Efficient and Exact Multi-Marginal Optimal Transport with Pairwise Costs

Bohan Zhou and Matthew Parno
Bohan.Zhou@Dartmouth.edu, Matthew.D.Parno@Dartmouth.edu

DARTMOUTH

Problem

We provide an exact and efficient method to solve Multimarginal Optimal Transport (MMOT) under a family of cost functions:

$$\inf_{P\in\Gamma(\mu_1,\cdots,\mu_m)} \int c(x_1,\cdots,x_m)\mathrm{d}P(x_1,\cdots,x_m), \tag{1}$$

for the space $\boldsymbol{X} = X_1\times\cdots\times X_m$ and prescribed marginal probability measures $(\mu_i)_{i=1}^m$. The set of transport plans $\Gamma(\mu_1,\cdots,\mu_m)$ is defined by

$$\Gamma(\mu_1,\cdots,\mu_m) \overset{\text{def}}{=} \{P\in\mathbb{P}(\boldsymbol{X}) \mid (\pi_i)_\# P = \mu_i\}.$$

We assume the cost function $c(x_1,\cdots,x_m)$ is a summed pairwise cost functions

$$c(x_1,\cdots,x_m) = \sum_{1\leqslant i<j\leqslant m} c_{ij}(x_i,x_j);$$

where $c_{ij}(x_i,x_j) = h_{ij}(x_i - x_j)$ for some strictly convex function h_{ij}.

Preliminary: Duality Theory

The dual problem to (1) is given by

$$\sup_{(f_1,\cdots,f_m)} \sum_{i=1}^m \int_{X_i} f_i(x_i)\mathrm{d}\mu_i, \tag{2}$$

where $\sum_{i=1}^m f_i(x_i) \leqslant c(x_1,\cdots,x_m)$.
[Kel84] provided a general duality theorem: there exists a c-conjugate solution to (2) under mild assumption.

Current Methods

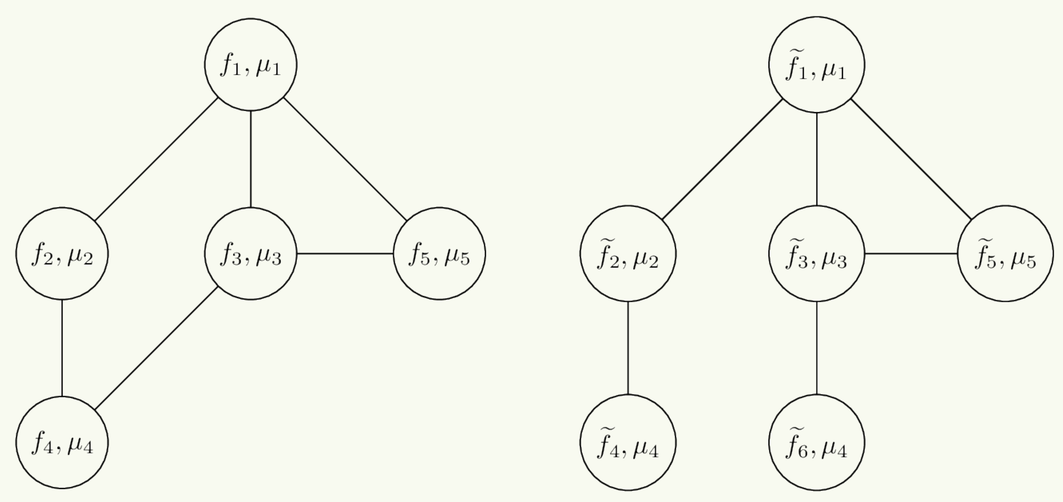
[BCC+15]: Entropy-regularized MMOT on primal variables.
[HRCK21]: Entropy-regularized MMOT with structure on dual variables.
[ABA22]: Solving exact MMOT with structure via ellipsoid algorithm with oracle.
[NX22]: LP-based method to approximate MMOT with controllable level of sub-optimality.
In general, entropy-regularized based methods may suffer from numerical instability and blurring issues. LP based methods may not be practical in solving large-scale problems.

References

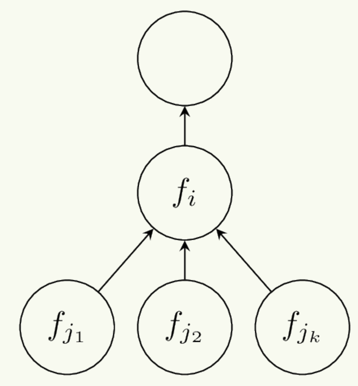
[ABA21] Altschuler and Boix-Adserà. Polynomial-time algorithms for multimarginal optimal transport problems with structure, *Math. Program.*, 2022.
[BCC+15] Benamou, Carlier, Cuturi, Nenna and Peyré. Iterative Bregman projections for regularized transportation problems, *SIAM J. Sci. Comput.*, 2015.
[HRCK21] Haasler, Singh, Zhang, Karlsson and Chen. Multi-marginal optimal transport and probabilistic graphical models, *IEEE Trans. Inform. Theory*, 2021.
[JL20] Jacobs and Léger. A fast approach to optimal transport: the back-and-forth method, *Numerische Mathematik*, 2020.
[NX22] Neufeld and Xiang. Numerical method for feasible and approximately optimal solutions of multi-marginal optimal transport beyond discrete measures, *ArXiv*, 2022.
[ZP22] – and Parno. Exact and Efficient Multi-marginal Optimal Transport with Pairwise Cost. *ArXiv*, 2022.

