Supplementary Material

1. Proof of Thm. 6

**Theorem 6** Given $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^d)$, the auxiliary measure is $\mu$, $T_k : \mu \rightarrow \nu_k$ are the OT maps with $k=1,2$. Suppose the distance from $\mu$ to the geodesic connecting $\nu_1$ and $\nu_2$ is $d$, then $T_2 \circ T_1^{-1} : \nu_1 \rightarrow \nu_2$ is measure preserving and its transport cost $C$ is bounded by

$$W_c(\nu_1, \nu_2) \leq C(T_2 \circ T_1^{-1}) \leq W_c(\nu_1, \nu_2) + 2d \quad (1)$$

**Proof.** Suppose the geodesic connecting $\nu_1$ and $\nu_2$ is $\gamma$, $\mu^*$ is the closest point to $\mu$ on $\gamma$. By definition, $(T_k)_{\#} \mu = \nu_k$, then we have

$$(T_2 \circ T_1^{-1})_{\#} \nu_1 = (T_2)_{\#}(T_1^{-1})_{\#} \nu_1 = (T_2)_{\#} \mu = \nu_2. \quad (2)$$

Thus, $T_2 \circ T_1^{-1}$ is measure preserving, but it may not be optimal. Since here we assume that the cost function is given by the $L^2$ distance, we have $C(T_k) = W_c^2(\mu, T_k)$. Then

$$C(T_2 \circ T_1^{-1}) \geq W_c^2(\nu_1, \nu_2). \quad (3)$$

$T_k$’s are the optimal transport maps, according to the triangle inequality, we have

$$C(T_1) + C(T_2) \leq W_c(\nu_1, \nu_2) + 2d. \quad (4)$$

Assume the cell decomposition of $T_1$ and $T_2$ is given by $\{W_i\}$ and $\{W_j\}$, and the refined cell decomposition of $\{W_i\}$ and $\{W_j\}$ is $\{W_{ij}\}$ with $W_{ij} := W_i \cap W_j$. If we set $d(x, y) = \|x - y\|_2$ and by Minkowski inequality,
\[ C^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \]
\[ = \left[ \sum_{i,j=1}^{m,n} \int_{W_{ij}} d(y^1_i, y^2_j)^2 d\mu(x) \right]^{\frac{1}{2}} \]
\[ \leq \left[ \sum_{i,j=1}^{m,n} \int_{W_{ij}} (d(x, y^1_i) + d(x, y^2_j))^2 d\mu(x) \right]^{\frac{1}{2}} \]
\[ \leq \left[ \sum_{i,j=1}^{m,n} \int_{W_{ij}} d(x, y^1_i)^2 d\mu(x) \right]^{\frac{1}{2}} + \left[ \sum_{i,j=1}^{m,n} \int_{W_{ij}} d(x, y^2_j)^2 d\mu(x) \right]^{\frac{1}{2}} \]
\[ = \left[ \sum_{i=1}^{m} \int_{W_i} \|x - y^1_i\|^2 d\mu(x) \right]^{\frac{1}{2}} + \left[ \sum_{j=1}^{n} \int_{W_j} \|x - y^2_j\|^2 d\mu(x) \right]^{\frac{1}{2}} \]
\[ = C^{\frac{1}{2}}(T_1) + C^{\frac{1}{2}}(T_2) \]

Thus,
\[ C^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \leq \mathcal{W}_c(\nu_1, \nu_2) + 2d. \tag{6} \]

Combining the above estimates, we obtain the bounds
\[ \mathcal{W}_c(\nu_1, \nu_2) \leq C^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \leq \mathcal{W}_c(\nu_1, \nu_2) + 2d \tag{7} \]

2 Proof of Proposition 7

**Proposition 7** Given \( \mu = \sum_{i=1}^{m} \nu^1_i N(x_i, \sigma^2 I_d) \) and \( \nu_1 = \sum_{i=1}^{m} \nu^1_i \delta(x - x_i) \), then we have \( \mathcal{W}_c(\mu, \nu_1) \leq \sigma \) under the quadratic Euclidean cost. Moreover, if \( \sigma \) is small enough, then the cell \( W_i \) of the cell decomposition induced by the semi-discrete OT map from \( \mu \) to \( \nu_1 \) should cover \( x_i \) itself.

**Proof.** If we transport all the mass corresponding to \( N(x_i, \sigma^2 I_d) \) to \( x_i \) of \( \nu_1 \), then we get a transport plan from \( \mu \) to \( \nu_1 \). By defining \( f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left( \frac{-x^2}{2\sigma^2} \right) \), the transport cost of such a transport plan is given by
\[ \mathcal{C} = \sum_{i=1}^{m} \nu^1_i \int \|x\|^2 f(d)dx = \sigma^2 \tag{8} \]

Thus, the optimal transport cost from \( \mu \) to \( \nu_1 \), namely \( \mathcal{W}^2_c(\mu, \nu_1) \), should be no more than \( \sigma^2 \). This gives
\[ \mathcal{W}_c(\mu, \nu_1) \leq \sigma \tag{9} \]

