A Biconvex Formulation for Stable Transport of Mixture Models with a Unique Solution

Yeganeh Marghi, Kelly Jin, Uygar Sümbül

Allen Institute, Brain Science

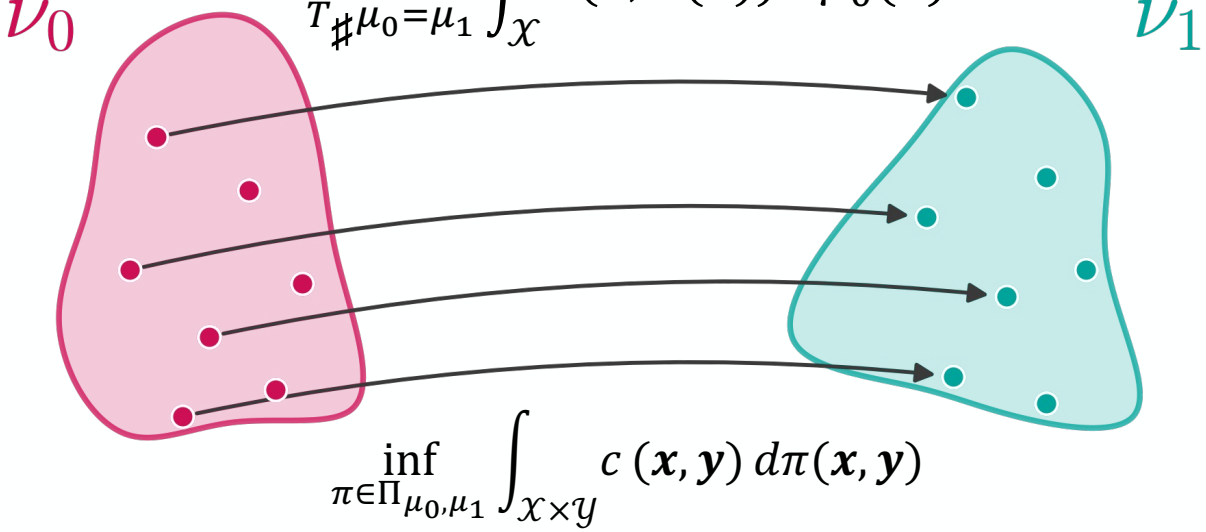
NeurIPS 2026



Optimal Transport (OT)

Monge Problem

$$\inf_{T_\sharp \mu_0 = \mu_1} \int_{\mathcal{X}} c\,(\boldsymbol{x}, T(\boldsymbol{x}))\, d\mu_0(\boldsymbol{x})$$



$$\inf_{\pi \in \Pi_{\mu_0,\mu_1}} \int_{\mathcal{X}\times\mathcal{Y}} c\,(\boldsymbol{x}, \boldsymbol{y})\, d\pi(\boldsymbol{x}, \boldsymbol{y})$$

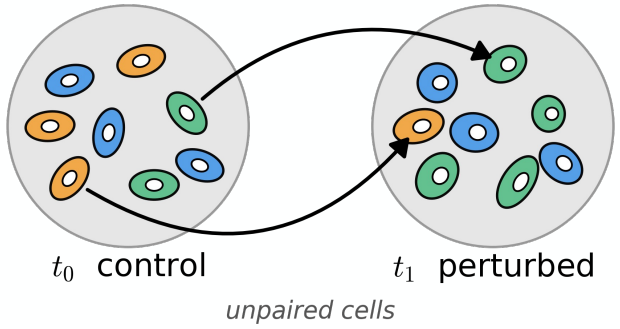
source target

Kantorovich Problem

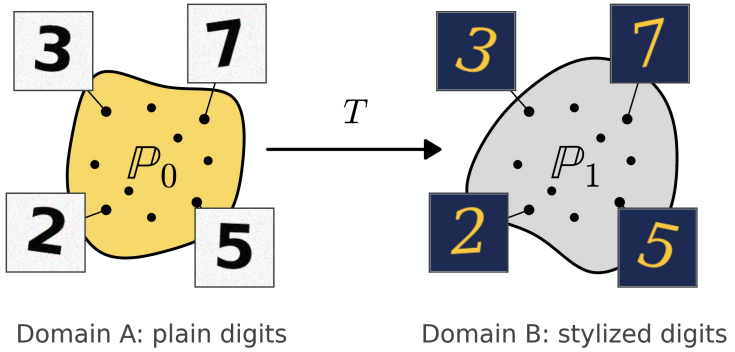
What can we do with OT?

