Error analysis of target measure diffusion maps with applications to the committor problem

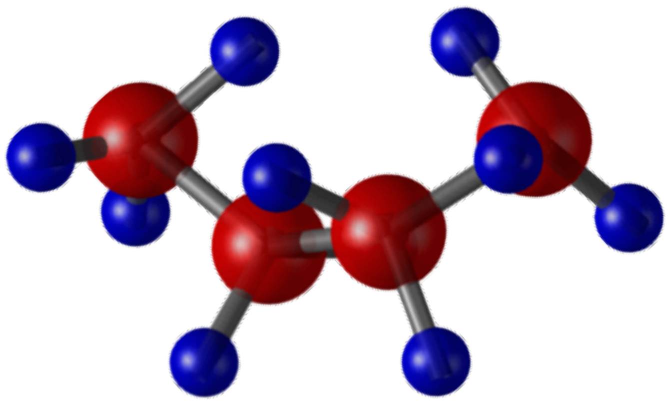
Shashank Sule (UMD, College Park)

Joint work with: Maria Cameron (UMD), Luke Evans (Flatiron Institute)

SS, Luke Evans
Maria Cameron
arXiv:2312.14418

Motivation: molecular dynamics

Butane:



Tasks:

1. Find the transition rates
2. Find low dimensional collective variables.

The overdamped Langevin dynamics

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dW_t$$

The invariant density is the Gibbs density

$$\mu(x) = Z^{-1}e^{-\beta V(x)}$$

The generator

$$\mathscr{L} = \beta^{-1}e^{\beta V(x)}\nabla \cdot \left(e^{-\beta V(x)}\nabla\right) = \beta^{-1}\Delta - \nabla V \cdot \nabla$$

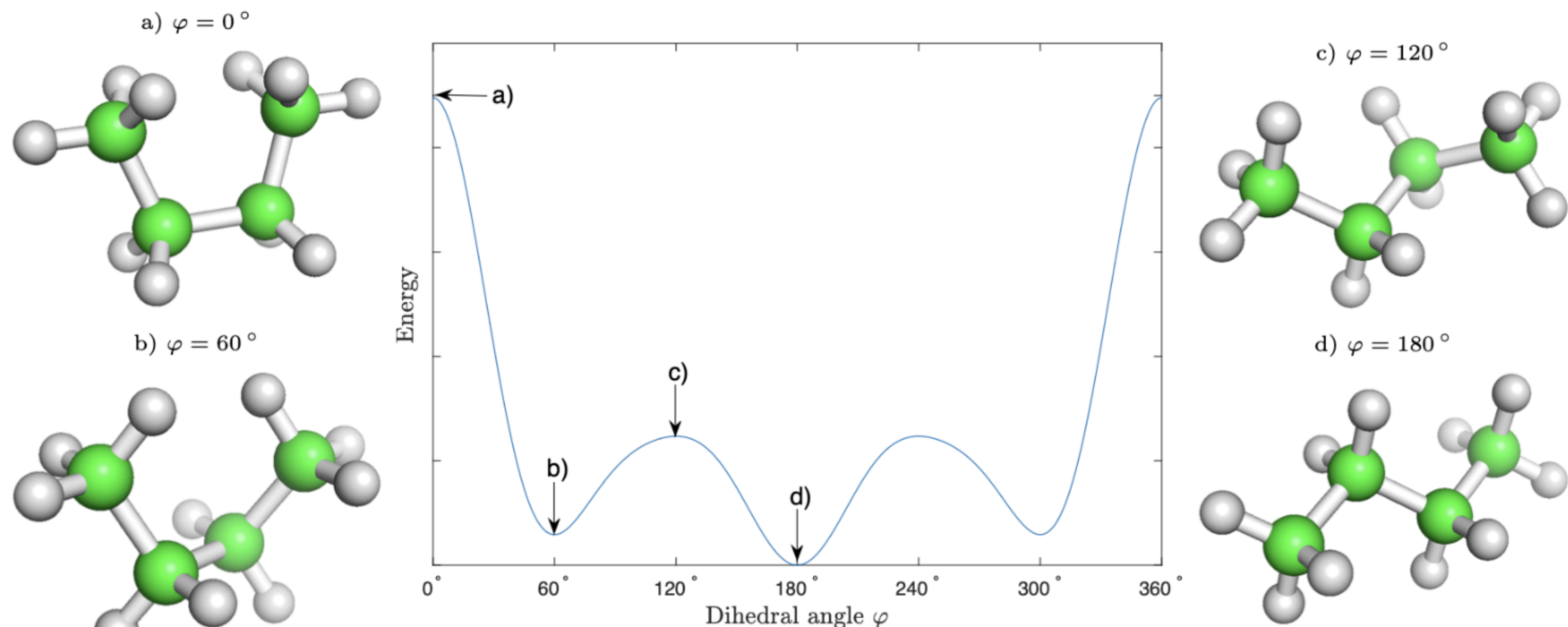


Figure from Klus Koltai Schutte '18

Typically, molecular systems exhibit (1) metastability, (2) ergodicity, and (3) low intrinsic dimensionality.

Tasks:

1. Find the transition rates:

(Vanden-Eijnden & E, 2006) The transition rate is given by

$$\nu_{AB} = \lim_{T\to\infty} \frac{N_{AB}}{T} = \int_{\Sigma_{AB}} J(x)\cdot n(x)\, d\sigma(x) = \beta^{-1}\int_{\mathcal{M}_{AB}} \|\nabla q(x)\|_2^2 \mu(x) d\mathrm{vol}(x)$$

Here q is the committor function which satisfies the **committor problem**

$$\mathcal{L}q(x) = 0, \quad x \in \mathcal{M}_{AB}, \quad q = 0, x \in \partial A, \quad q = 1, x \in \partial B$$

2. Find low dimensional coordinates

(Coifman, Kevrekedis, Maggioni, Nadler '08) The optimal low dimensional coordinates are given by the backward **Fokker-Planck eigenfunctions** $\{\psi_i\}_{i=1}^m$ satisfying

$$\mathcal{L}\psi_i = \lambda_i \psi_i, \quad \partial_n \psi_i = 0$$

New Task: Approximate $\mathcal{L}$ on data $\mathcal{X} \subseteq \mathcal{M}$ and discretize the relevant PDE.

New Task: Approximate $\mathscr{L}$ on data $\mathscr{X} \subseteq \mathscr{M}$ and discretize the relevant PDE.

