SIMULATION RESULTS FOR MINIMIZING ENTROPY

We check the application of our method by pushing finite number of particles through the dynamic system described by the SB scheme for 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 for minimizing $\mathcal{F}$ is known analytically and is given by $\rho_t = \mathcal{N}(0, \eta^2 + t)$ for $t \geq 0$. The barycentric projection with Gaussian marginals is also known in closed form; thanks to [JMPC20]. Therefore, the discrete-time SB sequence $(\mathrm{SB}_\varepsilon^k(\rho_0); k \in [T\varepsilon^{-1}])$ is available in closed form.

Let $(x_1, \ldots, x_n)$ be the n i.i.d. observed particles distributed as ρ_0. Then $\hat{\rho}_0 = n^{-1}\sum_{i=1}^n \delta_{x_i}$ is the initial empirical distribution. Since SB steps $(\mathrm{SB}_\varepsilon^k(\hat{\rho}_0); k \in [T\varepsilon^{-1}])$ cannot be computed explicitly in most practical implementations, $\mathrm{SB}_\varepsilon^k(\hat{\rho}_0)$ is approximated by its empirical version, denoted by $\hat{\rho}_k^\varepsilon$. The sample approximation $\hat{\rho}_k^\varepsilon$ is the empirical distribution of particles at step k. That is,

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

where $(x_1^{(k)}, \ldots, x_n^{(k)})$ denotes the n particles at kth step. The pushforward for particles at step $k-1$ is approximated via the empirical conditional mean of discrete Schrodinger bridge with equal marginals $\hat{\rho}_{k-1}^\varepsilon$. For any $i \in [n]$ and $k \in [T\varepsilon^{-1}]$, the evolution of particles is given by

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

where $\hat{\pi}_{\hat{\rho}_{k-1}^\varepsilon,\varepsilon}$ is the discrete Schrödinger bridge with equal marginals $\hat{\rho}_{k-1}^\varepsilon$ and $\hat{\pi}^{(i,j)}_{\hat{\rho}_{k-1}^\varepsilon,\varepsilon}$ denotes its (i,j)th element. Algorithm 1 describes the SB scheme dynamics for any starting finite set of particles.

While the algorithm works for empirically known starting distribution via a finite set of 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 $\varepsilon = 0.01$, and total time $T = 5$ units, we push the particles drawn randomly from $\rho_0 = \mathcal{N}(0,1)$ via (1). In Figure 1, we present the histogram of $\hat{\rho}_k^\varepsilon$ 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 depicts the density curve for the piecewise constant interpolation of the SB scheme.

Date: July 24, 2024.

Algorithm 1 Schrödinger bridge scheme for minimizing entropy

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

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

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

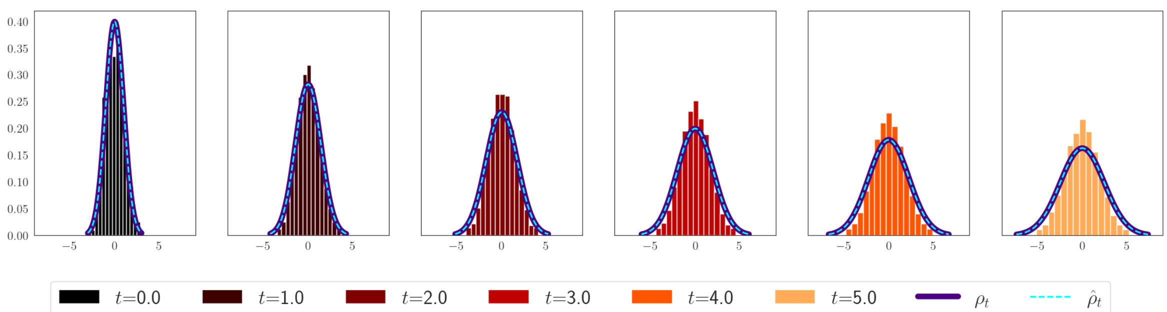

FIGURE 1. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho_t; t \in [0,T])$ is the true gradient flow and $(\hat{\rho}_t; t \in [0,T])$ is the piece-wise constant interpolation of the analytically known scheme $(\rho_k^{\varepsilon}; k \in [T\varepsilon^{-1}])$.

