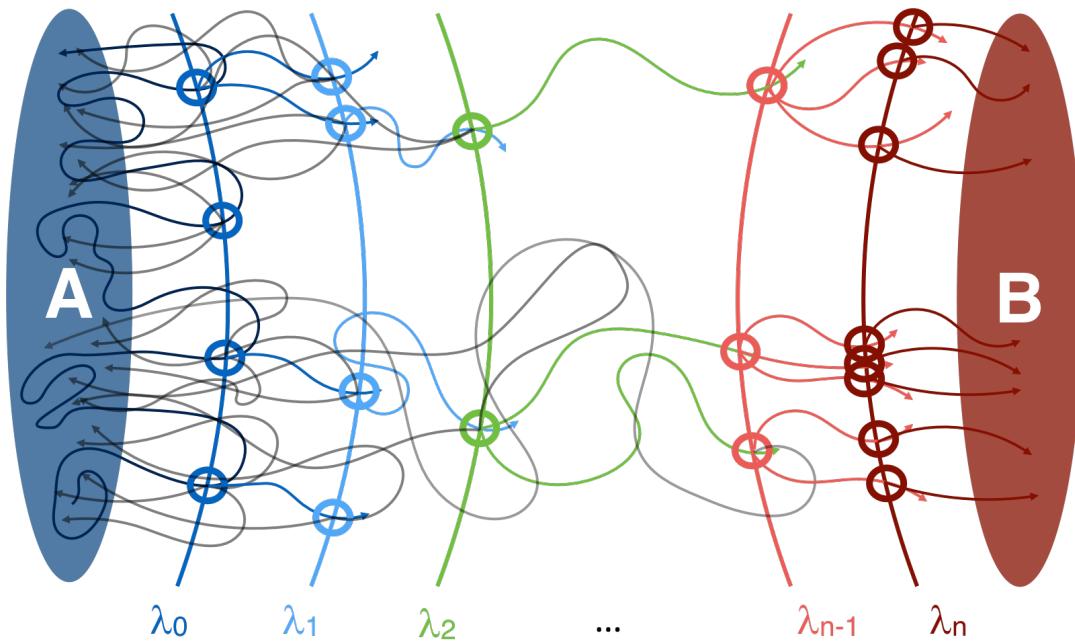
- Need to be able to decide interfaces on-the-fly
- Need to avoid any modification to the source codes
- Need to keep track of the status of each job
- Need to keep track of all the files
- Need to be able to distribute these jobs efficiently

Implementation of FFS Sarupria automated version!



- Need to calculate the next interface on-the-fly
- Specify the length of the each simulation and decide its fate

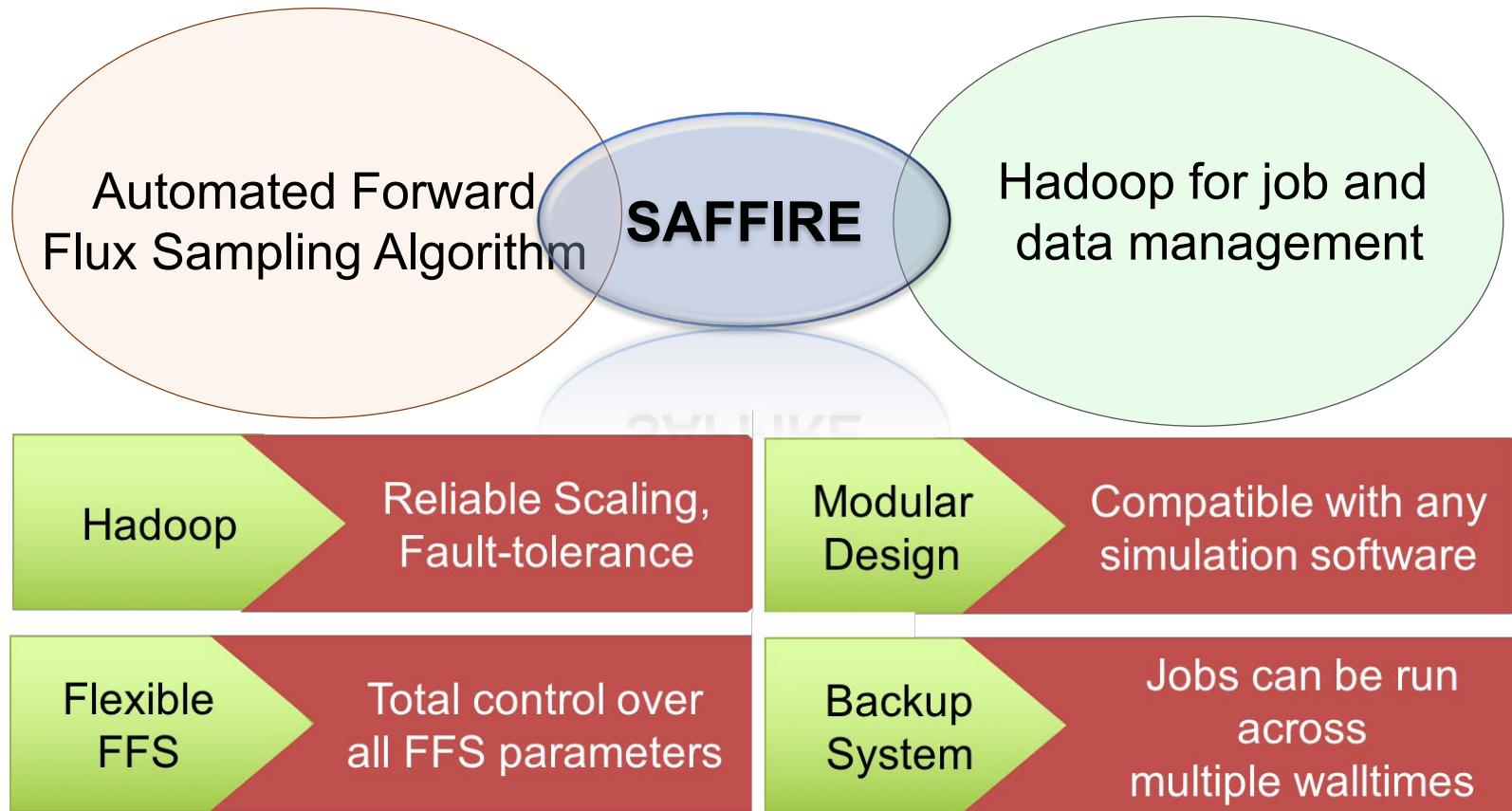
FFS: Implementation Challenge



- Need to be able to decide interfaces on-the-fly
- Need to avoid any modification to the source codes
- Need to keep track of the status of each job
- Need to keep track of all the files
- Need to be able to distribute these jobs efficiently

Rare events: SAFFIRE

Sarupria group software for FFS



Walter Judge



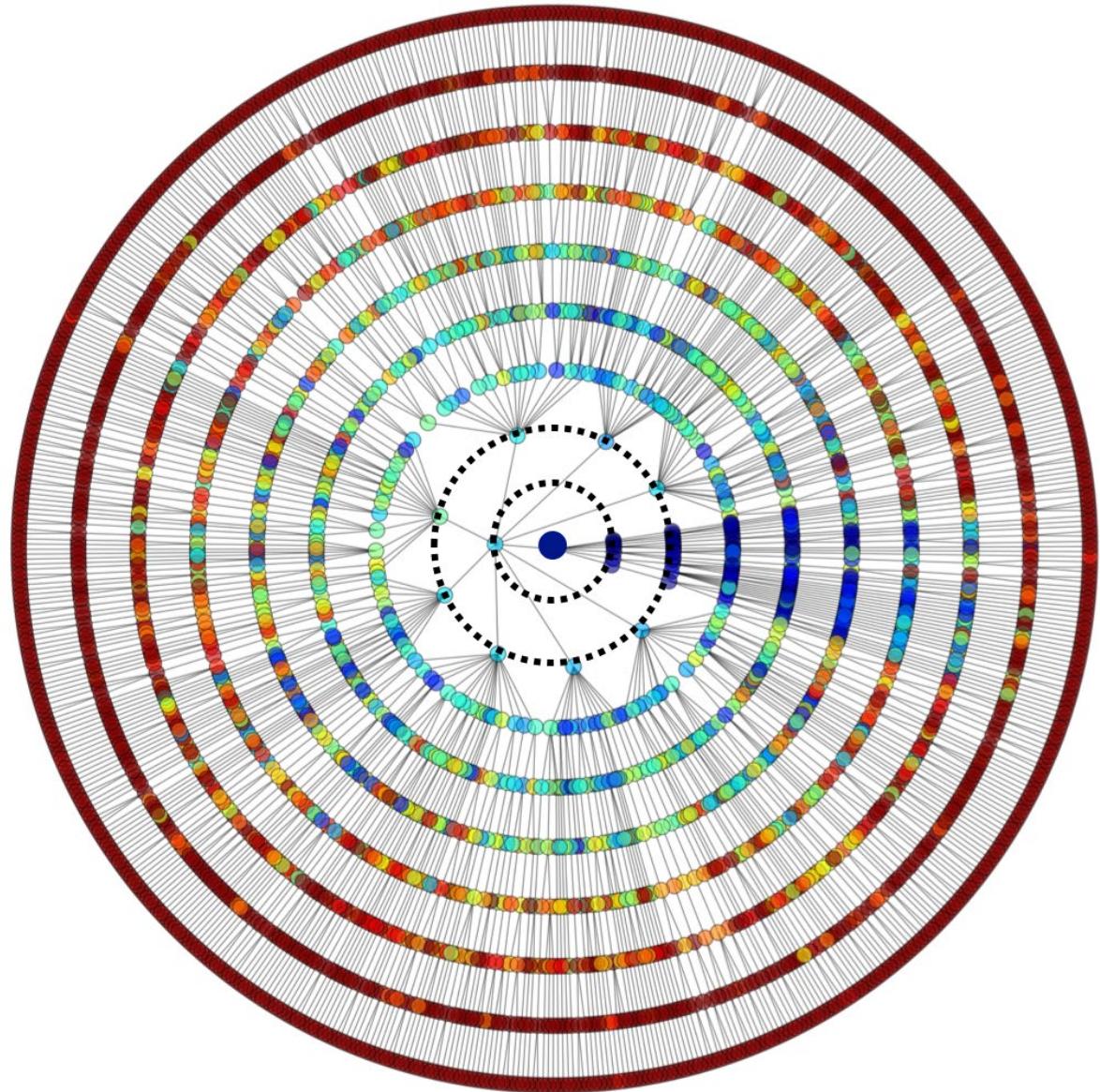
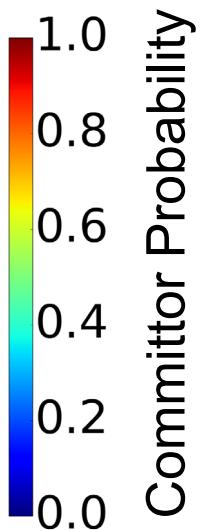
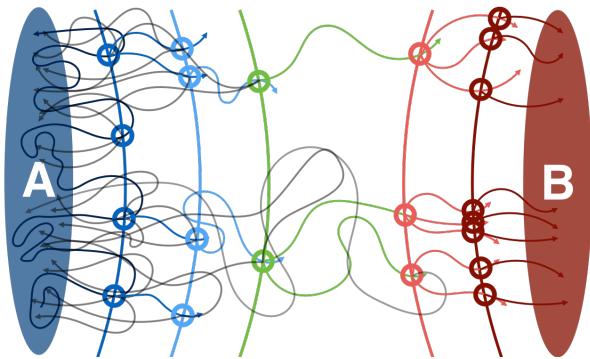
Ryan DeFever



Amy Apon



Linh Ngo



We performed large scale FFS consisting of over 250,000 individual MD simulations and requiring over 300,000 CPU-hours

FFS: How do you “really” perform FFS?