Approach: Use Diffusion Map (Coifman and Lafon '06)

The Gaussian kernel

$$k_\epsilon(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}$$

The kernel density estimate

$$\rho_\epsilon^{(n)}(x_i) = \frac{1}{n}\sum_{j=1}^{n} k_\epsilon(x_i, x_j)$$

The reweighted kernel

$$k_{\epsilon,\alpha}(x_i, x_j) = \frac{e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}}{\left(\rho_\epsilon^{(n)}\right)^\alpha (x_j)}$$

The Markov operator

$$\mathscr{P}_{\epsilon,\alpha} f(x) = \frac{\int_{\mathscr{M}} k_{\epsilon,\alpha}(x_i, y) f(y)\rho(y)\, dy}{\int_{\mathscr{M}} k_{\epsilon,\alpha}(x_i, y)\rho(y)\, dy}$$

The generator

$$\mathscr{L}_{\epsilon,\alpha} f(x) = \frac{\mathscr{P}_{\epsilon,\alpha} f(x) - f(x)}{\epsilon}$$

$$= \frac{f(x)\rho(x) + \frac{\epsilon}{2}\Delta(f(x)\rho(x)) + O(\epsilon^2)}{\rho(x) + \frac{\epsilon}{2}\Delta\rho(x) + O(\epsilon^2)}$$

$$(\pi\epsilon)^{-d/2}\int_{\mathbb{R}^d} k_\epsilon(x, y)\, g(y)\, dy = g + \frac{\epsilon}{4}\Delta g + O(\epsilon^2)$$

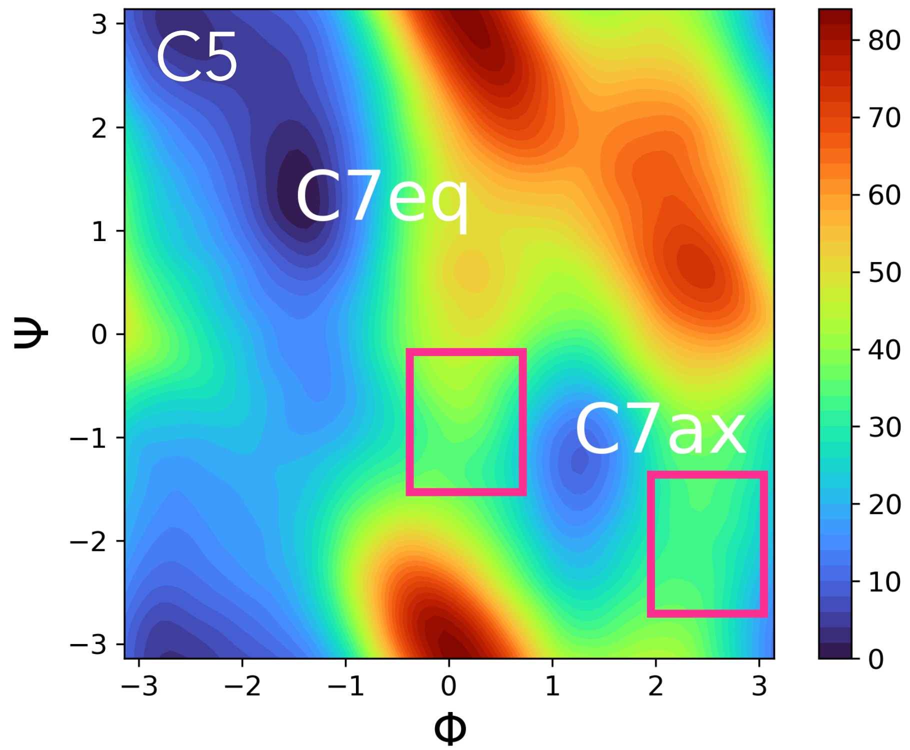
$$\lim_{\epsilon\to 0}\lim_{n\to\infty} L_{\epsilon,\alpha} = \Delta f + (2 - 2\alpha)\langle \nabla \log \rho, \nabla f \rangle$$

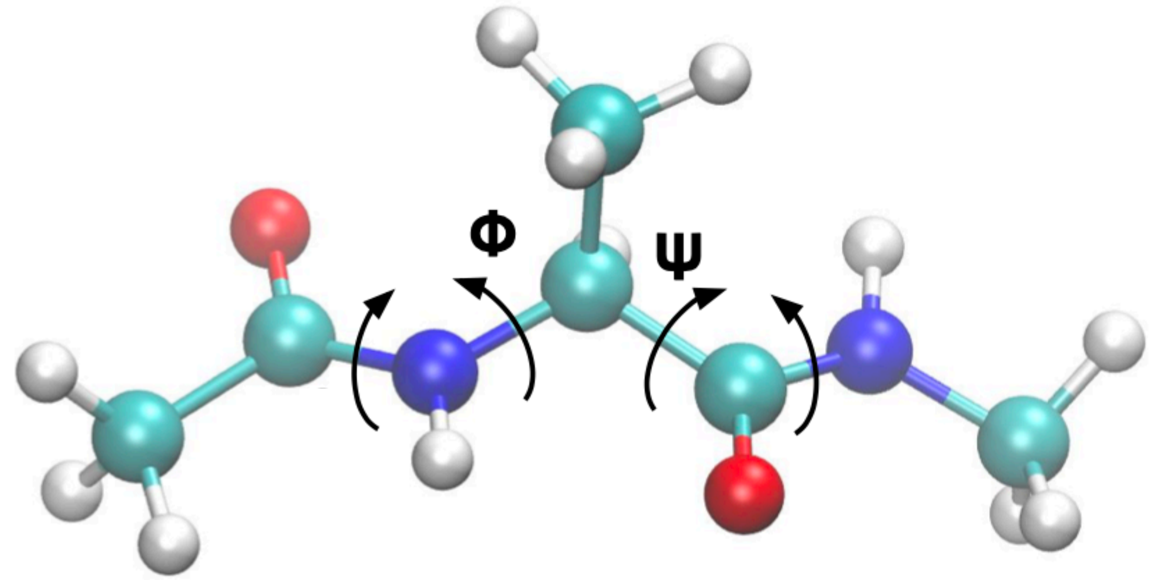
$$\alpha = \frac{1}{2}, \quad \rho = Z^{-1} e^{-\beta V} \quad \Rightarrow \quad \lim_{\epsilon\to 0}\lim_{n\to\infty} L_{\epsilon,\frac{1}{2}} f = \frac{\beta}{4}\left(\beta^{-1}\Delta f - \nabla V \cdot \nabla f\right)$$