Starting from a Thin Gaussian. Consider a thin Gaussian starting distribution $\rho_0 = \mathcal{N}(0, 0.25)$. For $n = 500$ samples, regularizing parameter $\varepsilon = 0.01$, and total time $T = 5$ units, we push the particles drawn randomly from $\rho_0 = \mathcal{N}(0, 0.25)$ via (1). In Figure 2, we present the histogram of $\hat{\rho}_k^{\varepsilon}$ at time increments of 1 unit, or equivalently iteration increments of 10 steps. Similar to the previous case, the true gradient flow density curve and piecewise constant interpolation of SB scheme density is also plotted.

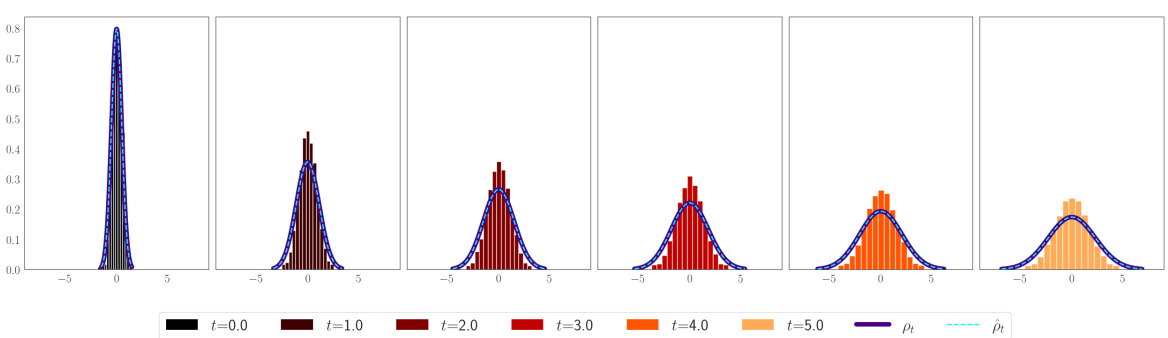

FIGURE 2. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho_t; t \in [0,T])$ is the true gradient flow and $(\hat{\rho}_t; t \in [0,T])$ is the piece-wise constant interpolation of the analytically known scheme $(\rho_k^{\varepsilon}; k \in [T\varepsilon^{-1}])$.

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

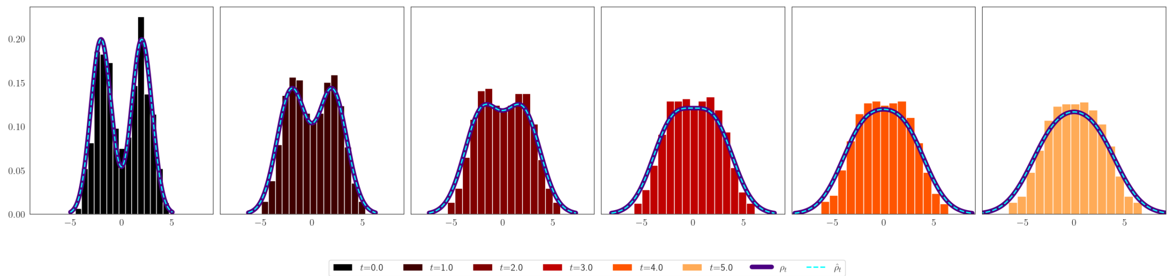

FIGURE 3. Histograms of $n = 500$ particles at increments of one time unit. Here $(\rho_t; t \in [0,T])$ is the true gradient flow and $(\hat{\rho}_t; t \in [0,T])$ is the piece-wise constant interpolation of the analytically known scheme $(\rho_k^{\varepsilon}; k \in [T\varepsilon^{-1}])$.

The simulations suggests 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

[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.