ON MULTI-MARGINAL PARTIAL OPTIMAL TRANSPORT: RECTIFYING INFEASIBLE EXTENSION STRATEGIES AND EFFICIENT PRIMAL-DUAL METHODS

Quang Nguyen Dinh Thien   Dung Hoang Duc

University of Information Technology, Vietnam National University, Ho Chi Minh City, Vietnam

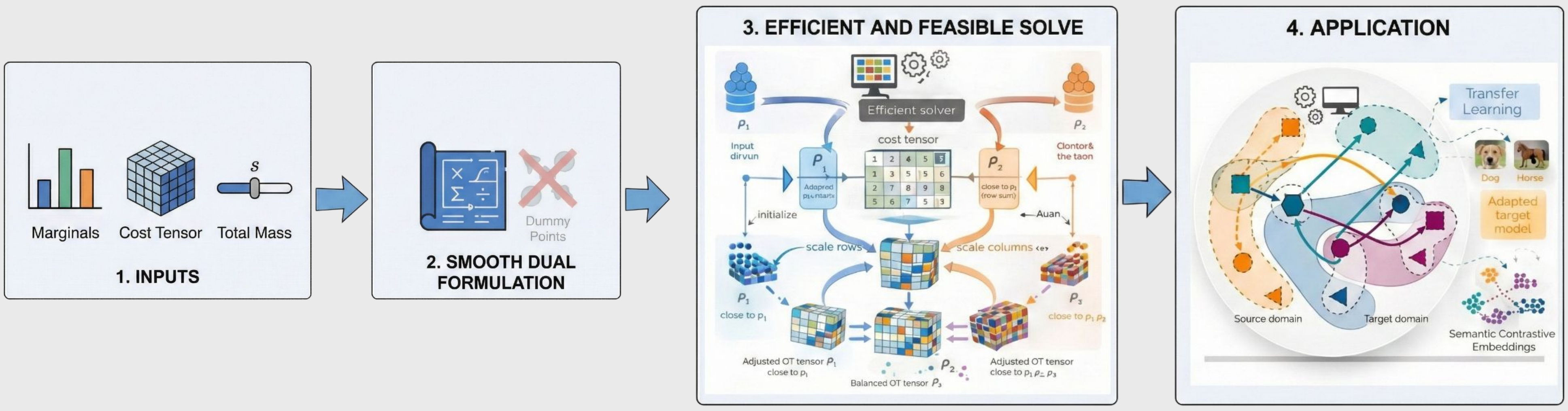
Motivation

- **Robustness:** MMPOT handles **noisy and unbalanced data** significantly better than standard Optimal Transport.
- **Research Gap:** Current "dummy point" extension strategies are mathematically flawed, leading to **infeasible solutions** that violate total mass constraints.
- **Computational Cost:** Previous methods suffer from high complexity $\mathcal{O}(1/\epsilon^4)$, making them impractical for large-scale problems.

Contribution

- **Novel Framework:** A new **Primal-Dual formulation** for MMPOT that eliminates the need for dummy point extensions.
- **Feasibility:** Theoretically rectifies the infeasibility of previous strategies, ensuring strict adherence to mass constraints.
- **Fast Algorithms:** Introduces **GreenkhornMMPOT, PDAAM, and APDAGD** solvers achieving an optimal convergence rate of $\mathcal{O}(1/\epsilon)$.

Overview



Description

1. The Problem with Previous Methods

- **Infeasibility of Extension Strategies:** Current methods rely on adding "dummy points" to balance mass. However, the non-linearity of **Entropic Regularization** mathematically prevents accurate mass adjustment at the marginals.
- **Consequence:** This results in **infeasible solutions** that strictly violate the total mass transportation constraint (s).
- **Complexity Explosion:** To minimize approximation error, the extended cost entries must be extremely large ($A \approx \mathcal{O}(1/\epsilon)$). This causes the Sinkhorn complexity to surge to $\mathcal{O}(1/\epsilon^4)$ (compared to the standard $\mathcal{O}(1/\epsilon^2)$), or even grow exponentially when marginals $m \geq 4$.