WARNING. DMap approximates the generator of the overdamped Langevin dynamics **ONLY WITH data ~ Gibbs density**

DMap approximates the generator of the overdamped Langevin dynamics

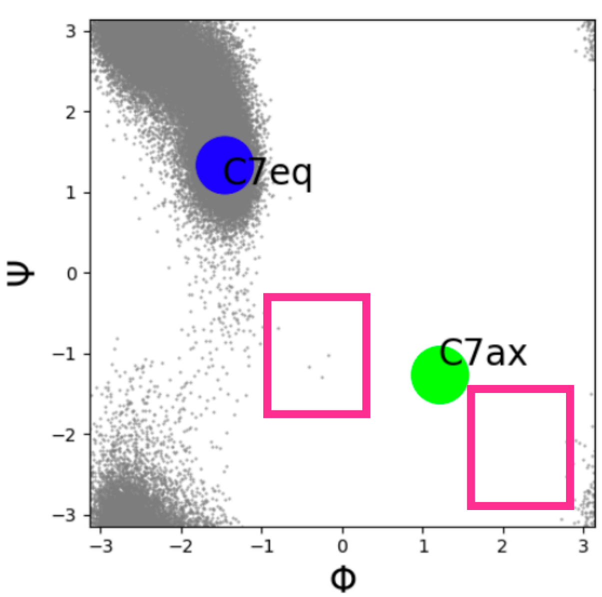
ONLY WITH data ~ Gibbs density



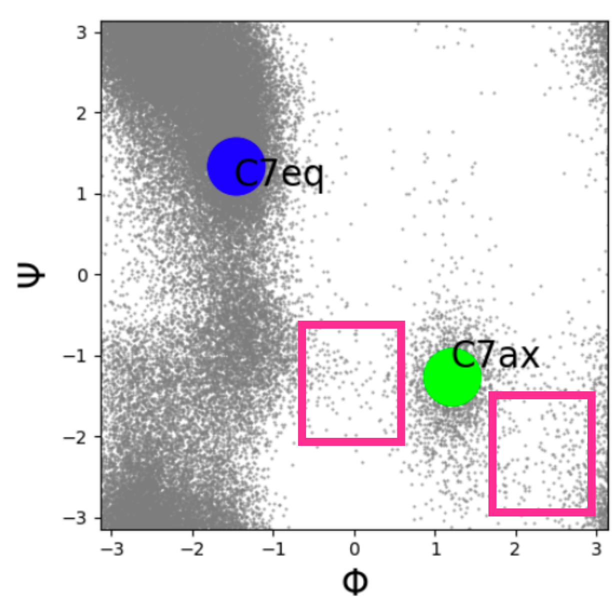


MD simulations by Luke Evans

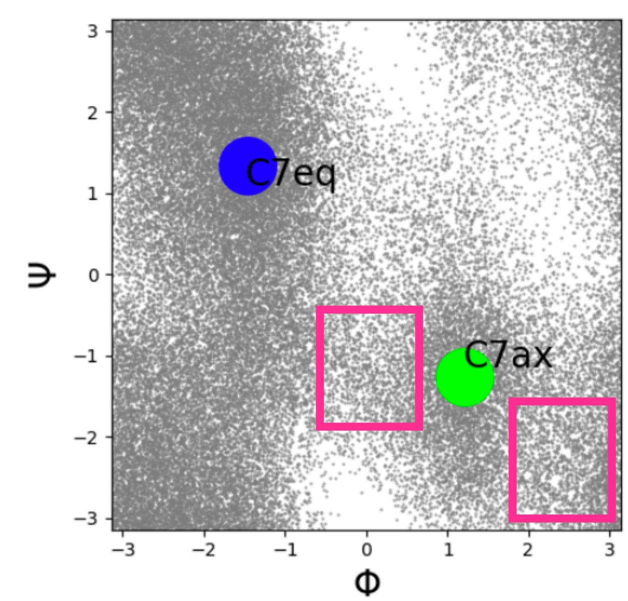
10ns trajectory, 300K



10ns trajectory, 500K



10ns trajectory, Metadynamics



Samples from Gibbs density will result in poor accuracy in the transition region!

We need to use enhanced sampling data

We need to use enhanced sampling data

How to combine this with diffusion map?

Answer: Target Measure Diffusion (TMD) Maps (Banisch et. al 2020)!

The Gaussian kernel

$$k_\epsilon(x, y) = e^{-\frac{\|x-y\|^2}{\epsilon}}$$

The reweighted kernel

$$k_{\epsilon,\alpha}(x_i, x_j) = \frac{e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}}{\left(\rho_\epsilon^{(n)}\right)^\alpha (x_j)}$$

$$k_{\epsilon,\mu}(x_i, x_j) = \frac{e^{-\frac{\|x_i - x_j\|^2}{\epsilon}}}{\rho_\epsilon^{(n)}(x_j)} \mu^{1/2}(x_j)$$

The kernel density estimate

$$\rho_\epsilon^{(n)}(x_i) = \frac{1}{n} \sum_{y_i} k_\epsilon(x, y_i)$$

The Markov operator

$$P_{\epsilon,\mu}^{(n)} f(x) = \frac{n^{-1} \sum_{j=1}^n k_{\epsilon,\mu}(x_i, x_j) f(x_j)}{n^{-1} \sum_{j=1}^n k_{\epsilon,\mu}(x_i, x_j)}$$

The generator

$$L_{\epsilon,\mu}^{(n)} f(x) = \frac{P_{\epsilon,\mu} f(x) - f(x)}{\epsilon}$$

Theorem [Banisch et al. 2020] As $n \to \infty$, $\forall x \in \Omega$

$$4\beta^{-1} L_{\epsilon,\mu}^{(n)} f(x) \to \mathscr{L}_{\epsilon,\mu} = \mathscr{L} + O(\epsilon).$$

TMD map = Dmap + Importance sampling

Key advantage of TMD map: the sampling density ρ can be arbitrary!

