Stability of Entropic Optimal Transport and Convergence of Sinkhorn’s Algorithm

Marcel Nutz
Columbia University

June 2022
Joint work with

Espen Bernton  Stephan Eckstein  Promit Ghosal  Johannes Wiesel
E. Bernton, P. Ghosal, and M. Nutz.
Enterropic optimal transport: Geometry and large deviations.

S. Eckstein and M. Nutz.
Quantitative stability of regularized optimal transport and convergence of Sinkhorn’s algorithm.
_Preprint arXiv:2110.06798_.

P. Ghosal, M. Nutz, and E. Bernton.
Stability of entropic optimal transport and Schrödinger bridges.
_J. Funct. Anal., to appear_.

M. Nutz.
*Introduction to Entropic Optimal Transport.*
Lecture notes, Columbia University.

M. Nutz and J. Wiesel.
Entropic optimal transport: Convergence of potentials.
_Probab. Theory Related Fields, to appear_.

M. Nutz and J. Wiesel.
Stability of Schrödinger potentials and convergence of Sinkhorn’s algorithm.
_Preprint arXiv:2201.10059_.

Outline

1. Entropic Optimal Transport
2. Weak Stability
3. $W_p$ Stability
4. Sinkhorn’s Algorithm
5. Dual Picture
Monge–Kantorovich Optimal Transport

Given:
- Probability spaces \((X, \mu)\) and \((Y, \nu)\), Polish
- Cost function \(c : X \times Y \to \mathbb{R}_+\), continuous

Problem:
- Find a coupling \(\pi\) of the marginals \(\mu, \nu\) such as to minimize the cost:

\[
C_0(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int c(x, y) \pi(dx, dy)
\]

with \(\Pi(\mu, \nu) = \{\pi : (\text{proj}_X)_\# \pi = \mu, (\text{proj}_Y)_\# \pi = \nu\}\)
Entropic Optimal Transport

- Regularization parameter $\varepsilon > 0$
- Entropic optimal transport (EOT) problem:

$$
C_\varepsilon := \inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times Y} c \, d\pi + \varepsilon H(\pi | P), \quad P := \mu \otimes \nu
$$

- $H(\cdot | P)$ is relative entropy (Kullback–Leibler divergence) wrt. $P$,

$$
H(\pi | P) := \begin{cases}
\int \log \left( \frac{d\pi}{dP} \right) d\pi, & \pi \ll P, \\
\infty, & \pi \not\ll P.
\end{cases}
$$

- Call problem finite if $C_\varepsilon < \infty$
- In that case, unique minimizer $\pi_\varepsilon$, and $\pi_\varepsilon \sim P$
- EOT is tradeoff between transport cost and entropy
- “Interpolates” between $P$ and optimal transport
Entropic Regularization

(Figure from Peyré–Cuturi 2019)
Properties of EOT

- Computation through Sinkhorn’s algorithm (IPFP)
- Solve EOT for small $\varepsilon$ to approximate OT

EOT has many desirable properties related to smoothness:
EOT as cost allows for gradient descent, improved sampling complexity, ... Sinkhorn divergence, differentiable ranks, ...

EOT can also be written as pure entropy minimization problem:
the static Schrödinger bridge problem (Föllmer, Léonard, ...)

$$\pi_\varepsilon = \arg \min_{\Pi(\mu,\nu)} H(\cdot | R) \quad \text{for} \quad dR := \frac{e^{-c/\varepsilon}}{EP[e^{-c/\varepsilon}]} dP$$
Outline

1. Entropic Optimal Transport
2. Weak Stability
3. $W_p$ Stability
4. Sinkhorn’s Algorithm
5. Dual Picture
Geometry of Optimal Transport

- **c-cyclical monotonicity** captures geometry
- \( \pi \) is an optimal transport iff \( \text{spt} \pi \) is \( c \)-cyclically monotone
- A cornerstone of modern OT theory: stability, Monge solutions, etc.

**Definition:** A set \( \Gamma \subset X \times Y \) is \( c \)-cyclically monotone if for all \((x_i, y_i) \in \Gamma, 1 \leq i \leq k,\)

\[
\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{i+1}) \quad \text{where} \quad y_{k+1} := y_1
\]
Definition: $\pi \in \Pi(\mu, \nu)$ is $(c, \varepsilon)$-cyclically invariant if $\pi \sim P$ and

$$\prod_{i=1}^{k} \frac{d\pi}{dP}(x_i, y_i) = \exp \left( -\frac{1}{\varepsilon} \left[ \sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{i+1}) \right] \right) \prod_{i=1}^{k} \frac{d\pi}{dP}(x_i, y_{i+1})$$

for all $k \in \mathbb{N}$ and $(x_i, y_i)_{i=1}^{k} \subset X \times Y$, where $y_{k+1} := y_1$.

