Computing rare event properties by discretization and optimal control

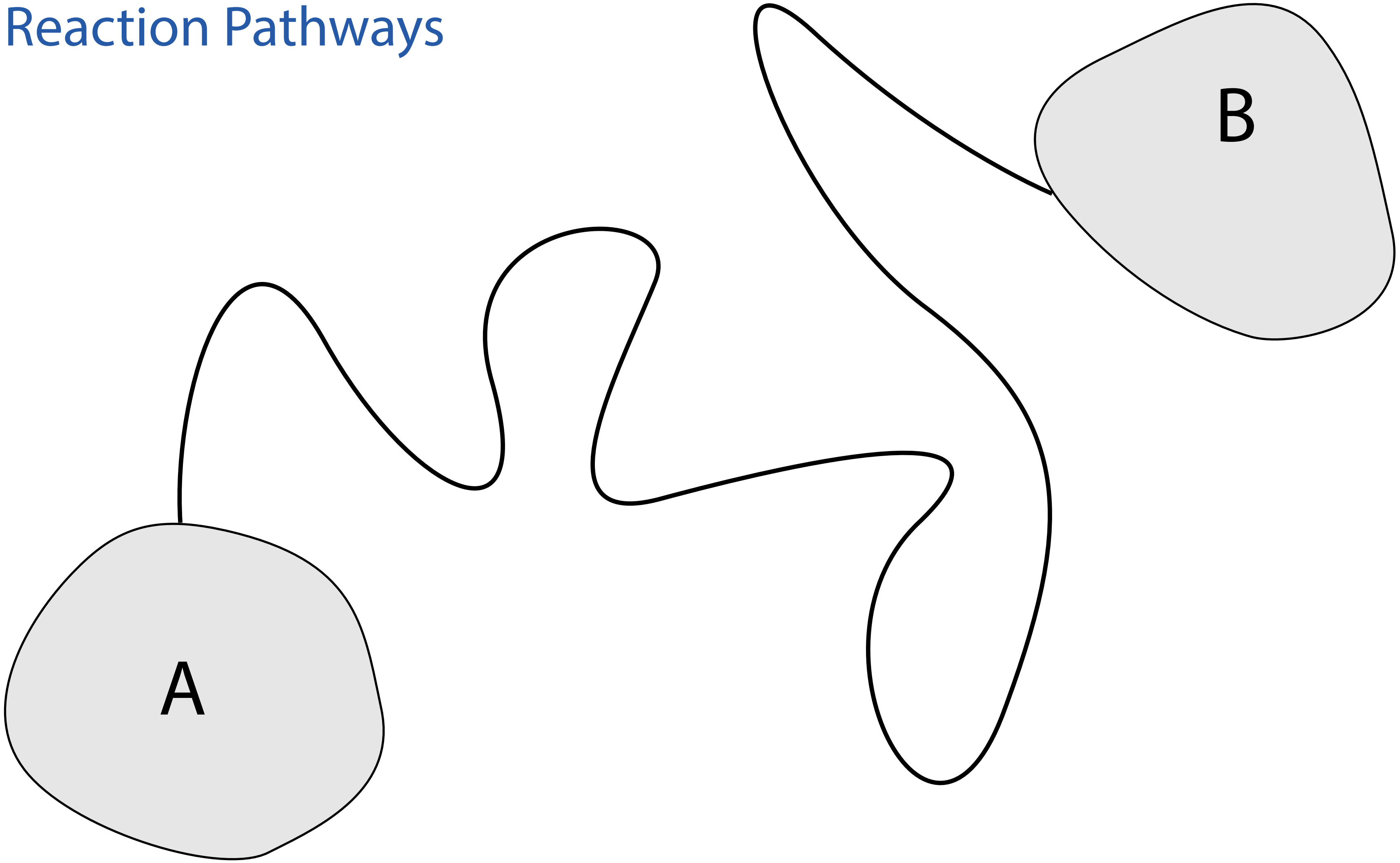
Marco Sarich

Freie Universität Berlin

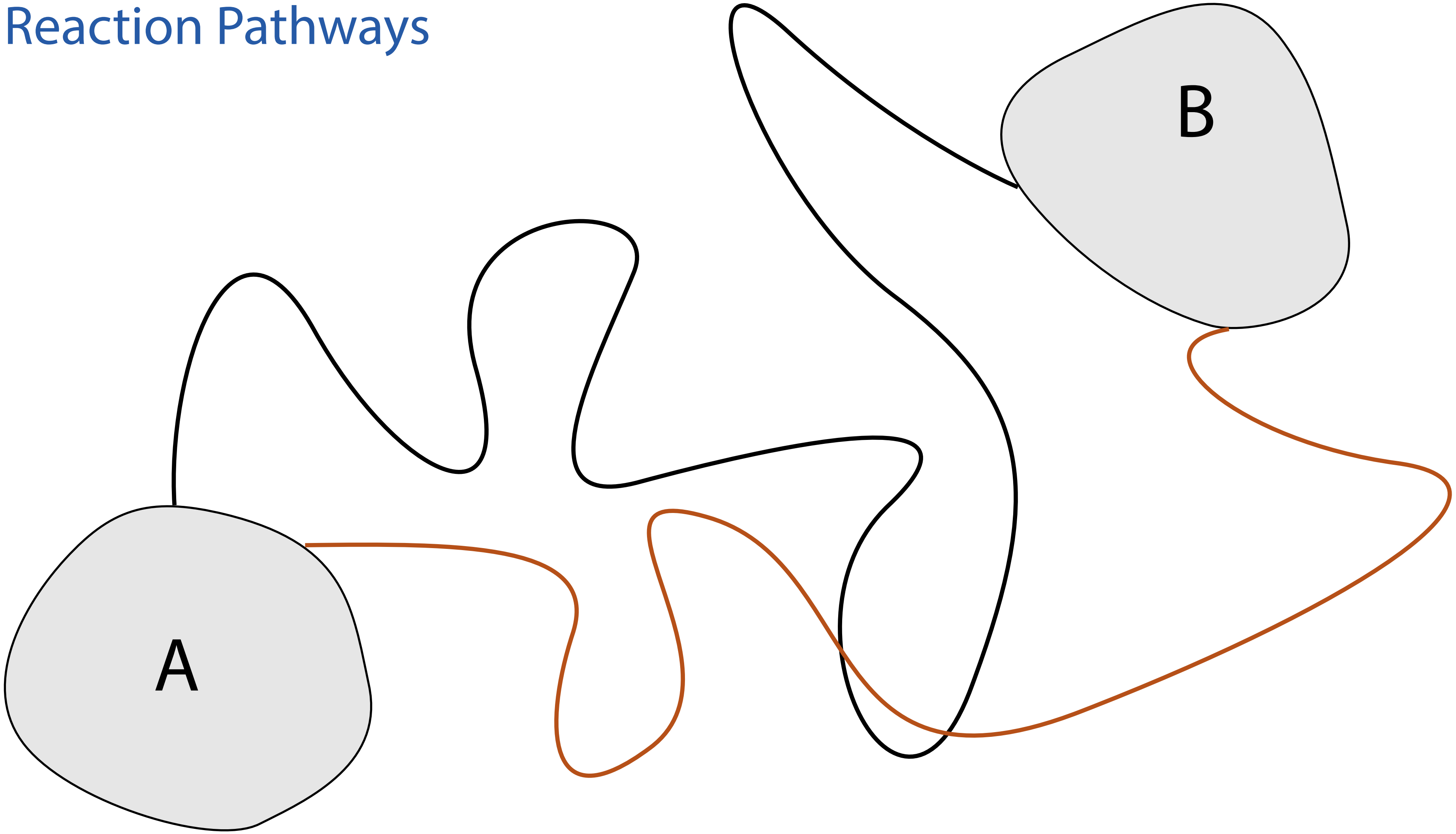
DFG Research Center MATHEON
Mathematics for key technologies

