Error Analysis of Target Measure Diffusion Maps on $\mathbb{R}^d$ and applications to Transition Path Theory

Maria Cameron, Luke Evans, Shashank Sule

University of Maryland, College Park

Introduction

Target Measure Diffusion Maps (TMD Maps) is a non-linear dimensionality reduction technique where a dataset $\mathcal{X} \subset \mathcal{M}$ drawn from a sampling measure dq is normalized to approximate the generator $\mathcal{L}$ of an SDE admitting a "target" invariant measure $d\mu$. In this work we use a prefactor analysis to justify using a **quasi-uniform sampling density to reduce the pointwise error** between the approximate generator $L_{\epsilon,\mu}$ and the Fokker-Planck operator $\mathcal{L}$ in the case when $\mathcal{M} = \mathbb{R}^d$. We confirm this fact using numerical experiments on elliptic boundary value problems arising in Transition Path Theory.

Target Measure Diffusion Maps: Algorithm

Goal: Approximate infinitesimal generator $\mathcal{L}$ for the time-reversible dynamics governed by the SDE

$$dX_t = -\nabla V(x)dt + \sqrt{2\beta^{-1}}dW \quad (1)$$

We recall that the generator is given by

$$\mathcal{L} = \beta^{-1}\Delta - \nabla V \cdot \nabla. \quad (2)$$

Input: Dataset $\mathcal{X} \subset \mathbb{R}^d \sim dq$, target measure μ, kernel bandwidth ϵ, and kernel matrix

$$[K_\epsilon]_{ij} = \exp\left(\epsilon^{-1}\|x_i - x_j\|_2^2\right) := k_\epsilon(x_i, x_j) \quad (3)$$

Step 1 (Kernel Density estimation): M = diag$([\mu(x_i)]_{x_i \in \mathcal{X}})$, $D_\epsilon = (1/N)$diag$(K_\epsilon \mathbf{1})$. Note that $[D_\epsilon]_i \to q_\epsilon(x_i)$ where

$$q_\epsilon(x) := \int_{\mathbb{R}^d} k_\epsilon(x,y)q(y)dy \quad (4)$$

Step 2 (Renormalization): $K_{\epsilon,\mu} = K_\epsilon D_\epsilon^{-1} M^{1/2}$. Here we define

$$\mathcal{K}_{\epsilon,\mu} f(x) = \int_{\mathbb{R}^d} k_{\epsilon,\mu}(x,y)f(y)q(y)dy \quad (5)$$

Step 3 (Markov Process): $D_{\epsilon,\mu} = (1/N)$diag$(K_{\epsilon,\mu}\mathbf{1})$, $P_{\epsilon,\mu} = D_{\epsilon,\mu}^{-1}K_{\epsilon,\mu}$. Define

$$\mathcal{P}_{\epsilon,\mu} f(x) = \frac{\mathcal{K}_{\epsilon,\mu} f(x)}{\mathcal{K}_{\epsilon,\mu} 1(x)} \quad (6)$$

Step 4/Output (Generator): $L_{\epsilon,\mu} = \epsilon^{-1}(I - P_{\epsilon,\mu})$. Let

$$\mathcal{L}_{\epsilon,\mu} f(x) = \epsilon^{-1}\mathcal{P}_{\epsilon,\mu} f(x) - f(x) \quad (7)$$

Banisch et al. [1] prove that the generator $\mathcal{L}_{\epsilon,\mu} f \to \mathcal{L}f$. We refine this result in Theorem 1 by computing the rate of convergence.

Transition Path Theory

Transition Path Theory (TPT) aims to study the rare events associated with systems observing dynamics according to Equation 1. In particular let $A, B \subset \Omega$ be disjoint subsets of a state space and let the *forward comittor function* $u(x)$ be the probability that X_t enters B before A such that $X(0) = x$. Then Vanden-Eijnden et al. [2] show that u satisfies

$$\mathcal{L}u = 0, \quad u\mid_{\partial A} = 0, \quad u\mid_{\partial B} = 1 \quad (8)$$

TMD maps can be used as a meshless algorithm for numerically solving the elliptic BVP (8) since $L_{\epsilon,\mu}$ is a discrete approximation to $\mathcal{L}$. Consequently it is important to study the (pointwise) error rate for $|L_{\epsilon,\mu} f - \mathcal{L}f(x)|$ as $\epsilon \to 0$ and $|\mathcal{X}| \to \infty$

Error Analysis: Theory

Fix $x \in \mathbb{R}^d$, assume the point cloud $\mathcal{X} \subseteq \mathbb{R}^d (N := |\mathcal{X}|)$ has been sampled through a measure $dq = q dx$ with density q. Let $f \in C^\infty(\mathbb{R}^d)$, $V : \mathbb{R}^d \to \mathbb{R}$ a coercive potential, and $d\mu = \mu dx$ be the target measure with density μ. Note that $L_{\epsilon,\mu} f(x) \to \mathcal{L}_{\epsilon,\mu} f(x)$ and $\mathcal{L}_{\epsilon,\mu} f(x) \to \mathcal{L}f(x)$. The approximation error with respect to N is the *Variance Error* and with respect to ϵ is the *Bias Error*.