- Equivalently, $\prod_{i=1}^{k} \frac{d\pi}{dR}(x_i, y_i) = \prod_{i=1}^{k} \frac{d\pi}{dR}(x_i, y_{i+1})$
- Equivalently, density admits a factorization $\frac{d\pi}{dR}(x, y) = a(x)b(y)$
  Borwein–Lewis (1992), Rüschendorf–Thomsen (1997), …

- **Main novelty:** tool used along the lines of $c$-cyclical monotonicity
Relation to Optimality

If EOT problem is finite:

- $\pi$ cyclically invariant $\iff$ $\pi$ is the minimizer.

In OT, geometry can single out a unique coupling even if optimization is not meaningful. McCann (1995), ... 

General EOT problem:

- Uniqueness: There exists are most one cyclically invariant coupling
- Existence: See below
Stability Theorem for EOT

- Marginals \((\mu_n, \nu_n) \to (\mu, \nu)\) converging weakly
- Measurable cost functions \(c_n \to c\) locally uniformly
- \(\varepsilon_n \to \varepsilon > 0\)
- Stability: associated EOT solutions satisfy \(\pi_n \to \pi\) weakly

If \(X, Y\) are Euclidean spaces, we can show:

**Theorem**

Let \(\pi_n\) be cyclically invariant wrt. \((c_n, \varepsilon_n, \mu_n, \nu_n)\). Then \(\pi_n\) converges weakly and the limit \(\pi\) is cyclically invariant wrt. \((c, \varepsilon, \mu, \nu)\).

- If the EOT problems are all finite, this states the stability of the optimizers
- Implies existence of cyclically invariant coupling: approximate \((\mu, \nu)\) with discrete marginals. Alternative proof (cf. OT).
Remarks on the Proof

- Imitate $c$-cyclical monotonicity: fix $(x_i, y_i)$ and pass to limit in

$$\prod_{i=1}^k \frac{d\pi_n}{dR_n}(x_i, y_i) = \prod_{i=1}^k \frac{d\pi_n}{dR_n}(x_i, y_{i+1})$$

- Weak convergence and densities are not immediately compatible
- **Blow up** points to balls, pass to limit of integrals, **shrink** back

→ Condition: spaces $X, Y$ satisfy a version of Lebesgue’s theorem on differentiation of measures

- **Step 1:** $\pi_n \to \pi$ and $R_n \to R$ imply $\pi \ll R$. Uses only a local boundedness condition on $dR_n/dP_n$. Based on rigidity:

$$\prod_{i=1}^k \pi_n(A_i \times B_i) \approx \prod_{i=1}^k \pi_n(A_i \times B_{i+1})$$

- **Step 2:** Pass to limit in the display
Outline

1. Entropic Optimal Transport
2. Weak Stability
3. $W_p$ Stability
4. Sinkhorn’s Algorithm
5. Dual Picture
Setting

- \( c \) continuous with growth of order \( p \)
- \( N \) marginals \( \mu_1, \ldots, \mu_N \) on Polish spaces \( X_1, \ldots, X_N \)
- Marginals have finite \( p \)-th moments
- Distance of marginals measured by

\[
W_p(\mu_1, \ldots, \mu_N; \tilde{\mu}_1, \ldots, \tilde{\mu}_N) := \begin{cases} 
\left( \sum_{i=1}^N W_p(\mu_i, \tilde{\mu}_i)^p \right)^{1/p}, & p < \infty, \\
\max_{i=1, \ldots, N} W_\infty(\mu_i, \tilde{\mu}_i), & p = \infty.
\end{cases}
\]

- Aim: estimate distance of value functions and optimizers
Continuity of Value

EOT value (for $\varepsilon = 1$) is

$$C(\mu_1, \ldots, \mu_N) := \inf_{\pi \in \Pi(\mu_1, \ldots, \mu_N)} \int c \, d\pi + H(\pi|P), \quad P := \mu_1 \otimes \cdots \otimes \mu_N$$

Theorem

Let $p \in [1, \infty]$.