TMD map = Dmap + Importance sampling

Theorem [Banisch et al.] As $\epsilon \to 0$ and $n \to \infty$, $\forall x \in \Omega$

$$\mathscr{L}_{\epsilon,\mu} f(x) \to \frac{1}{4}\left(\Delta f + \nabla \log \mu \cdot \nabla f\right) \equiv \frac{\beta}{4}\mathscr{L}.$$

Questions:

1. How to choose ρ, ϵ?

2. Can we quantify the **consistency error**

$$|\,4\beta^{-1}L^{(n)}_{\epsilon,\mu} f(x) - \mathscr{L}f(x)\,| = E(n, \epsilon; \rho, \mu, f, x)$$

3. Can we quantify the **solution error** $|\,q_{n,\epsilon} - q\,|$ where

$$4\beta^{-1}L^{(n)}_{\epsilon,\mu} q(x_i) = 0, \quad x_i \in \mathscr{M}\backslash(A \cup B),$$
$$q(x_i) = 0, \quad x_i \in A,$$
$$q(x_i) = 1, \quad x_i \in B$$

Bias error and variance error

$$|4\beta^{-1}L^{(n)}_{\epsilon,\mu}f(x) - \mathscr{L}f(x)| = E(n,\epsilon;\rho,\mu,f,x)$$

Bias error: $4\beta^{-1}\mathscr{L}_{\epsilon,\mu}f(x) - \mathscr{L}f(x)$

Variance error: $L^{(n)}_{\epsilon,\mu}f(x) - \mathscr{L}_{\epsilon,\mu}f(x)$

Discrete kernel density estimate: $$\rho^{(n)}_{\epsilon}(x) = \frac{1}{n}\sum_{j=1}^{n} k_{\epsilon}(x,x_j)$$

Discrete TMDmap generator: $$[L^{(n)}_{\epsilon,\mu}f](x_i) = \frac{1}{\epsilon}\left(\frac{\sum_{j=1}^{n}\frac{k_\epsilon(x_i,x_j)\mu^{1/2}(x_j)f(x_j)}{\rho^{(n)}_\epsilon(x_j)}}{\sum_{j=1}^{n}\frac{k_\epsilon(x_i,x_j)\mu^{1/2}(x_j)}{\rho^{(n)}_\epsilon(x_j)}} - f(x_i)\right)$$

Continuous TMDmap generator: $$[\mathscr{L}_{\epsilon,\mu}f](x_i) = \frac{1}{\epsilon}\left(\frac{\int_{\mathscr{M}}\frac{k_\epsilon(x_i,y)\mu^{1/2}(y)f(y)}{\rho_\epsilon(y)}dy}{\int_{\mathscr{M}}\frac{k_\epsilon(x_i,y)\mu^{1/2}(y)}{\rho_\epsilon(y)}dy} - f(x_i)\right)$$

Exact generator: $$[\mathscr{L}f](x) = \beta^{-1}\Delta f(x) + \nabla\log\mu(x)\cdot\nabla f(x)$$

The **bias** **and** **variance** **errors** **including** **prefactors.**

Scaling between ε **and** n

SS, Luke Evans
Maria Cameron
arXiv:2312.14418

If the manifold is locally flat,

$$\mathcal{Q} = \partial_i^i \partial_j^j f(x) + 2\partial_{ii}^{ii} f(x).$$

The main theorem

Let $\mathcal{M}$ be a compact d-dimensional manifold without boundary. Let $\mathcal{X}(n) \subset \mathcal{M}$ be a point cloud sampled i.i.d. with density ρ, $0 < \rho_{\min} \le \rho(x) \le \rho_{\max} < \infty$. Let $x \in \mathcal{M}$ be an arbitrary point. Let $f \in C^2(\mathcal{M})$ be an arbitrary function. Furthermore, let ϵ be the kernel bandwidth and μ be the target density used for constructing the TMDmap generator $L^{(n)}_{\epsilon,\mu}$. Then as $n \to \infty$ and $\epsilon \to 0$ so that

$$\lim_{\substack{n\to\infty \\ \epsilon \to 0}} \frac{n\epsilon^{2+d/2}}{\log n} = \infty,$$

with probability greater than $1 - 2n^{-3}$, we have:

$$|4\beta^{-1} L^{(n)}_{\epsilon,\mu} f(x) - \mathcal{L}f(x)| \le \underbrace{\frac{\alpha\epsilon}{\rho^{1/2}(x_i)}\left(2\|\nabla f(x)\|\epsilon^{1/2} + 11|f(x)|\right)}_{\text{variance error}}$$
$$+ \underbrace{\epsilon\,|\mathcal{B}_1[f,\mu] + \mathcal{B}_2[f,\mu,\rho] + \mathcal{B}_3[f,\mu,\rho]| + O(\epsilon^2)}_{\text{bias error}}.$$

The expressions for $\alpha, \mathcal{B}_1, \mathcal{B}_2$, and $\mathcal{B}_3$ are given by:

$$\alpha = \frac{1}{(2\pi)^{d/4}}\sqrt{\frac{\log n}{n\epsilon^{4+d/2}}},$$

$$\mathcal{B}_1[f,\mu] := \frac{1}{4}\left[\mathcal{Q}\left(f\mu^{1/2}\right) - f\mathcal{Q}(\mu^{1/2})\right] + \frac{1}{16}\left(2\nabla f\cdot\nabla\left(\mu^{1/2}\omega\right) + (\mu^{1/2}\omega)\Delta f\right),$$

$$\mathcal{B}_2[f,\mu,\rho] := -\frac{1}{16}\left(2\nabla f\cdot\nabla\left(\mu^{1/2}\frac{\Delta\rho}{\rho}\right) + \left(\mu^{1/2}\frac{\Delta\rho}{\rho}\right)f\right),$$

$$\mathcal{B}_3[f,\mu,\rho] := \frac{1}{16}\left[\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \left(\frac{\Delta\rho}{\rho} - \omega\right)\right]\left[f\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \frac{\Delta(\mu^{1/2}f)}{\mu^{1/2}}\right].$$

Here, $\mathcal{Q}$ is a non-linear differential operator and ω is a smooth function on $\mathcal{M}$.

Why is the uniform sampling density good?