- Basin simulations
- Flux through interfaces
- Probability convergence



Gas Hydrates/Clathrate Hydrates

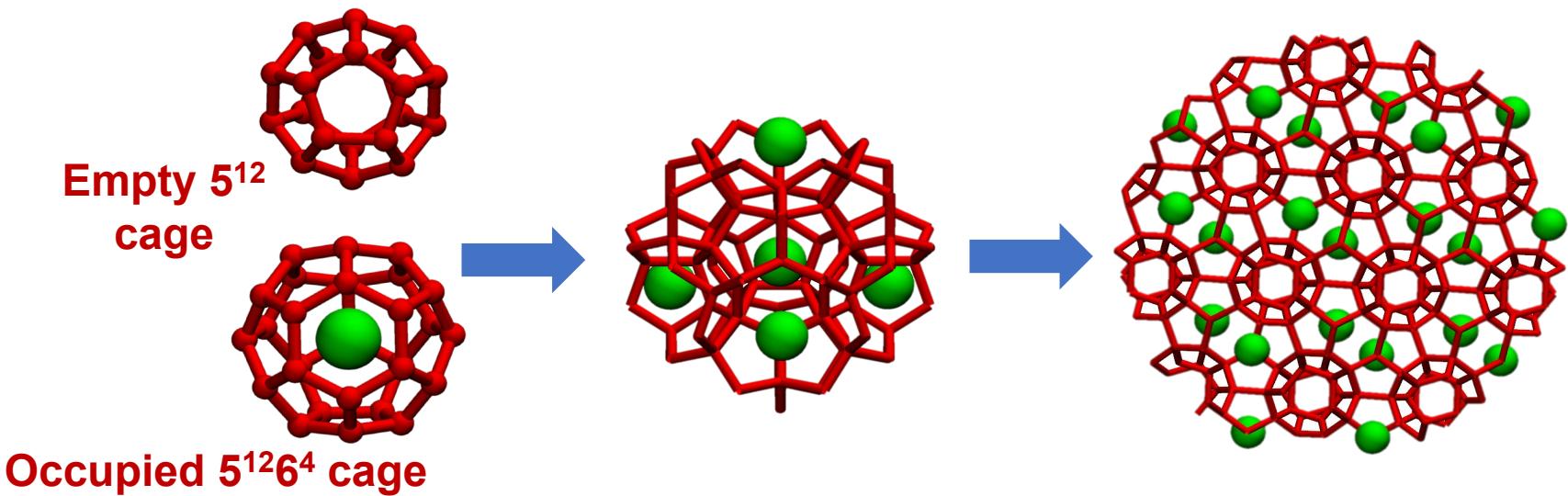
Hydrate Plug



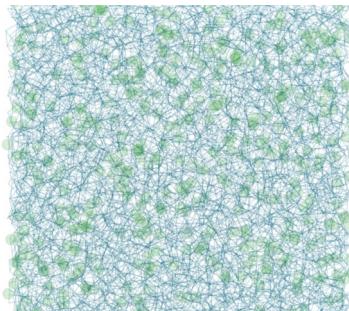
[http://www.hydrafact.com/
Gas_Hydrates.html](http://www.hydrafact.com/Gas_Hydrates.html)



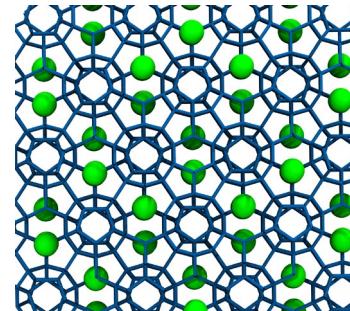
Methane hydrate burning
<http://www.fossil.energy.gov/programs/oilgas/hydrates/>



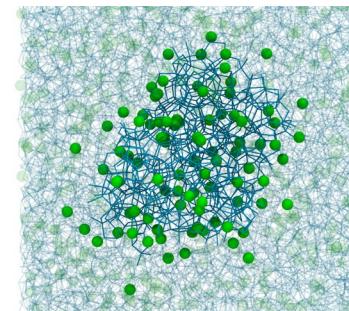
Gas Hydrates/Clathrate Hydrates



Supercooled liquid



Structure II (sII) hydrate



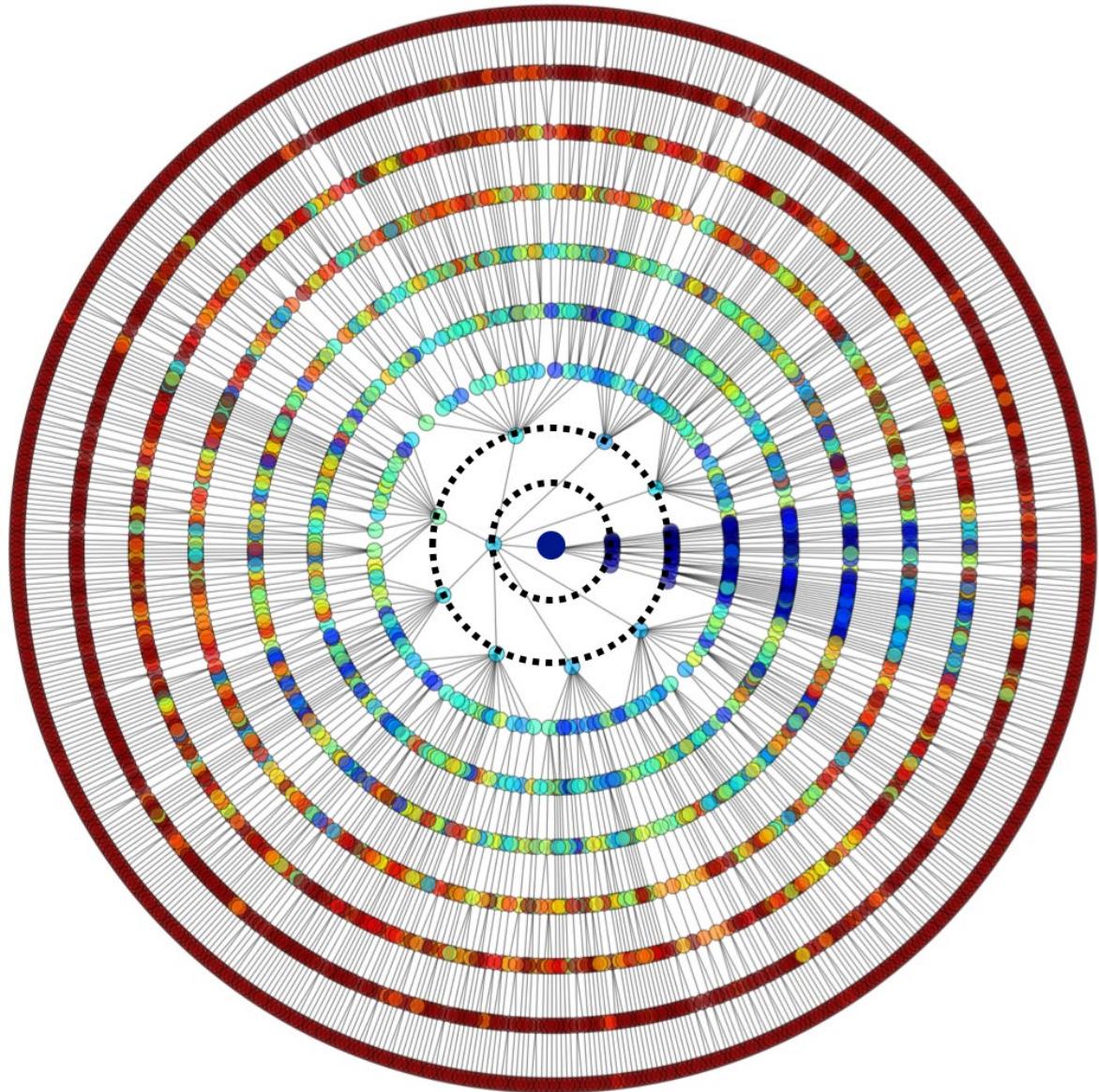
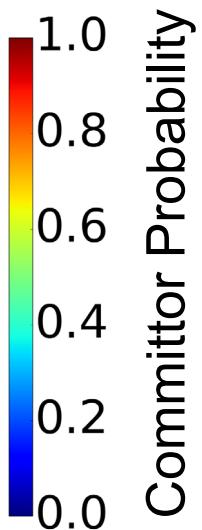
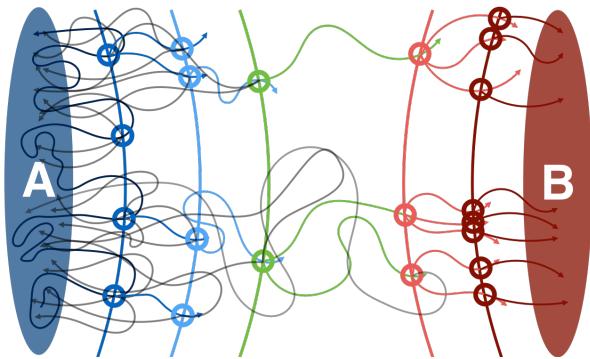
Hydrate nucleus

Hydrate phase forms from liquid through nucleation and growth.

Soluble additives (e.g. THF) promoted hydrate nucleation and growth in methane-THF-water systems.

How do gas hydrates of soluble guest molecules form?

DeFever, R.S., Sarupria, S. Nucleation mechanism of clathrate hydrates of water-soluble guest molecules, J. Chem. Phys., 147, 204503 (2017)



We performed large scale FFS consisting of over 250,000 individual MD simulations and requiring over 300,000 CPU-hours

Visualizing nucleation: FFS Trajectories



Simulation #1
Order parameter #1

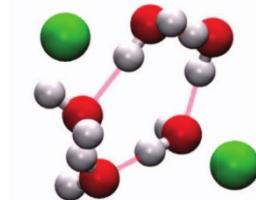
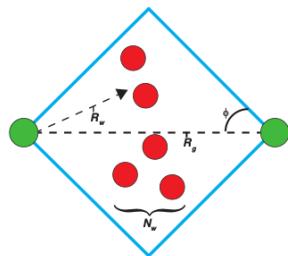


Simulation #1
Order parameter #2

How should we visualize hydrate nucleation in simulations?

Reaction coordinate for hydrate nucleation

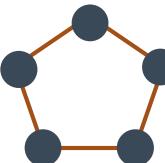
Mutually Coordinated Guest (MCG)



Red - water oxygen
Green - guest

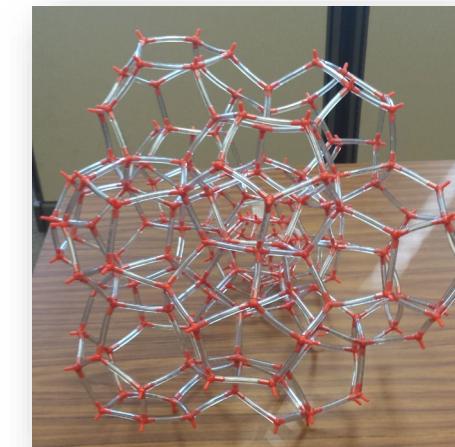
Barnes, B. C. et al. J. Chem. Phys. **2014**, 140, 164506.