(i) Let $\mu_i, \mu_i^n \in \mathcal{P}_p(X_i)$ satisfy $\lim_n W_p(\mu_i, \mu_i^n) = 0$ for $i = 1, \ldots, N$. Then $C(\mu_1^n, \ldots, \mu_N^n) \to C(\mu_1, \ldots, \mu_N)$ and the associated optimal couplings converge in $W_p$.

(ii) Let $\mu_i, \tilde{\mu}_i \in \mathcal{P}_p(X_i)$ for $i = 1, \ldots, N$ and let $c$ satisfy $(A_L)$. Then

$$|C(\mu_1, \ldots, \mu_N) - C(\tilde{\mu}_1, \ldots, \tilde{\mu}_N)| \leq LW_p(\mu_1, \ldots, \mu_N; \tilde{\mu}_1, \ldots, \tilde{\mu}_N).$$
Lipschitz-type Condition \((A_L)\)

**Definition**

Let \(p \in [1, \infty]\) and \(\mu_i, \tilde{\mu}_i \in \mathcal{P}_p(X_i), \ i = 1, \ldots, N\). For a constant \(L \geq 0\), we say that \(c\) satisfies \((A_L)\) if

\[
\left| \int c \, d(\pi - \tilde{\pi}) \right| \leq LW_p(\pi, \tilde{\pi}) \tag{A_L}
\]

for all \(\pi \in \Pi(\mu_1, \ldots, \mu_N)\) and \(\tilde{\pi} \in \Pi(\tilde{\mu}_1, \ldots, \tilde{\mu}_N)\).

**Example:** \((A_L)\) holds for \(c(x_1, x_2) = \|x_1 - x_2\|^2\) on \(\mathbb{R}^d \times \mathbb{R}^d\) and \(p = 2\), with constant

\[
L := \sqrt{2} \left[ M(\mu_1) + M(\tilde{\mu}_1) + M(\mu_2) + M(\tilde{\mu}_2) \right]
\]

where \(M(\mu) := \left( \int \|x\|^2 \mu(dx) \right)^{1/2}\). More generally, it holds for \(\|x_1 - x_2\|^p\).
Lipschitz-type Condition ($A_L$) (cont’d)

**Idea:** Relax requirements on $c$ by making constant depend on marginals

**Lemma**

Let $p \in [1, \infty)$ and

$$c(x) = c_1(x)c_2(x)$$

where $c_1, c_2$ are Lipschitz and have growth of order at most $p - 1$.

Then ($A_L$) holds with a constant $L$ depending only on the Lipschitz and growth constants of $c_1, c_2$ and the $p$-th moments of $\mu_i, \tilde{\mu}_i, i = 1, \ldots, N$.

For $p = \infty$, the analogue holds with dependence on the bounds of $c_1, c_2$.

Generalizes to product $c(x) = c_1(x) \cdots c_m(x)$ of $m$ Lipschitz functions satisfying a suitable growth condition.
Stability of Optimizers

**Theorem**

Let \( p \in [1, \infty] \) and \( q \in [1, \infty) \) with \( q \leq p \) and let \( \mu_i, \tilde{\mu}_i \in \mathcal{P}_p(X_i) \).

Let \( \mu_1, \ldots, \mu_N \) satisfy \((T'_q)\) with constant \( C'_q \), and let \( c \) satisfy \((A_L)\).

The optimizers \( \pi^*, \tilde{\pi}^* \) of \( \mu_1, \ldots, \mu_N \) and \( \tilde{\mu}_1, \ldots, \tilde{\mu}_N \) satisfy

\[
W_q(\pi^*, \tilde{\pi}^*) \leq N \left( \frac{1}{q} - \frac{1}{p} \right) \Delta + C'_q \left[ (2L\Delta) \frac{1}{q} + (L \Delta) \frac{1}{2q} \right],
\]

\[
\Delta := W_p(\mu_1, \ldots, \mu_N; \tilde{\mu}_1, \ldots, \tilde{\mu}_N).
\]

In particular, \((\mu_1, \ldots, \mu_N) \mapsto \pi^*\) is \( \frac{1}{2p} \)-Hölder continuous in \( W_p \) when restricted to a bounded set of marginals satisfying \((A_L)\) and \((T'_p)\) with given constants.
Transport Inequality \((T'_{q})\)

