The numerical approximation of the weak optimal transport problem

Benjamin JOURDAIN [1,3] **Kexin SHAO** [1,2,3]

INRIA Paris Équipe Mathrisk [1] CEREMADE, Université Paris Dauphine-PSL [2] CERMICS, École Nationale des Ponts et Chaussées, IP Paris [3]

Introduction

For $\mu, \nu \in \mathcal{P}_\rho(\mathbb{R})$:

- **Classical Optimal Transport (OT) problem:**

$$OT(\mu,\nu,c) := \inf_{\pi \in \Pi(\mu,\nu)} \int c(x,y)\pi(dx,dy), \tag{OT}$$

where $\Pi(\mu,\nu) := \left\{\pi \in \mathcal{P}_\rho(\mathbb{R}\times\mathbb{R}) \mid \pi(dx,\mathbb{R}) = \mu(dx) \text{ and } \pi(\mathbb{R},dy) = \nu(dy)\right\}$.

- **Weak optimal transport (WOT):**

$$V_C(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \int C(x,\pi_x)\mu(dx), \tag{WOT}$$

where $\pi(dx,dy) = \mu(dx)\,\pi_x(dy)$.

- **Weak martingale optimal transport (WMOT):**

$$V_C^M(\mu,\nu) := \inf_{\substack{\pi\in\Pi(\mu,\nu)\\ \int y\,\pi_x(dy)=x\ \mu\text{-a.s.}}} \int C(x,\pi_x)\mu(dx). \tag{WMOT}$$

Numerical Approach

For empirical measures $\mu_I = \frac{1}{I}\sum_{i=1}^I \delta_{x_i}$, $\nu_J = \frac{1}{J}\sum_{j=1}^J \delta_{y_j}$ with (x_i,y_j) i.i.d. from $(\mu,\nu)\in\mathcal{P}_2(\mathbb{R})^2$, denote $p_{j|i} := \pi_{i,j}/\mu_i$ the conditional probability of y_j given x_i.

- **WOT discretization :**

$$\min \sum_{i=1}^I \mu_i C\left(x_i, \sum_{j=1}^J p_{j|i}\delta_{y_j}\right)$$

$$\text{s.t.}\quad p_{j|i}\ge 0,\quad \sum_{j=1}^J p_{j|i} = \mathbf{1},\quad \sum_{i=1}^I \mu_i p_{j|i} = \nu_j.$$

- **WMOT discretization :**

$$\min \sum_{i=1}^I \mu_i C\left(x_i, \sum_{j=1}^J p_{j|i}\delta_{y_j}\right)$$

$$\text{s.t.}\quad p_{j|i}\ge 0,\quad \sum_{j=1}^J p_{j|i} = \mathbf{1},\quad \text{and}\quad \sum_{i=1}^I \mu_i p_{j|i} = \nu_j,\quad \sum_{j=1}^J p_{j|i}\,y_j = x_i.$$

Frank-Wolfe Algorithm

The Frank-Wolfe (conditional gradient) method solves:

$$\min_{x\in\mathcal{C}} f(x)$$

for convex function f and compact convex $\mathcal{C}$, via linear approximation of the objective function.

Algorithm 1 Vanilla Frank-Wolfe Algorithm

Input: initial point $x_0 \in \mathcal{C}$, function $f:\mathbb{R}\to\mathbb{R}$ convex, step sizes $\gamma_t > 0$, smoothness L

Output: Iterates $x_1,\cdots\in\mathcal{C}$

```
for t = 1 to n do
    v_t ← argmin_{v∈C} ⟨∇f(x_t), v⟩
    γ_t ← min{ ⟨∇f(x_t), x_t − v_t⟩ / (L‖x_t − v_t‖²), 1 }
    x_{t+1} ← x_t + γ_t(v_t − x_t)
end for
```

Key features:

- **Projection-free** (uses linear minimization oracle, LMO)
- **Preserves constraints via convex combinations**
- Ideal for our (WOT) and (WMOT) problems since $(p_{j|i}) \mapsto \sum_{i=1}^I \mu_i C\left(x_i, \sum_{j=1}^J p_{j|i}\delta_{y_j}\right)$ is convex when $C(x,\cdot)$ is convex.