SS, Luke Evans
Maria Cameron
arXiv:2312.14418

If the manifold is locally flat,

$\mathcal{Q} = \partial_i^i \partial_j^j f(x) + 2\partial_{ii}^{ii} f(x).$

The main theorem

Let $\mathcal{M}$ be a compact d-dimensional manifold without boundary. Let $\mathcal{X}(n) \subset \mathcal{M}$ be a point cloud sampled i.i.d. with density ρ, $0 < \rho_{\min} \le \rho(x) \le \rho_{\max} < \infty$. Let $x \in \mathcal{M}$ be an arbitrary point. Let $f \in C^2(\mathcal{M})$ be an arbitrary function. Furthermore, let ϵ be the kernel bandwidth and μ be the target density used for constructing the TMDmap generator $L_{\epsilon,\mu}^{(n)}$. Then as $n \to \infty$ and $\epsilon \to 0$ so that

$$\lim_{\substack{n\to\infty \\ \epsilon\to 0}} \frac{n\epsilon^{2+d/2}}{\log n} = \infty,$$

If f is the committor, the manifold is flat, and the sampling density ρ is uniform:

with probability greater than $1 - 2n^{-3}$, we have:

$$|4\beta^{-1} L_{\epsilon,\mu}^{(n)} f(x) - \mathcal{L}f(x)| \le \underbrace{\frac{\alpha\epsilon}{\rho^{1/2}(x_i)} \left(2\|\nabla f(x)\|\epsilon^{1/2} + 11|f(x)|\right)}_{\text{variance error}}$$

← Minimized!

$$+ \underbrace{\epsilon\,|\mathcal{B}_1[f,\mu] + \mathcal{B}_2[f,\mu,\rho] + \mathcal{B}_3[f,\mu,\rho]| + O(\epsilon^2)}_{\text{bias error}}.$$

The expressions for $\alpha, \mathcal{B}_1, \mathcal{B}_2$, and $\mathcal{B}_3$ are given by:

$$\alpha = \frac{1}{(2\pi)^{d/4}} \sqrt{\frac{\log n}{n\epsilon^{4+d/2}}},$$

If ω=0

$$\mathcal{B}_1[f,\mu] := \frac{1}{4}\left[\mathcal{Q}\left(f\mu^{1/2}\right) - f\mathcal{Q}(\mu^{1/2})\right] + \frac{1}{16}\left(2\nabla f \cdot \nabla\left(\mu^{1/2}\omega\right) + (\mu^{1/2}\omega)\Delta f\right),$$

$$\mathcal{B}_2[f,\mu,\rho] := -\frac{1}{16}\left(2\nabla f \cdot \nabla\left(\mu^{1/2}\frac{\Delta\rho}{\rho}\right) + \left(\mu^{1/2}\frac{\Delta\rho}{\rho}\right)f\right),$$

If ρ=const

$$\mathcal{B}_3[f,\mu,\rho] := \frac{1}{16}\left[\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \left(\frac{\Delta\rho}{\rho} - \omega\right)\right]\left[f\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \frac{\Delta(\mu^{1/2}f)}{\mu^{1/2}}\right].$$

If f=committor

Here, $\mathcal{Q}$ is a non-linear differential operator and ω is a smooth function on $\mathcal{M}$.

Error in the TMDmap solution to BVPs

$$\begin{cases} \mathscr{L}u = f, & x \in \Omega \\ u = g, & x \in \partial\Omega \end{cases}$$

u = The exact solution
$v_{n,\varepsilon}$ = The TMDmap solution

Theorem [SS, Evans, Cameron, 2023]

$$\lim_{\substack{n\to\infty \\ \epsilon\to 0}} \frac{n\epsilon^{2+d/2}}{\log n} = \infty.$$

Then $\exists \epsilon_0 > 0$ and $\exists n_0 \in \mathbb{N}$: $\forall \epsilon \leq \epsilon_0$ and $n \geq n_0$ with probability $\geq 1 - 4n^{-2} - \exp\left(-n\mathbb{E}[\mathbb{I}_\Omega]\right)$

$$|v_{n,\epsilon}(x_i) - u(x_i)| \leq \epsilon \left[\max_{x\in\mathcal{M}} C_{u,\mu,\rho}(x) + 1\right] |\phi(x_i)|$$

where ϕ is the exact solution to $\mathcal{L}\phi = 1$, $x \in \mathcal{M}\backslash\Omega$ and $\phi = 0$, $x \in \Omega$, and

$$C_{u,\mu,\rho}(x) := \max_{\substack{\epsilon\leq\epsilon_0 \\ n\geq n_0}} \left[\frac{\alpha\left(2\|\nabla f(x)\|\epsilon^{1/2} + 11|f(x)|\right)}{\rho^{1/2}(x)}\right.$$
$$\left. + \left|\mathcal{B}_1[u,\mu](x) + \mathcal{B}_2[u,\mu,\rho](x) + \mathcal{B}_3[u,\mu,\rho](x)\right| + O(\epsilon)\right].$$

Proof: the maximum principle and the method of comparison functions

The 2nd-order kernel expansion formula

The key tool for computing the bias error

Lemma [SS, Evans, Cameron 2023]

Let $f \in C^\infty(\mathcal{M})$ and $\mathcal{G}_\epsilon$ be the integral operator defined by

$$\mathcal{G}_\epsilon f(x) = \int_{\mathcal{M}} k_\epsilon(x, y) f(y) \, d\mathrm{vol}(y).$$

Then, for small enough ϵ, $\mathcal{G}_\epsilon f$ admits the following expansion at x:

$$(\pi\epsilon)^{-d/2}\mathcal{G}_\epsilon f(x) = f(x) + \frac{\epsilon}{4}(\Delta f(x) - \omega(x) f(x)) + \frac{\epsilon^2}{4}\mathcal{Q} f(x) + O(\epsilon^3),$$

where $\mathcal{Q}$ is a fourth-order differential operator on $\mathcal{M}$.
In particular, if $\mathcal{M}$ is isometric to $\mathbb{R}^d$ in a neighborhood of x
(i.e $\mathcal{M}$ is locally flat) then

$$\mathcal{Q} = \partial_i^i \partial_j^j + 2\partial_{ii}^{ii}.$$

The 2nd-order kernel expansion formula

The key tool for computing the bias error

Goal: Expand the following integral:

$$\mathcal{G}_\epsilon f(x) = \int_{\mathcal{M}} k_\epsilon(x,y) f(y)\, d\mathrm{vol}(y).$$

Key idea 1: Keep track of fourth order correction factor to intrinsic/extrinsic distance (Jost '08)

$$\|x-y\|_2^2 = (d_{\mathscr{M}}(x,y))^2 - \frac{(d_{\mathscr{M}}(x,y))^4}{12}\|H(\dot{\gamma}_{xy}(0),\dot{\gamma}_{xy}(0))\|_2^2$$
$$+o((d_{\mathscr{M}}(x,y))^4)$$

Key idea 2: Keep track of third order correction to volume form:

$$g_{ij}(u) = \delta_{ij} + \frac{1}{3}R_{i\alpha\beta j}u^\alpha u^\beta + O(\|u\|^3)$$

$$d\,\mathrm{vol}(y) = \sqrt{\det g_{ij}}\,du := 1 + C_x^2(u),\, C_x^2 = O(\|u\|_2^2)$$

Variance error: key ideas

Bernstein's inequality: $$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq t\right) \leq \exp\left[-\frac{t^2}{2\sum_{i=1}^{n}\mathbb{E}[X_i^2] + \frac{2}{3}Mt}\right] \quad \forall t > 0.$$

Amplification: $$\mathbb{P}_{x_i}\left(\rho_\epsilon^{(n)}(x_i) - \rho_\epsilon(x_i) \geq t\right) = \mathbb{P}_{x_i}\left(\sum_{j\neq i} X_j \geq t - X_i\right) < \mathbb{P}_{x_i}\left(\sum_{j\neq i} X_j \geq t\right)$$

Unlike Singer (2006), or Berry & Harlim (2015), we don't remove the point from the point cloud!