Baez and Clancy (BC)

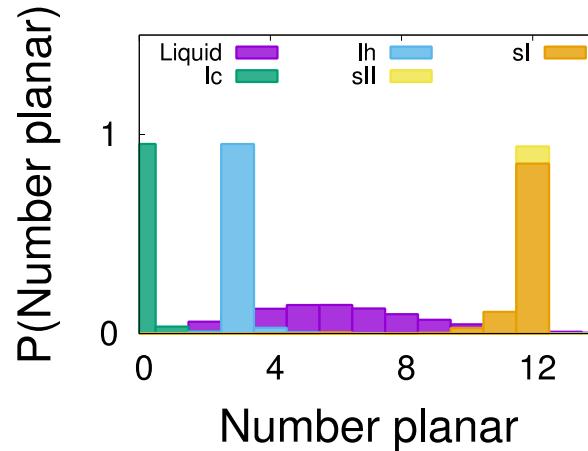


- 4-6 water pentamers
- Tetrahedral

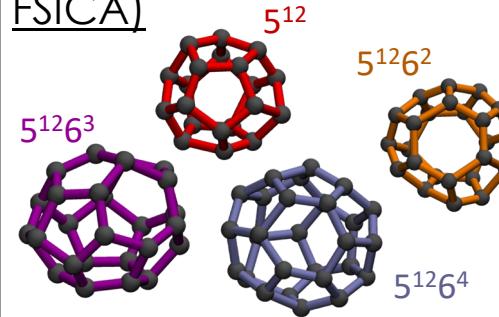
Baez, L. A., Clancy, P. Ann. N.Y. Acad. Sci. **1994**, 715, 177-186.



Planar Dihedrals (DHOP)



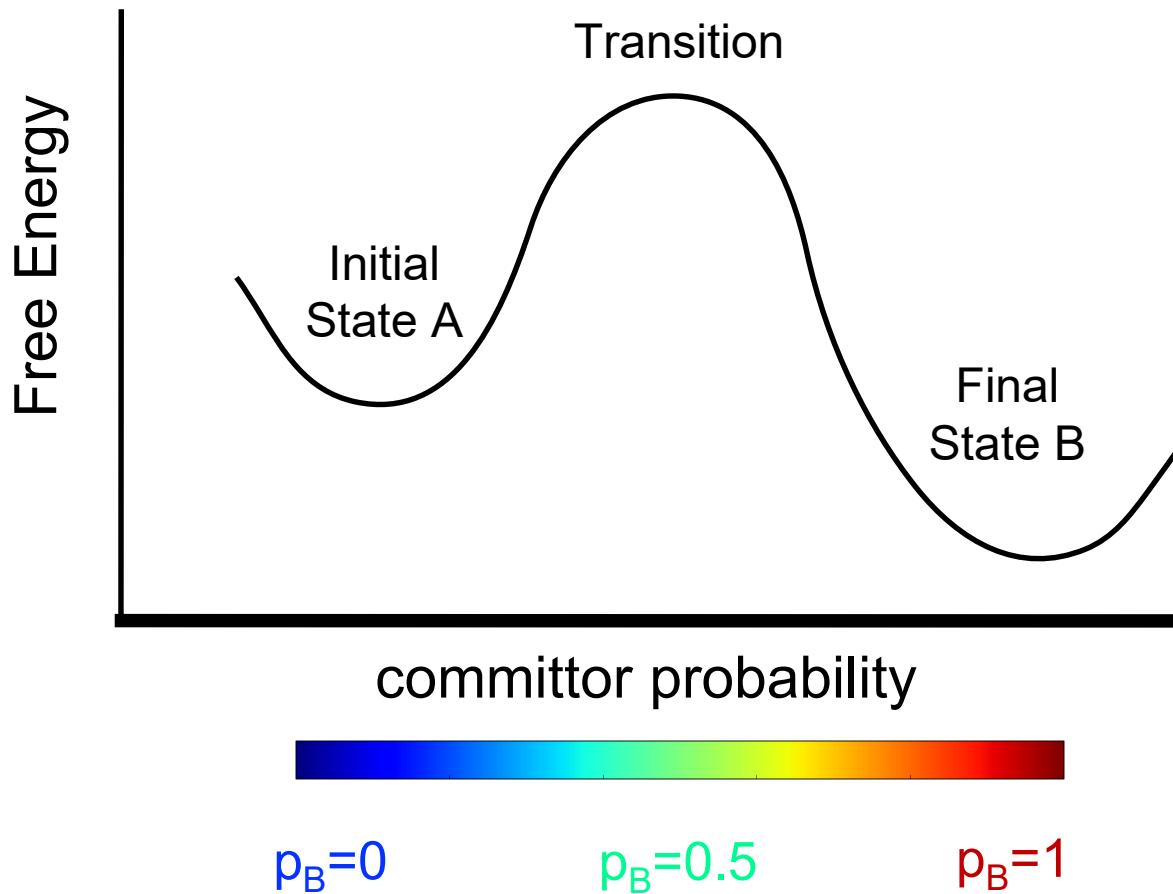
Cage-based (CG, HCOP, FSICA)



Bi, Y. and Li, T. J. Phys. Chem. B, 2014, 118, 13324-13332

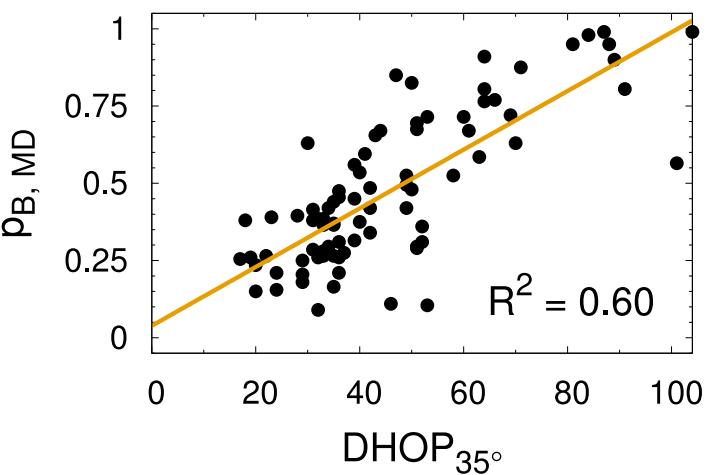
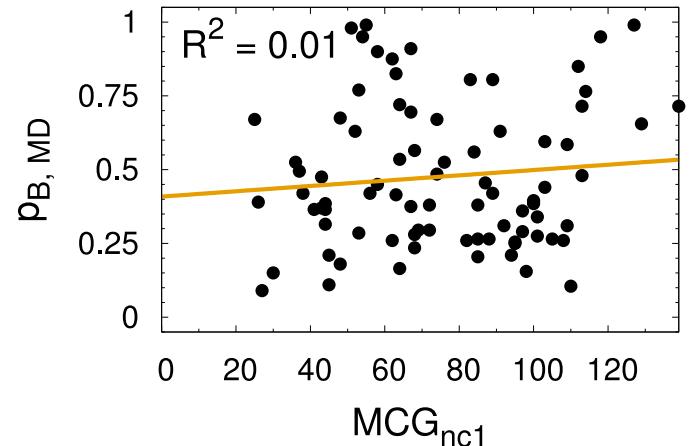
Guo, G.-J. et al. Phys. Chem. Chem. Phys., 2011, 13, 12048-12057

Rare events: Committor Probability



Committer probability is the best order order parameter to describe a transition.

Reaction coordinate for hydrate nucleation



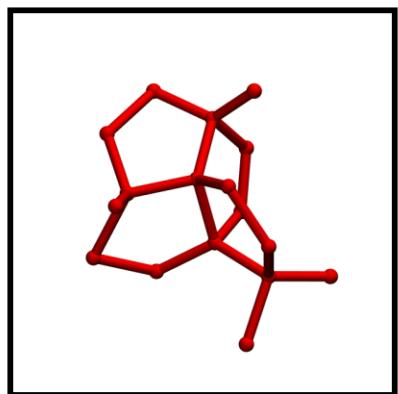
Order parameters based upon water structure better approximate of the reaction coordinate than guest-based.

Primarily water-structure-based

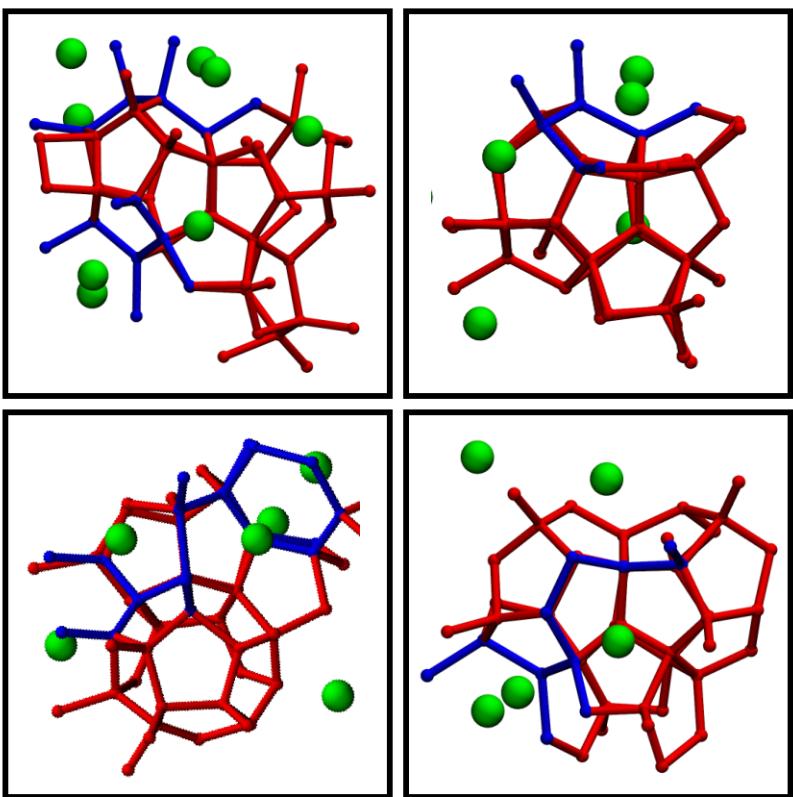
Rank	Model	R^2
1	DHOP _{35°}	0.60
2	BC _{planar}	0.55
3	DHOP _{30°}	0.52
4	BC-MCG ₁	0.50
5	BC-MCG ₂	0.50
6	RNGOP	0.47
7	FSICA _{FS}	0.43
8	BC-MCG ₃	0.42
9	BC	0.40
10	HCOP	0.40
11	MCG-6R ₃	0.38
12	BC ₅₆	0.38
13	MCG-6R ₂	0.36
14	FSICA _{CC}	0.30
15	MCG ₃	0.28
16	MCG ₂	0.28
17	CG-5 ⁶	0.26
18	LSSGOP	0.26
19	CG-full	0.24
20	CG-half	0.23
21	FSICA _{STD}	0.14
22	FSICA _{SII}	0.12
23	CG-5 ¹²	0.12