Theorem 1 (Bias Error)**.** Fix a sampling density q, target density μ and let

$$\mathcal{Q}(f) := 2\sum_{i<j}\frac{\partial^4 f(x)}{\partial x_i^2 \partial x_j^2} + 3\sum_i \frac{\partial^4 f(x)}{\partial x_i^4}$$

$$\mathcal{R}_{\mu,q} f := \mathcal{Q}\left(f\mu^{1/2}\right) - \frac{1}{4}\Delta\left(f\mu^{1/2}\frac{\Delta q}{q}\right) + f\mu^{1/2}\left(\frac{1}{4}\left(\frac{\Delta q}{q}\right)^2 - \frac{\mathcal{Q}(q)}{q}\right).$$

Then we have the following error formula:

$$\mathcal{L}_{\mu,\varepsilon} - \frac{1}{4}\mathcal{L} = \frac{\epsilon}{4}\left[\frac{\mathcal{R}_{\mu,q} f}{\mu^{1/2}} - f\frac{\mathcal{R}_{\mu,q} 1}{\mu^{1/2}}\right] + \frac{\epsilon}{16}\left(f\left[\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \frac{\Delta q}{q}\right]^2 - \left[\frac{\Delta(\mu^{1/2} f)}{\mu^{1/2}} - f\frac{\Delta q}{q}\right]\left[\frac{\Delta(\mu^{1/2})}{\mu^{1/2}} - \frac{\Delta q}{q}\right]\right) + O(\epsilon^2)$$

Theorem 2 (Variance error)**.** Let

$$t(x) := \frac{1}{2}(1-f)\nabla(\mu/q)\cdot\nabla f + \frac{1}{4}(\mu/q)|\nabla f|_2^2, \quad s(x) := \frac{\Delta\mu^{1/2}}{\mu^{1/2}} - \frac{\Delta q}{q}$$

Then, letting $p(N,\alpha) := P(|L_{\epsilon,\mu} f(x) - \mathcal{L}_{\epsilon,\mu} f(x)| > \alpha)$,

$$p(N,\alpha) \leq 2\exp - \frac{(\pi\epsilon)^{d/2}(N-1)\alpha^2(\mu + \epsilon s(x) + O(\epsilon^2))}{2\epsilon(t(x) + O(\epsilon))} \quad (9)$$

Consequently, $|L_{\epsilon,\mu} f(x) - \mathcal{L}_{\epsilon,\mu} f(x)| = O\left(\frac{\epsilon^{-(1/2+d/2)}}{\sqrt{N}}\right)$

Theorems 1 and 2 can be used to show that **the BVP (8) is a rather good problem to solve using TMD maps** when $q(x) \equiv C$ on $\Omega \subseteq \mathbb{R}^d$. In particular, we can improve the variance error by $\sqrt{\epsilon}$ and kill a whole term in the bias error:

Corollary. Borrowing all notation from Theorems 1 2, let $q(x) \equiv C$ for $x \in \Omega$. Then we have the following approximation rate:

$$4L_{\epsilon,\mu} f(x) - \mathcal{L}f(x) = O\left(\frac{\epsilon^{-d/2}}{\sqrt{N}}\right) + \frac{\epsilon}{4}\left[\frac{\mathcal{Q}(f\mu^{1/2})}{\mu^{1/2}} - f\frac{\mathcal{Q}\mu^{1/2}}{\mu^{1/2}}\right] \quad (10)$$

Proof. The estimate follows by noting that when $\mathcal{L}f = 0$ and $\Delta q = 0$, the second $O(\epsilon)$ term in the bias error formula is zero. Furthermore, when q is locally constant, $t(x) \approx 0$ because $\nabla f \approx 0$ around the regions A and B. . □

Error Analysis: Numerics

We the committor problem numerically using TMD Maps by solving $L_{\epsilon,\mu} u = 0$ such that $u(I_A) = 0$ and $u(I_B) = 1$ where I_A and I_B are indices of points in $\mathcal{X}$ that intersect with A and B respectively. We visualize the RMS error between u and a FEM-computed ground truth value u^* as a function of kernel bandwidth ϵ for two 2-dimensional choices of V: **Muller's Potential** and **Two well Potential**. In both cases we use two choices of q:

- By sampling the SDE 1 through the Euler-Maruyama method (this is equivalent to sampling by the Gibbs density $\alpha \exp -\beta V(x)$)
- By sampling the SDE using Euler-Maruyama with modifications through well-tempered metadynamics, which periodically modifies the density of sampling in Euler-Maruyama.