11/05/2013

Reaction Pathways

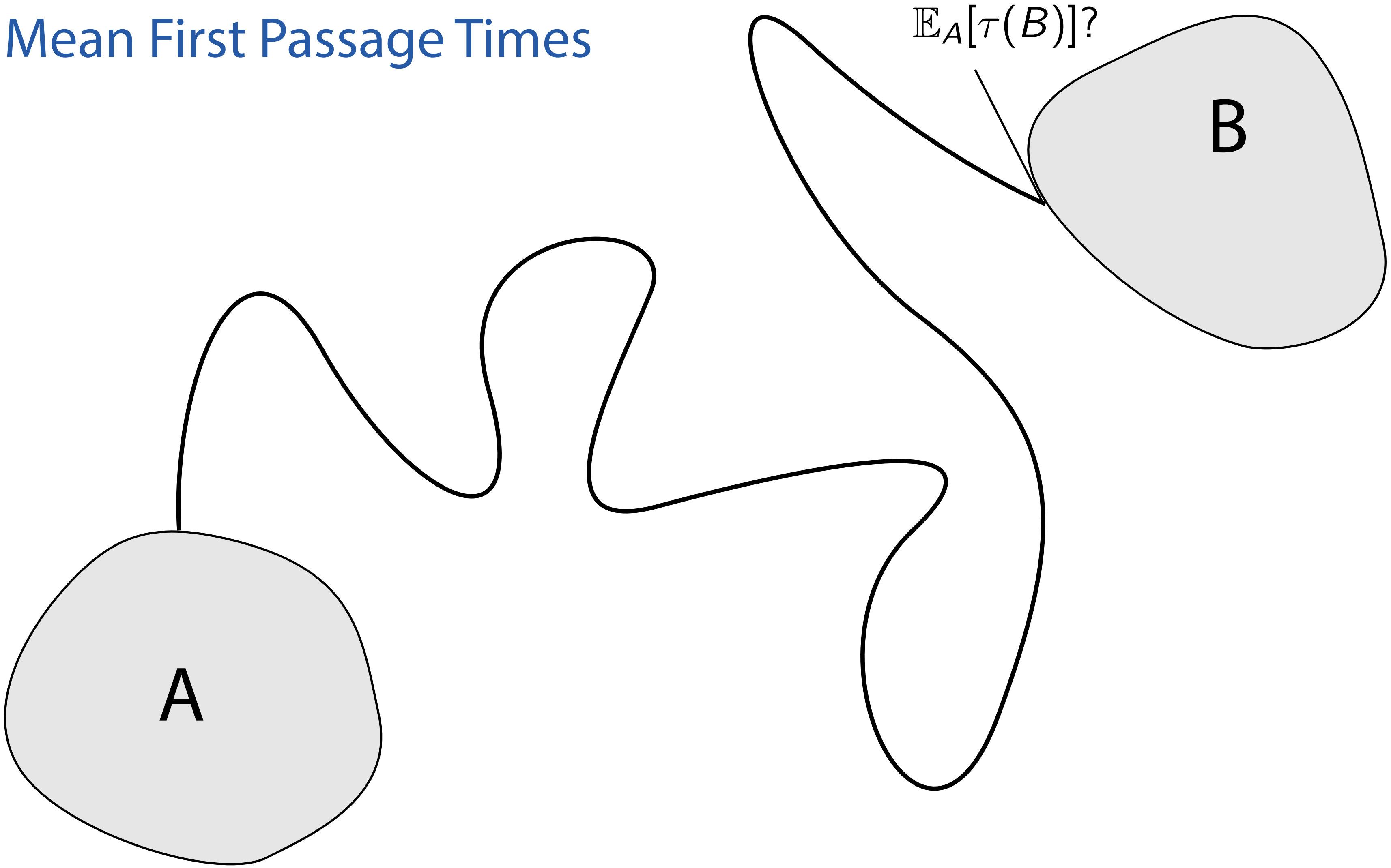


Reaction Pathways



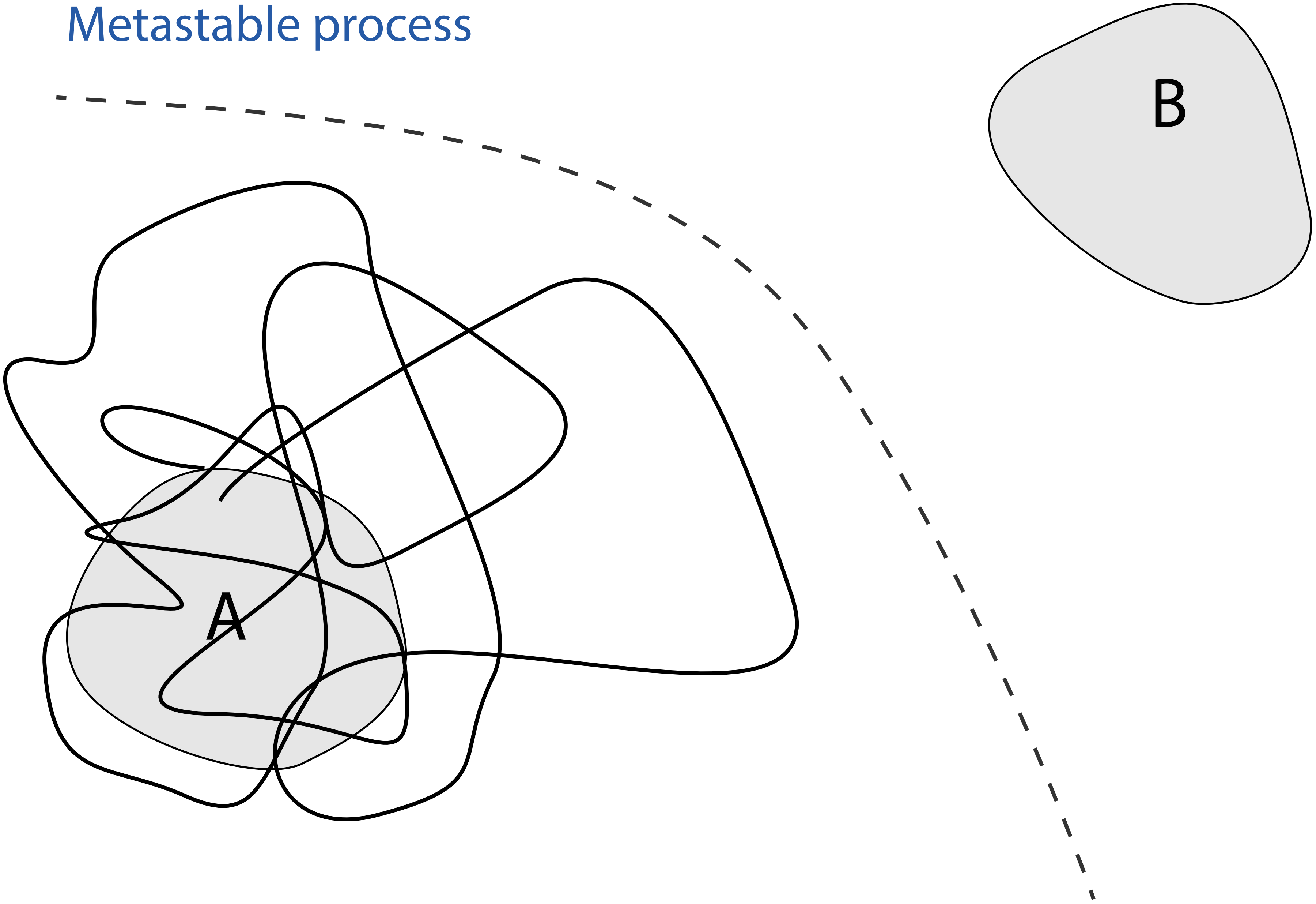
What are the dominant reaction pathways from set A to set B?

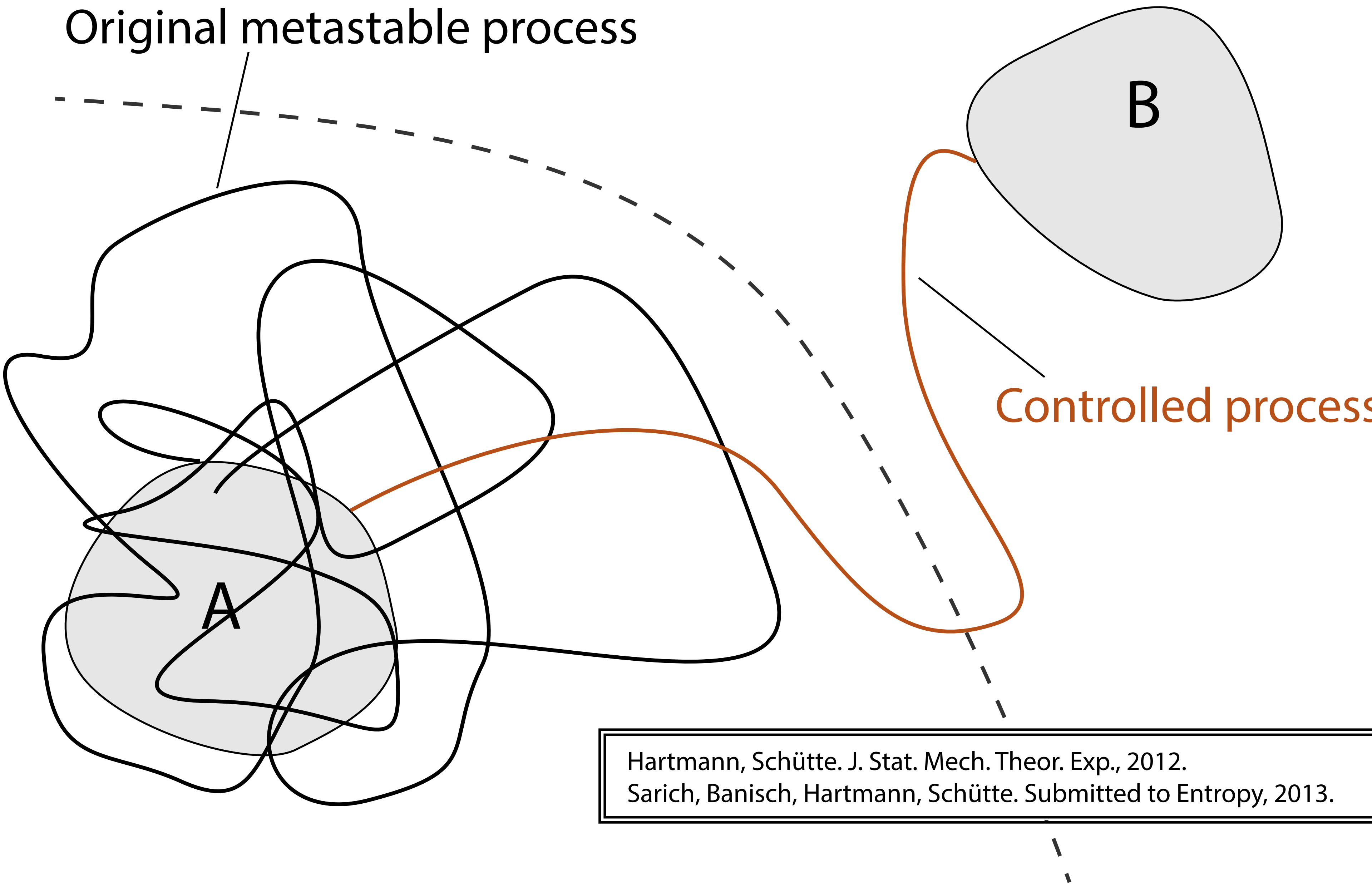
Mean First Passage Times



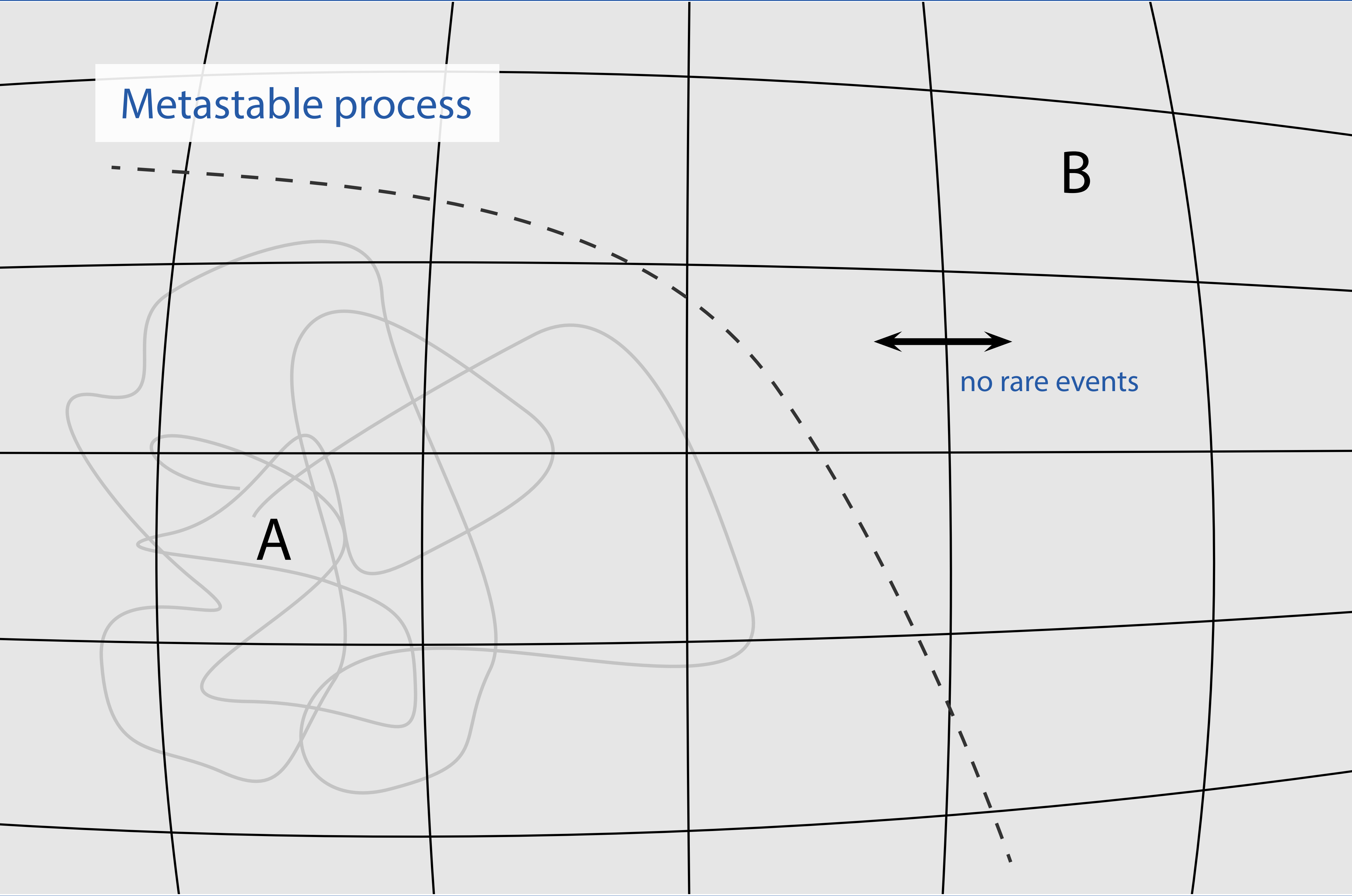
$$\tau(B) = \inf\{t \geq 0, X_t \in B\}$$

