

VẬN CHUYỂN TỐI ƯU MỘT PHẦN ĐA BIÊN: KHẮC PHỤC CÁC CHIẾN LƯỢC MỞ RỘNG BẤT KHẢ THI VÀ CÁC PHƯƠNG PHÁP PRIMAL-DUAL HIỆU QUẢ

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

Nguyễn Đình Thiên Quang - 23521285

Hoàng Đức Dũng - 23520328

- Lớp: CS519.Q11.KHTN
- Link Github của nhóm: <https://github.com/Binbonk5/CS519.Q11.KHTN>
- Link YouTube video: <https://www.youtube.com/watch?v=rEUJsG6ulhM>



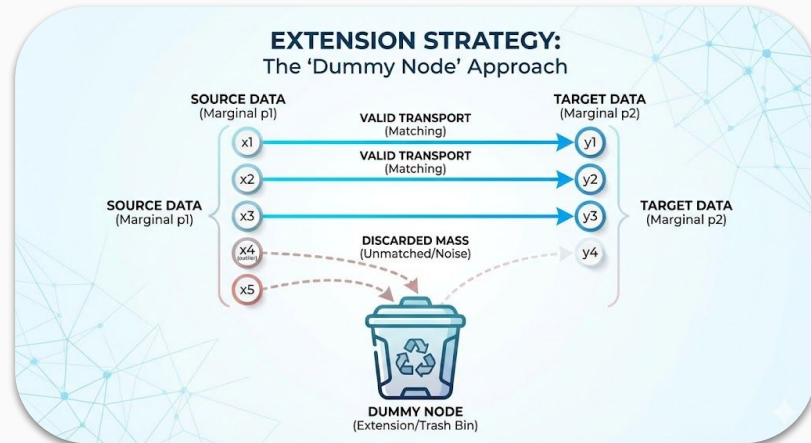
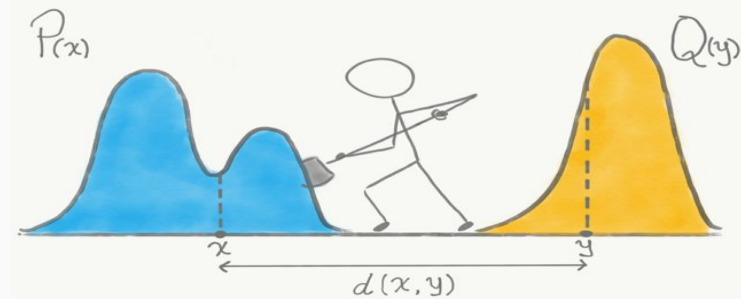
Hoàng Đức Dũng



Nguyễn Đình Thiên Quang

- **Bối cảnh:**
 - **Optimal Transport (OT):** là tiêu chuẩn vàng trong so sánh phân phối (Domain Adaptation, Generative Models).
 - **Hạn chế của OT:** OT bắt buộc bảo toàn khối lượng \rightarrow Nhạy cảm với outliers và dữ liệu không cân bằng.
- **Giải pháp hiện hành:**
 - Multi-Marginal Partial OT (MMPOT) cho phép vận chuyển một phần khối lượng $s < \min \|p_i\|_1$.
 - **Cách tiếp cận SOTA:** Dùng "Dummy Points" để đưa MMPOT về bài toán OT cân bằng và giải bằng Sinkhorn.
- **Vấn đề tồn đọng:**
 - **Tính bất khả thi (Infeasibility):** Chiến lược mở rộng dẫn đến sai số lớn về tổng khối lượng vận chuyển khi dùng điều chuẩn Entropy.
 - **Chi phí tính toán đắt đỏ:** Độ phức tạp bị đẩy lên tới $\mathcal{O}(1/\epsilon^4)$ thay vì mức tối ưu.