Better approximation of reaction coordinate

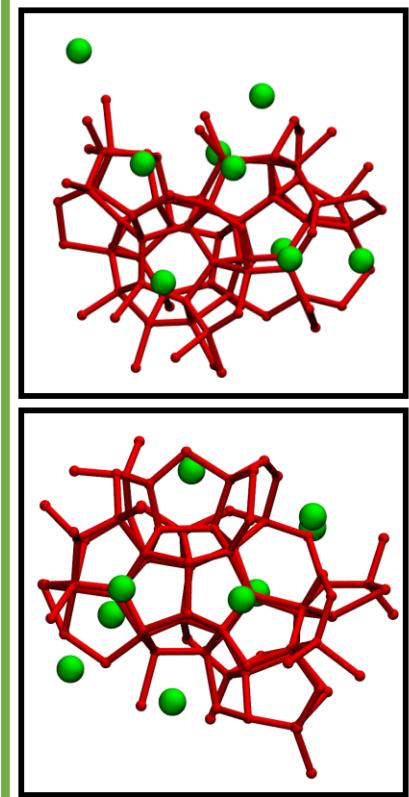
(a) Before TS



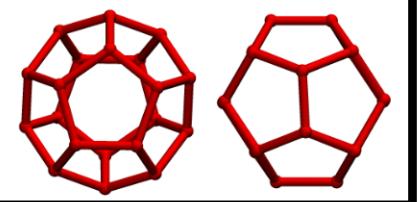
(b) At TS



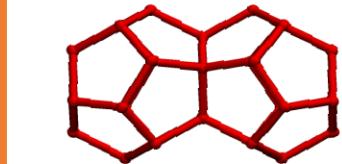
(c) After TS



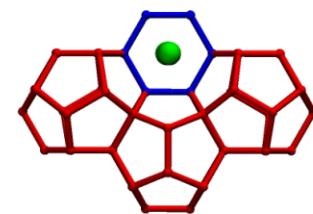
(d) 5^{12} cage



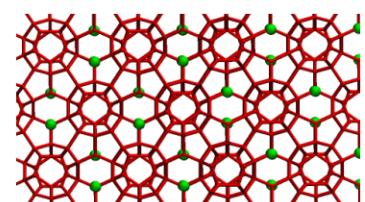
**(e) Face sharing
 5^{12} cages**



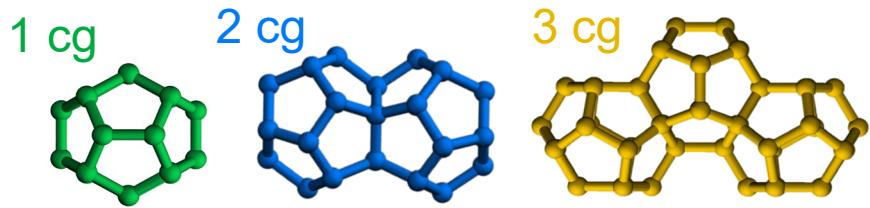
(f) sll motif



(g) sll crystal



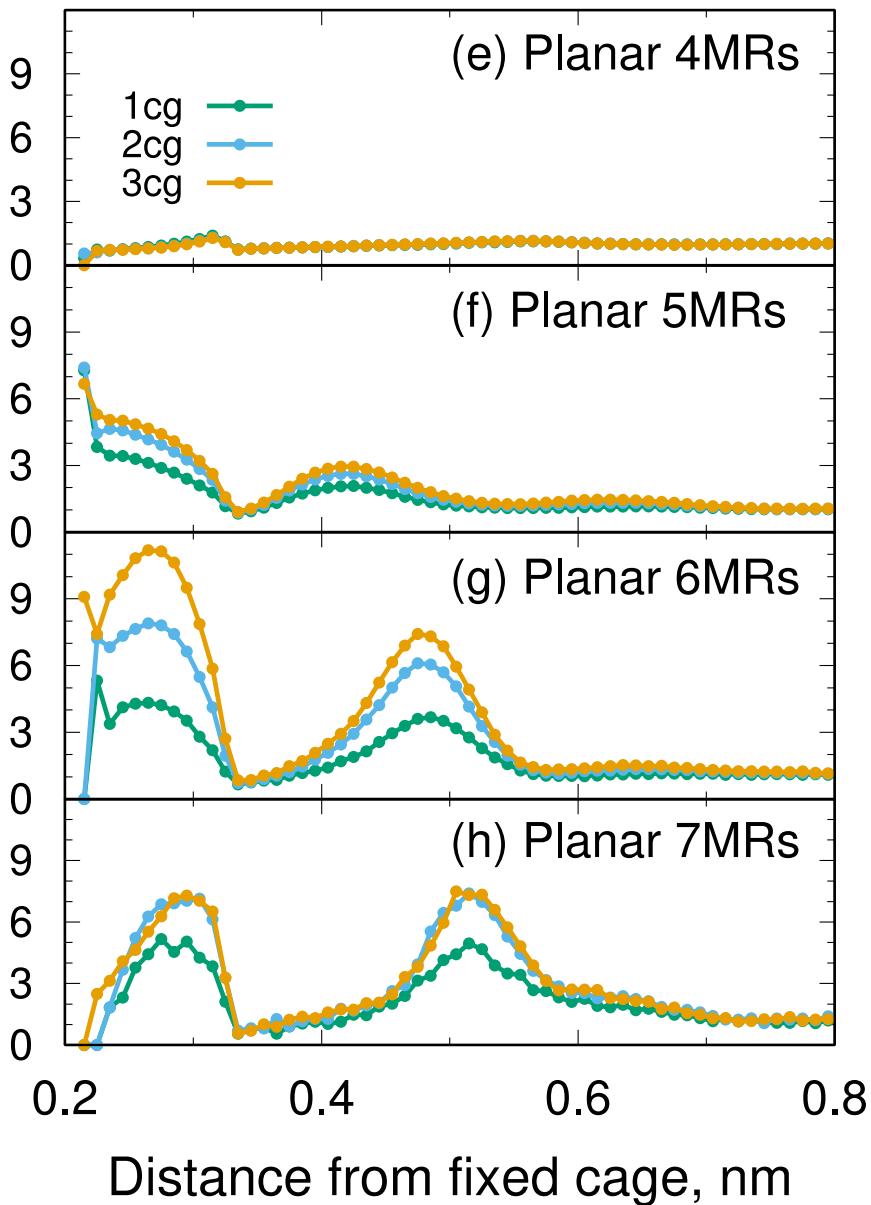
Role of face sharing cages in hydrate nucleation



$\langle R_W \rangle$ average number of planar primitive rings that water molecules at a given distance participate in

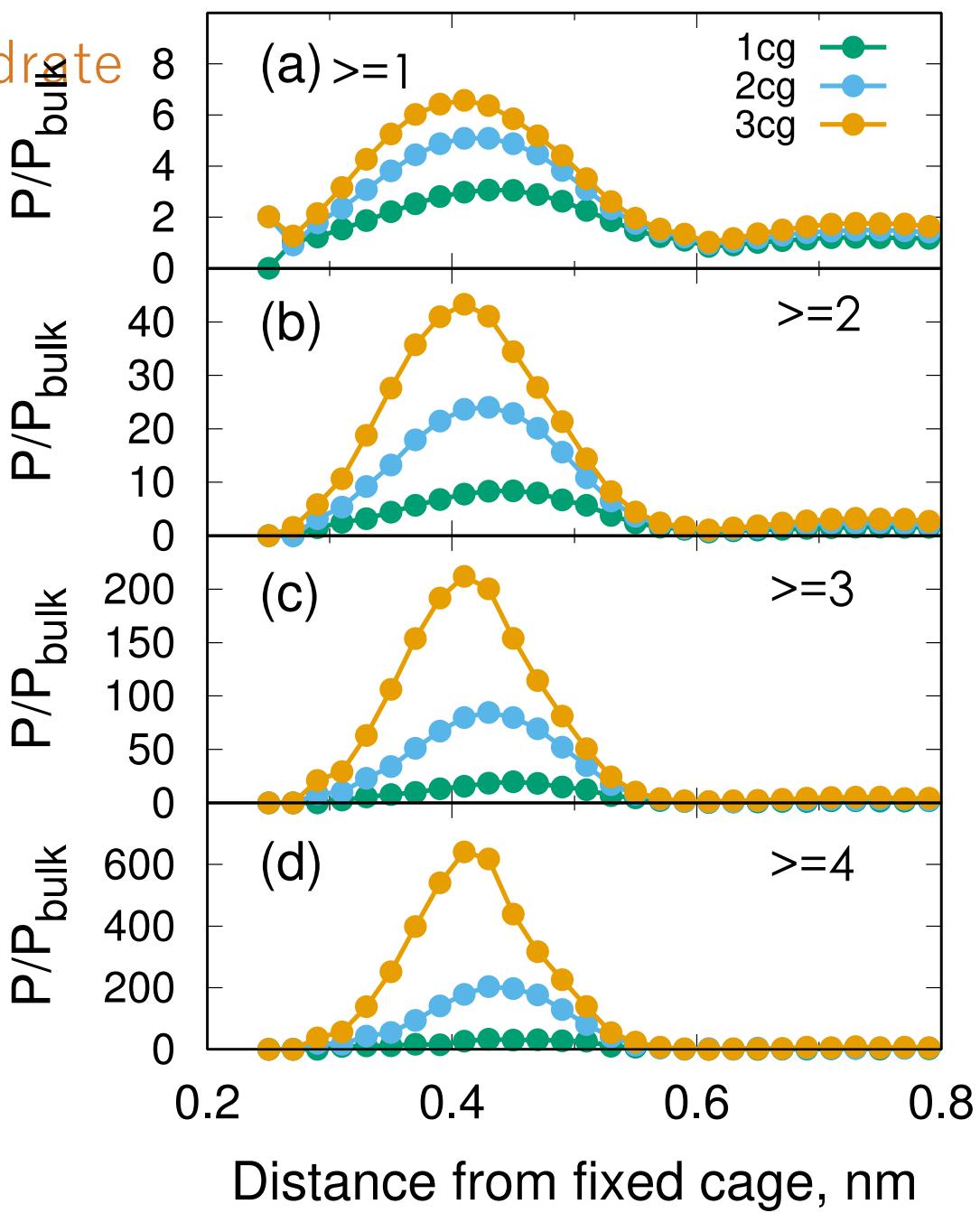
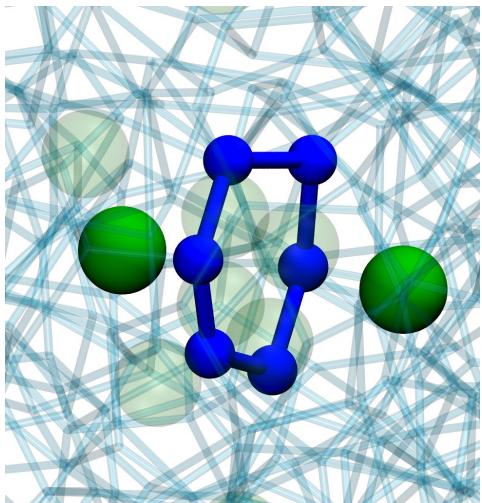
Face sharing cages promote formation of planar 6-membered rings.