Metastable process





Hartmann, Schütte. J. Stat. Mech. Theor. Exp., 2012.
Sarich, Banisch, Hartmann, Schütte. Submitted to Entropy, 2013.


Metastable process
B
no rare events
A

For a Markov chain $(\hat{X}_n)$ with transition matrix P:

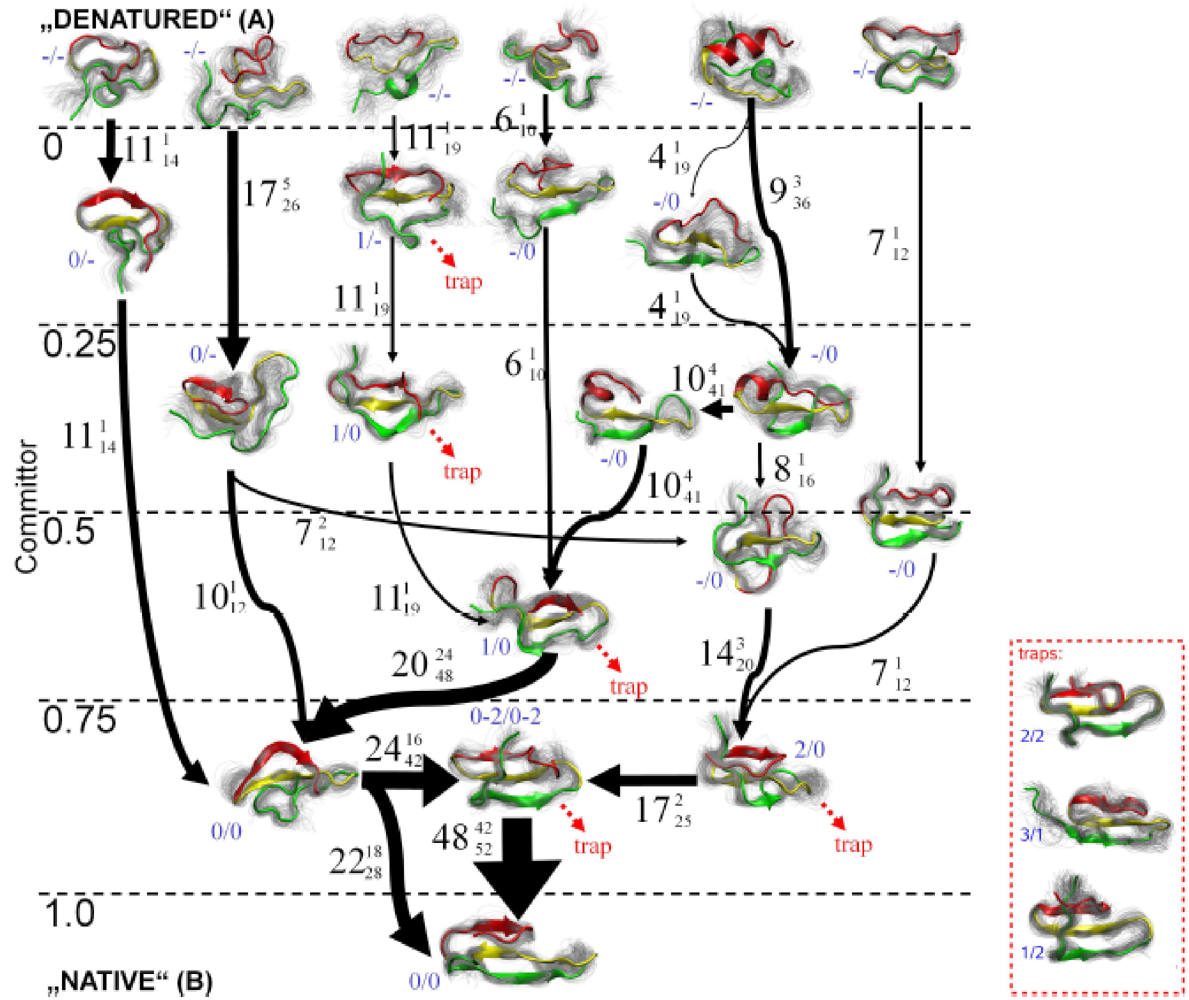
The mean first passage times

$$m(i) = \mathbb{E}[\tau(B) \mid X_0 = i]$$

can be computed without sampling

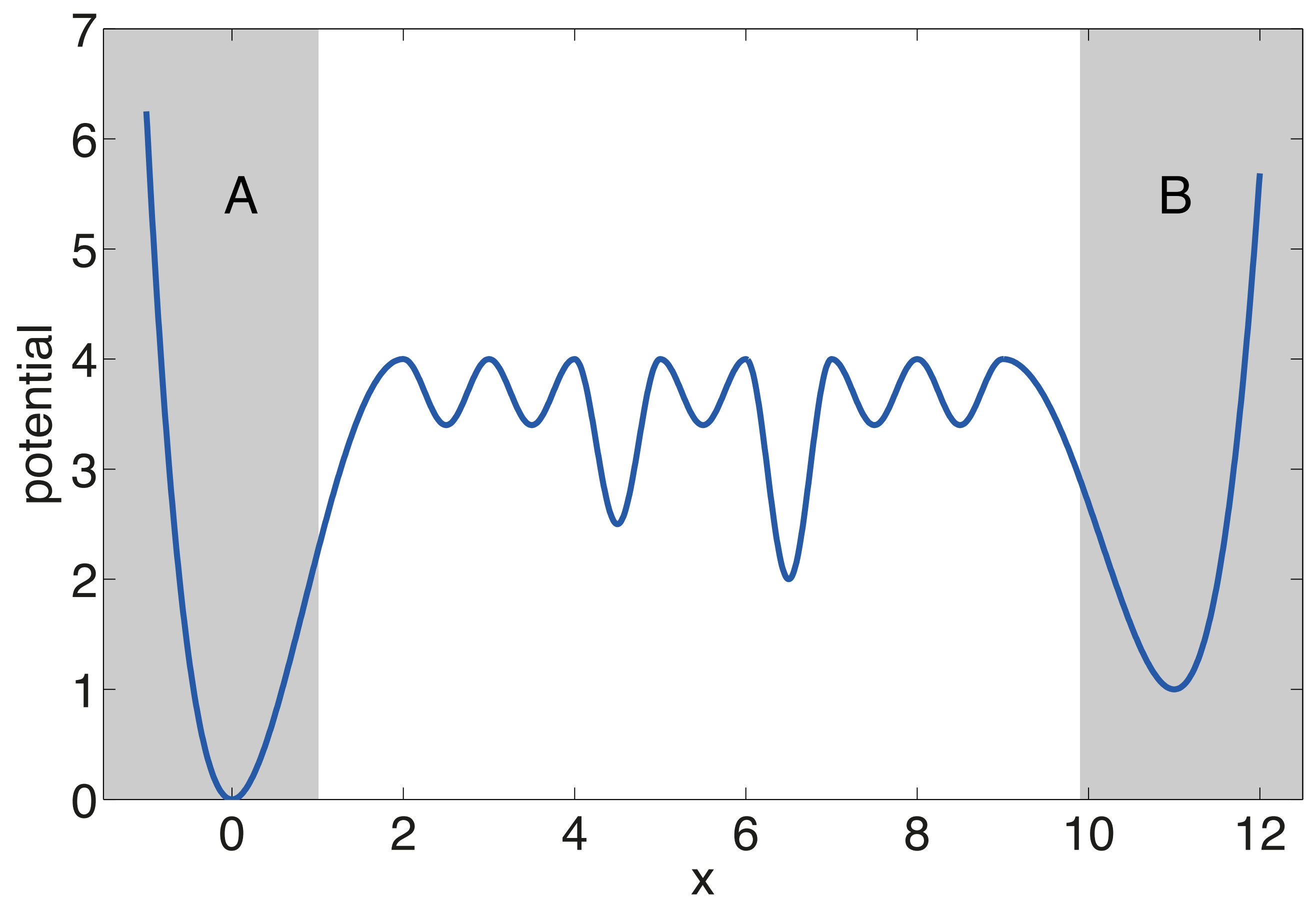
$$\begin{aligned} (P - I)m &= -\mathbb{1}, \qquad && \text{on } B^c, \\ m &= 0, && \text{on } B. \end{aligned}$$

Also the statistics of transition pathways can be computed without sampling by using transition path theory (TPT).