➡ Cần một phương pháp giải trực tiếp thay vì mở rộng

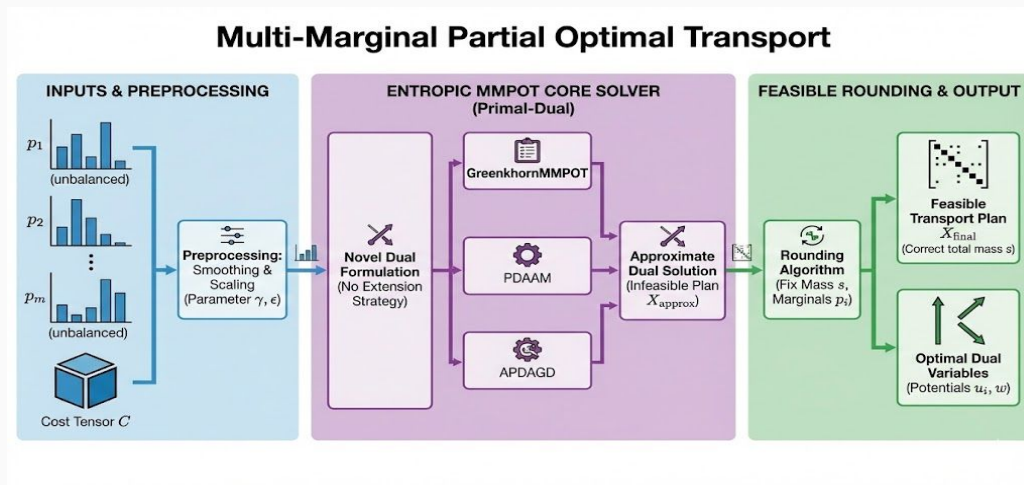


- **Định nghĩa bài toán:**

- **Mục tiêu:** Tối ưu hóa chi phí vận chuyển giữa m phân phối rời rạc p_1, \dots, p_m .
- **Ràng buộc:** Tổng khối lượng vận chuyển $= s$ (với $s \leq \min \|p_i\|_1$).

- **Luồng xử lý:**

- **INPUT:** m Marginals (unbalanced), Cost Tensor $C \in \mathbb{R}^{n^m}$, Mass constraint s .
- **MODEL:** Entropic MMPOT (Primal-Dual Formulation).
- **OUTPUT:** Feasible Transport Plan X & Dual Variables.



Phân tích lý thuyết

Chứng minh tính **infeasibility** và **sự kém hiệu quả** của các chiến lược mở rộng hiện có trong việc giải bài toán Entropic MMPOT.

Phát triển thuật toán

Xây dựng **Dual Formulation mới** và thiết kế các thuật toán tối ưu dựa trên kỹ thuật **đối ngẫu Primal-Dual** (GreenkhornMMPOT, PDAAM, APDAGD) cùng thuật toán làm tròn mới để đảm bảo nghiệm khả thi.

Đánh giá thực nghiệm

Cài đặt các thuật toán và so sánh hiệu năng (thời gian chạy, sai số hội tụ) với các phương pháp SOTA hiện có trên dữ liệu tổng hợp và ứng dụng **Partial Barycenter**.

Nội dung 1: Phân tích sự bất khả thi của các chiến lược mở rộng (Extension Strategies)

Phương pháp thực hiện:

- **Chứng minh tính bất khả thi (Infeasibility):** Sử dụng phân tích toán học để chỉ ra rằng tính phi tuyến của **Entropic Regularization** trong chiến lược thêm điểm giả (dummy points) ngăn cản việc điều chỉnh khối lượng chính xác, dẫn đến lời giải vi phạm trực tiếp ràng buộc tổng khối lượng vận chuyển s .
- **Phân tích bùng nổ độ phức tạp:** Chứng minh rằng để giảm sai số, các phần tử tensor chi phí mở rộng phải đạt ngưỡng cực lớn $A \approx \mathcal{O}(1/\epsilon)$. Điều này khiến độ phức tạp của thuật toán Sinkhorn tăng vọt từ mức chuẩn lên $\mathcal{O}(1/\epsilon^4)$, thậm chí tăng theo **hàm mũ** đối với chiến lược mở rộng thứ 2 khi số lượng biên $m \geq 4$

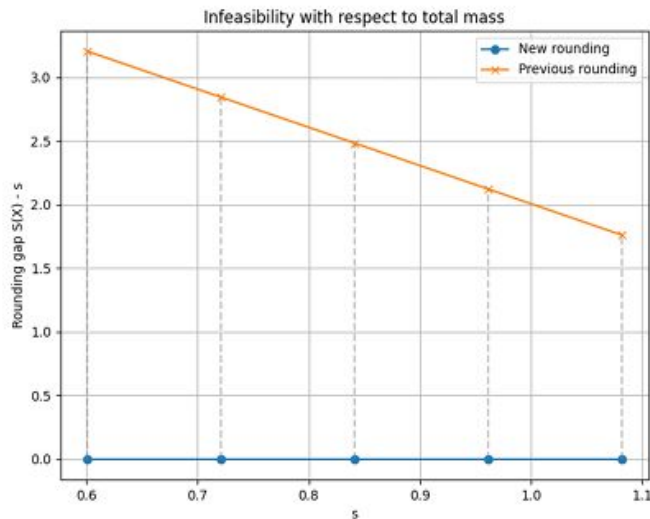


Figure 1. We compare the total transported mass obtained using our previous method(which combines an extension strategy and a multimarginal Sinkhorn algorithm, followed by a rounding step) with the actual target mass.

Nội dung 2: Xây dựng công thức Đối ngẫu mới (Novel Dual Formulation)

Phương pháp thực hiện

- **Thiết lập bài toán gốc (Primal Problem with Slack Variables):**
 - Mô hình hóa MMPOT dưới dạng Quy hoạch tuyến tính (Linear Programming) bằng cách thêm các **biến bù (slack variables)** q_i để xử lý ràng buộc bất đẳng thức của các biên.
 - Hàm mục tiêu được điều chuẩn bởi Entropy: $\min \langle C, X \rangle - \gamma H(X)$
- **Chuyển đổi sang bài toán Đối ngẫu (Dual Formulation):**
 - Thiết lập bài toán tối ưu đối ngẫu **trực tiếp** mà không cần mở rộng tensor chi phí, giúp giữ nguyên kích thước bài toán.
 - **Tính chất hội tụ:** Chứng minh được hàm mục tiêu đối ngẫu là **lồi mạnh** và **trơn** trên miền xác định
- **Chứng minh đóng góp lý thuyết then chốt:**
 - Chứng minh được **Cận trên mới (Novel Upper Bound)** cho chuẩn L_∞ của nghiệm tối ưu. Đây là cơ sở toán học quan trọng để đảm bảo các thuật toán đề xuất (Greenkhorn, PDAAM) hội tụ nhanh.