- Uncovering correspondences between unpaired samples across distributions is a fundamental challenge.

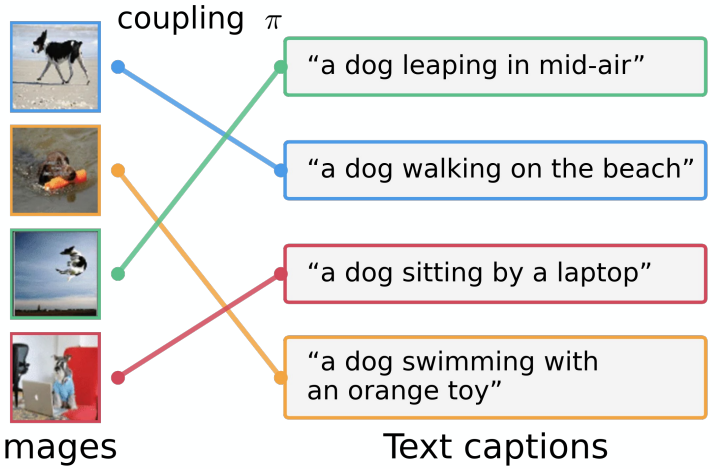
Single-cell analysis



Domain translation



Alignment



The bottleneck: Pointwise transport

- Entropic solvers

$$\inf_{\pi \in \Pi_{\mu_0,\mu_1}} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} \| \boldsymbol{x} - \boldsymbol{y} \|_2^2 \; d\pi(\boldsymbol{x}, \boldsymbol{y}) - 2\varepsilon H(\pi) \right\}$$

- Sinkhorn makes EOT solvable
  - Time and memory drop from $O(n^3 \log n)$ to $O(n^2)$
  - But transport is still pointwise: every sample is coupled to every other sample

Solvers	Scalable (memory)	Unique solution	Theoretical stability
EOT (Sinkhorn)	✗	✓	✓
MB-OT	✓	✗	✗
ProgOT	✗	✓	✓
LR-OT	✓	✗	✗
HiRef	✓	✗	✗
Transport Clustering	✗	✗	✗
Neural OT	✓	✗	✗
GMM-OT	✓	✗	unexplored

Our solution: Optimal Mixture Transport (OMT)

- Moving subpopulations, not individual points

Given two measures $\nu_0 \in M_{K_0}(R^d)$ and $\nu_1 \in M_{K_1}(R^d)$, we define

$$\mathcal{L}_{\varepsilon_1,\varepsilon_2}(\Omega, P) = \sum_{i,j}^{K} \omega_{ij} \left(\int_{\mathcal{X}\times\mathcal{Y}} \| \boldsymbol{x} - \boldsymbol{y} \|_2^2 \; dp_{ij}(\boldsymbol{x}, \boldsymbol{y}) - \varepsilon_1 H(p_{ij}) \right) - \varepsilon_2 H(\Omega)$$