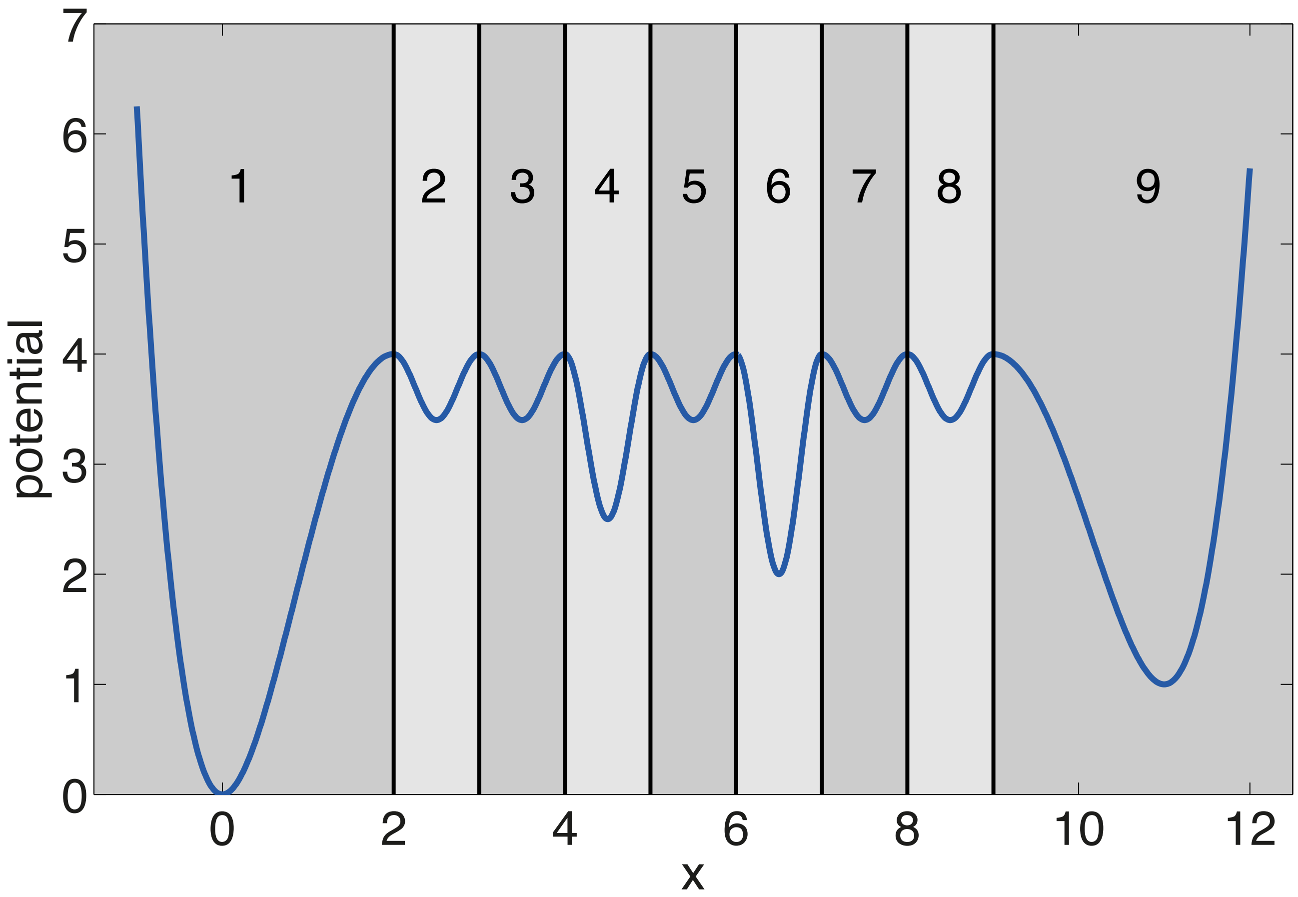


Noé, Schütte, Vanden-Eijnden, Reich, Weikl, PNAS, 2009.

$$dX_t = -\nabla V(X_t)dt + \sigma dB_t$$



Cluster state space into sets $A_1, \ldots A_n$ and consider $P_{ij} = \mathbb{P}_\mu[X_\tau \in A_j | X_0 \in A_i]$.



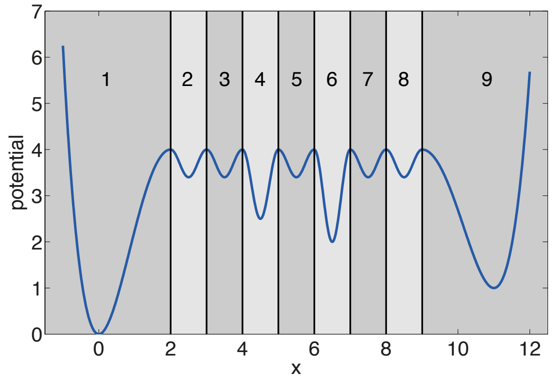
$D = span\{\mathbb{1}_{A_1}, ..., \mathbb{1}_{A_n}\}$

Q - orthogonal projection onto D

P can be considered as projection of the original transfer operator of the system

$$QT : D \to D,$$

$$T := T_\tau = e^{L\tau}, \qquad (Lv)(x) = \frac{1}{2}\sigma^2 \Delta v(x) - \nabla V(x) \cdot \nabla v(x).$$



Example: Implied timescales

$$\eta_i = -\frac{\tau}{\log \lambda_i}$$

λ_i - largest eigenvalues of transfer operator or its projection

	η_1	η_2	η_3	η_4
original	17.5267	3.1701	0.9804	0.4524
full partition	16.5478	2.9073	0.8941	0.4006

Theorem

Let T be a self-adjoint transfer operator and Q the orthogonal projection to a subspace D with $\mathbb{1} \in D$. Let λ be an eigenvalue of T, u a corresponding normalized eigenvector, and set $\delta = \|u - Qu\|$. Then there exists an eigenvalue $\hat{\lambda}$ of the projected transfer operator QT with

$$|\lambda - \hat{\lambda}| \leq \lambda_1 \delta (1 - \delta^2)^{-\frac{1}{2}},$$

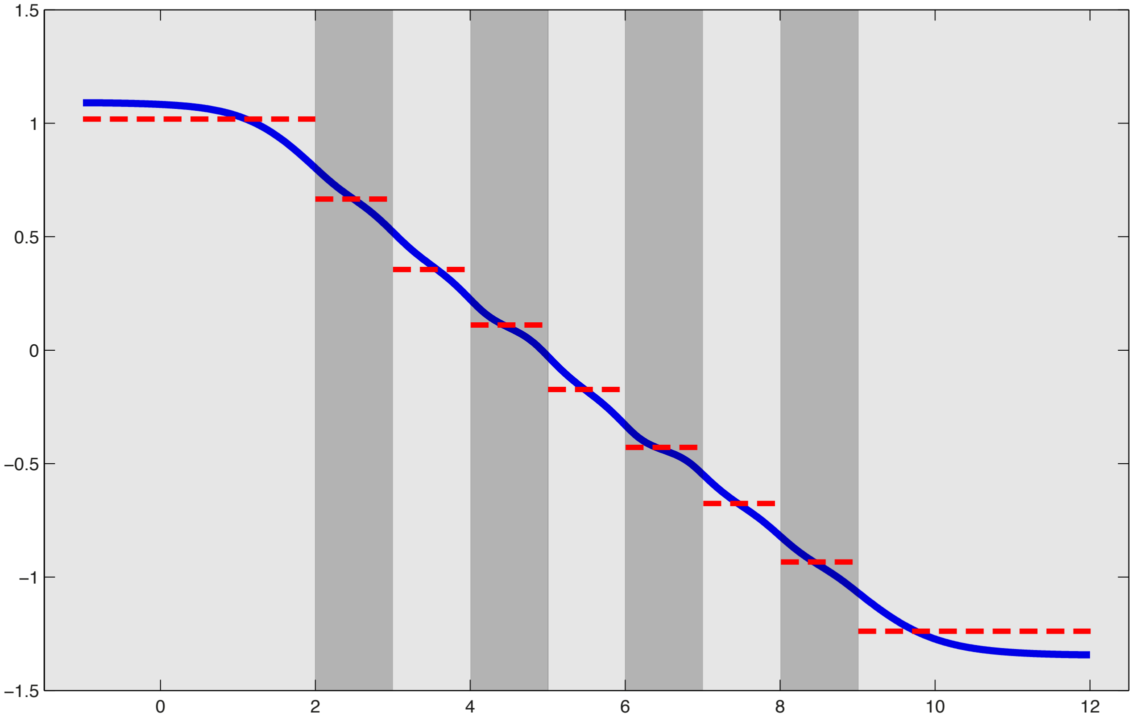
where $\lambda_1 < 1$ is the largest non-trivial eigenvalue of T.

In particular, for $\delta \leq \frac{3}{4}$ one can simplify the equation to

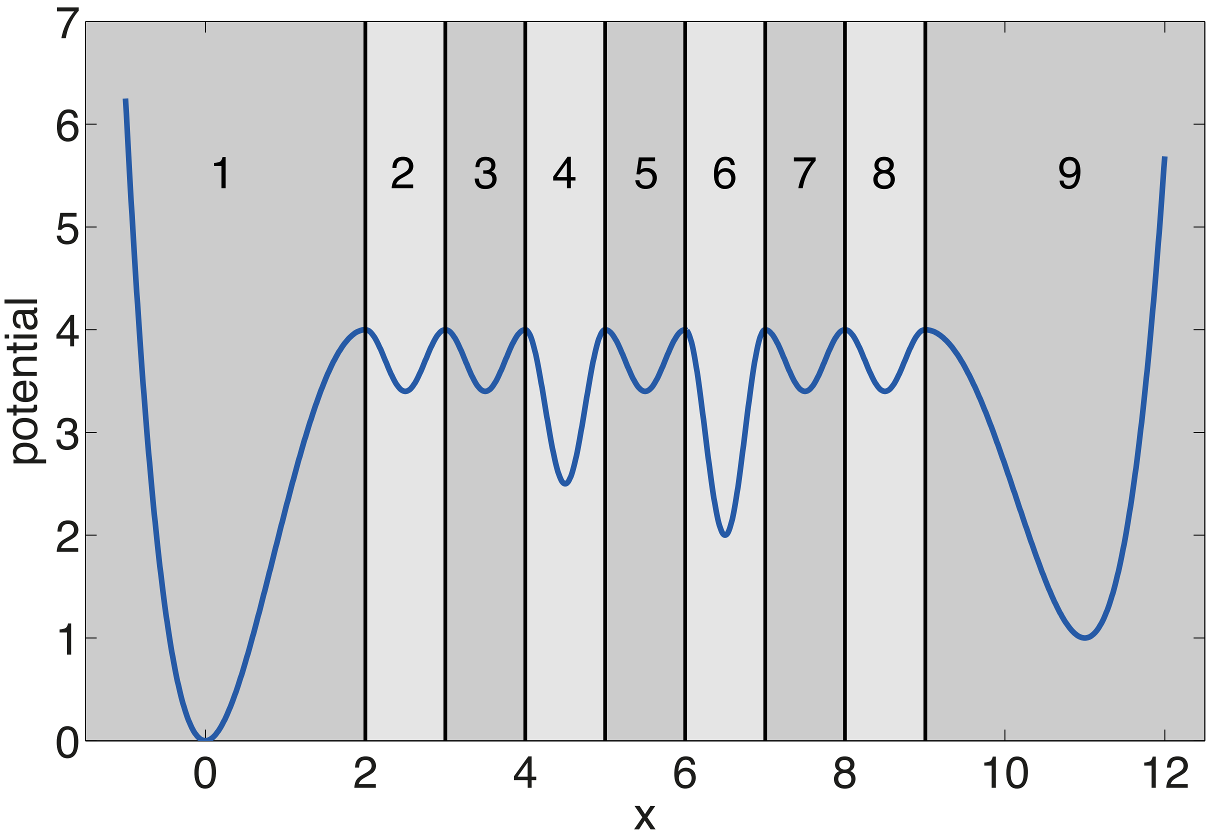
$$|\lambda - \hat{\lambda}| \leq 2\lambda_1 \delta.$$

Sarich, Schütte. Comm. Math. Sci., 2012.