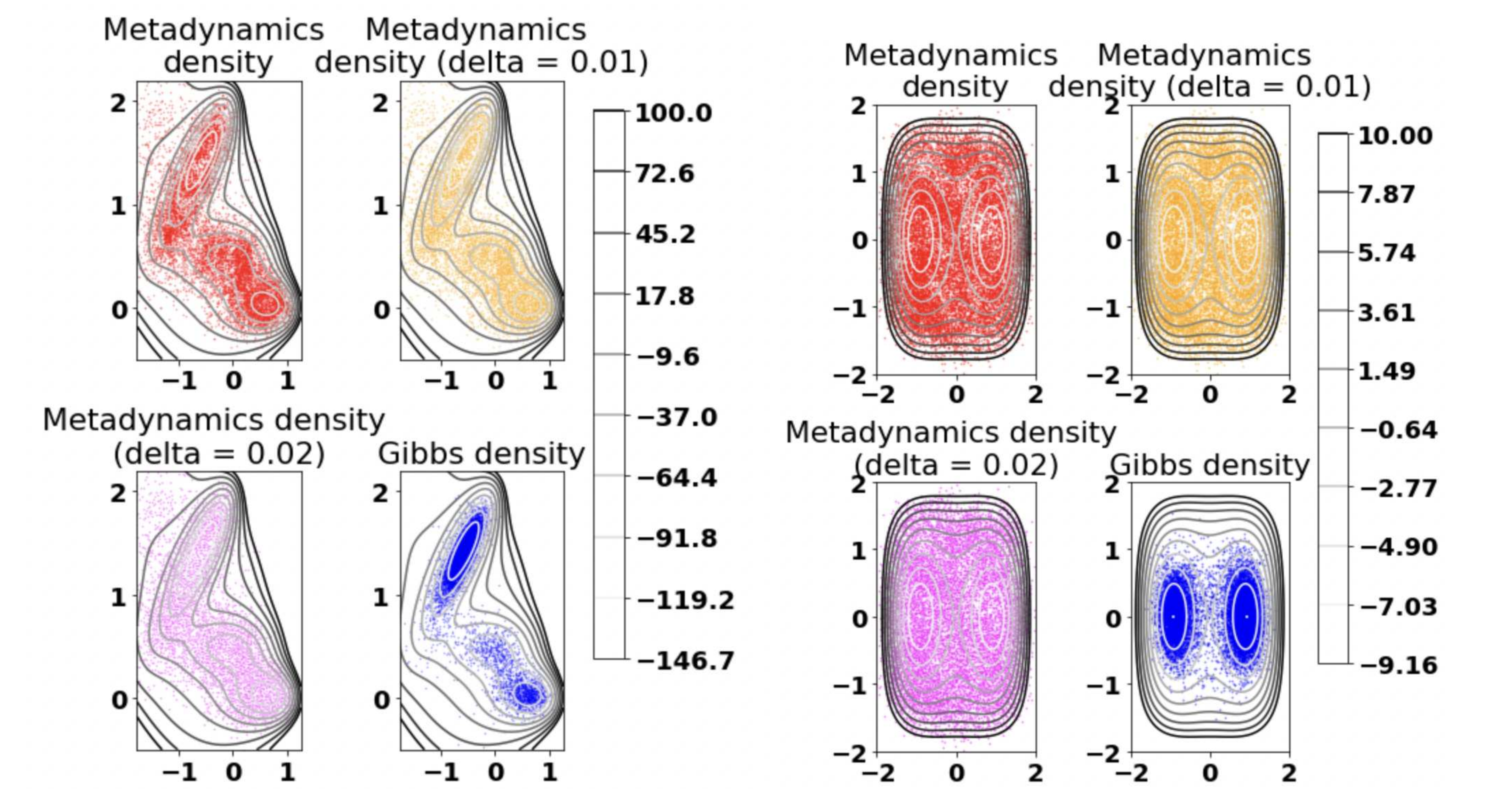

Figure 1: Sample datasets $\mathcal{X}$ for Muller's Potential (Left) and Twowell Potential (right). Note how in each case sampling by the Gibbs measure leaves the transition region underresolved. Metadynamics sampling is in general much better at covering these regions, but it leads to a different sampling density q. These datasets have also been uniformized by the δ-nets where we greedily delete a ball of size δ around each point.