Equivalent Theorem and Strategy

Step 1: For summed pairwise costs, we may present it into a graph with possible cycle.
Step 2: We proved an equivalent theorem: for any MMOT that has a graphical representation with possible cycles, there exists an equivalent MMOT that has a tree representation. The dual solution of tree-MMOT induces a dual solution of graph-MMOT.
Step 3: Inspired by [JL22]'s back-and-forth method on 2-marginal OT, we maximize (2) by gradient ascent on $(m-1)$ dual variables in $\dot{H}^1$, and c-transform on the root dual variable.



Step 2: Get a tree (Right) from a graph (Left) iteratively



Step 3: Updating a downstream node

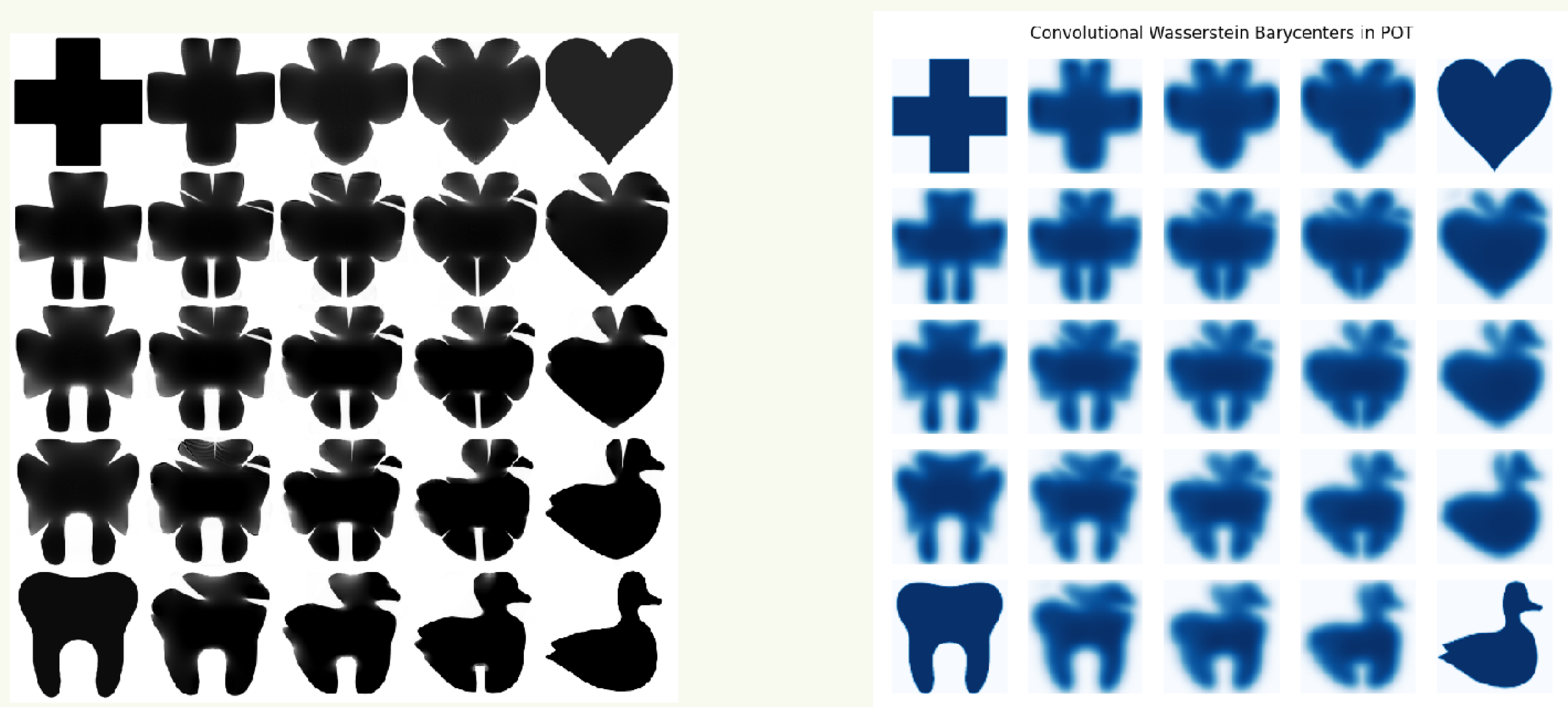
Gradient-ascent on Rooted Tree

Define $I_r(f_1,\ldots,f_{r-1},f_{r+1},\ldots,f_m) \overset{\text{def}}{=} I(f_1,\ldots,f_{r-1},(\sum_{i\neq r} f_i)^c, f_{r+1},\ldots,f_m)$. The updates are:

$$\begin{cases} f_i \leftarrow f_i - \sigma\nabla_{\dot{H}^1} I_r(f_i); \\ f_r \leftarrow (\sum_{i\neq r} f_i)^c. \end{cases} \quad \text{where} \quad \begin{cases} \nabla_{\dot{H}^1} I_r(f_i) = (-\Delta)^{-1}\left(\mu_i - (S_{f_i'})_\#\mu_{N^+(i)}\right) \\ f_r(x_r) = \sum_{i\in N^-(r)} f_i'(x_r). \end{cases}$$

The net potential f_i' at edge $(i, N^+(i))$ we introduced, are recursively defined by $f_i' = \left(f_i - \sum_{j\in N^-(i)} f_j'\right)^{c_{iN^+(i)}}$. $N^-(i)$ are the collections of upsteam nodes of node i and $N^+(i)$ are the downsteam node of node i.

Numerical Results



Left plot: sharp Wasserstein barycenter via our method. Right plot: blurred Wasserstein barycenter via entropy-regularized based method in POT package, regularization parameter is 0.004. Both 4-marginals are given at four corners.