Theorem [Discrete Kernel Density Estimate]

Let $\mathcal{X}(n) = \{x_j\}_{j=1}^{n} \sim \rho(x)$. Let $x \in \mathcal{M}$. Then for $\epsilon \to 0$ and $n \to \infty$ so that

$$\lim_{\substack{n\to\infty \\ \epsilon\to 0}} \frac{n\epsilon^{2+d/2}}{\log n} = \infty,$$

with probability at least $1 - n^{-4}$,

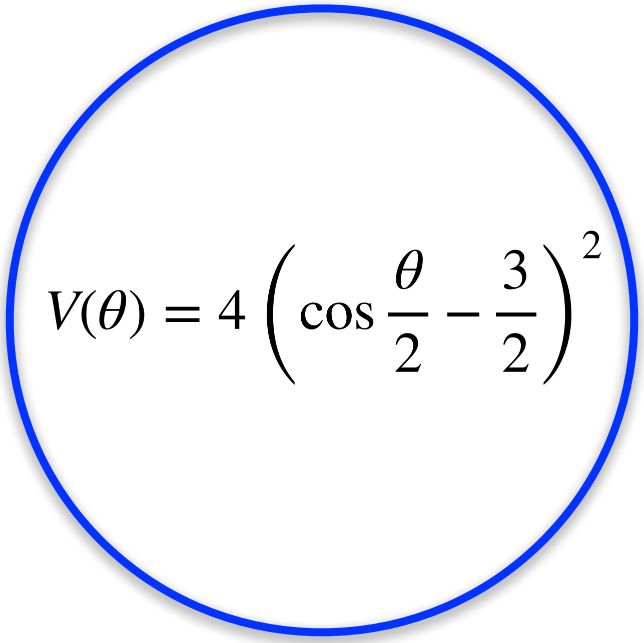
$$\left|\rho_\epsilon^{(n)}(x) - \rho_\epsilon(x)\right| < 5(\pi\epsilon)^{d/2}\rho^{1/2}(x)\epsilon^2\alpha$$

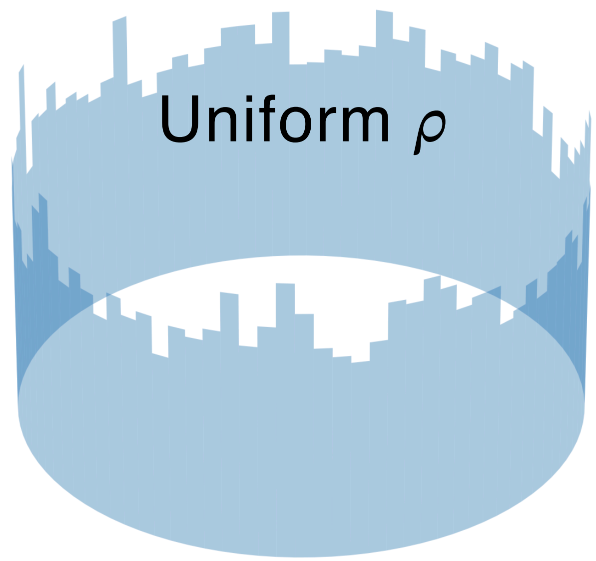
where

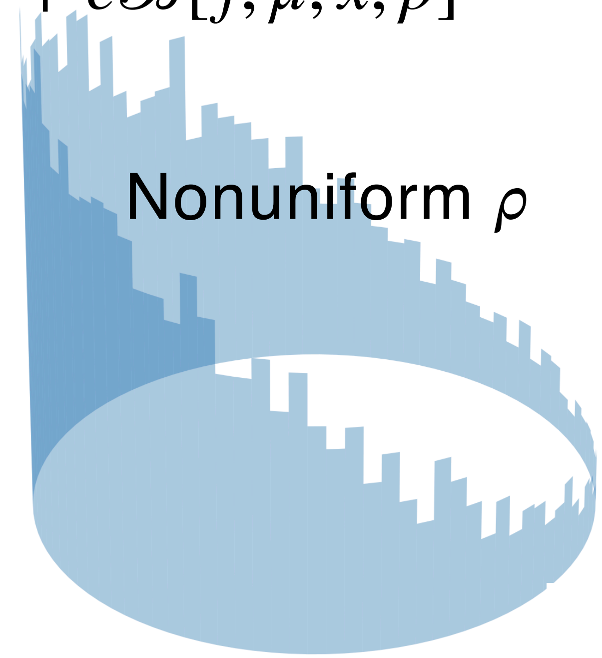
$$\alpha = \frac{1}{(2\pi)^{d/4}}\sqrt{\frac{\log n}{n\epsilon^{4+d/2}}}.$$

Two-well potential on the unit circle

Error $\sim \sqrt{\dfrac{\log n}{n\epsilon^{2+d/2}}}\mathscr{V}[f,\rho,x] + \epsilon\mathscr{B}[f,\mu,x,\rho]$