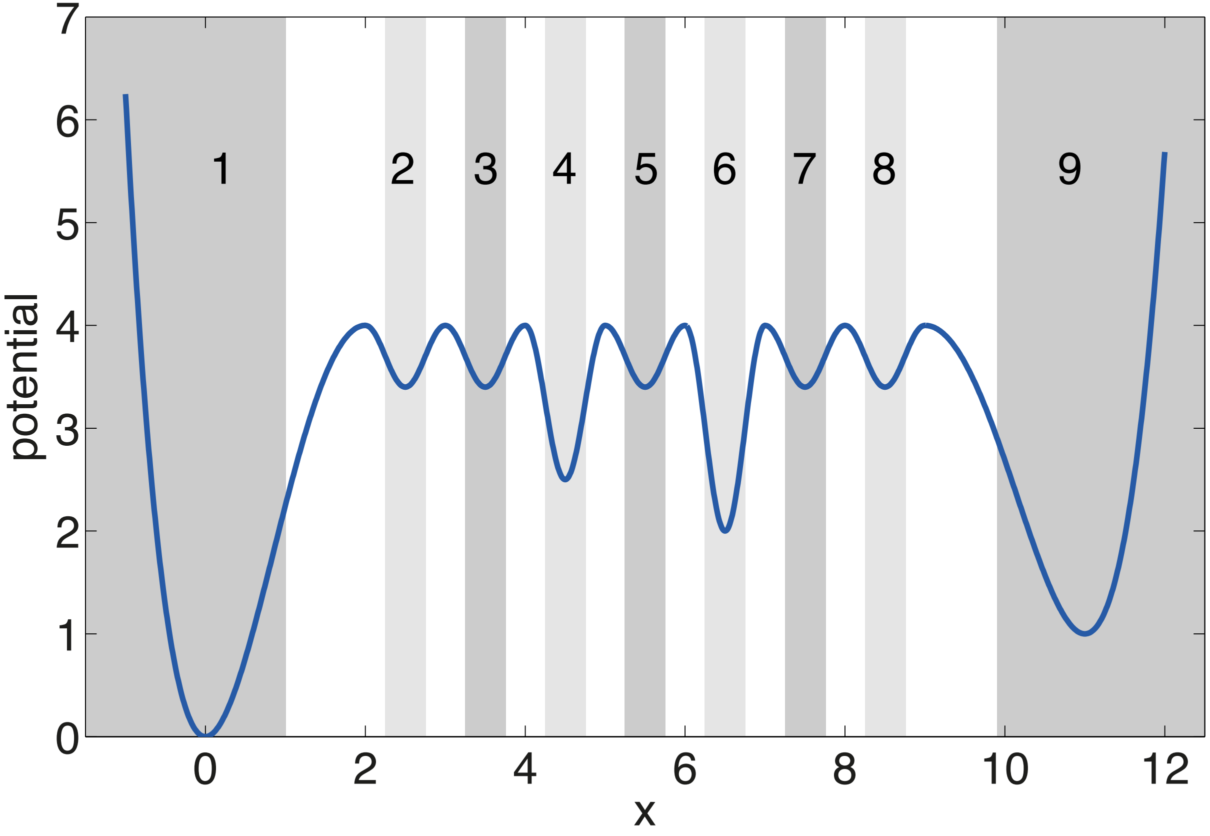
	η_1	η_2	η_3	η_4
original	17.5267	3.1701	0.9804	0.4524
full partition	16.5478	2.9073	0.8941	0.4006

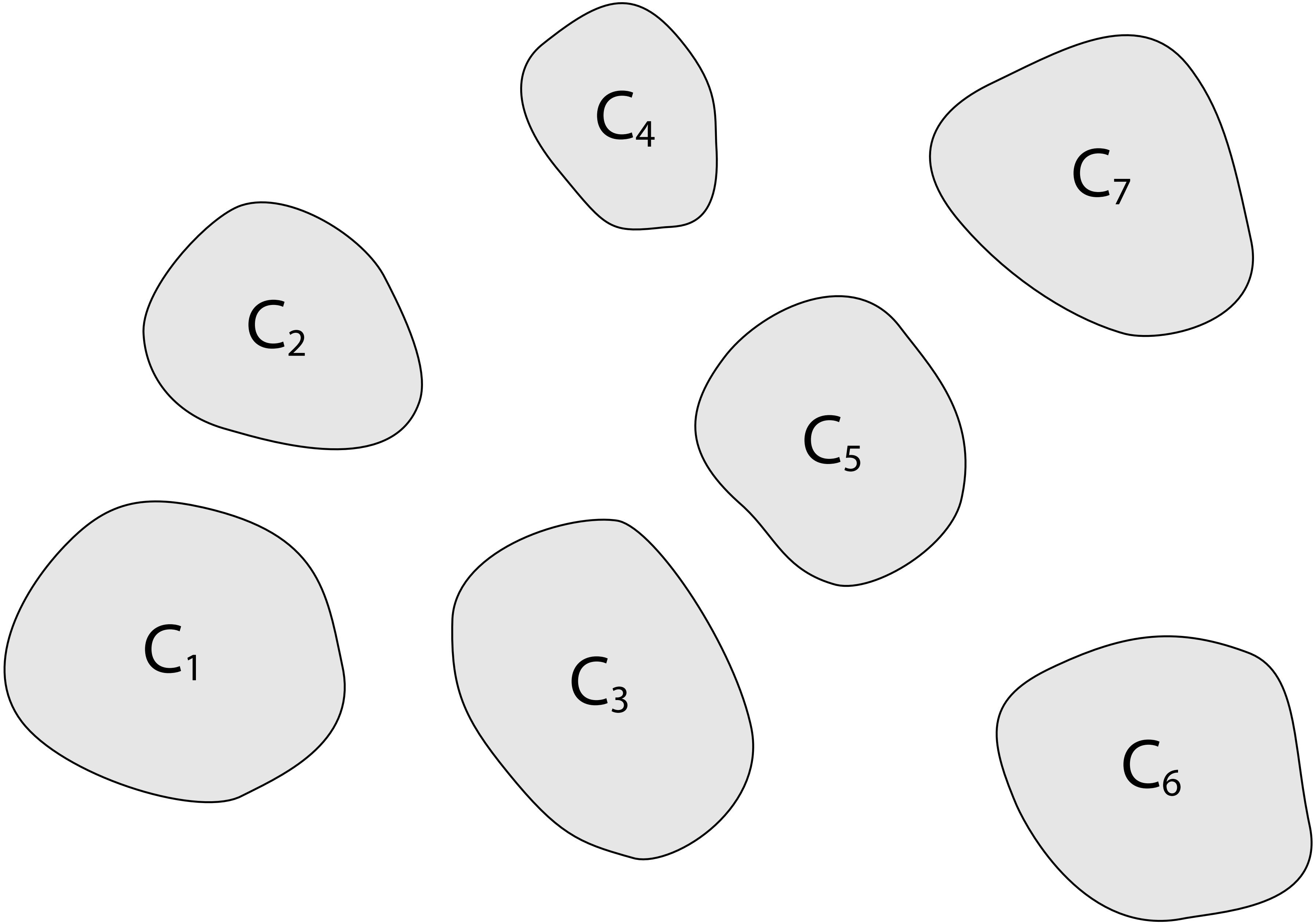


Use instead of a full partition....

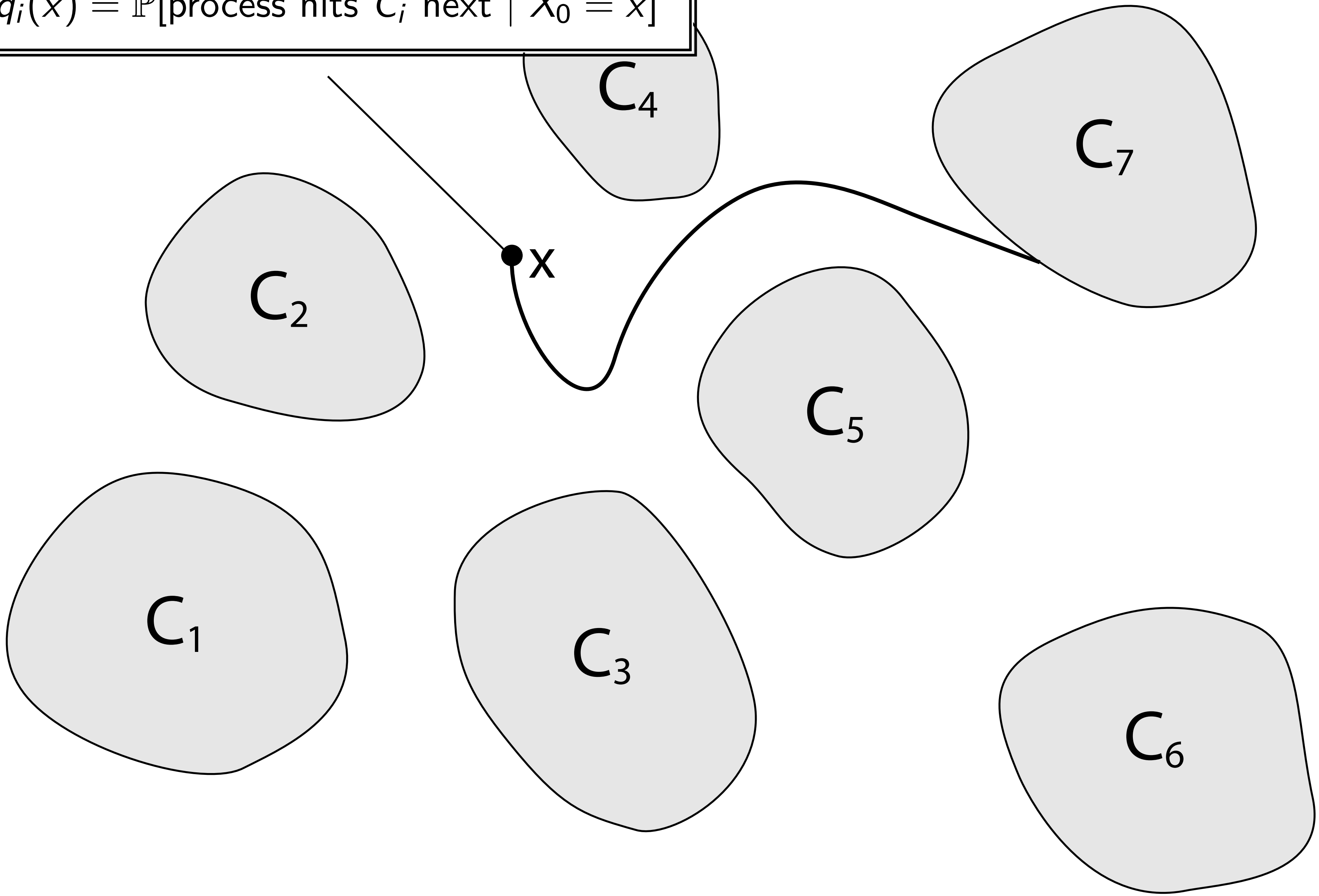


...so called **core sets** by cutting out a transition region.