Let \(q \in [1, \infty)\). We say that \(\mu_i \in \mathcal{P}_q(X_i), i = 1, \ldots, N\) satisfy \((T'_{q})\) with constant \(C'_q\) if for all \(\pi, \theta \in \Pi(\mu_1, \ldots, \mu_N)\),

\[
W_q(\theta, \pi) \leq C'_q \left[ H(\theta|\pi)^{\frac{1}{q}} + \left( \frac{H(\theta|\pi)}{2} \right)^{\frac{1}{2^q}} \right]
\]

\((T'_{q})\)

Lemma (Based on Bolley–Villani 2005)

If \(\mu_i \in \mathcal{P}(X_i)\) satisfy \(\int \exp(\alpha \, d_{X_i}(\hat{x}_i, x_i)^q) \, \mu_i(dx_i) < \infty\) for some \(\alpha \in (0, \infty)\) and \(\hat{x}_i \in X_i\), then \((T'_{q})\) holds with constant

\[
C'_q = 2 \inf_{\hat{x} \in X, \alpha > 0} \left( \frac{1}{\alpha} \sum_{i=1}^{N} \left( \frac{3}{2} + \log \int \exp(\alpha \, d_{X_i}(\hat{x}_i, x_i)^q) \, \mu_i(dx_i) \right) \right)^{\frac{1}{q}}.
\]
Outline

1. Entropic Optimal Transport
2. Weak Stability
3. $W_p$ Stability
4. Sinkhorn’s Algorithm
5. Dual Picture
Sinkhorn’s Algorithm (Primal Formulation)

$N = 2$ marginals $\mu_1, \mu_2$.

Algorithm (Sinkhorn)

Set $\pi^0 := R$ and define $\pi^n$, $n \geq 1$ recursively via

\[
\frac{d\pi^n}{d\pi^{n-1}}(x) := \begin{cases} 
\frac{d\mu_1}{d\pi^{n-1}_1}(x_1) & \text{for } n \text{ odd,} \\
\frac{d\mu_2}{d\pi^{n-1}_2}(x_2) & \text{for } n \text{ even,}
\end{cases}
\]

where $\pi^{n-1}_i$ is the $i$-th marginal of $\pi^{n-1}$.

- $\pi^n_1 = \mu_1$ for $n$ odd (and $\pi^n_2 = \mu_2$ for $n$ even)
- $\pi^n = \arg\min_{\Pi(\mu_1, \ast)} H(\cdot | \pi^{n-1})$ for $n$ odd
- $\frac{d\pi^n}{dR}(x) = \frac{d\pi^n}{d\pi^{n-1}} \cdots \frac{d\pi^1}{d\pi^0} = a(x_1)b(x_2)$
- $\pi^n$ is solution of EOT problem with its own marginals
Sinkhorn Marginals

Suppose $C < \infty$ and let $\pi^* \in \Pi(\mu_1, \mu_2)$ be the optimal coupling

- Key identity: $H(\pi^*|\pi^n) = H(\pi^*|R) - \sum_{t=0}^{n} H(\pi^t|\pi^{t-1})$
- Hence $H(\pi^t|\pi^{t-1}) \to 0$ and thus
- marginals converge in entropy: $H(\pi^n_i|\mu_i) \to 0$, $i = 1, 2$
- Implies $\pi^n_i \to \mu_i$ in total variation
- Léger (2021): sublinear rate $H(\pi^n_i|\mu) \leq H(\pi^*|R)/n$

Summary:

- $\pi^n$ are optimizers of EOT problems with marginals $(\mu_n, \nu_n) \to (\mu, \nu)$
- Sinkhorn convergence is an instance of EOT stability
Weak Sinkhorn Convergence

- Sinkhorn convergence is well understood when $c$ is bounded: linear convergence
- Slightly more general conditions in Rüschendorf (1995)
- We are interested in results for unbounded $c$
- Especially quadratic cost and Gaussian-like marginals

Sinkhorn marginals converge in TV, hence weakly. If $X_i = \mathbb{R}^d$ (or any space with differentiation of measures), weak EOT stability yields:

**Theorem**

Let $c$ be continuous and $\mathcal{C}(\mu_1, \mu_2) < \infty$. Then $\pi^n \to \pi^*$ weakly.
Sinkhorn Convergence in $W_p$

- $F(\pi) := \int c \, d\pi + H(\pi | \mu_1 \otimes \mu_2)$

**Theorem**

Let $p \in [1, \infty)$. For $i = 1, 2$, let $\mu_i \in \mathcal{P}(X_i)$ satisfy

$$\int \exp(\alpha \, d\chi_i(\hat{x}_i, x_i)^p) \, \mu_i(dx_i) < \infty$$

for some $\alpha \in (0, \infty)$ and $\hat{x}_i \in X_i$.