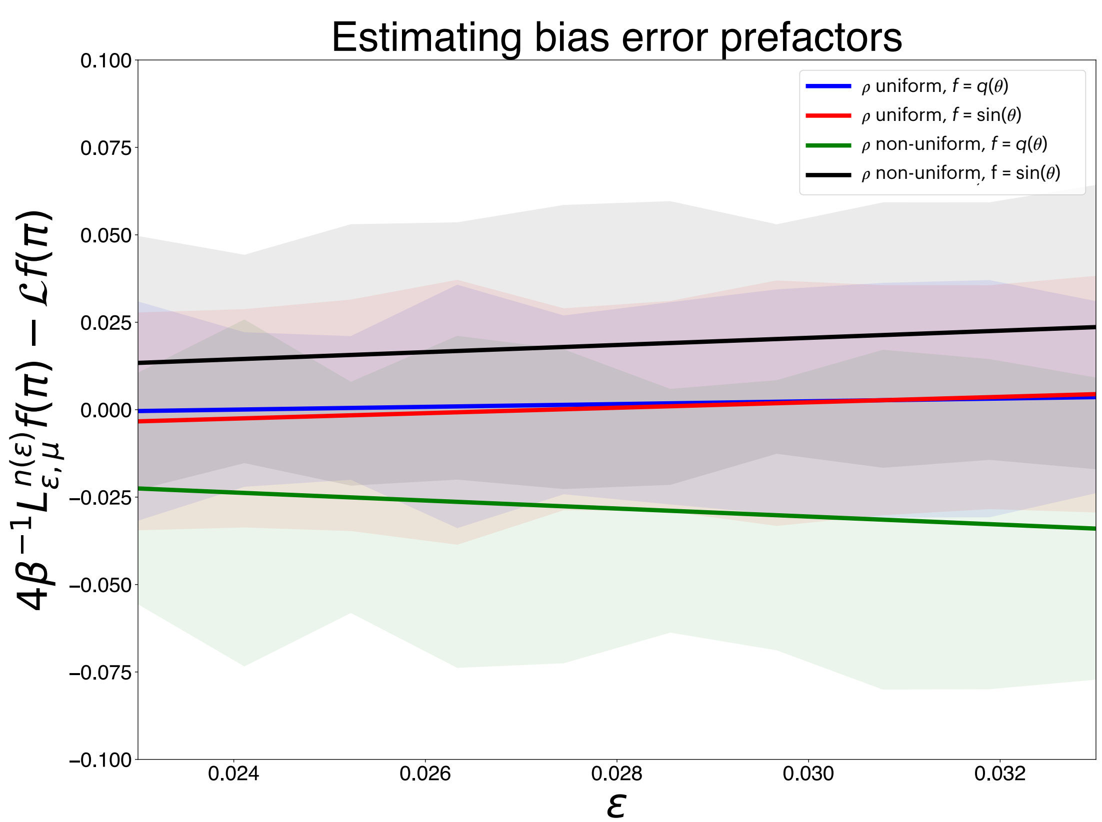


Test functions:

Committor, q

$t(\theta) = \sin\theta$

If $n(\epsilon)/\log(n(\epsilon)) = 0.25\epsilon^{-5/2}$, the variance error ~ const, then the total error $4\beta^{-1}L_{\epsilon,\mu}^{(n(\epsilon))}f(x) - \mathscr{L}f(x) \sim a + \epsilon b$.

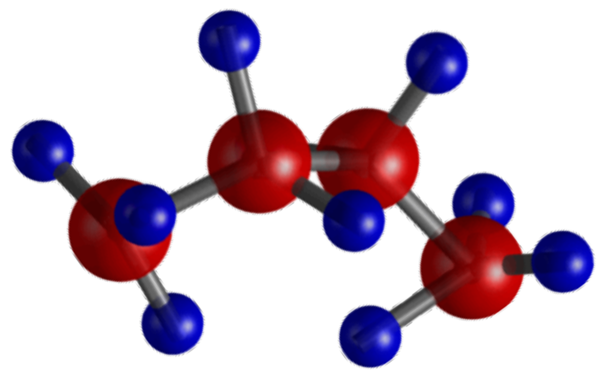


Prefactor |*b*| of the bias error

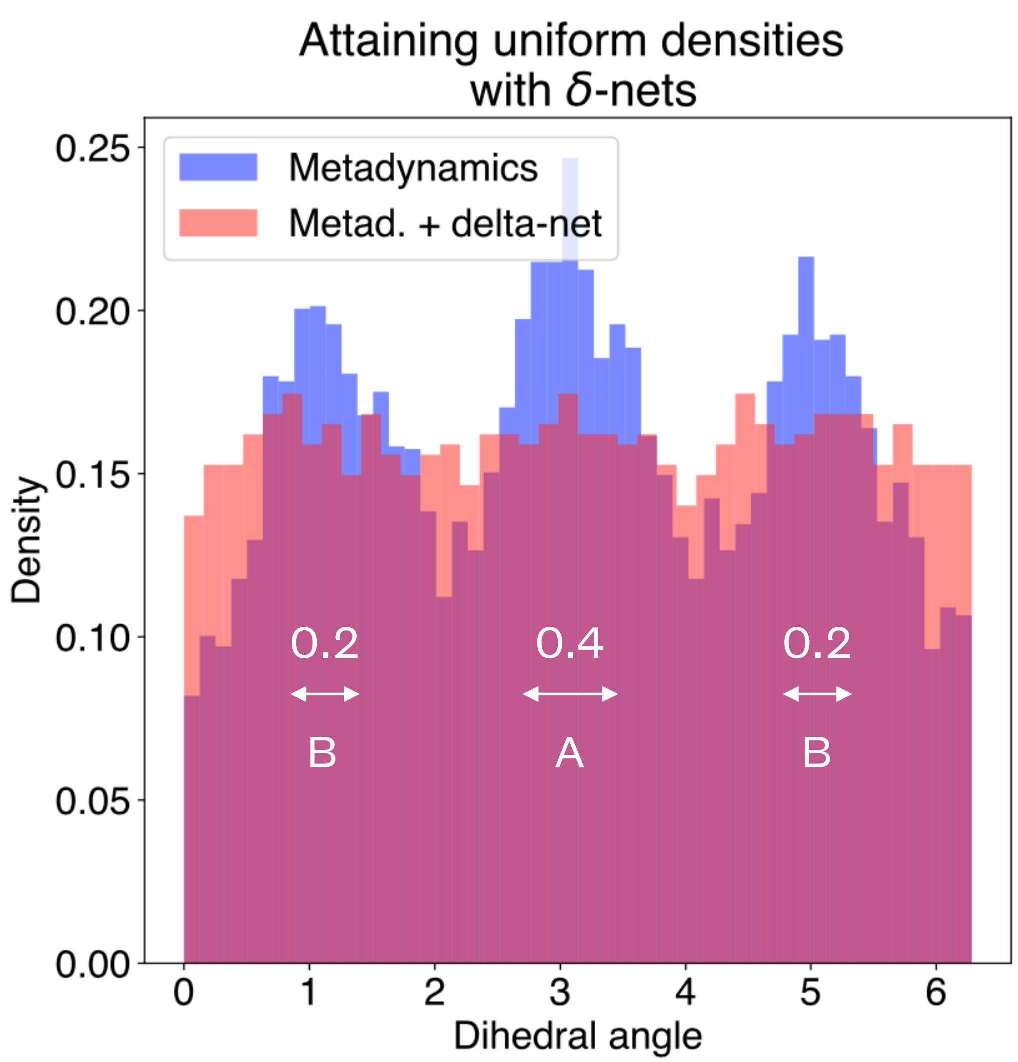
	sinθ	Committor
Uniform density	0.778	**0.398**
Nonuniform density	1.024	1.148