C_4
C_7
C_2
C_5
C_1
C_3
C_6

$$q_i(x) = \mathbb{P}[\text{process hits } C_i \text{ next} \mid X_0 = x]$$

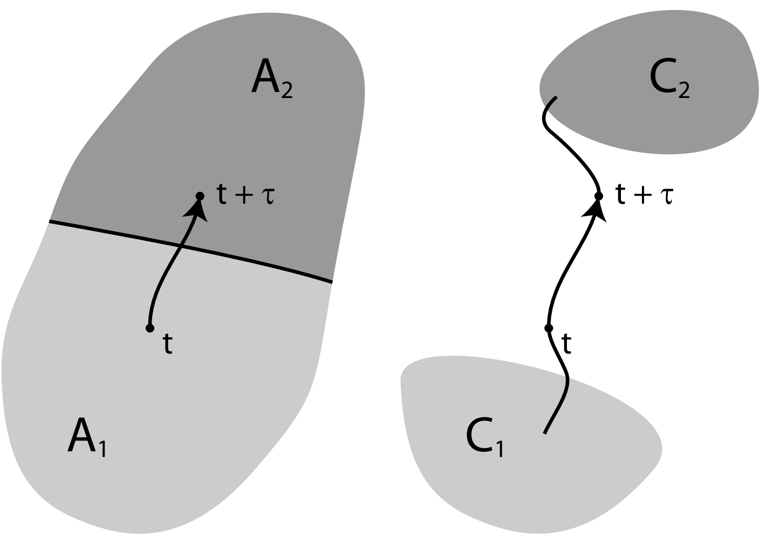


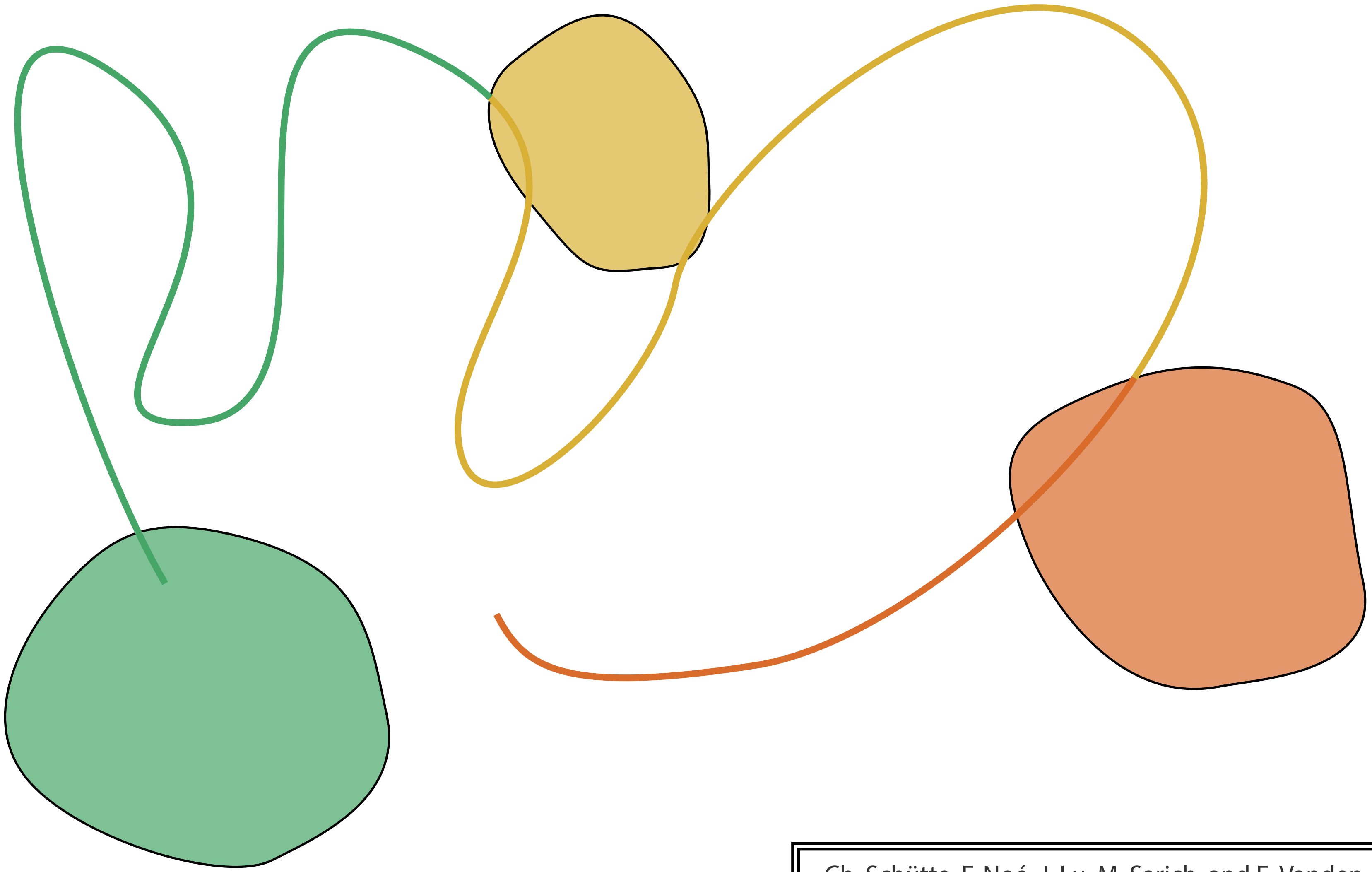
Projection of transfer operator QT onto $D = span\{q_1, ...q_n\}$ leads to matrix $P = \hat{T}M^{-1}$ with

$$M_{ij} = \mathbb{P}[\text{after time } t \text{ process will hit next } C_j \mid \text{at time } t \text{ process came last from } C_i]$$

and

$$\hat{T}_{ij} = \mathbb{P}[\text{after time } t+\tau \text{ process will hit next } C_j \mid \text{at time } t \text{ process came last from } C_i].$$





Ch. Schütte, F. Noé, J. Lu, M. Sarich, and E, Vanden-Eijnden. Markov State Models Based on Milestoning. J. Chem. Phys., 134 (20). 204105, 2011.

...and it can be proven to be more accurate!

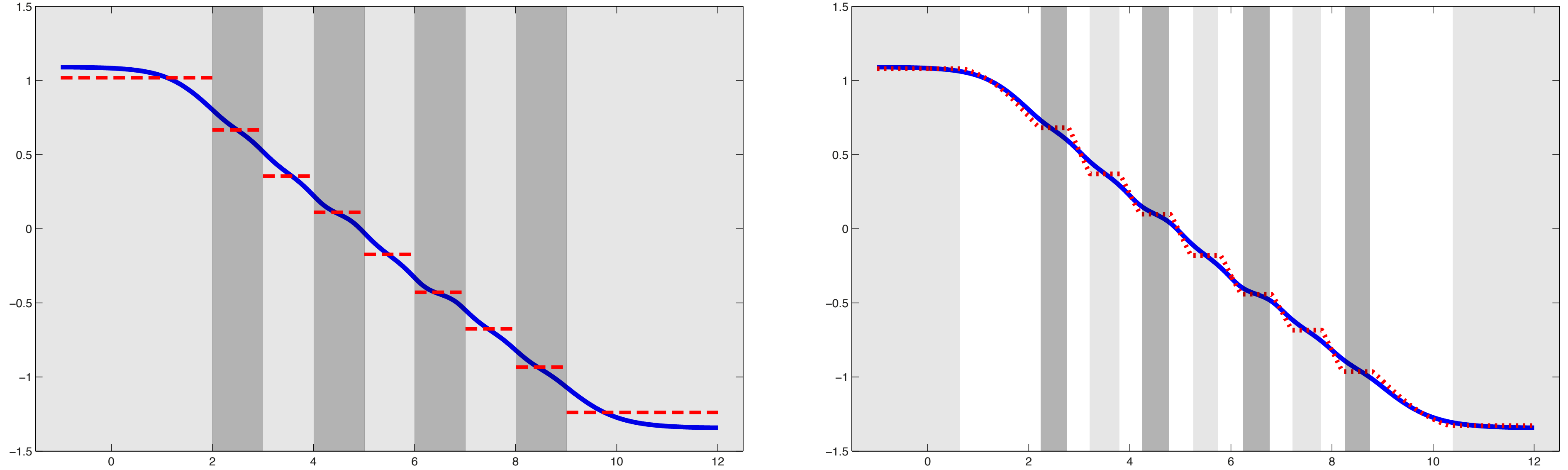
$$\|u - Q_{cores} u\| \leq \|u - Q_{full} u\| - \|(u - Q_{full} u)\big|_C\| + \frac{\rho}{\eta},$$

- $\rho \leq \max_{x \in C} \mathbb{E}_x[\tau(C^c)|]$
- η implied timescale belonging to u
- C region that is cut out

Sarich, Schütte. Comm. Math. Sci., 2012.

Example

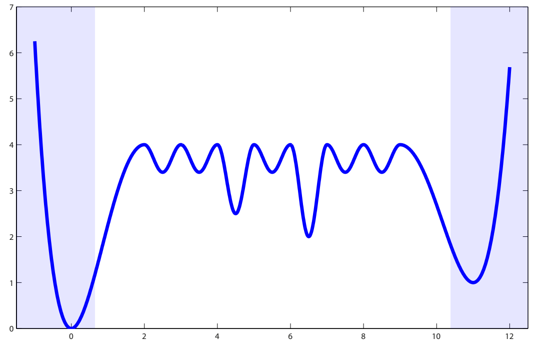
	η_1	η_2	η_3	η_4
original	17.5267	3.1701	0.9804	0.4524
core sets	17.3298	3.1332	0.9690	0.4430
full partition	16.5478	2.9073	0.8941	0.4006



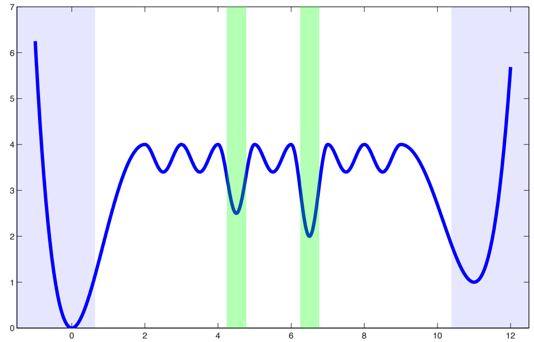
Using a multigrid

Let $\mathcal{C}_j = \{C_1^j, \ldots, C_{n_j}^j\}$ be an increasing sequence of core set discretizations, i.e.