$$\langle R_W \rangle / \langle R_W \rangle_{\text{bulk}}$$

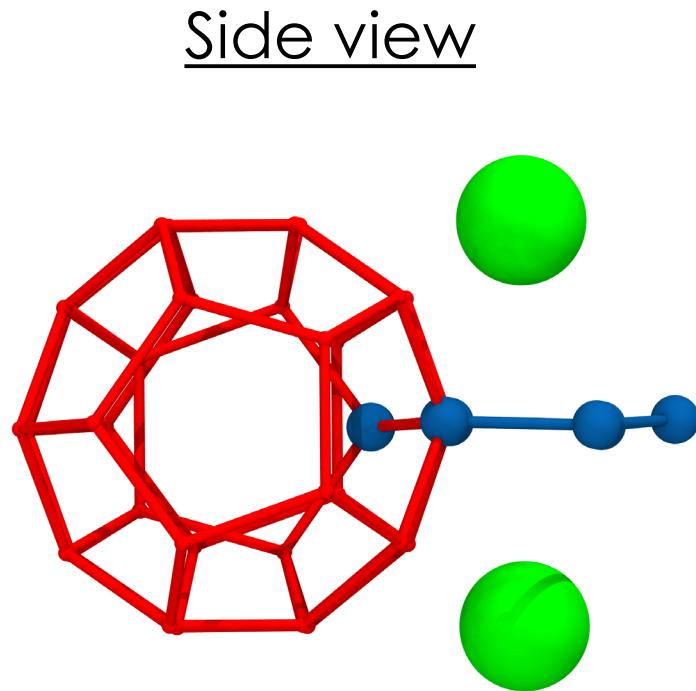
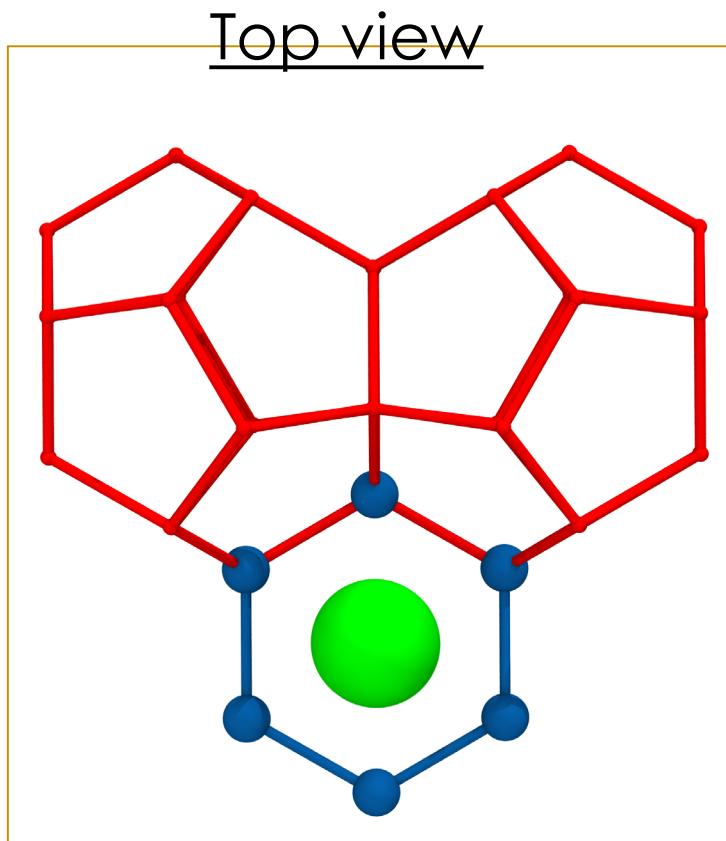


Face sharing cages in hydrate nucleation

Probability that a guest molecule is part of MCG-6R pairs as a function of distance from the nearest vertex of the 1, 2, or 3 fixed cages



Hydrate nucleation: Transition state



Face-sharing 5^{12} cages provide structural anchoring for the formation of planar 6-membered rings.

This provides the template for growing sII hydrate structure motifs.

Order parameters based upon water structure, (e.g. DHOP, BC) performed better than those based upon guest structure.

The transition state

- is characterized by a core of 2-3 face-sharing 5^{12} cages.
- enhances the planarity of surrounding water molecules and promote guest ordering into sII-like configurations.

DeFever, R.S., Sarupria, S. Nucleation mechanism of clathrate hydrates of water-soluble guest molecules, J. Chem. Phys., 147, 204503 (2017)



FFS is (in practice) somewhat sensitive to the choice of order parameter

“Among the advanced sampling techniques that can be used for direct calculation of nucleation rates, forward flux sampling (FFS) is the least sensitive to the proper selection of the order parameter.”

Phys. Chem. Chem. Phys. 2014, 1, (47), 25916-25927.

“…choosing a good order parameter (one which is close to the true reaction coordinate) will increase the computational efficiency…”

J. Phys. Condens. Matter. 2009, 21, 463102

“FFS-type methods are more dependent on a good order parameter than methods such as TIS and TPS (which only use λ to distinguish the stable states of the system).”

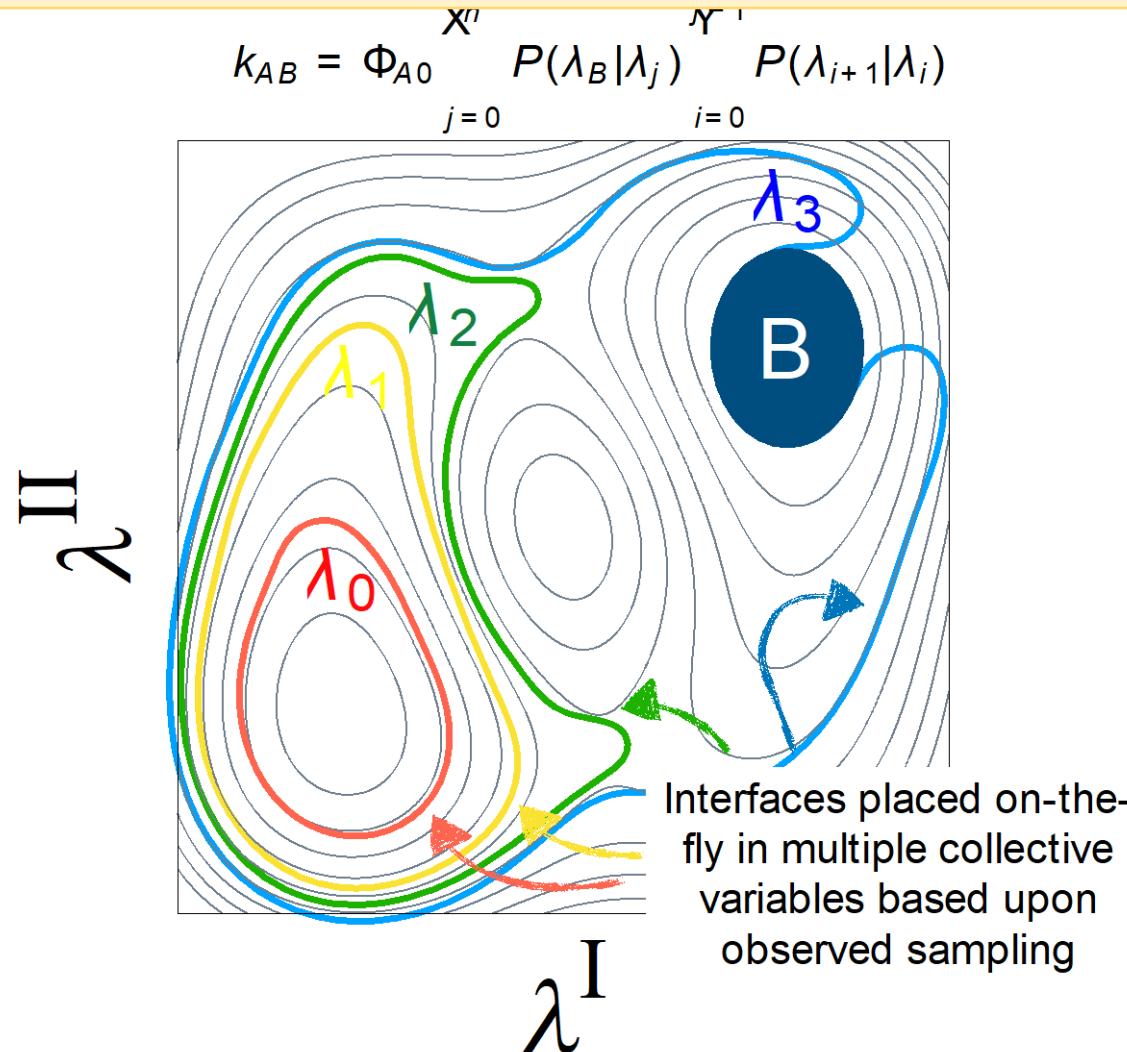
J. Phys. Condens. Matter. 2009, 21, 333101

“…FFS is more sensitive to the RC [order parameter] than TIS [transition interface sampling] or even RF [reactive flux].”

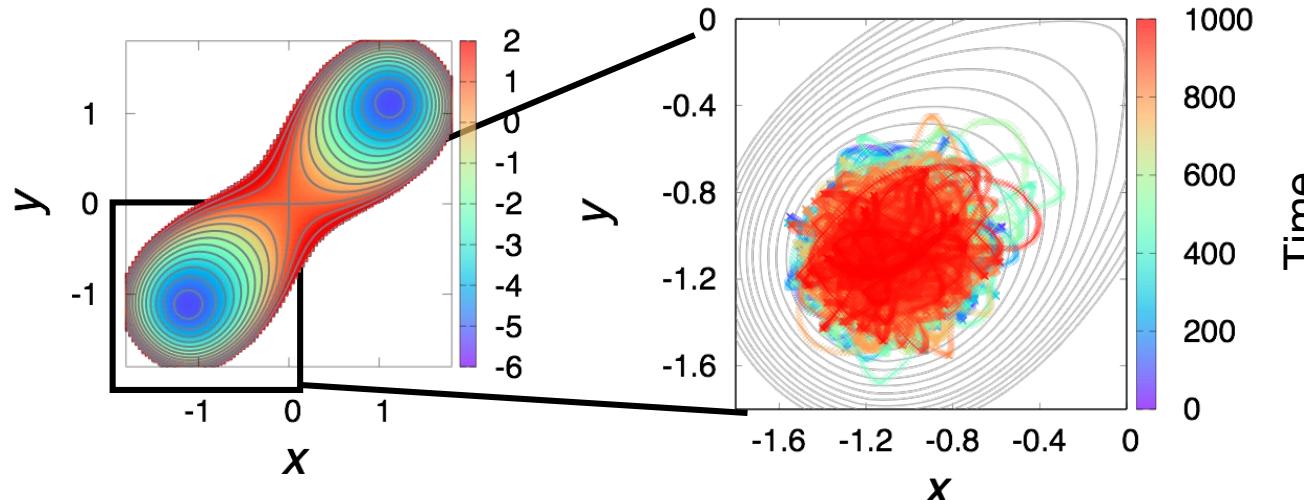
Dynamical Rare Event Simulation Techniques for Equilibrium and Nonequilibrium Systems. *Adv. Chem. Phys.*, 2012

n-dimensional FFS: contour FFS (cFFS)