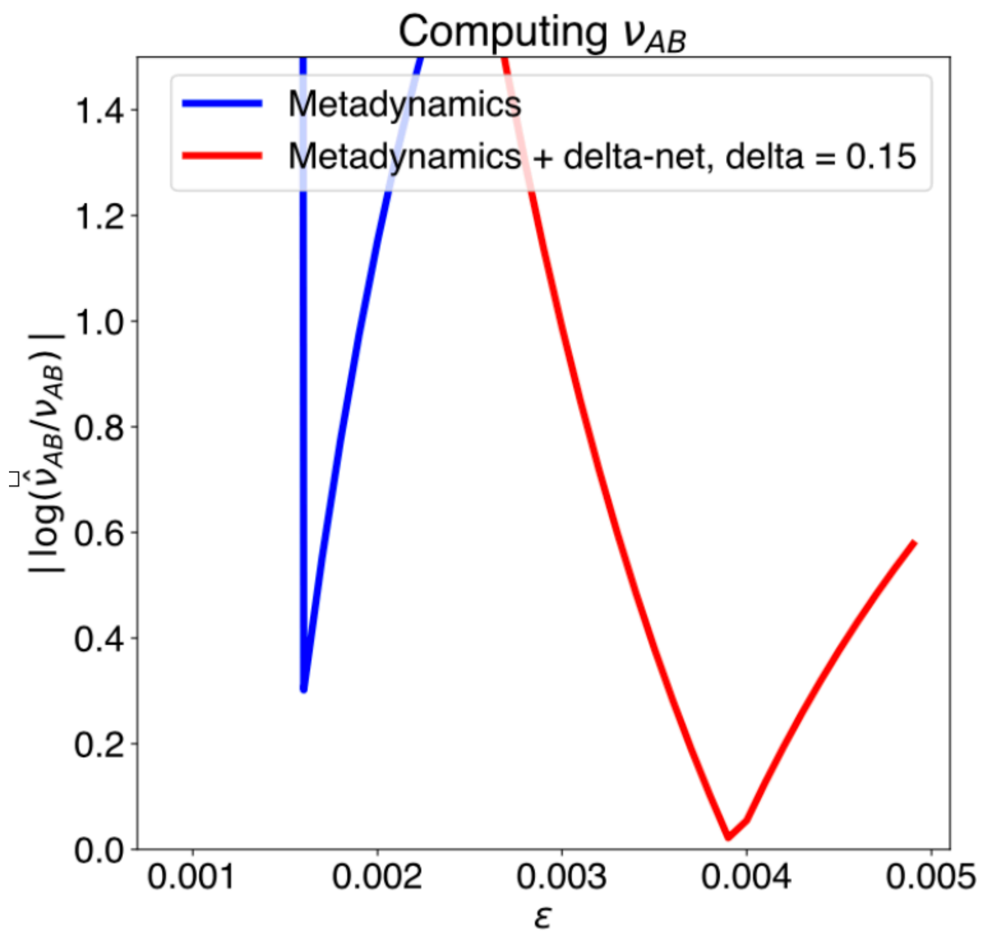
Conformal changes in butane

$$\text{Error} \sim \sqrt{\frac{\log n}{n\epsilon^{2+d/2}}}\mathscr{V}[f,\rho,x] + \epsilon\mathscr{B}[f,\mu,x,\rho]$$



12D data mapped to the dihedral angle space





	ε*	Rate A↔B
Metadynamics	0.0016	0.0109
Metad + δ-net	0.0039	0.0112
Ground truth		0.0114

δ-net: Crosskey and Maggioni (2017)

S. Sule, L. Evans, MC, 2024

Key conclusions

Sharp **estimates reveal that:**

1. Bias error at a point is simplified when **(1) $\mathscr{M}$ is flat**, the **(2) sampling density ρ is uniform**, or **(3) when the test function is the committor.**

2. **Worst case variance error** is **minimized** when $\rho \propto \text{vol}\,(\mathscr{M})^{-1}$, i.e the **uniform density**

3. The **solution error** is **controlled by** an $O(\epsilon)$ perturbation of the **consistency error**, so **improving consistency will improve solution accuracy.**

TMD map admits an arbitrary sampling density, so we can **post process our pointcloud to improve the accuracy** of the committor/eigenfunction by uniformizing.

In particular, **spatially uniform subsampling** improves both accuracy and stability to the bandwidth parameter. These gains in accuracy are practically realizable!

Future work

1. We post-process datasets generated using metadynamics into δ-nets.
   - δ-nets are spatially quasi-uniform random sets where points are dependent!
   - Numerical experiments suggest that δ-nets lead to a better scaling between n and ϵ.
   - What is the kernel density estimate and the variance error for δ-nets?
2. **Spectral convergence:** Belkin and Niyogi (2006) showed that control on the norm of the residual operator in the bias error leads to spectral convergence. Here we have a much more quantitative result. **Spectral convergence rate?**
3. The sampling question: TMD map accepts arbitrary sampling densities, but in theory practice, the uniform density does best. **Can we train a generative model to sample only the support of a distribution?**

A δ-net, PC: Maria Cameron

Appendix A: Estimating the transition rate (Evans, Cameron, Tiwary 2021)

$$\nu_{AB} = \beta^{-1} \int_{\mathcal{M}_{AB}} \|\nabla q(x)\|_2^2 \mu(x) d\mathrm{vol}(x)$$

$$\hat{\nu}_{AB} = \frac{1}{|I_{AB}|} \sum_{i \in I_{AB}} \left[\hat{\Gamma}([q],[q])\right]_i = \frac{1}{|I_{AB}|} \sum_{i \in I_{AB}} \sum_{j=1}^{n} L_{ij}([q]_i - [q]_j)^2,$$

Appendix B: Langevin on Manifold (Hsu '02)

$\{Z_t\}_{t\geq 0}$ is a M-valued diffusion process generated by L if $\{Z_t\}_{t\geq 0}$ is an $\mathcal{F}_*$-adapted semimartingale and

$$M^f(Z)_t := f(Z_t) - f(Z_0) - \int_0^t Lf(Z_s)\, ds\,,$$

is an $\mathcal{F}_*$-adapted local martingale for all $f \in C^\infty(M)$.