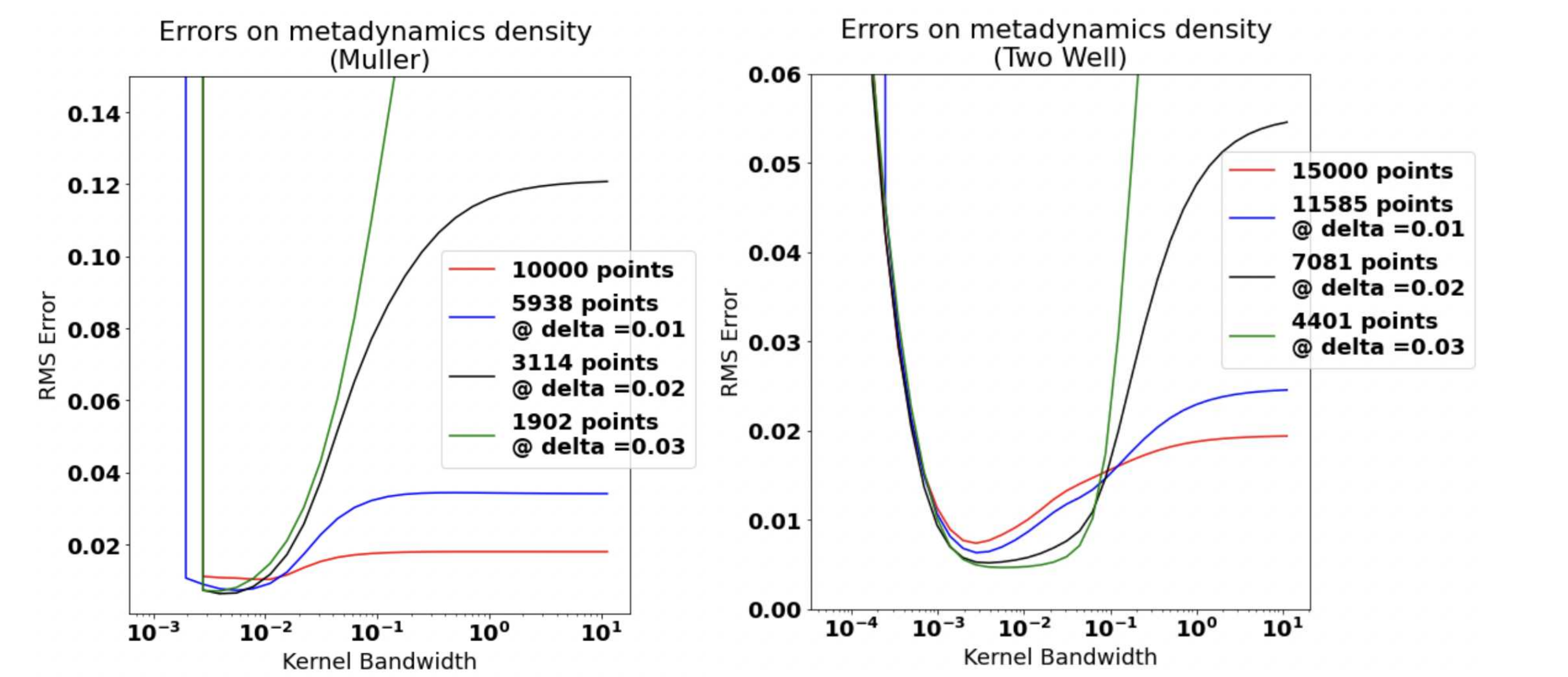

Figure 2: RMS error performance on the uniformizations of the Metadynamics-sampled datasets on Muller's Potential (Left) and Twowell Potential (right). Note that uniformization *decreases* minimum RMS error over ϵ and increases the region of sensitivity of error to ϵ. These two phenomena are predicted by our Theorems 1 and 2: the improved minimum error is due to smaller bias error and the flatter basin of sensitivity is due to reduced variance error by a factor of ϵ when $\epsilon < 1$.

References

[1] Ralf Banisch et al. "Diffusion maps tailored to arbitrary non-degenerate Itô processes". In: *Applied and Computational Harmonic Analysis* 48.1 (2020), pp. 242–265.

[2] Eric Vanden-Eijnden et al. "Towards a theory of transition paths". In: *Journal of statistical physics* 123.3 (2006), pp. 503–523.