Combining Wasserstein-1 and Wasserstein-2 proximals: robust manifold learning via well-posed generative flows

Hyemin Gu[1], Markos A. Katsoulakis[1] and Luc Rey-Bellet[1], Benjamin J. Zhang[2]
[1] *University of Massachusetts Amherst*, [2] *Brown university*

Objective

- To formulate well-posed continuous-time generative flows through MFG for **learning distributions that are supported on low-dimensional manifolds**

Generative flows as Mean Field Games (MFG)

$\mathcal{F} : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$: terminal cost (ex. f-divergence $D_f(\rho(\cdot, T)\|\pi)$ between target distribution π and approximating distribution $\rho(\cdot, T)$)
$L(x, v) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$: convex running cost (ex. OT cost with $L = \frac{1}{2}|v|^2$)

- **MFG** that learns the velocity $v(x, t)$ and measure $\rho(x, t)$ is defined as

$$\min_{v,\rho} \mathcal{F}(\rho(\cdot, T)) + \int_0^T \int_{\mathbb{R}^d} L(x, v(x, t))\rho(x, t)dxdt \quad \text{s.t. } \partial_t\rho + \nabla \cdot (v\rho) = 0 \quad \rho(x, 0) = \rho_0(x). \tag{1}$$

$H(x, p) = \sup_v\{-p^T v - L(x, v)\}$: Hamiltonian of (1)

- **Optimality conditions** for the optimizers of (1) are

$$-\frac{\partial U}{\partial t} + H(x, \nabla U) = 0, \quad U(x, T) = \frac{\delta\mathcal{F}(\rho(\cdot, T))}{\delta\rho(\cdot, T)}(x), \tag{2}$$

$$\frac{\partial\rho}{\partial t} - \nabla \cdot (\nabla_p H(x, \nabla U)\rho) = 0, \quad \rho(x, 0) = \rho_0(x). \tag{3}$$

Wasserstein $\mathcal{W}$ proximal regularization

$\mathcal{P}_p(\mathbb{R}^d)$: the Wasserstein-p space

- **Wasserstein-p proximal regularized cost function** is written as

$$\inf_{R\in\mathcal{P}_p(\mathbb{R}^d)} \left\{\mathcal{F}(R) + \theta \cdot \mathcal{W}_p^p(P, R)\right\} \tag{4}$$

where $\theta > 0$ is a weighting parameter.

$\mathcal{W}_1$ proximal of D_f

$D_f^{\Gamma_L}(\rho\|\pi)$: $\mathcal{W}_1$ proximal regularized f-divergence in dual formulation (6) as terminal cost $\mathcal{F}(\rho)$

$$\inf_{\nu\in\mathcal{P}_1(\mathbb{R}^d)} \{D_f(\nu\|\pi) + L \cdot \mathcal{W}_1(\rho, \nu)\} \tag{5}$$

$$= \sup_{\phi\in\Gamma_L} \{\mathbb{E}_\rho[\phi] - \mathbb{E}_\pi[f^\star(\phi)]\} \tag{6}$$

$\mathcal{W}_2$ proximal of $D_f^{\Gamma_L}$

Running cost $L(x, v) = \frac{1}{2}|v|^2$ results in $\mathcal{W}_2$ proximal of $D_f^{\Gamma_L}$ using Bernamou-Brenier dynamic OT formulation (7)

$$\inf_{\rho_T,v} \int_0^T \int_{\mathbb{R}^d} \frac{1}{2}|v(x, t)|^2\rho(x, t)dxdt \quad \text{s.t. } \partial_t\rho + \nabla \cdot (v\rho) = 0,\ \rho(x, 0) = \rho_0(x) \tag{7}$$

Combination of Wasserstein proximals $\mathcal{W}_1 \oplus \mathcal{W}_2$

We combine $D_f^{\Gamma_L} = \mathcal{W}_1$ proximal of D_f and $\mathcal{W}_2$ proximal of $D_f^{\Gamma_L}$

$$\inf_{\rho_T} \left\{D_f^{\Gamma_L}(\rho_T\|\pi) + \frac{\lambda}{2T} \cdot \mathcal{W}_2^2(\rho_0, \rho_T)\right\} \tag{8}$$

$$= \inf_{\rho_T} \left\{\inf_\sigma \{D_f(\sigma\|\pi) + L \cdot \mathcal{W}_1(\rho_T, \sigma)\} + \frac{\lambda}{2T} \cdot \mathcal{W}_2^2(\rho_0, \rho_T)\right\}. \tag{9}$$

- Our **learning objective** is of the form

$$\inf_{v,\rho} \left\{\sup_{\phi\in\Gamma_L} \{\mathbb{E}_{\rho(\cdot,T)}[\phi] - \mathbb{E}_\pi[f^\star(\phi)]\} + \lambda \int_0^T \int_{\mathbb{R}^d} \frac{1}{2}|v(x, t)|^2\rho(x, t)dxdt\right\} \tag{10}$$

$$\frac{dx}{dt} = v(x(t), t),\ x(0) \sim \rho_0,\ t \in [0, T].$$

MFG analysis

- $\mathcal{W}_2$ proximal yields a well-posed Hamilton-Jacobi dynamics in (2).
- $\mathcal{W}_1$ proximal provides a well-defined terminal condition in (2).
- Optimal trajectories are linear.
- The solution for (10) is unique.