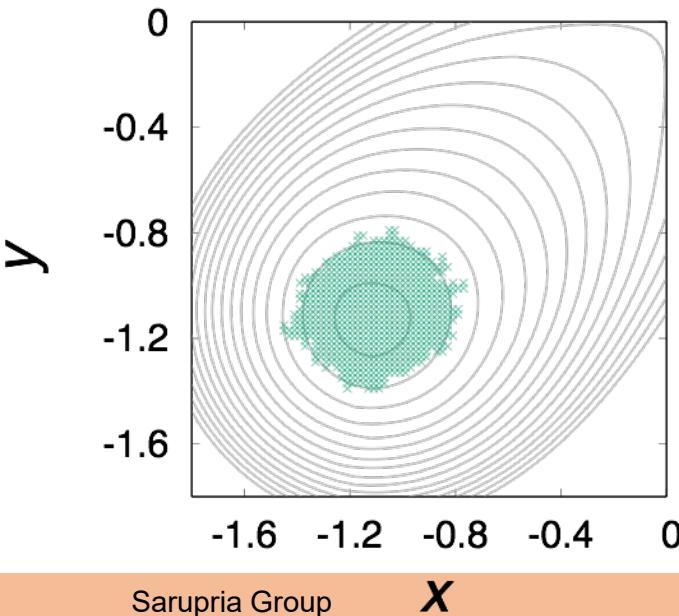
cFFS allows us to perform FFS along multiple order parameters without knowing the best combination of those *a priori*



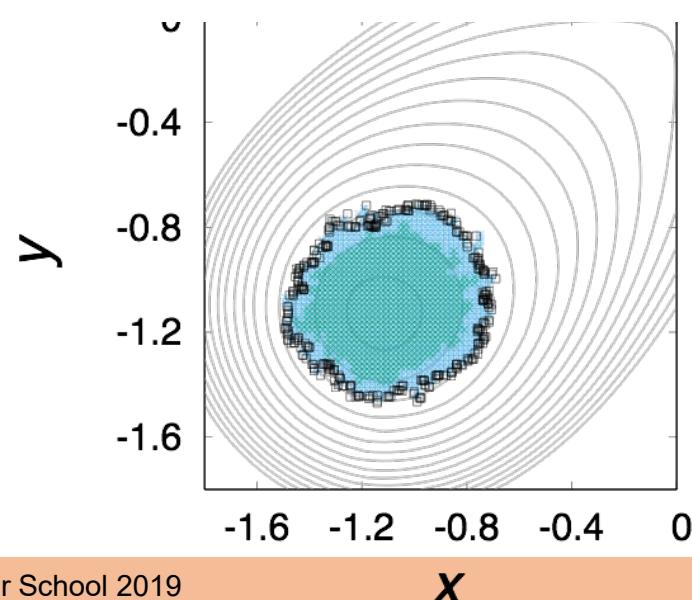
n -dimensional FFS: contour FFS (cFFS)



Closed set describing basin
created from sampling

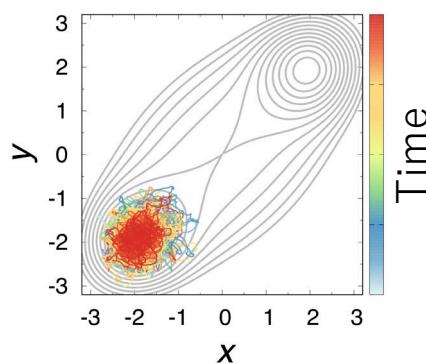


Closed set describing first interface
with initial configs shown

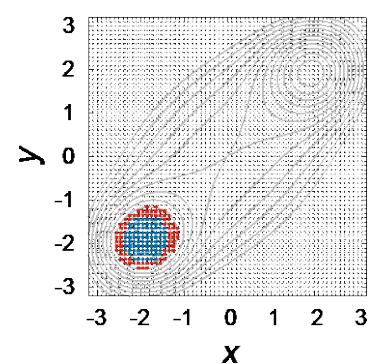


Contour forward flux sampling (cFFS)

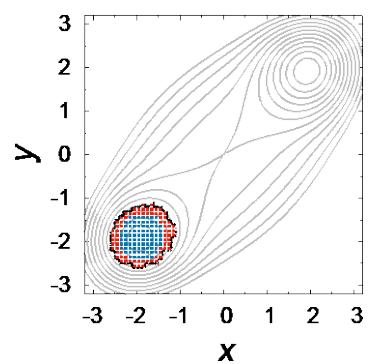
Run simulation in basin A to determine sampling in x,y



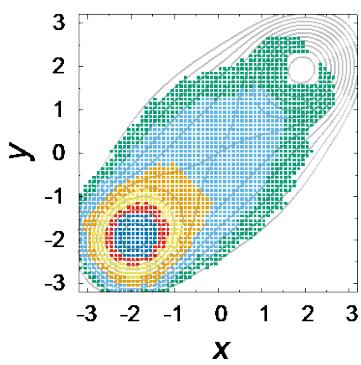
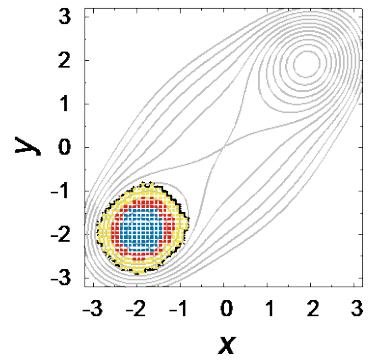
Create set of grid sites describing basin and first interface from extent of sampling



Collect configs from trajectories crossing the first interface (i.e., exiting the set)



Run simulations from first interface. Create next interface set from extent of sampling.



Repeat until reaching state B

Blue = basin A set, Red = interface 0 set, Yellow = interface 1 set, Black points = configurations

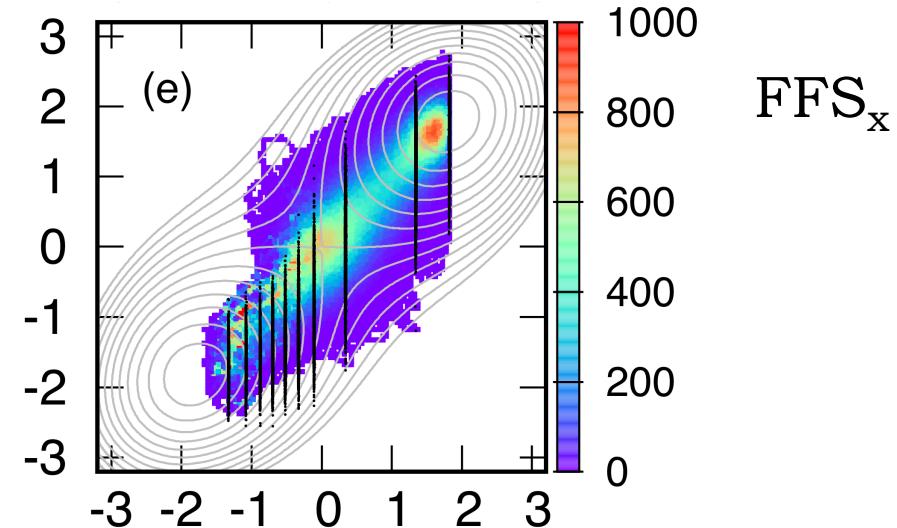
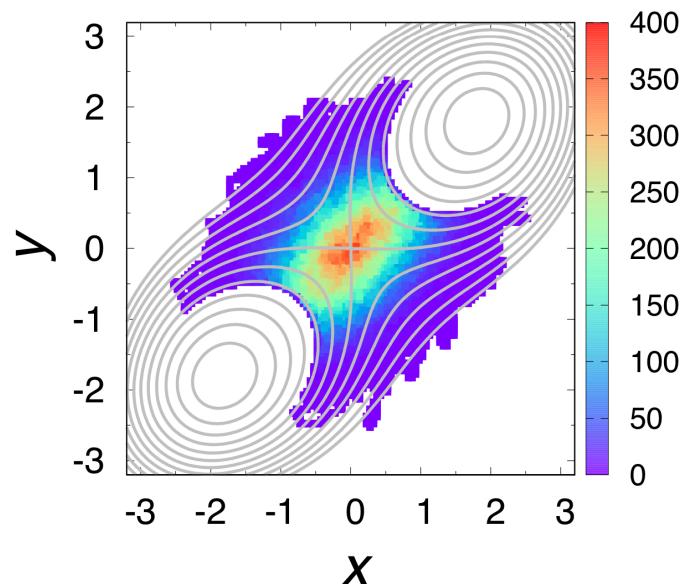
$$k_{AB} = \Phi_{A0} \sum_{j=0}^n P(\lambda_B | \lambda_j) \prod_{i=0}^{j-1} P(\lambda_{i+1} | \lambda_i)$$

cFFS defines each interface on the fly from the extent of sampling from simulations initiated at the previous interface.

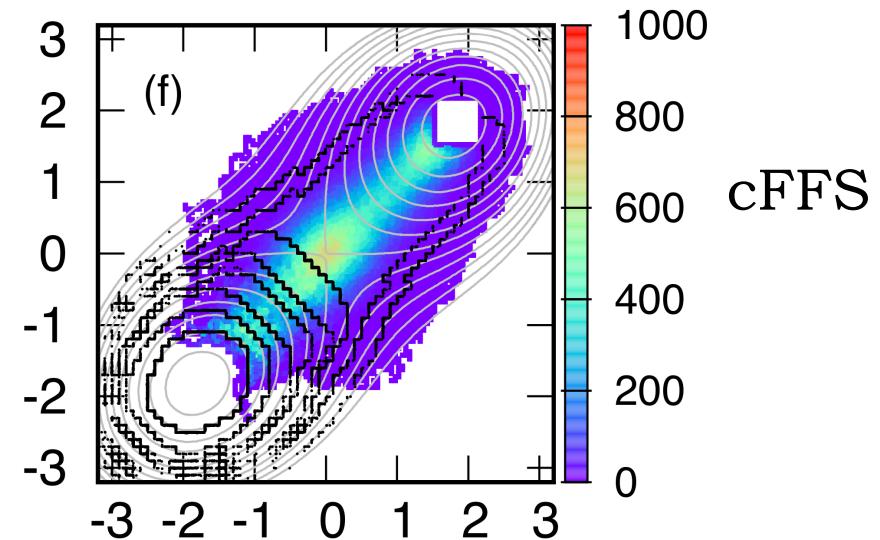
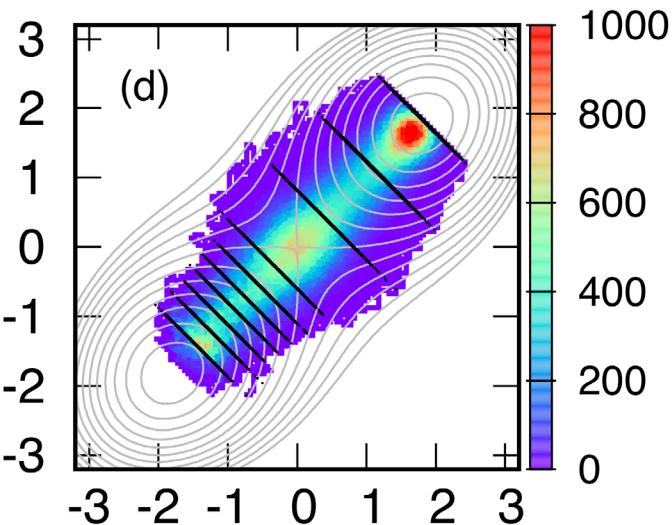
n -dimensional FFS: contour FFS (cFFS)