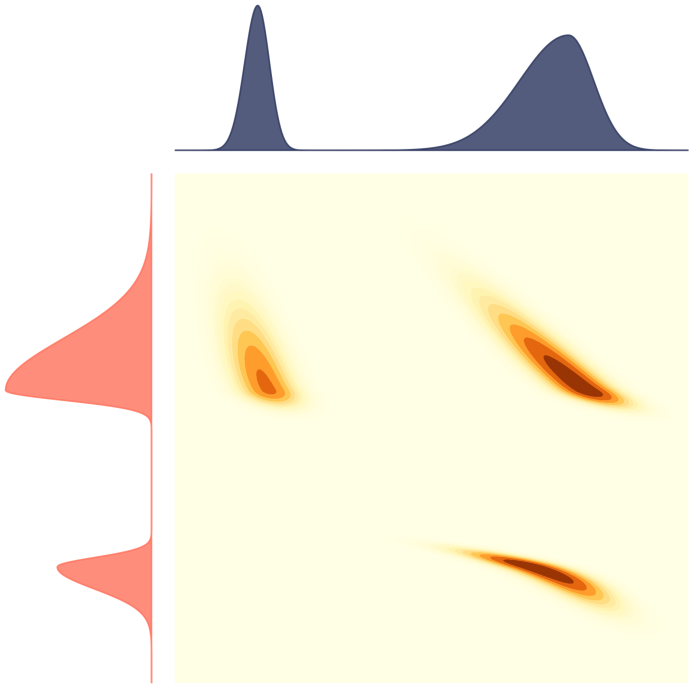
the OMT coupling, defined as $\pi_{\mathrm{OMT}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j} \omega_{ij}\, p_{ij}(\boldsymbol{x}, \boldsymbol{y})$, is obtained by the following constrained optimization.

$$\inf_{\Omega, P} \mathcal{L}_{\varepsilon_1,\varepsilon_2}(\Omega, P)$$

$$\text{s.t.} \quad \mathbf{1}\Omega = \boldsymbol{\alpha}_0, \quad \Omega^T \mathbf{1} = \boldsymbol{\alpha}_1,$$

$$\int_{\mathcal{Y}} \sum_j \frac{\omega_{ij}}{\alpha_{0_i}} \, dp_{ij}(\boldsymbol{x}, \boldsymbol{y}) = \mu_{0_i}(\boldsymbol{x}), \quad \int_{\mathcal{X}} \sum_i \frac{\omega_{ij}}{\alpha_{1_j}} \, dp_{ij}(\boldsymbol{x}, \boldsymbol{y}) = \mu_{1_j}(\boldsymbol{y})$$

where $\Omega \in \mathcal{S}^{K-1}$ and $P \in \mathcal{P}^K(\mathcal{X}\times\mathcal{Y})$.



Theoretical guarantees

- OMT loss is strictly biconvex with a unique solution

Lemma

For any $\varepsilon_1, \varepsilon_2 > 0$, $\mathcal{L}_{\varepsilon_1,\varepsilon_2}(\Omega, P)$ is strictly biconvex.

Theorem

For the regularized OMT problem, the GOP algorithm converges to a unique solution in a single iteration.

Theoretical guarantees

- OMT offers stable transport maps

Theorem (Stability of OMT under perturbation)

Let $\nu_0 \in M_{K_0}(\mathbb{R}^d)$, $\nu_0' \in M_{K_0'}(\mathbb{R}^d)$, and $\nu_1 \in M_{K_1}(\mathbb{R}^d)$ be mixture probability densities that satisfy the density regularity conditions. Then there exist constants $a_0, b_0 \in \mathbb{R}_+$ such that

$$\mathbb{E}_{\nu_0}\left[\left\| T_{\mathrm{OMT}}^{\nu_0 \to \nu_1}(\boldsymbol{x}) - T_{\mathrm{OMT}}^{\nu_0' \to \nu_1}(\boldsymbol{x}) \right\|\right] \leq a_0\, \mathcal{W}_{\mathcal{M}}(\nu_0, \nu_0') + b_0\, \mathcal{W}_{\mathcal{M}}^{1/2}(\nu_0, \nu_0')$$

$$\mathcal{W}_{\mathcal{M}}^2(\nu_0, \nu_1) = \min_{\Omega, P} \sum_{i,j}^{K} \omega_{ij} \int_{\mathcal{X}\times\mathcal{Y}} \| \boldsymbol{x} - \boldsymbol{y} \|_2^2 \; dp_{ij}(\boldsymbol{x}, \boldsymbol{y})$$

$$\text{s.t.} \quad \mathbf{1}\Omega = \boldsymbol{\alpha}_0, \quad \Omega^T \mathbf{1} = \boldsymbol{\alpha}_1,$$

$$\int_{\mathcal{Y}} \sum_j \frac{\omega_{ij}}{\alpha_{0_i}}\, dp_{ij}(\boldsymbol{x}, \boldsymbol{y}) = \mu_{0_i}(\boldsymbol{x}), \quad \int_{\mathcal{X}} \sum_i \frac{\omega_{ij}}{\alpha_{1_j}}\, dp_{ij}(\boldsymbol{x}, \boldsymbol{y}) = \mu_{1_j}(\boldsymbol{y})$$

