SIMULATION RESULTS FOR MINIMIZING ENTROPY

We demonstrate the application of the Schrödinger Bridge (SB) scheme in providing a discrete approximation of the heat flow in a simple Gaussian example. Consider the problem of minimizing the entropy functional

$$\mathcal{F}(\rho) = \frac{1}{2}\mathrm{Ent}(\rho) = \begin{cases} \frac{1}{2}\int \rho(x)\log\rho(x)dx & \text{if } \rho \text{ admits a density,} \\ +\infty & \text{otherwise.} \end{cases}$$

We work with the simple case where the starting measure is $\rho_0 = \mathcal{N}(0, \eta^2)$. In this case, the gradient flow minimizing $\mathcal{F}$ is the heat flow known analytically and given by $\rho(t) = \mathcal{N}(0, \eta^2 + t)$ for $t \geq 0$. Let us construct a discrete approximation of the gradient flow $\big(\rho(t), t \in [0, T]\big)$ using our SB scheme with temperature $\varepsilon > 0$.

Schrödinger Bridge Scheme. Let $\pi_{\rho,\varepsilon}$ denote the Schrödinger bridge with equal marginals ρ and temperature parameter ε. Then recall from [AHMP24], the SB step for approximating the forward heat flow is given by $\mathrm{SB}^1_\varepsilon(\rho) = \big(2\,\mathrm{Id} - \mathcal{B}_{\rho,\varepsilon}\big)_{\#}\,\rho$ where $\mathcal{B}_{\rho,\varepsilon}(x) = \mathbb{E}_{\pi_{\rho,\varepsilon}}\big[Y|X = x\big]$ is the barycentric projection. The barycentric projection with Gaussian marginals is known in closed form; thanks to [JMPC20]. Therefore, the SB step is available in closed form and derived in [AHMP24] as

$$\mathrm{SB}^1_\varepsilon(\rho_0) = \mathcal{N}\left(0, \left(2 - C^\varepsilon_{\eta^2}\right)^2 \eta^2\right) \qquad \text{where } C^\varepsilon_{\eta^2} = \frac{1}{\eta^2}\left(\sqrt{\eta^4 + \frac{\varepsilon^2}{4}} - \frac{\varepsilon}{2}\right). \tag{1}$$

Since the SB scheme operates via linear pushforwards, all steps of the SB scheme are mean-zero Gaussian distributed. Denote $\mathrm{SB}^k_\varepsilon(\rho_0) = \mathcal{N}(0, \alpha_k^2)$, then in [AHMP24, Section 5.1], we show that the SB scheme evolves as $\alpha^2_{k+1} = \left(2 - C^\varepsilon_{\alpha^2_k}\right)^2 \alpha^2_k$. Now let us consider the finite particles setting and construct a sample approximation of the SB scheme. We demonstrate that this sample approximation of the SB scheme approximates the heat flow in our analytically tractable Gaussian setting.

Sample Approximation of Schrödinger Bridge Scheme. Let $(x_1, \ldots, x_n)$ be the n i.i.d. observed particles distributed as ρ_0. Then $\hat{\rho}_0 = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ is the initial empirical distribution. When ρ_0 is not known, the SB steps $\left(\mathrm{SB}^k_\varepsilon(\rho_0); k \in [N_\varepsilon]\right)$, with $N_\varepsilon := T\varepsilon^{-1}$, cannot be computed explicitly in most practical implementations. Therefore, $\mathrm{SB}^k_\varepsilon(\rho_0)$ is approximated by its sample approximation, that is $\mathrm{SB}^k_\varepsilon(\rho_0) \approx \mathrm{SB}^k_\varepsilon(\hat{\rho}_0)$. For ease of notation, denote

$$\hat{\rho}_k = \frac{1}{n}\sum_{i=1}^n \delta_{x_i^{(k)}},$$

where $(x_1^{(k)}, \ldots, x_n^{(k)})$ is the position of particles at the kth step evolving via the SB scheme

$$x_i^{(k)} = 2x_i^{(k-1)} - \mathcal{B}_{\hat{\rho}_{k-1},\varepsilon}\left(x_i^{(k-1)}\right) \qquad \text{for all } i \in [n].$$

The above pushforward, $2\,\mathrm{Id} - \mathcal{B}_{\hat{\rho}_{k-1},\varepsilon}$, is approximated via two levels of sample approximation. First, the barycentric projection $\mathcal{B}_{\hat{\rho}_{k-1},\varepsilon}$ is constructed with respect to an approximation of the discrete Schrodinger bridge $\pi_{\hat{\rho}_{k-1},\varepsilon}$ computed using the celebrated Sinkhorn algorithm [SK67]. This is denoted by $\hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}$. Second, for $(X, Y) \sim \hat{\pi}_{\hat{\rho}_{k-1}}$, we consider an empirical