PES-2

SLD

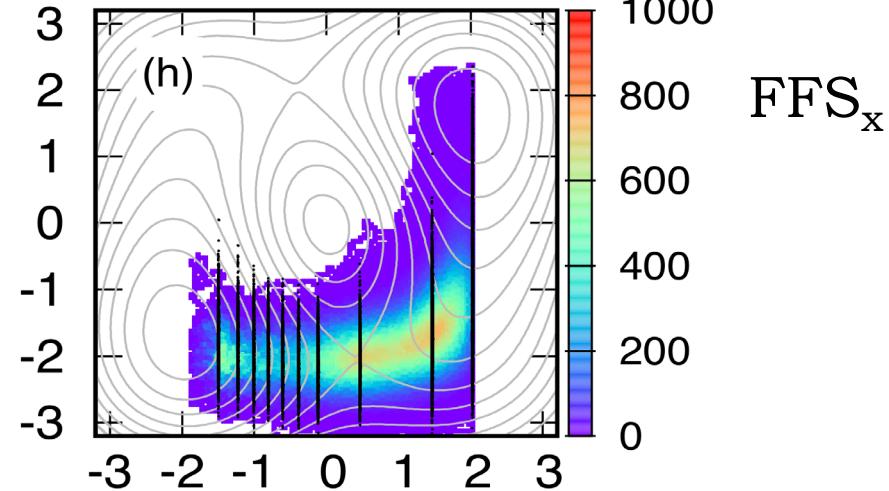
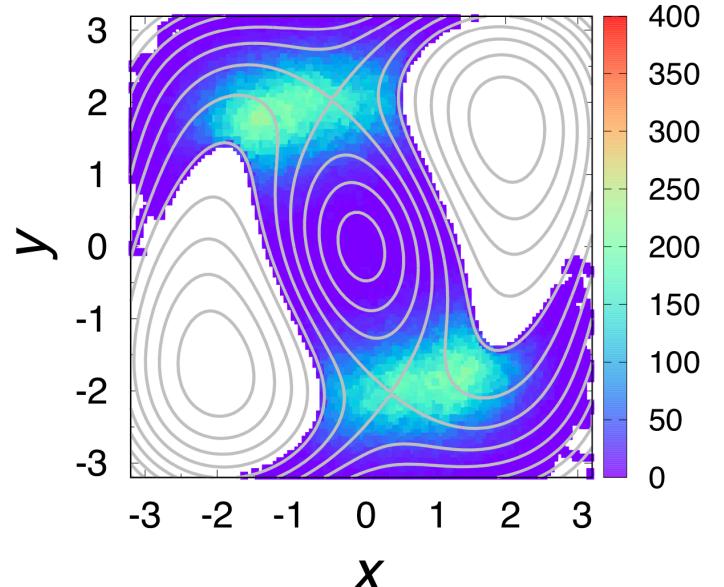
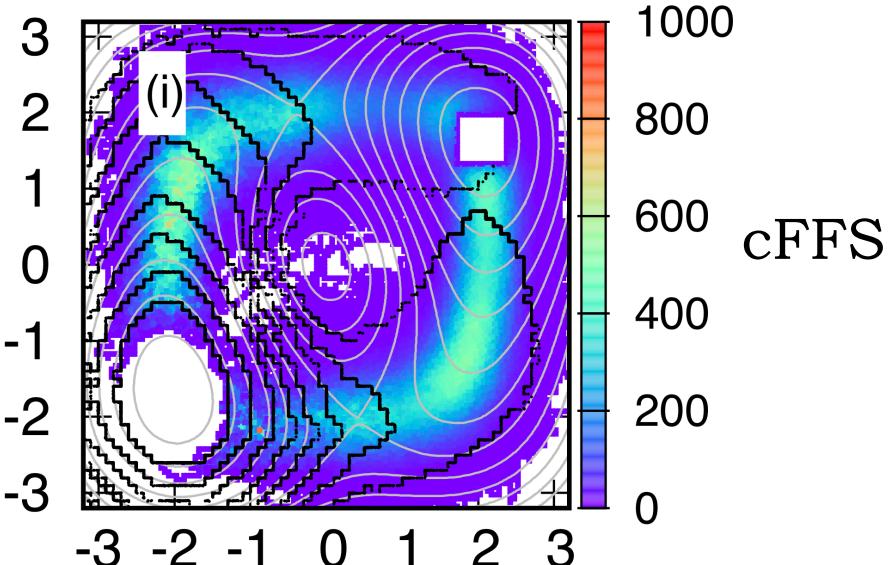
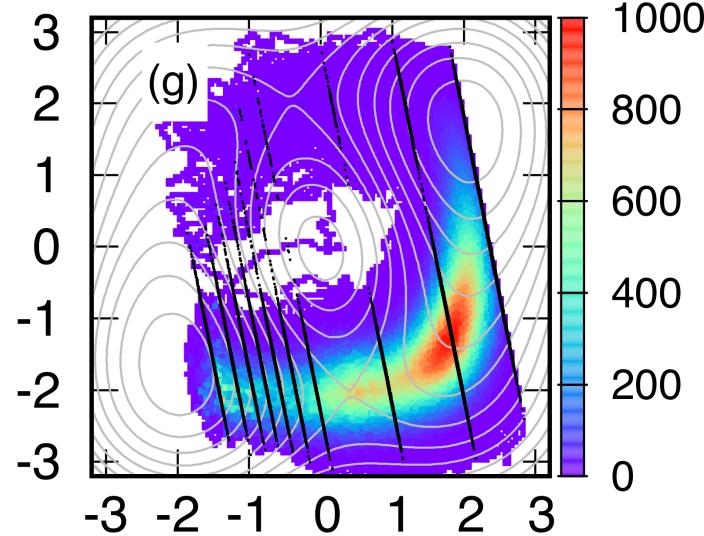


FFS_{opt}



n-dimensional FFS: contour FFS (cFFS)

SLD

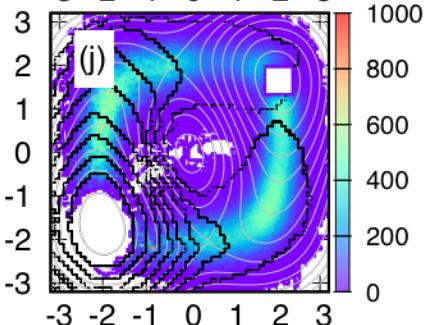
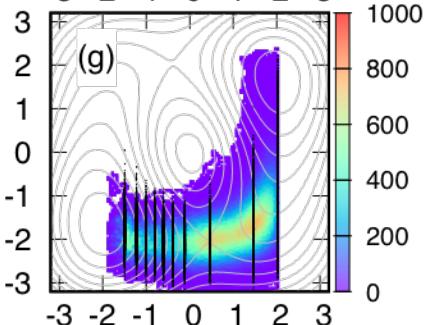
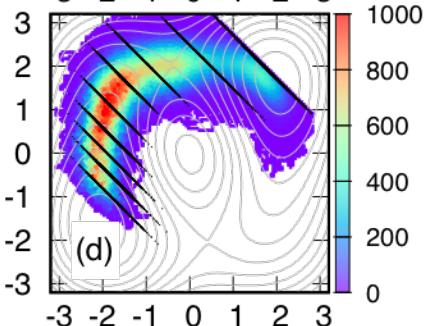
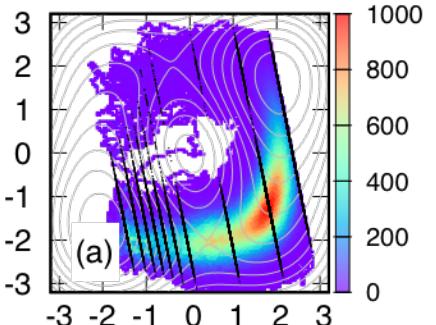
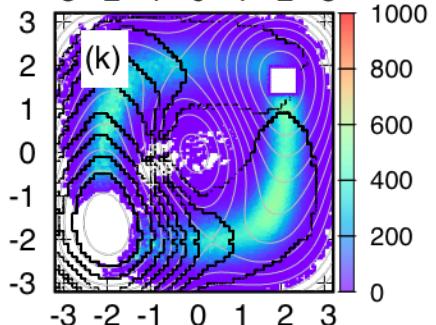
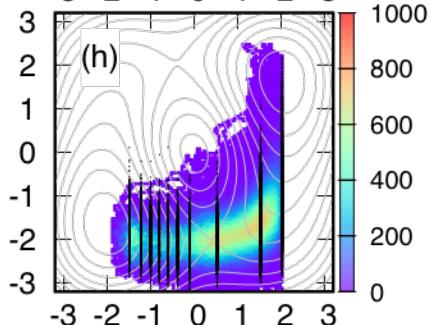
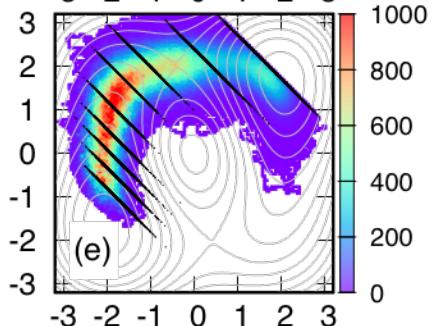
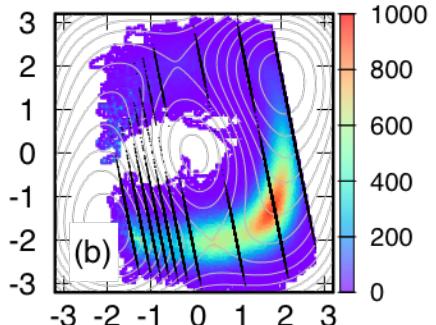
 FFS_{opt} 

Run 1

Run 2

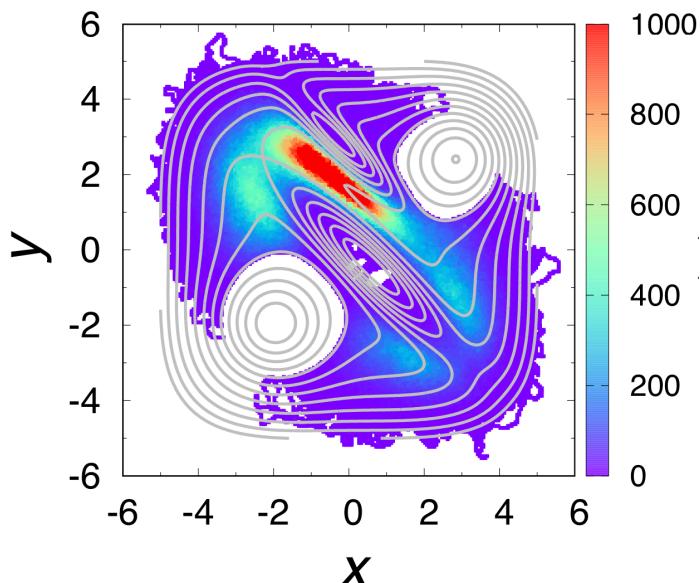
Run 3

PES-3

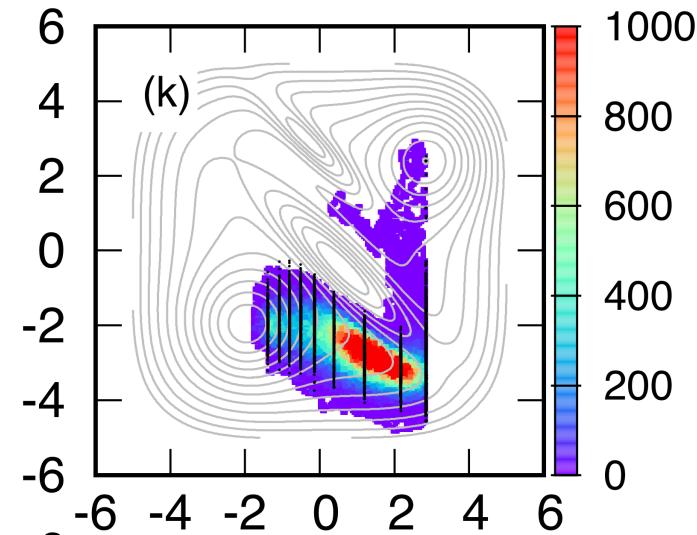
 y  y  X X X FFS_{opt} FFS_{x+y} FFS_x $cFFS$

n -dimensional FFS: contour FFS (cFFS)