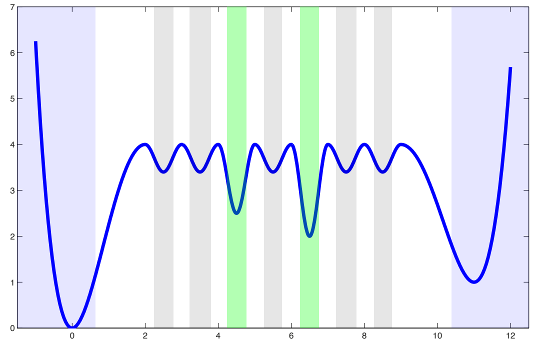
$$\mathcal{C}_i \subset \mathcal{C}_j \text{ for all } j \geq i.$$



Level 1



Level 2

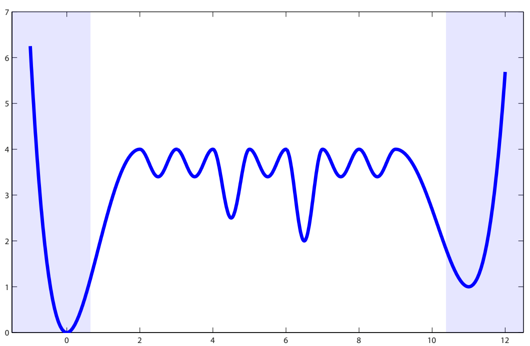


Level 3

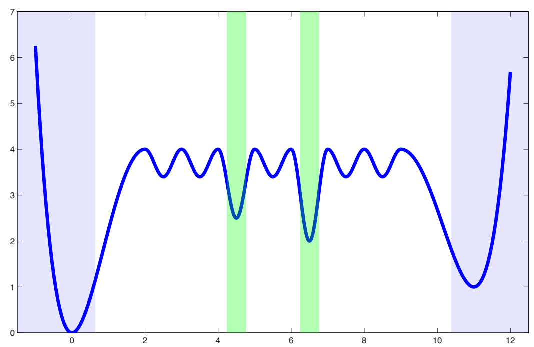
Using a multigrid

Let $\mathcal{C}_j = \{C_1^j, \ldots, C_{n_j}^j\}$ be an increasing sequence of core set discretizations, i.e.

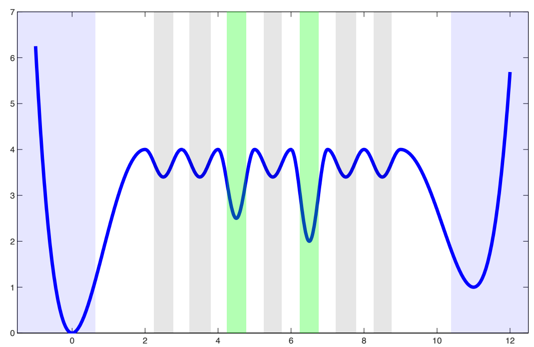
$$\mathcal{C}_i \subset \mathcal{C}_j \text{ for all } j \geq i.$$



Level 1



Level 2



Level 3

Iterative Space Construction

Level 1: $D_1 = span\{q_1, \ldots, q_n\}$, q_i usual committors.

Level $k+1$: $D_{k+1} = \hat{D}_{k+1} \cap D_k^\perp$, where $\hat{D}_{k+1}$ is the usual committor space for the cores on level $k+1$.

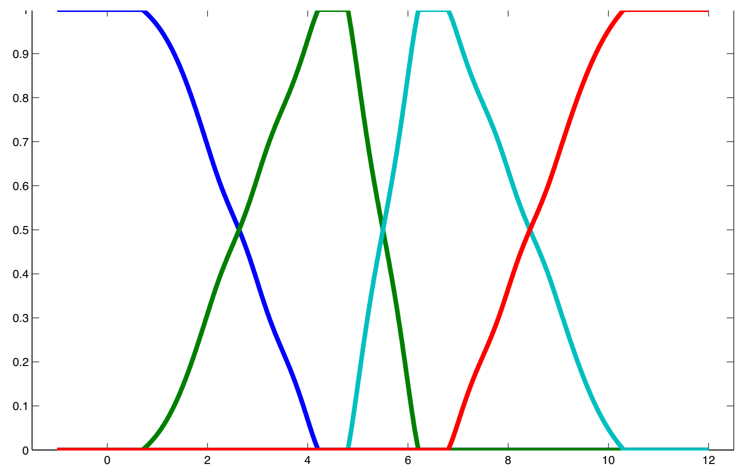
Total space for projection with m levels: $D = D_1 + \ldots + D_m$

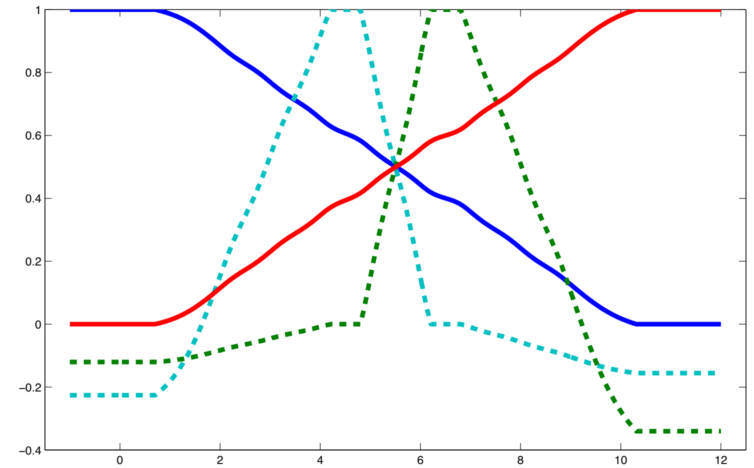
Computing the model

A matrix representation $P = \hat{T} M^{-1}$ can again be computed from stochastic quantities.

Assume C_i was first introduced on level k and C_j was first introduced on level l, then

$$M_{ij} = \mathbb{P}[\text{hit } C_j \text{ next on level } l \mid \text{came from } C_i \text{ on level } k].$$



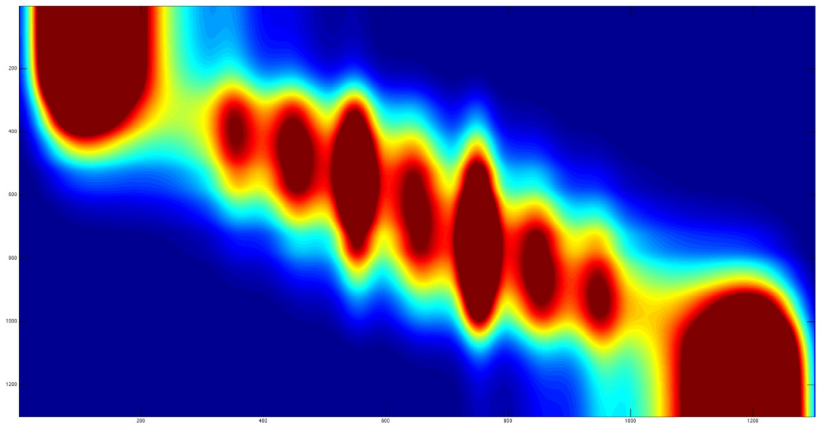


Mutilevel committors for 4 sets and 2 levels.

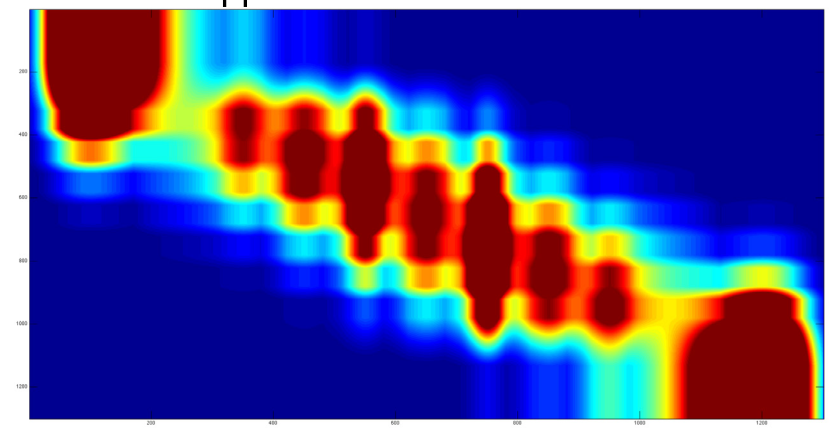
Example

	T_1	T_2	T_3	T_4
original	17.5267	3.1701	0.9804	0.4524
9 cores multigrid	17.5043	3.1579	0.9703	0.4441
9 core sets	17.3298	3.1332	0.9690	0.4430
full partition	16.5478	2.9073	0.8941	0.4006
2 core sets	17.5043	-	-	-