(i) Let $c$ be continuous with growth of order $p$. As $n \to \infty$, we have

$$F(\pi^n) \to F(\pi^*), \quad \pi^n \to \pi^* \quad \text{in} \quad W_p.$$ 

(ii) Let $1 \leq q \leq p$ and $c(x) = c_1(x)c_2(x)$ where $c_1, c_2$ are Lipschitz with
growth of order $p - 1$. For all $n \geq 2$, with a known constant $C$,

$$|F(\pi^*) - F(\pi^n)| \leq Cn^{-\frac{1}{2p}}, \quad W_q(\pi^*, \pi^n) \leq Cn^{-\frac{1}{4pq}}.$$ 

- Covers quadratic cost and subgaussian marginals
- The constant $C$ does not grow exponentially in $c$
Outline

1. Entropic Optimal Transport
2. Weak Stability
3. $W_p$ Stability
4. Sinkhorn’s Algorithm
5. Dual Picture
Dual Problem and Potentials

Let $c \in L^1(\mu \otimes \nu)$. The dual EOT problem (for $\varepsilon = 1$) is

$$
\sup_{f \in L^1(\mu), \ g \in L^1(\nu)} G(f, g),
$$

$$
G(f, g) := \mu(f) + \nu(g) - \int e^{f(x) + g(y) - c(x,y)} \mu(dx)\nu(dy) + 1.
$$

- Unique (up to constant) solution $(f, g) \in L^1(\mu) \times L^1(\nu)$
- $f, g$ are called (Schrödinger, EOT) potentials
- Potentials give the density of the optimal coupling:

$$
\frac{d\pi_*}{dP}(x, y) = e^{f(x) + g(y) - c(x,y)}
$$

- Normalize potentials, e.g., symmetrically: $\mu(f) = \nu(g)$
Stability of Potentials

One incarnation, for absolutely continuous marginals:

**Theorem**

Let $c$ be continuous, $\mu_n \rightarrow \mu$ and $\nu_n \rightarrow \nu$ weakly, and

$$\int c \, d(\mu_n \otimes \nu_n) \rightarrow \int c \, d(\mu \otimes \nu).$$

Suppose $\mu \ll \mu_n$ and $\nu \ll \nu_n$, with densities bounded in probability:

$$\lim_{K \to \infty} \sup_{n \in \mathbb{N}} \mu \left\{ \frac{d\mu}{d\mu_n}(x) \geq K \right\} = 0,$$

$$\lim_{K \to \infty} \sup_{n \in \mathbb{N}} \nu \left\{ \frac{d\nu}{d\nu_n}(y) \geq K \right\} = 0.$$

Then the potentials converge: $f_n \rightarrow f$ in $L^0(\mu)$ and $g_n \rightarrow g$ in $L^0(\nu)$. 
Stability of Potentials for TV Convergence

**Remark:** if \( \sup_n \int e^{(1+\epsilon)c} \, d(\mu_n \otimes \nu_n) < \infty \), the potentials admit versions that are equicontinuous and uniformly bounded. Uniform convergence on compacts follows, without any additional conditions.

Boundedness in probability clearly holds if \( \mu_n \rightarrow \mu \) and \( \nu_n \rightarrow \nu \) in TV.

**Corollary**

Let \( c \) be continuous and \( \int c \, d(\mu_n \otimes \nu_n) \rightarrow \int c \, d(\mu \otimes \nu) \).

If \( \mu_n \rightarrow \mu \) and \( \nu_n \rightarrow \nu \) in total variation, then \( f_n \rightarrow f \) in \( L^0(\mu) \) and \( g_n \rightarrow g \) in \( L^0(\nu) \).

In particular: optimal couplings \( \pi_n \rightarrow \pi_* \) in total variation.
Sinkhorn in Dual Formulation

Algorithm (Sinkhorn in Dual)

Set $g^0 := 0$, $f^0 := 0$. For $n \geq 1,$

\[
  f^n(x) = -\log \int_Y e^{g^{n-1}(y) - c(x,y)} \nu(dy),
\]

\[
  g^n(y) = -\log \int_X e^{f^n(x) - c(x,y)} \mu(dx).
\]

Link to primal formulation:

\[
  d\pi(f, g) := e^{f \oplus g - c} d(\mu \otimes \nu),
\]

\[
  \pi^{2n} := \pi(f^n, g^n), \quad \pi^{2n-1} := \pi(f^n, g^{n-1}), \quad n \geq 1.
\]
Convergence of Sinkhorn in Dual Formulation

- Sinkhorn marginals \((\mu_n, \nu_n) := (\pi^n_1, \pi^n_2)\) are equivalent to \((\mu, \nu)\) and converge in total variation.

