**Optimization over probability distribution** is a joint optimization fundamental algorithmic primitive in data science and machine learning. Distributions in statistics, images in computer vision, object meshes in computer graphics, fMRI scans in neuroscience, …

**Wasserstein Barycenters**

Barycenter: canonical notion of average, given distance.

\[ \arg\min_{\mu \in \mathcal{P}(X)} \sum_{i=1}^{k} d^{2}(\mu_{i}, \nu_{i}) \]

Using the Wasserstein distance captures the geometry.

Integrate vertical distances

Integrate horizontal distances

Optimal transport (aka Wasserstein distance)

\[ L_{2} \text{ Barycenter} \]

Wasserstein Barycenter (computed with our algorithm)

**Exponential-size LP reformulation**

- **Multimarginal Transportation Polytope**: tensors with fixed marginals

\[ \mathcal{M}(\mu_{1}, \ldots, \mu_{k}) = \{ P \in \mathbb{R}^{d}_{+}^{k} : m_{i}(P) = \mu_{i} \} \]

- **Multimarginal Optimal Transport**: LP over this polytope

\[ \min_{P \in \mathcal{M}(\mu_{1}, \ldots, \mu_{k})} \sum_{i=1}^{k} \sum_{j=1}^{n} P_{i,j} C_{i,j} \]

- **Fact**: Wasserstein Barycenter optimization is equivalent to MOT with cost

\[ C_{i,j} = \min_{\pi \in \Pi(x_{i}, y_{j})} \sum_{i} \| x_{i} - y_{j} \|^{2} \]

**Algorithm for fixed dimension (simplified)**

Dual separation oracle (simplified)

- **Given**: \( k \) sets of \( n \) points in \( \mathbb{R}^{d} \)
- **Compute**: \( \min_{x \in \mathcal{P}(\mathbb{R}^{d})^{k}} \min_{y \in \mathbb{R}^{d}} \sum_{i=1}^{k} \| x_{i} - y_{i} \|^{2} \)
- **But, how to optimize non-convex \( F(y) \)?**
- **Polytope convex on finitely many “pieces”**
- **Naive bound is \( n^{k} \) pieces (1 per tuple \( i_{1}, \ldots, i_{k} \))
- **Key lemma**: For fixed \( d \), only \( \text{poly}(n) \) pieces!
- **Key proof technique**: Hyperplanes partition \( \mathbb{R}^{d} \) into few regions.
- **Algorithm**: Enumerate pieces. Easily optimize \( y \) on each piece. Return best.

**Algorithm for fixed dimension (simplified)**

**Solution strategy**

**Classical**

1. LP formulation

- **Wasserstein barycenter**

\[ \text{Pro: finite size LP} \]

\[ \text{Con: } n^{k} \text{ variables} \]

**Steps for solution**

2. Implicit LP ideas

- **MOT with barycenter cost**

**Separation oracle for dual MOT LP**

- **Poly time in fixed dimension**

- **NP-hard in high dimension**

**Key algorithmic insights**: MOT is not a generic LP. Can solve separation oracle efficiently by exploiting the structure of low-dimensional power diagrams.

**Key insight**: geometric convexity of \( \sqrt{\sum_{i=1}^{k} \lambda_{i} C_{i,j}} \) in Bures-Wasserstein space

**Key obstacle**: different separation oracles require very different algorithms

**Key insight**: unified algorithmic framework that captures nearly all applications

**References**