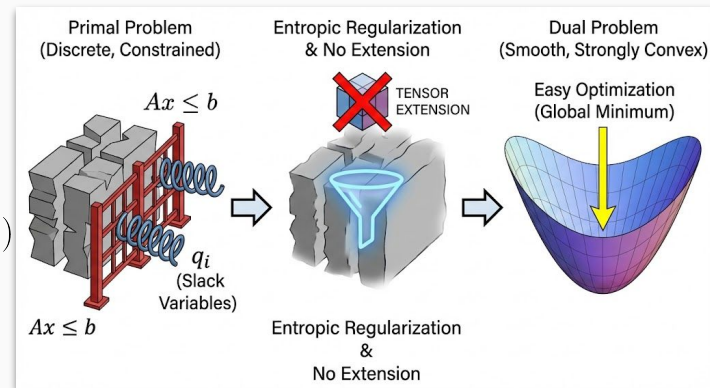


Figure 2. The pipeline of

Nội dung 3: Thiết kế và Cài đặt các thuật toán tối ưu (Solvers) và Làm tròn (Rounding)

Phương pháp thực hiện:

- **Thuật toán Rounding:** Cài đặt quy trình **Enforcing Procedure (EP)** để chiếu nghiệm gần đúng (từ bài toán entropy) về nghiệm khả thi thỏa mãn các ràng buộc biên và tổng khối lượng.
- **GreenkhornMMPOT:** Thiết kế thuật toán Greenkhorn dựa trên nguyên lý **coordinate minimization**, từ đó giải quyết vấn đề về việc thiếu thuật toán giải MMPOT với độ chính xác thấp trong thời gian rất nhanh.
- **PDAAM & APDAGD:** Áp dụng các phương pháp **Accelerated** dựa trên **Gradient Descent** và **Alternating Minimization** để thiết kế thuật giải cho lời giải với độ chính xác cao hơn, tuy nhiên trong thời gian lâu hơn.

Nội dung 4: Thực nghiệm và Đánh giá hiệu năng

Phương pháp thực hiện:

- Sử dụng bộ dữ liệu Gaussian tổng hợp (Synthetic Gaussian) để kiểm tra tính chất hội tụ và độ chính xác lý thuyết.
- Thử nghiệm trên bộ dữ liệu ảnh chữ viết tay MNIST (bài toán Partial Barycenter) để đánh giá khả năng ứng dụng thực tế.
- Đo lường và so sánh dựa trên các metrics: khoảng cách hàm mục tiêu (Objective value gap), thời gian thực thi (runtime), và sai số ràng buộc.

- ★ Báo cáo chứng minh toán học về giới hạn của các phương pháp cũ và tính đúng đắn của phương pháp mới.
- ★ Mã nguồn Python cài đặt hoàn chỉnh 03 thuật toán: GreenkhornMMPOT, PDAAM, APDAGD và thuật toán Rounding.
- ★ Các biểu đồ so sánh cho thấy các thuật toán cho ra lời giải thỏa ràng buộc, xấp xỉ gần đúng với lời giải tối ưu trong thời gian nhanh hơn so với thuật toán SOTA là dummy point strategy.
- ★ Ứng dụng: Giải nhanh bài toán tìm Wasserstein Barycenter, ứng dụng trong domain adaptation, dictionary learning, representation learning, ..., .

- [1]. Anh Duc Nguyen, Khang Le, Quang Nguyen, Tung Pham, Nhat Ho. On partial optimal transport: Revising the infeasibility of sinkhorn and efficient gradient methods. AAAI 2024: 14387-14395.
- [2]. Tianyi Lin, Nhat Ho, Marco Cuturi, Michael I. Jordan. On the Complexity of Approximating Multimarginal Optimal Transport. J. Mach. Learn. Res. 23: 65:1-65:43 (2022).
- [3]. Khang Le, Huy Nguyen, Khai Nguyen, Tung Pham, Nhat Ho. On Multimarginal Partial Optimal Transport: Equivalent Forms and Computational Complexity. AISTATS 2022: 4397-4413.
- [4]. Marco Cuturi. Sinkhorn Distances: Lightspeed Computation of Optimal Transport. NIPS 2013: 2292-2300.
- [5]. Pavel Dvurechensky, Alexander Gasnikov, Alexey Kroshnin. Computational Optimal Transport: Complexity by Accelerated Gradient Descent Is Better Than by Sinkhorn's Algorithm. ICML 2018: 1366-1375