Example: Wasserstein projection (WP) in 1D

For $\mu,\nu\in\mathcal{P}_\rho(\mathbb{R})$ and $\rho\in[1,+\infty)$, we consider

$$V^\rho_{WP,\rho}(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}} \left| x - \int_{\mathbb{R}} y\,\pi_x(dy)\right|^\rho \mu(dx).$$

Alfonsi, Corbetta and Jourdain [1] introduced the Wasserstein projections in the convex order:

$$\mathcal{I}_\rho(\mu,\nu) := \text{argmin}\left\{\mathcal{W}_\rho(\mu,\eta) : \eta \le_{cx} \nu\right\}.$$

The quantile functions of $\mathcal{I}(\mu,\nu)$ is obtained for all $u\in(0,1)$ by:

$$F^{-1}_{\mathcal{I}(\mu,\nu)} = F^{-1}_\mu - \partial_-\text{co}(G)(u),$$

where $G(u) := \int_0^u (F_\mu^{-1} - F_\nu^{-1})(v)dv$, co denotes the convex hull, and ∂_- the left-hand derivative. According to [1, 3], the values of $V_{\text{WP},\rho}(\mu,\nu)$ and $\mathcal{W}_\rho(\mu,\mathcal{I}(\mu,\nu))$ coincide.

Example 1 (Constructed by [5]). *Let*

$$F_\mu^{-1}(u) = u\mathbb{1}_{(0,\frac12]}(u) + \frac{12+5u}{18}\mathbb{1}_{(\frac12,1)}(u), \qquad F_\nu^{-1}(u) = \frac{u}{3}\mathbb{1}_{(0,\frac12]}(u) + \frac{u}{2}\mathbb{1}_{(\frac12,1)}(u).$$

It has been checked that

$$F^{-1}_{\mathcal{I}(\mu,\nu)}(u) = \frac{u}{3}\mathbb{1}_{(0,\frac12]}(u) + \frac{3+5u}{18}\mathbb{1}_{(\frac12,1)}(u).$$