When \( \sigma \ll \min_{i \neq j} \|x_i - x_j\|_2 \), the cell \( W_i \) of the cell decomposition induced by the semi-discrete OT map from \( \mu \) to \( \nu_1 \) should cover the corresponding \( x_i \)s, namely nearly all mass of \( \nu^1_i N(x_i, \sigma^2 I_d) \) should be transported to \( x_i \). If \( W_i \) does not cover \( x_i \), some mass of \( N(x_j, \sigma^2 I_d) \) with \( x_j \neq x_i \) will be transported to \( x_i \), as a result \( \mathcal{W}_c(\mu, \nu_1) \) will be larger than \( \sigma \). This corresponds to the cyclical monotonicity of the optimal transport (Chapter 5 of [4]).
Algorithm 1 Semi-discrete OT Map
1: **Input**: the absolutely continuous source measure $\mu$ and the discrete target measure $\nu = \sum_{i=1}^{n} \nu_i \delta(x - x_i)$, number of Monte Carlo samples $N$, positive integer $s$ and the measure accuracy $\theta$.
2: **Output**: Optimal transport map $T(\cdot)$.
3: Initialize $h = (h_1, h_2, \ldots, h_{|I|}) \leftarrow (0, 0, \ldots, 0)$.
4: repeat
5: Sample $N$ samples $\{z_j\}_{j=1}^{N} \sim \mu$.
6: Calculate $\nabla h = (\hat{w}_i(h) - \nu_i)^T$.
7: $\nabla h = \nabla h - \text{mean}(\nabla h)$.
8: Update $h$ by Adam algorithm with $\beta_1 = 0.9, \beta_2 = 0.5$.
9: if $E(h)$ has not decreased for $s$ steps then
10: $N \leftarrow N \times 2$.
11: end if
12: until $\sum_{i=1}^{n} \text{abs}(\hat{w}_i(h) - \nu_i) < \theta$
13: OT map $T(\cdot) \leftarrow \nabla (\max_i \langle \cdot, x_i \rangle + h_i)$.

Algorithm 2 Construct the sparse matrix
1: **Input**: the absolutely continuous source measure $\mu$, the computed $h_1$ for $\nu_1$, and the computed $h_2$ for $\nu_2$.
2: **Output**: Sparse matrix $S$ of the transport plan.
3: Initialize $S = 0_{m \times n}$.
4: repeat
5: Sample $z \sim \mu$.
6: Find the cell $W^1_i$ in $\{W^1_i\}$ that contains $z$.
7: Find the cell $W^2_j$ in $\{W^2_j\}$ that contains $z$.
8: Set $S(i, j) = 1$
9: until converge

3 Algorithm Pipeline for the SDOT algorithm

Based on [1], we summarize the whole pipeline of the SDOT (semi-discrete optimal transport) algorithm in Alg. [1]

4 Algorithm Pipeline for constructing the spare matrix

We also summarize the whole pipeline of constructing and extending the sparse matrix $S$ in Alg. [2]

5 Algorithm for Discrete OT plan with continuous $\mu$ where the source measure is sampled from

In the section, we give the algorithm pipeline for computing the discrete OT plan with the continuous $\mu$ where the source measure $\nu_1$ is sampled from, as shown in Alg. [3]
Algorithm 3 Discrete Optimal Transport Plan

1: **Input:** The absolutely continuous source measure $\mu$, $\nu_1 = \sum_{i=1}^{m} \nu_i^1 \delta(x - x_i)$ and $\nu_2 = \sum_{j=1}^{n} \nu^2_j \delta(y - y_j)$, the $\mu$-volume distortion $\theta$ and the number $k$ of the nearest neighbours.
2: **Output:** The approximate OT plan.
3: Compute the semi-discrete OT map $T_1$ and $T_2$ from $\mu$ to $\nu_1$ and $\nu_2$ with the parameter $\theta$.
4: Initialize the sparse matrix $S$ according to Alg. 2.
5: Extend $S$ according to its $k$ nearest neighbours.
6: Solve the sparse LP problem Eqn. (7).

Algorithm 4 Discrete Optimal Transport Plan by GM model

1: **Input:** $\nu_1 = \sum_{i=1}^{m} \nu_i^1 \delta(x - x_i)$ and $\nu_2 = \sum_{j=1}^{n} \nu^2_j \delta(y - y_j)$, the measure accuracy $\theta$ and the nearest number of $k$.
2: **Output:** The transport plan.
3: Construct $\mu = \sum_{i=1}^{m} \frac{1}{\sigma} \mathcal{N}(x_i, \sigma I_d)$, with $\sigma = 0.1 \min_{i \neq k} d(x_i, x_k)$.
4: Compute the semi-discrete OT map $T_2$ from $\mu$ to $\nu_2$ with the parameter $\theta$ based on Alg. 1.
5: Initialize the sparse matrix $S$: for each sample $x_i$, find the cell $W^2_j$ covering it. Then set $S(i, j) = 1$.
6: Extend $S$ according to the $k$ nearest neighbours.
7: Solve the sparse LP problem of Eqn. (7).

6 Algorithm for Discrete OT plan with Gaussian Mixture $\mu$ defined by the source measure

In this section, we introduce the algorithm to compute the discrete OT plan with $\mu$ being Gaussian mixture model defined by the source measure $\nu_1$, as shown in Alg. 4.

7 More results of Color Transfer

In Fig. 1 we show the additional color transfer results of (i) autumn to communion; (ii) autumn to graffiti; (iii) autumn to rainbow-bridge; (iv) communion to graffiti; and (v) communion to rainbow-bridge. It is obvious that the results of the proposed method are sharper than those of Sinkhorn. And though the color transferred images of SOT are sharp, the color spaces of them are problematic, as shown in the first three images of the 4th column.
Fig. 1. Additional comparison of the results on color transfer tasks.
References