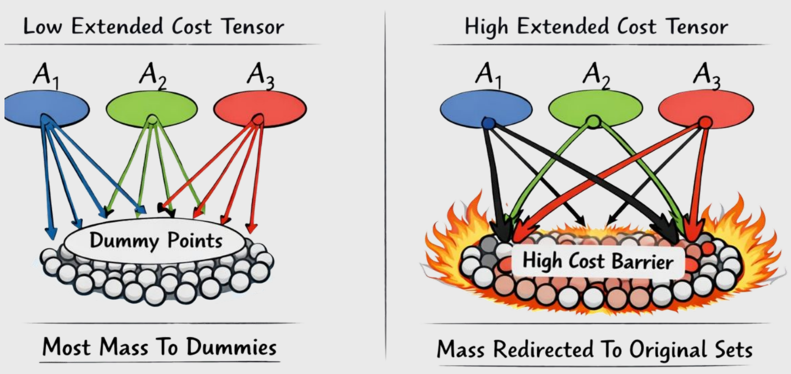


Figure 1. Illustration of mismatch mass behaviour in MMPOT

2. Proposed Algorithms

Novel Primal-Dual Framework: Instead of tensor extension, we establish a **smooth dual formulation** directly for MMPOT. This eliminates approximation errors caused by dummy points.

New Solvers: We introduce three algorithms:

- **Greenkhorn-MMPOT:** Greedy coordinate descent, highly efficient for early iterations.
- **PDAAM & APDAGD:** Accelerated algorithms that achieve the optimal theoretical convergence rate of $\mathcal{O}(1/\epsilon)$.
- **Rounding Algorithm:** A novel rounding method utilizing an "Enforcing Procedure" is proposed to project approximate solutions onto the feasible set, ensuring strict adherence to mass constraints.

Algorithm	Complexity	Extended Cost	Optimal n-dep.	Optimal ε-dep.	All Mass Transport
First Ext. Sinkhorn	$\mathcal{O}\left(\frac{n^m m^5 \|C\|_\infty^2 \log n}{\varepsilon^4}\right)$	✓	✓	×	×
Second Ext. Sinkhorn	$\mathcal{O}\left(\frac{n^m m^5 \|C\|_\infty^2 \log n}{\varepsilon^4}\right)$	✓	✓	×	✓
GreekhornMMPOT	$\mathcal{O}\left(\frac{n^m m^4 \|C\|_\infty \log n}{\varepsilon^2}\right)$	×	✓	×	✓
PDAAM	$\mathcal{O}\left(\frac{n^{m+0.5} m^3 \|C\|_\infty \log n}{\varepsilon}\right)$	×	×	✓	✓
APDAGD	$\mathcal{O}\left(\frac{n^{m+0.5} m^3 \|C\|_\infty \log n}{\varepsilon}\right)$	×	×	✓	✓

Figure 2. Comparison of entropic algorithm solvers for MMPOT

3. Applications & Results

- **Application:** Partial Barycenter computation on noisy datasets (e.g., MNIST). MMPOT is shown to effectively filter outliers by transporting only a specific fraction of the total mass, unlike standard MMOT.
- **Performance:** The proposed algorithms (PDAAM, APDAGD) demonstrate significantly **faster convergence** in terms of both runtime and iteration count compared to the Extension Sinkhorn baselines.
- **Accuracy:** The solutions achieve lower objective value gaps while strictly satisfying the total mass constraints, verifying the theoretical feasibility analysis.

NII

Quang Nguyen Dinh Thien - Dung Hoang Duc – University of Information Technology

TEL : 0833506988 - 0522757578   Email : 23521285@gm.uit.edu.vn - 23520328@gm.uit.edu.vn