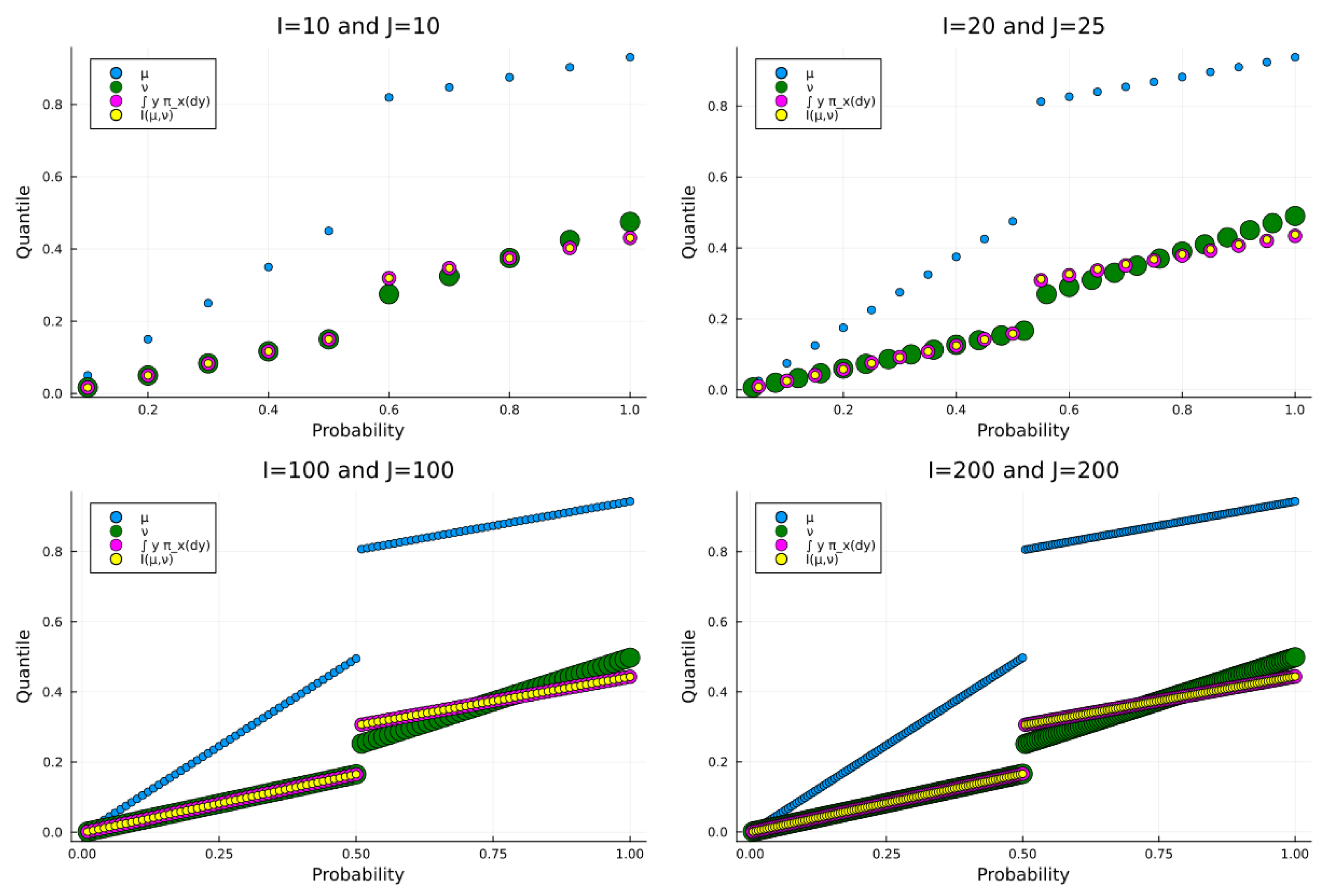

Figure 1: Wasserstein projection with $\rho = 3$ for different problem sizes.

Example: Stretched Brownian motion

For $\mu \le_{cx} \nu$, the martingale Benamou-Brenier problem is:

$$MT(\mu,\nu) := \sup \mathbb{E}\left[\int_0^1 \sigma_t dt\right] \tag{MBB}$$

over martingales $M_t = M_0 + \int_0^t \sigma_s dB_s$ with $M_0\sim\mu$, $M_1\sim\nu$. When ν is finite, (MBB) has a unique maximiser $(M_t^*)_{t\in[0,1]}$ called the stretched Brownian motion [2, Theorem 1.5].

Let $C_2:\mathbb{R}\times\mathcal{P}_2(\mathbb{R})\to\mathbb{R}$ be the cost function as $C_2(x,p) = \mathcal{W}_2^2(p,\mathcal{N}(0,1))$, $\mathcal{N}(0,1)$ denotes the standard normal distribution. Using the optimal Hoeffding-Fréchet coupling for the Wasserstein distance $\mathcal{W}_\rho(\mu,\nu)$, $C_2(x,p)$ can be expressed with quantile functions F_p^{-1} and F_γ^{-1} of p and γ:

$$C_2(x,p) = \mathcal{W}_2^2(p,\gamma) = \int_0^1 \left(F_p^{-1}(u) - F_\gamma^{-1}(u)\right)^2 du.$$

The key relation:

$$V_{C_2}^M(\mu,\nu) = 1 + \int y^2\nu(dy) - 2MT(\mu,\nu).$$

We take two re-centered log-normal distributions $\mu = \exp\#\mathcal{N}(0,0.24)$ and $\nu = \exp\#\mathcal{N}(-0.0104, 0.28)$ and apply Baker's construction [4] to preserve the convex order. Then, we apply the Frank-Wolfe to demonstrate the optimal coupling π^* that attains $V_{C_2}^M$.