Implementation

- The generative flows are learned through **adversarial training** of continuous-time flows.
- We monitor optimality indicators from (2) while training.

Result

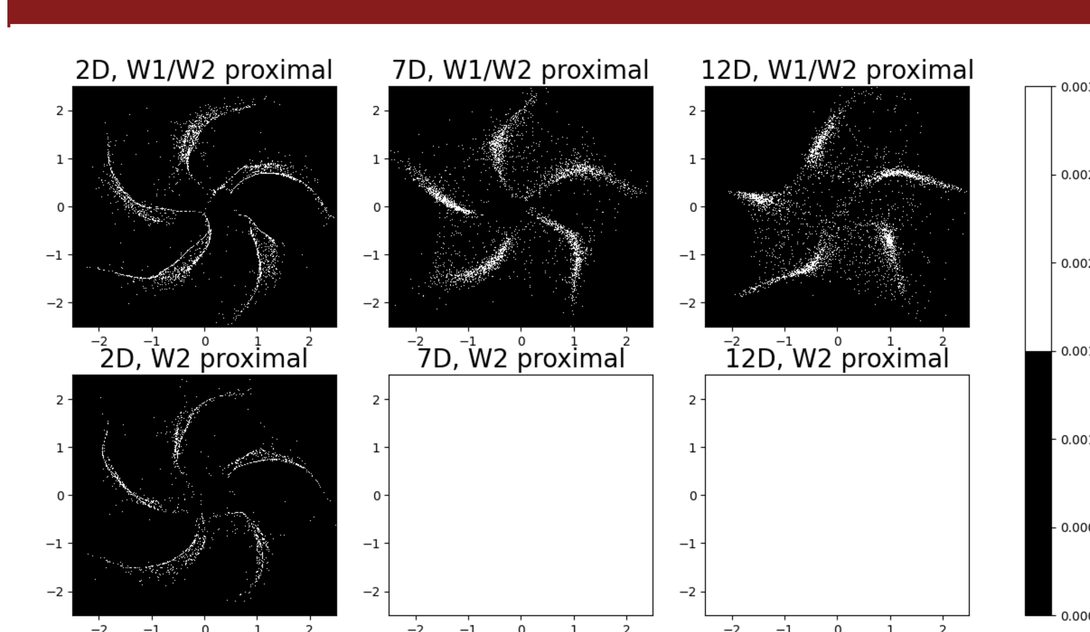

Fig. 1: Stable manifold learning via $\mathcal{W}_1$ proximal. $\mathcal{W}_1 \oplus \mathcal{W}_2$ flow (top), $\mathcal{W}_2$ flow (bottom).

	$\mathcal{W}_1 \oplus \mathcal{W}_2$ flow	Potential Flow GAN	OT flow
2D	**8.0e-03**	1.3e-02	1.9e-01
7D	**1.0e-02**	1.6e+01	4.5e+09
12D	**1.6e-02**	3.7e+00	7.9e+26

Table 1: Comparison with Potential Flow GAN (Yang et al.) and OT flow (Onken et al.). $\mathcal{W}_2$ distance between original and generated data manifolds.

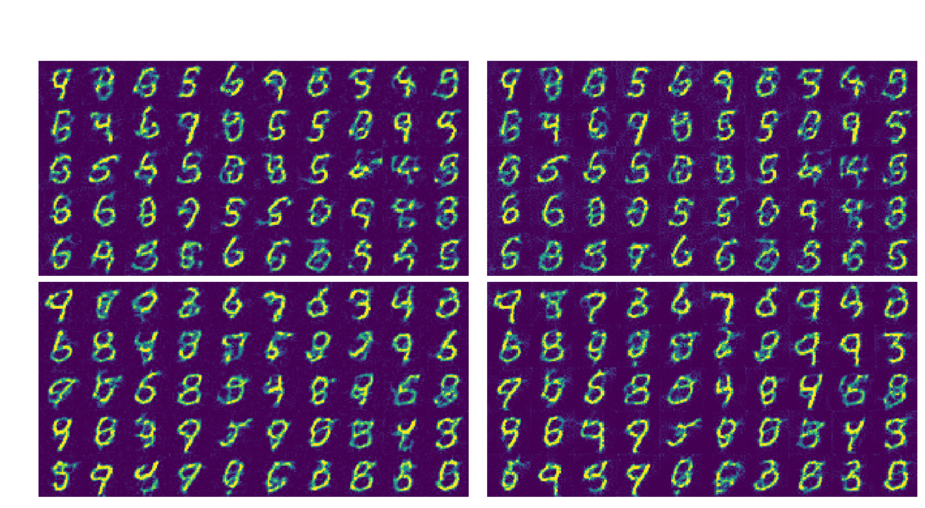

Fig. 2: Discretization-invariant flow learning via $\mathcal{W}_2$ proximal. $\mathcal{W}_1 \oplus \mathcal{W}_2$ flow (top), $\mathcal{W}_1$ flow (bottom) with different time step sizes $h = 2^0$ (left), $h = 2^{-6}$ (right).

Acknowledgement

The research was partially supported by AFOSR grant FA9550-21-1-0354, NSF DMS-2307115 and NSF TRIPODS CISE-1934846.

Link to Paper