Date: August 10, 2024.

approximation of the conditional expectation $\mathbb{E}_{\hat{\pi}_{\hat{\rho}_{k-1}}}\left[Y|X=.\right]$. The resultant sample approximation of the SB scheme gives the following update rule for the particles. r any $i \in [n]$ and $k \in [N_\varepsilon]$,

$$x_i^{(k)} = 2x_i^{(k-1)} - \frac{1}{n}\frac{\sum_{j=1}^n x_j^{(k-1)}\hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}^{(i,j)}}{\sum_{j=1}^n \hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}^{(i,j)}}, \tag{2}$$

where $\hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}^{(i,j)}$ denotes its (i,j)th element of $\hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}$. Algorithm 1 summarizes the finite sample SB scheme described above.

Algorithm 1 Schrödinger bridge scheme for minimizing entropy

1: **Input**: $(x_1,\ldots,x_n) \overset{i.i.d.}{\sim} \rho_0$, T total time to run the dynamic system, ε step size, M Sinkhorn iterations, τ Sinkhorn precision parameter.
2: Set $(x_1^{(0)},\ldots,x_n^{(0)}) \leftarrow (x_1,\ldots,x_n)$, $\hat{\rho}_0 \leftarrow \frac{\sum_{i=1}^n \delta_{x_i}}{n}$, and $N_\varepsilon = \frac{T}{\varepsilon}$.
3: **for** k **from** 1 **to** $\lfloor N_\varepsilon \rfloor$ **do**
4: Compute discrete Schrodinger bridge $\hat{\pi}_{\hat{\rho}_k,\varepsilon}$ for particles $(x_1^{(k-1)},\ldots,x_n^{(k-1)})$ using Sinkhorn iterations.
5: Compute sample barycentric projections for all particles. For all $i \in [n]$,

$$\hat{T}_{\hat{\rho}_{k-1},\varepsilon}(x_i^{(k-1)}) = \frac{1}{n}\frac{\sum_{j=1}^n x_j^{(k-1)}\hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}^{(i,j)}}{\sum_{j=1}^n \hat{\pi}_{\hat{\rho}_{k-1},\varepsilon}^{(i,j)}}.$$

6: Update position of particles. For all $i \in [n]$, $x_i^{(k)} \leftarrow 2x_i^{(k-1)} - \hat{T}_{\hat{\rho}_{k-1},\varepsilon}(x_i^{(k-1)})$.
7: Set $\hat{\rho}_k = \frac{1}{n}\sum_{i=1}^n \delta_{x_i^{(k)}}$.
8: **end for**

While the algorithm works for any empirically known starting distribution via a finite set of initial particles, we now study the examples where the gradient flow and SB scheme updates are known analytically.

Starting from Standard Gaussian. Consider $\rho_0 = \mathcal{N}(0,1)$. For $n = 500$ samples, regularizing parameter $\varepsilon = 0.1$, and total time $T = 5$ units, we push the particles drawn randomly from $\rho_0 = \mathcal{N}(0,1)$ via (2). In Figure 1, we present the histogram of $\hat{\rho}_k$ at time increments of 1 unit, or equivalently iteration increments of 100 steps. The bold solid line represents the density curve of the true gradient flow, while the dashed line represents the density curve for the piecewise constant interpolation of the SB scheme.

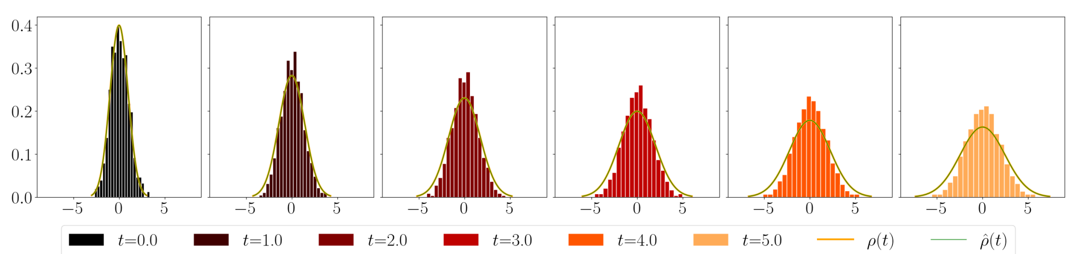

FIGURE 1. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho(t); t \in [0,T])$ is the true gradient flow and $(\hat{\rho}(t); t \in [0,T])$ is the piecewise constant interpolation of the known SB scheme $(\hat{\rho}_k; k \in [N_\varepsilon])$.