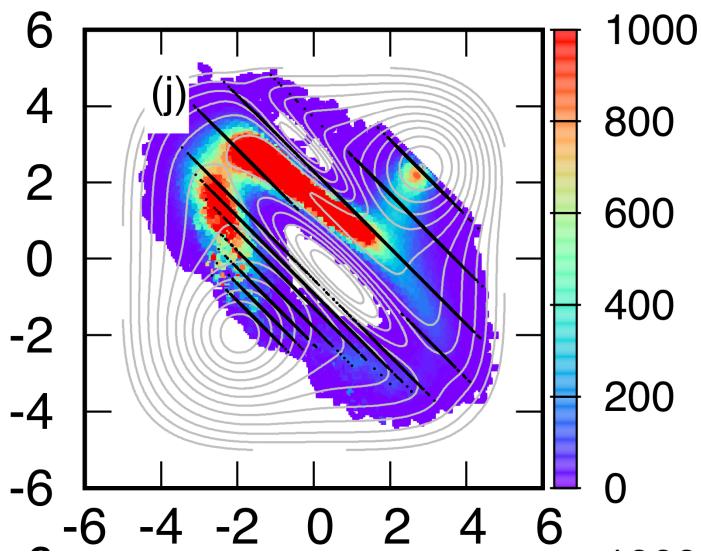
PES-4



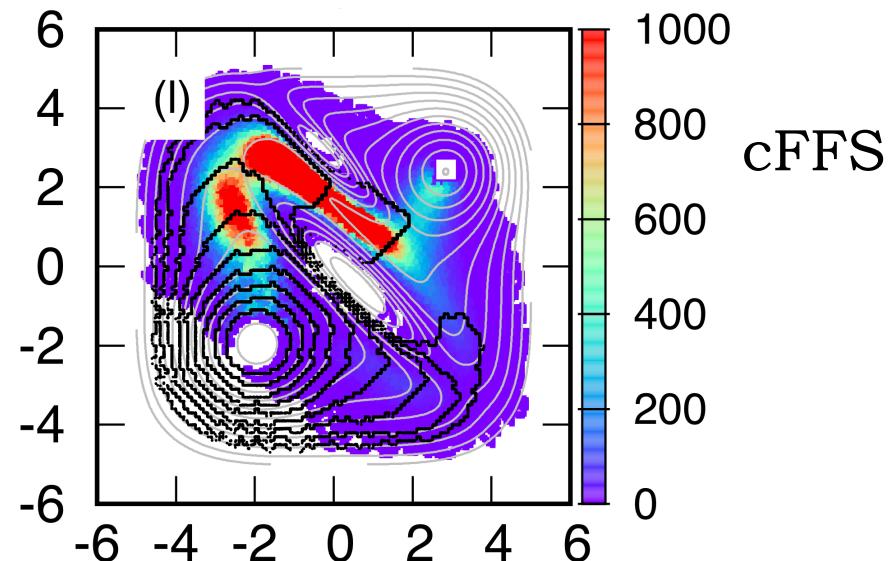
SLD



FFS_{opt}

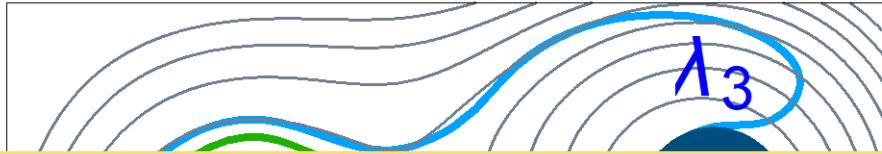


FFS_x

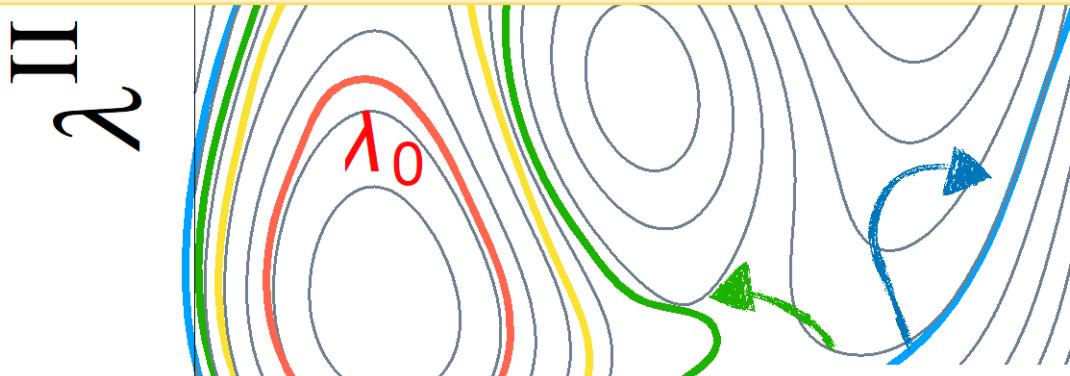


Summary

$$k_{AB} = \Phi_{A0} \prod_{j=0}^n P(\lambda_B | \lambda_j) \prod_{i=0}^{N-1} P(\lambda_{i+1} | \lambda_i)$$



cFFS places non-linear interfaces on-the-fly from the collective progress of the simulations, without any prior knowledge of the energy landscape or appropriate combination of order parameters.



Nonlinear combinations of order parameters have increased degeneracy compared with linear combinations of CVs in creating reaction coordinates (i.e., optimal order parameters).

λ₁

DeFever, R.S., Sarupria, S. Contour forward flux sampling: Sampling rare events along multiple collective variables , J. Chem. Phys., 150, 024103 (2019)

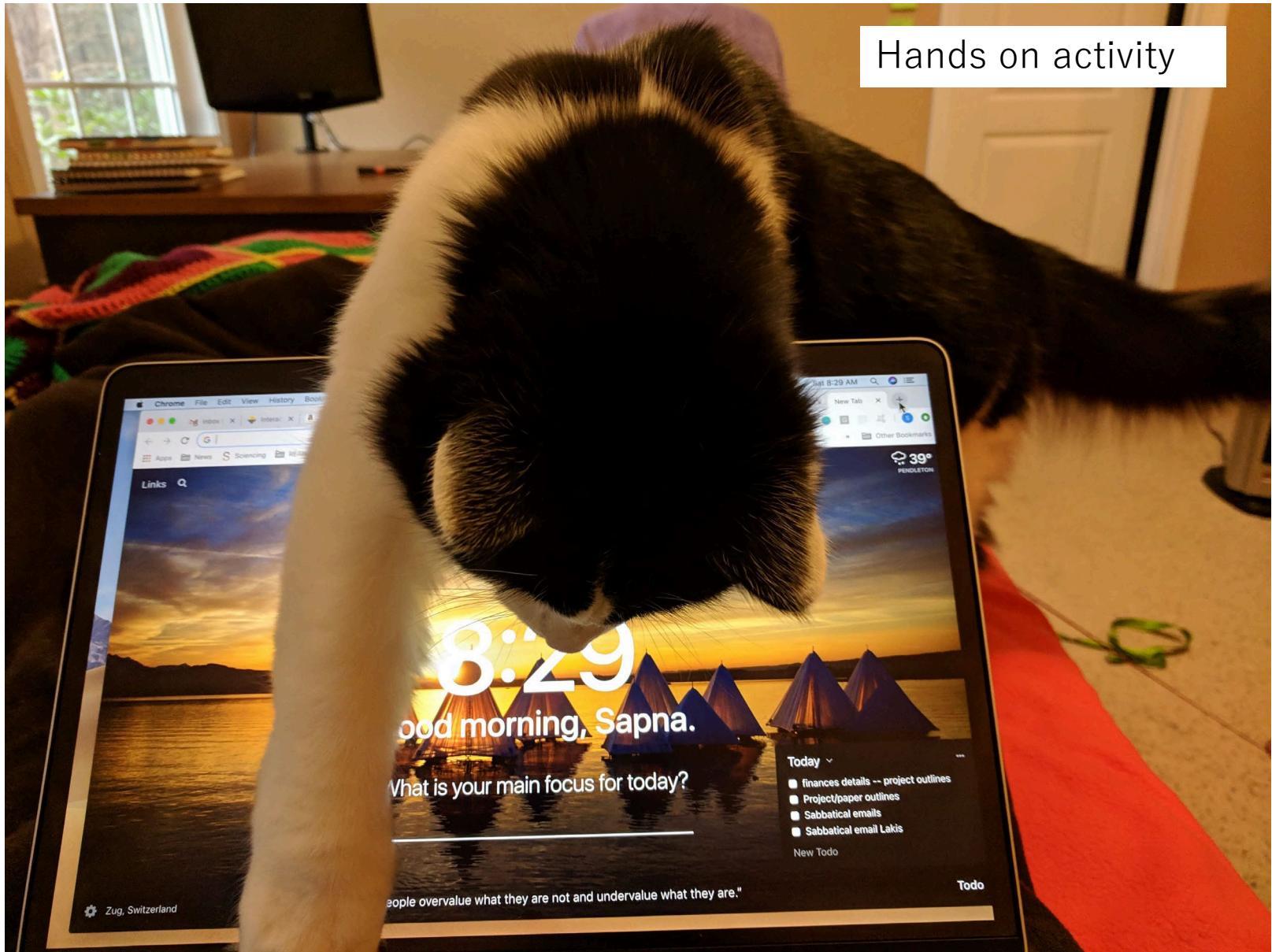
There are other approaches addressing this

- Escobedo – determination of reaction coordinates based on FFS and refining FFS

General thoughts

- TIS / RETIS / FFS / cFFS
- Software (TIS family): PyRETIS, Open Path Sampling (OPS)
- Software (FFS family): SAFFIRE, SSAGES
- Which method should you use?
- What is the fruit of the computational effort?

Hands on activity



Python notebooks

<https://github.com/SarupriaGroup/RareEventsSummerSchool2019>

- TIS, constrained TIS, FFS, cFFS
- Take a look at the algorithms
- Run each one and play with the parameters:
 - temperature (beta),
 - interfaces,
 - number of shooting moves,
 - number of simulations per interface,
 - Basin simulation length.
- To change potential energy surfaces – see `langevin_dynamics.py` in each folder.
- Compare the rate and the time it takes to get rates for each method