Closed-form exponential mixtures

We use a general class of mixture models parameterized as $\nu(\boldsymbol{x}) = \sum_i^n \alpha_i \, \mu_i(\boldsymbol{x})$, where $\boldsymbol{x} \in \mathbb{R}^d$.

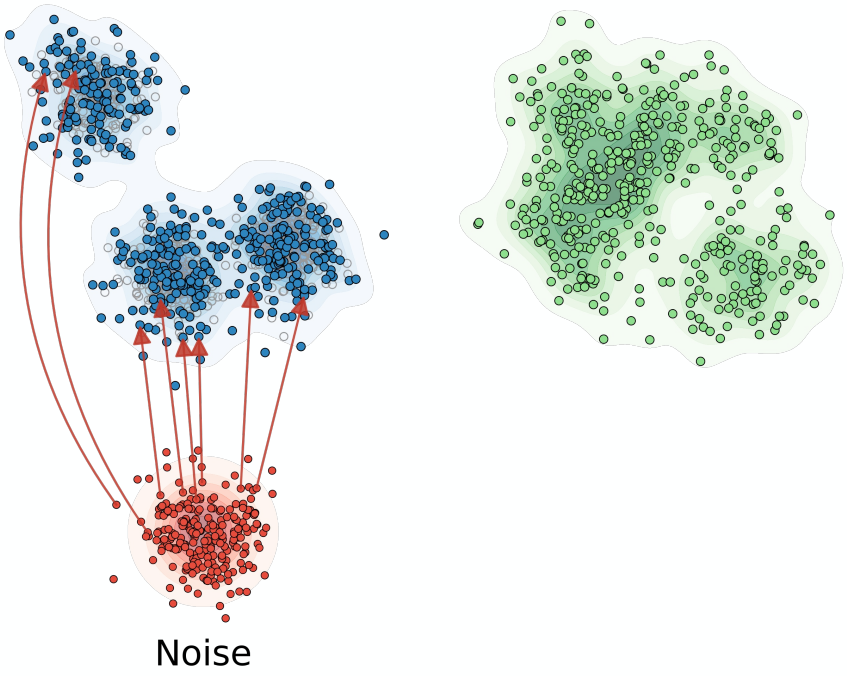
$$\mu_i(\boldsymbol{x}) = \begin{cases} a_i \, e^{-\|\boldsymbol{x}-\boldsymbol{b}_i\|^{p_i}_{\mathrm{C}_i^T \mathrm{C}_i}} & \text{if } \boldsymbol{u}_i^T(\boldsymbol{x}-\boldsymbol{b}_i) \geq 0 \\ a_i \, e^{-\|\boldsymbol{x}-\boldsymbol{b}_i\|^{q_i}_{\mathrm{D}_i^T \mathrm{D}_i}} & \text{otherwise} \end{cases}$$