Starting from a Thin Gaussian. Consider a thin Gaussian starting distribution $\rho_0 = \mathcal{N}(0, 0.25)$. For $n = 500$ samples, the regularizing parameter $\varepsilon = 0.1$, and total time $T = 5$ units, we push the particles drawn randomly from $\rho_0 = \mathcal{N}(0, 0.25)$ via (2). In Figure 2, we present the histogram of ρ_k^ε in time increments of 1 unit, or equivalently iteration increments of 10 steps. Similarly to the previous case, the true gradient flow density curve and piecewise constant interpolation of the SB scheme density are also plotted.

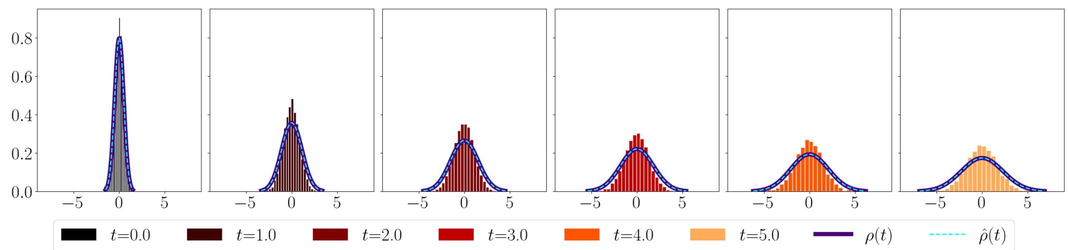

FIGURE 2. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho(t); t \in [0, T])$ is the true gradient flow and $(\hat{\rho}(t); t \in [0, T])$ is the piecewise constant interpolation of the known SB scheme $(\hat{\rho}_k; k \in [T\varepsilon^{-1}])$.

Starting from Mixture of Gaussians. Now choosing $\rho_0 = 0.5\,\mathcal{N}(-2, 1) + 0.5\,\mathcal{N}(2, 1)$, we again push the $n = 500$ i.i.d. particles sampled from ρ_0 via (2). The gradient flow and discrete-time SB steps are again known in closed form here. Figure 3 plots similar quantities for this starting distribution.

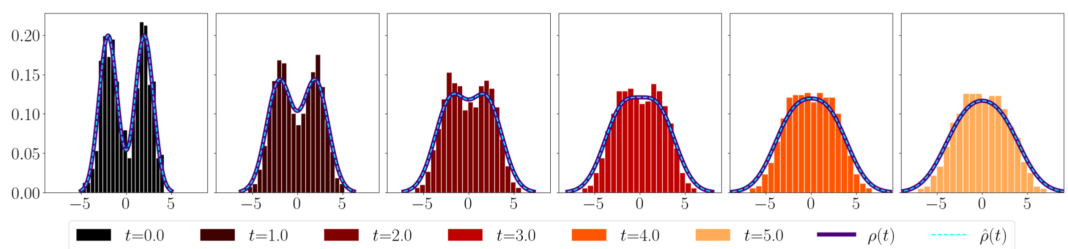

FIGURE 3. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho(t); t \in [0, T])$ is the true gradient flow and $(\hat{\rho}(t); t \in [0, T])$ is the piecewise constant interpolation of the known SB scheme $(\hat{\rho}_k; k \in [T\varepsilon^{-1}])$.

The simulations suggest that leveraging the discrete Schrödinger bridge offers a practical advantage, we can push a finite set of particles through an empirical approximation of the SB scheme. The aggregate distribution of these particles then provides a discrete-time approximation of the heat flow.

REFERENCES

[AHMP24] Medha Agarwal, Zaid Harchaoui, Garrett Mulcahy, and Soumik Pal. Iterated schr\" odinger bridge approximation to wasserstein gradient flows. *arXiv preprint arXiv:2406.10823*, 2024.

[JMPC20] Hicham Janati, Boris Muzellec, Gabriel Peyré, and Marco Cuturi. Entropic optimal transport between unbalanced Gaussian measures has a closed form. NIPS'20, Red Hook, NY, USA, 2020. Curran Associates Inc.

[SK67] Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific J. Math.*, 21:343–348, 1967.