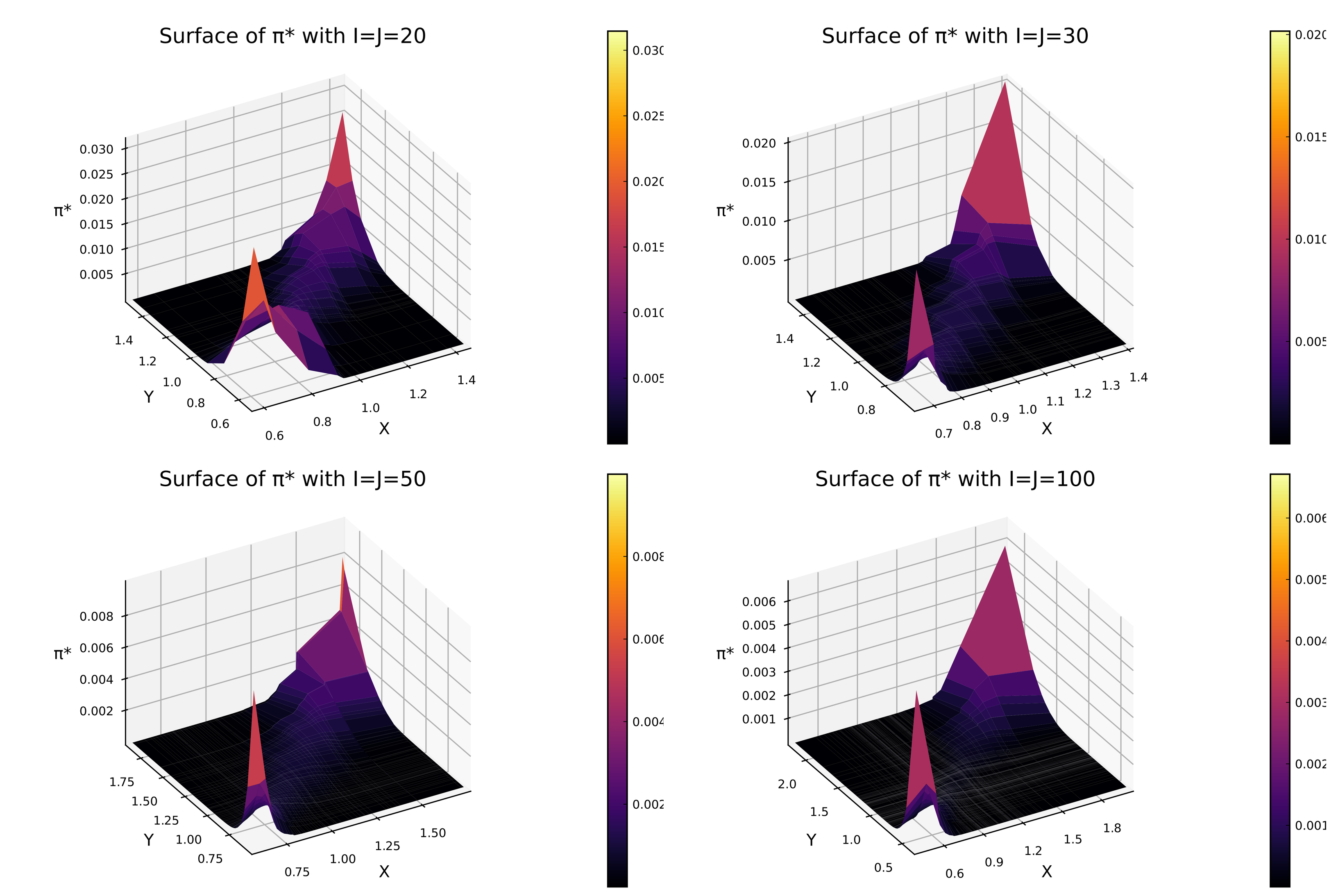

Figure 2: Optimal martingale coupling $\mu(dx)\pi_x^*(dy)$ for different problem sizes.

References

[1] Aurélien Alfonsi, Jacopo Corbetta, and Benjamin Jourdain. Sampling of probability measures in the convex order by Wasserstein projection. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56(3):1706 – 1729, 2020.

[2] Julio Backhoff-Veraguas, Mathias Beiglböck, Martin Huesmann, and Sigrid Källblad. Martingale Benamou–Brenier: A probabilistic perspective. *The Annals of Probability*, 48(5):2258 – 2289, 2020.

[3] Julio Backhoff-Veraguas, Mathias Beiglböck, and Gudmun Pammer. Existence, duality, and cyclical monotonicity for weak transport costs. *Calculus of Variations and Partial Differential Equations*, 58(6):203, 2019.

[4] Baker David. *Martingales with specified marginals.* Thèse, Université Pierre et Marie Curie - Paris VI, December 2012.

[5] Benjamin Jourdain, William Margheriti, and Gudmund Pammer. Lipschitz continuity of the wasserstein projections in the convex order on the line. *Electronic Communications in Probability*, 28:1–13, 2023.