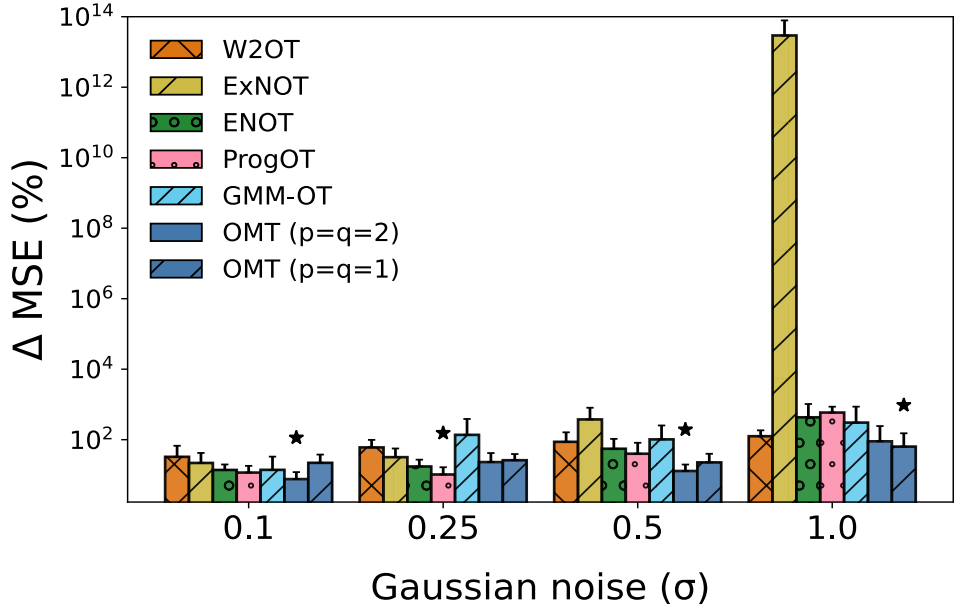
where $a_i \in \mathbb{R}$ is a normalization constant; $\boldsymbol{b}_i \in \mathbb{R}^d$ denotes the location parameter; $\boldsymbol{u}_i \in \mathbb{R}^d$ is a unit normal vector; and $C_i, D_i \in \mathbb{R}^{d \times d}$ are scaling matrices governing the right and left tails along the normal direction, respectively.

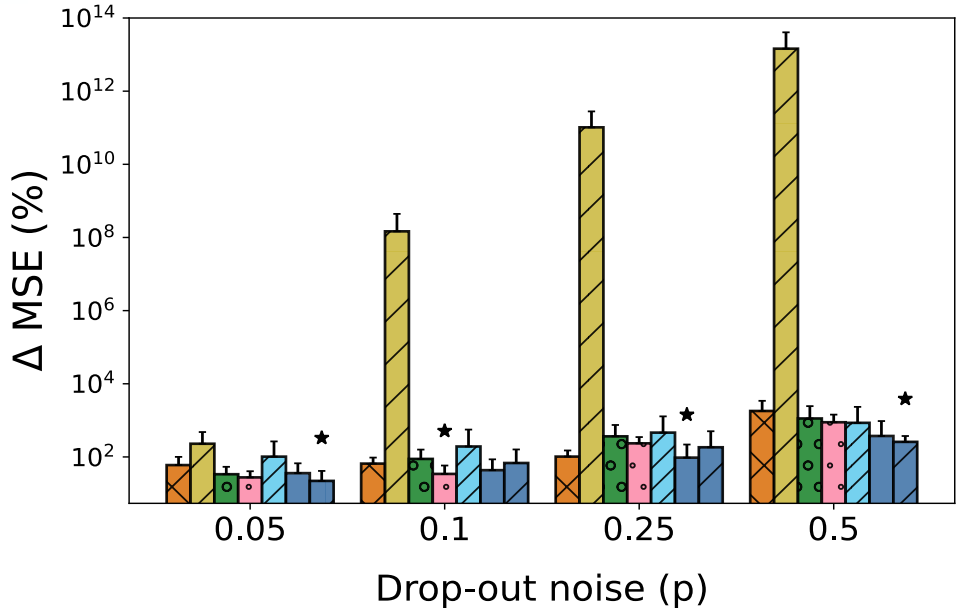
Empirical results

- Wassertian-2 benchmark

Noisy input $\mathbb{P}$ Ground truth output $\nabla\psi^* \sharp \mathbb{P}$







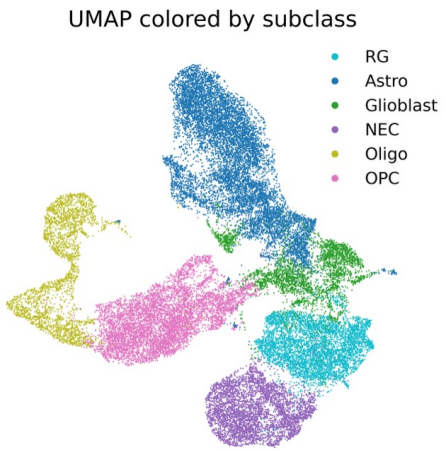
Drug perturbation: Sci-Plex (human single-cell transcriptomics)