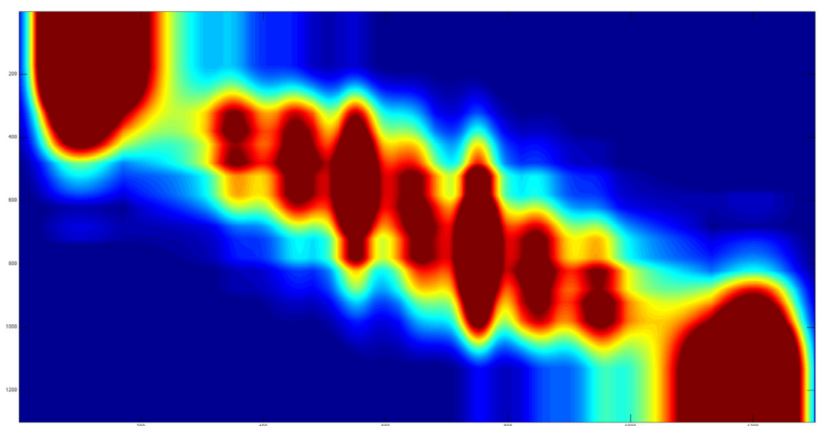
Original transfer operator



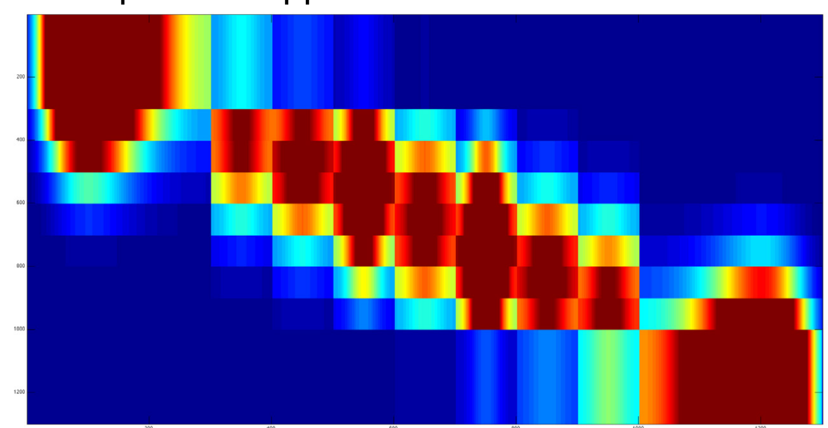
Core set approximation

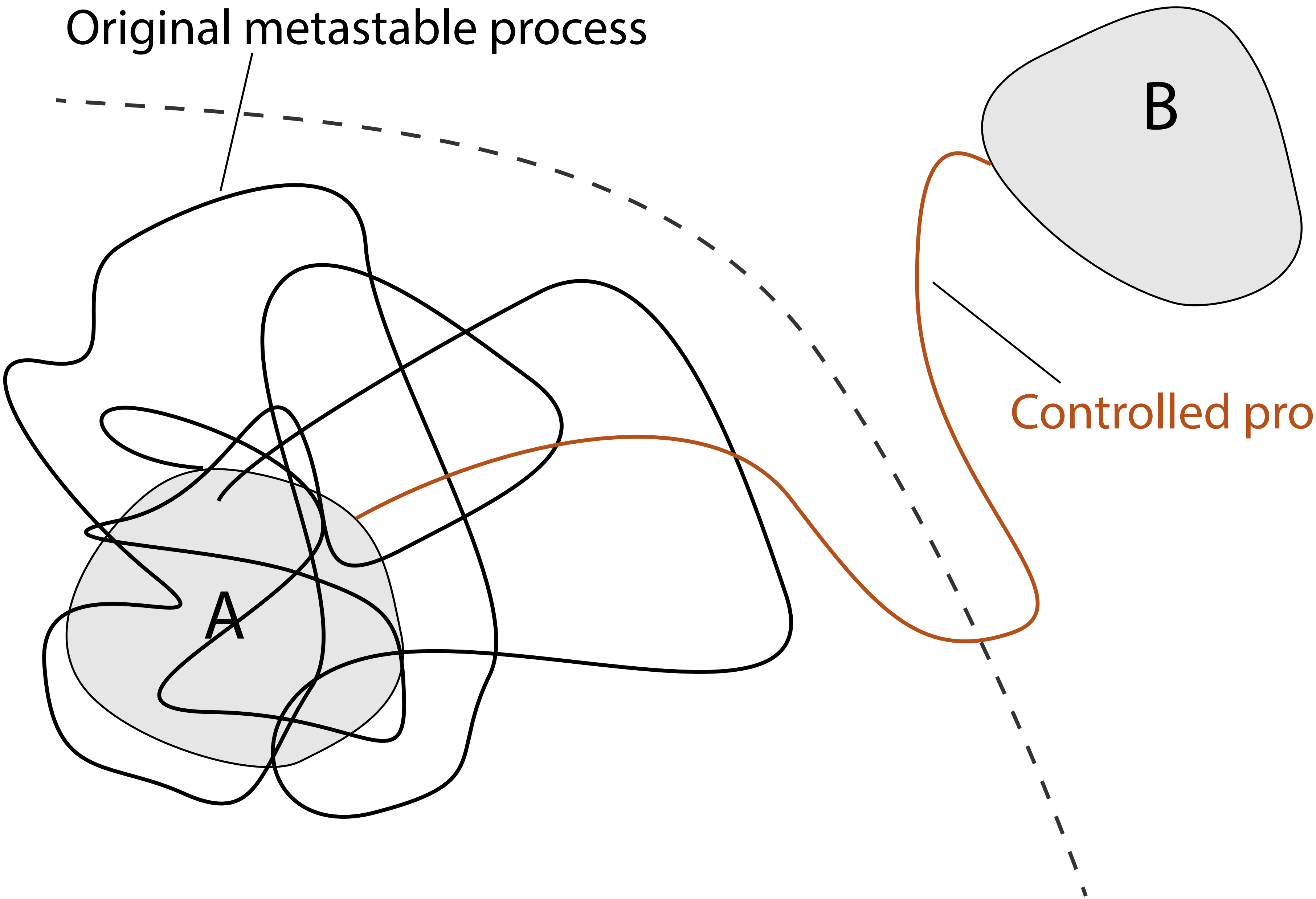


Multigrid approximation



Full partition approximation




Original metastable process
B
Controlled process
A

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}dB_t$$

Cumulant generating function

$$J(x) = -\epsilon \log \mathbb{E}_x[exp(-\tau(B)/\epsilon)] = \mathbb{E}_x[\tau(B)] + \epsilon$$

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon} dB_t$$

Cumulant generating function

$$J(x) = -\epsilon \log \mathbb{E}_x[exp(-\tau(B)/\epsilon)] = \mathbb{E}_x[\tau(B)] + \epsilon$$
$$= \min_{u \in U} \mathbb{E}_x^{\mathbb{P}_u} \left[\tau(B) + \int_0^{\tau(B)} \frac{1}{2} \|u_s\|^2 ds \right],$$

where $\mathbb{P}_u$ is the path measure belonging to

$$dX_t = (\sqrt{2} u_t - \nabla V(X_t))dt + \sqrt{2\epsilon} dB_t.$$

Solution has feedback form

If J is smooth enough, $u_t = -\sqrt{2}\nabla J(X_t)$ is the unique minimizer.

$$J(x) = \mathbb{E}_x^{\mathbb{P}_u}\left[\tau(B) + \int_0^{\tau(B)} \|\nabla J(X_s)\|^2 ds\right]$$

where $\mathbb{P}_u$ is the path measure belonging to

$$dX_t = -\nabla U(X_t)dt + \sqrt{2\epsilon}dB_t, \qquad U(x) = V(x) + 2J(x).$$

$$J(x) = \mathbb{E}_x^{\mathbb{P}_u}\left[\tau(B) + \int_0^{\tau(B)} \|\nabla J(X_s)\|^2 ds\right] =: (\mathcal{A}J)(x).$$

Idea

Fixed point iteration for the function $\mathcal{A}$ on a finite dimensional ansatz space, i.e. assuming

$$J(x) = \sum_{i=1}^{n} a_i \phi_i(x).$$

Another option

Writing

$$J(x) = -\epsilon \log \psi(x).$$

Then, ψ solves the linear system

$$\begin{aligned}(L - \epsilon^{-1}\mathbb{1}_B)\psi &= 0\\ \psi\big|_B &= 1.\end{aligned}$$

Another option

Writing

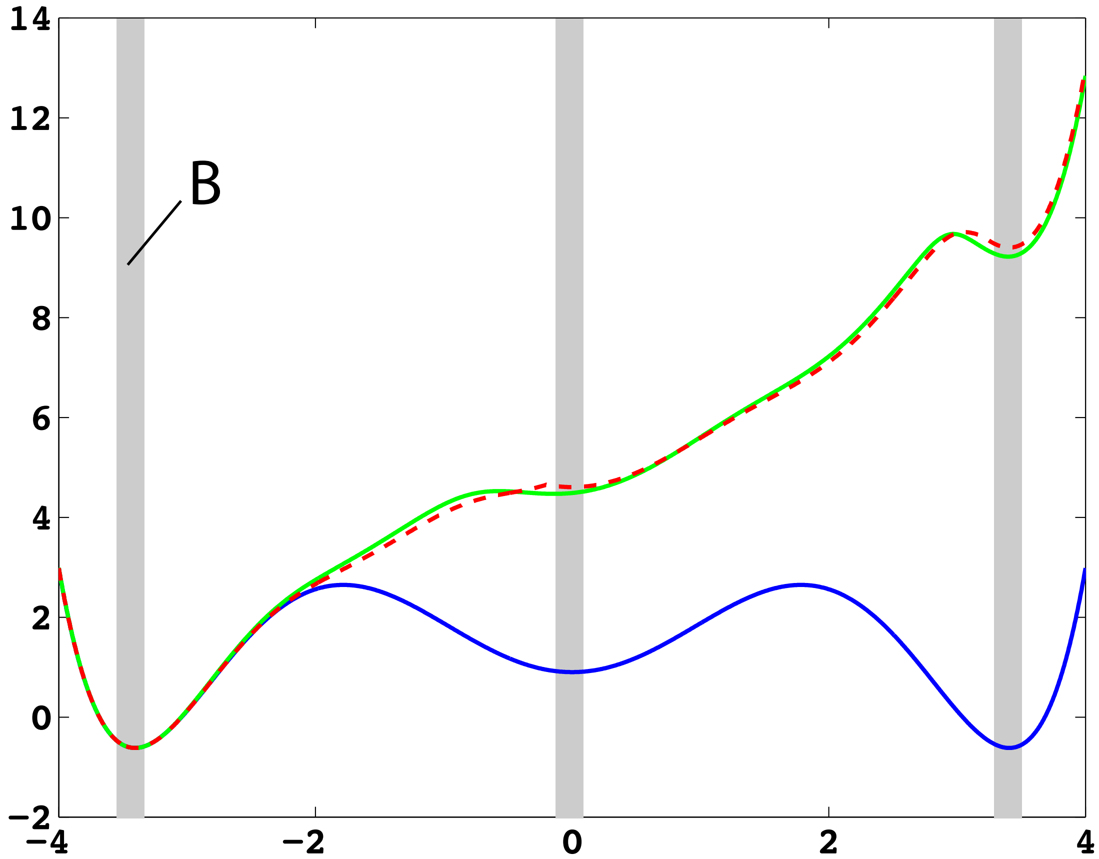
$$J(x) = -\epsilon \log \psi(x).$$

Then, ψ solves the linear system

$$(L - \epsilon^{-1}\mathbb{1}_B)\psi = 0$$
$$\psi\big|_B = 1.$$

Approximation by a Markov State Model possible!

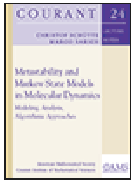
Hartmann, Schütte. J. Stat. Mech. Theor. Exp., 2012.
Sarich, Banisch, Hartmann, Schütte. Submitted to Entropy, 2013.



optimal control as lifted potential

computed via core set MSM using three core sets (grey)

Publications



Ch. Schütte and M. Sarich. Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches. Courant Lecture Notes 24, AMS, CIMS, 2013.

M. Sarich, R. Banisch, C. Hartmann, Ch. Schütte. Markov State Models for Rare Events in Molecular Dynamics. Entropy Special Issue „Molecular Dynamics Simulation“, 2013.

C. Hartmann, R. Banisch, M. Sarich, T. Badowski, and Ch. Schütte. Characterization of Rare Events in Molecular Dynamics. Entropy Special Issue „Molecular Dynamics Simulation“, 2013.

M. Sarich and Ch. Schütte. Approximating Selected Non-dominant Timescales by Markov State Models. Comm. Math. Sci., 10 (3). pp. 1001-1013, 2012.

Ch. Schütte, F. Noé, J. Lu, M. Sarich, and E, Vanden-Eijnden. Markov State Models Based on Milestoning. J. Chem. Phys., 134 (20). 204105, 2011.

J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Fischbach, M. Held, J. D. Chodera, Ch. Schütte, and F. Noé. Markov models of molecular kinetics: Generation and Validation. J. Chem. Phys., 134 . p. 174105, 2011.