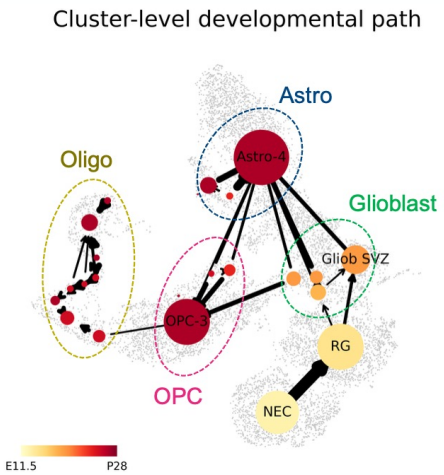
Method	Belinostat	Dacinostat	Givinostat	Quisinostat	Hesperadin
EOT	17.4 ± 0.01	18.4 ± 0.01	17.5 ± 0.01	17.6 ± 0.03	17.5 ± 0.01
HiRef	19.9 ± 0.21	24.0 ± 0.30	19.2 ± 0.55	22.8 ± 0.20	17.3 ± 0.54
ProgOT	8.4 ± 0.01	8.8 ± 0.03	8.8 ± 0.04	9.5 ± 0.01	8.1 ± 0.01
GMM-OT (K_0=3, K_1=5)	8.0 ± 0.02	8.9 ± 0.17	8.9 ± 0.06	8.8 ± 0.23	8.1 ± 0.03
OMT (K_0=3, K_1=5)	7.9 ± 0.06	8.1 ± 0.04	8.7 ± 0.14	8.7 ± 0.10	8.0 ± 0.03

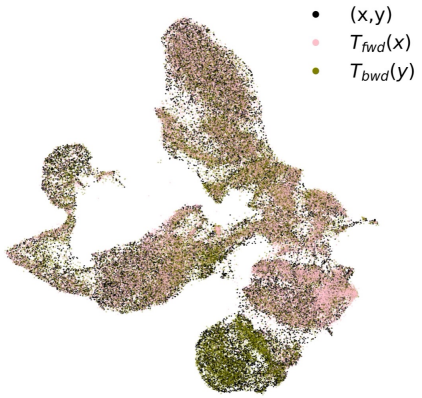
- Average Sinkhorn divergence (↓) for forward and backward OMT maps, compared with ProgOT, HiRef, GMM-OT and EOT on the human scRNA-seq dataset (Sci-Plex).
- Results are mean ± s.d. over 5 random seeds, with d_PCA = 16.
- OMT achieves the lowest divergence across all five drug perturbations.

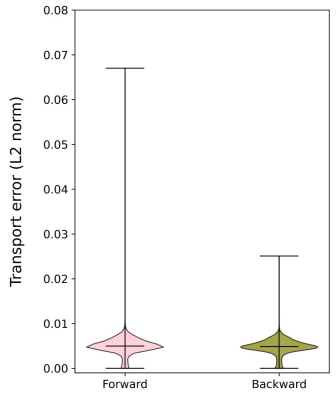
Single-cell trajectories: Mouse brain development and ageing

Developmental trajectory
(10x Genomics, Mouse VISp development)

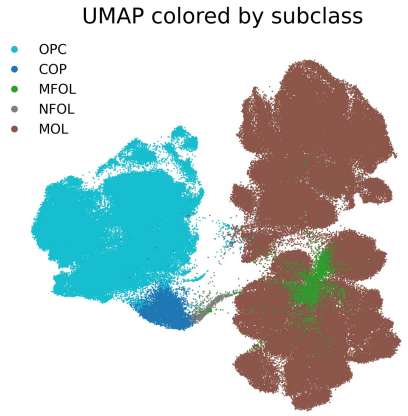


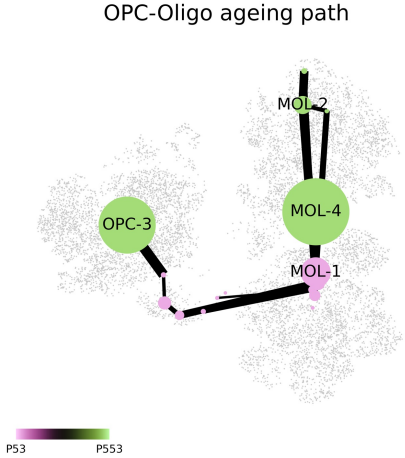


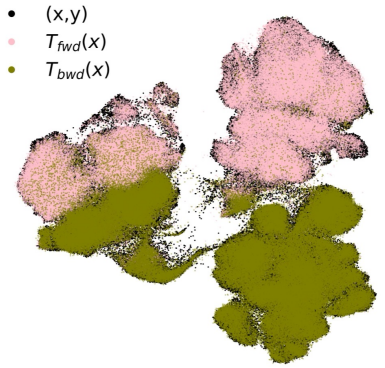


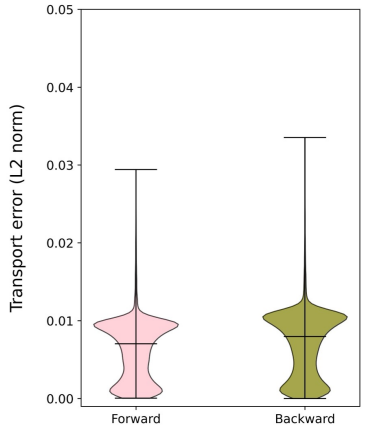


Ageing process
(10x Genomics, Mouse whole brain)



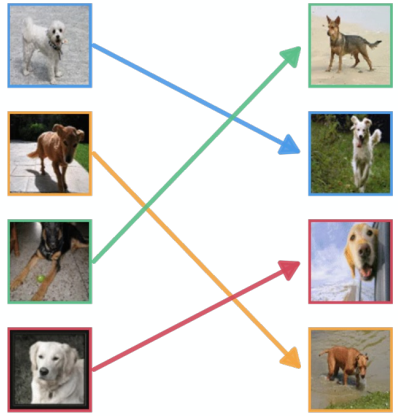






Other applications

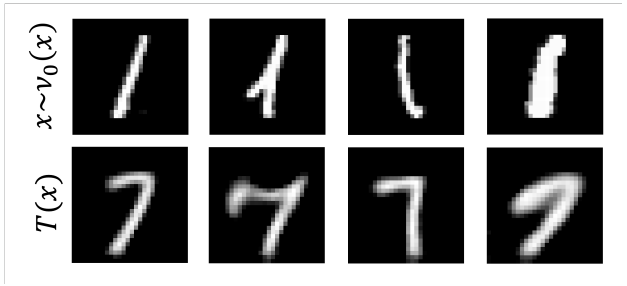
Alignment
(ImageNet)



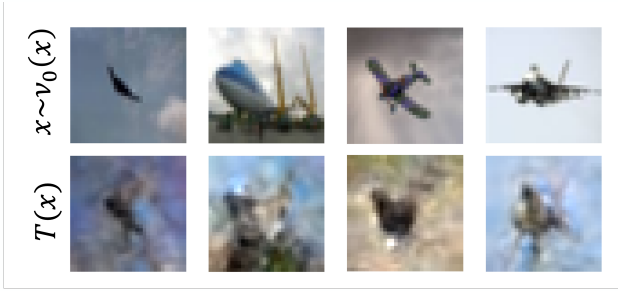
Method	Cost	kNN Acc top-1 / top-5 (%)
MB-OT	19.58	39.63 / 77.38
FRLC	24.12	8.45 / 31.50
HiRef	18.97	40.09 / 61.39
OMT	14.30	84.45 / 98.59

Translation
(MNIST and CIFAR10)

T: 1 → 7



T: airplane → bird



thank you!

See you at Sydney poster session 6

Wednesday, December 9, 5:00 – 8:00 PM