**Theorem**

Let \(c \in L^1(\mu \otimes \nu)\) be continuous. If \(c\) is such that the Sinkhorn marginals \((\mu_n, \nu_n) := (\pi^n_1, \pi^n_2)\) satisfy

\[
\int c \, d(\mu_n \otimes \nu_n) \rightarrow \int c \, d(\mu \otimes \nu),
\]

the iterates \((f^n, g^n)\) converge in probability to the potentials \((f, g)\).
Thank you!
Convergence as $\varepsilon \to 0$

- Fix marginals $\mu, \nu$ and continuous cost $c$ on Polish spaces $X, Y$
- For $\varepsilon > 0$, let $\pi_\varepsilon \in \Pi(\mu, \nu)$ be $(c, \varepsilon)$-cyclically invariant

**Proposition**

Let $\pi$ be a cluster point of $(\pi_\varepsilon)$ as $\varepsilon \to 0$. Then $spt \pi$ is $c$-cyclically monotone.

- Hence $\pi$ is an optimal transport, as soon as the OT problem is finite
- Compare: Gamma-convergence, Léonard (2012), Carlier at al. (2017), ...

**Corollary:** If OT problem has unique cyclically monotone coupling $\pi_*$, then $\pi_\varepsilon \to \pi_*$

- Convergence of $\pi_\varepsilon$ also known whenever an optimal transport with finite entropy exists. Also: 1D Monge problem, Di Marino–Louet (2018)
Proof

- Recall:

\[ \prod_{i=1}^{k} \frac{d\pi_{\varepsilon}}{dP}(x_i, y_i) = \exp \left( -\frac{1}{\varepsilon} \left[ \sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{i+1}) \right] \right) \prod_{i=1}^{k} \frac{d\pi_{\varepsilon}}{dP}(x_i, y_{i+1}) \]

- Let \( k \geq 2 \) and \( \delta \geq 0 \). Define set of \( \delta \)-improvable \( k \)-tuples:

\[ A = \left\{ (x_i, y_i)_{i=1}^{k} \in (X \times Y)^k : \sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{i+1}) \geq \delta \right\} \]

Then

\[ \pi_{\varepsilon}^k(A) \leq e^{-\delta/\varepsilon} \quad \text{for all} \quad \varepsilon > 0. \]

- If \( \pi_{\varepsilon} \to \pi \) along a subsequence, it follows that \( \text{spt} \pi \) is \( c \)-cyclically monotone
Convergence Rate as $\varepsilon \to 0$

- For simplicity of exposition: assume
  - OT problem has unique solution $\pi_* \in \Pi(\mu, \nu)$
  - its dual problem has a unique solution $(f_*(x), g_*(y))$
- As seen: $\pi_\varepsilon \to \pi_*$

**Finite-Dimensional Linear Programs:**
- If $\mu, \nu$ are discrete with finite support: Exponential convergence,
  \[ |\pi_\varepsilon - \pi_*|_{TV} \leq \alpha e^{-\beta/\varepsilon} \quad \text{for all} \quad \varepsilon \leq \varepsilon_0 \]

A Continuous Example:

- If \( \mu, \nu \) are centered Gaussians on \( \mathbb{R} \), cost \( c(x, y) = |x - y|^2 \):
- \( \pi_\varepsilon \) is Gaussian, (Monge) transport \( T \) is linear
- Linear convergence transport of cost
  \[
  \int c \, d\pi_\varepsilon - \int c \, d\pi_* = \varepsilon/2 + o(\varepsilon)
  \]
- Leading term: from mass at distance \( \sim \sqrt{\varepsilon} \) to \( \Gamma = \text{spt} \pi_* = \text{graph} \, T \)
- Expansion has been studied for more general marginals, including by Altschuler et al. (2021), Conforti–Tamanini (2019), Pal (2019)
- Local picture: Density
  \[
  \frac{d\pi_\varepsilon}{dP}(x, y) \propto e^{-\alpha|y - T(x)|^2/\varepsilon}
  \]
decays exponentially away from \( \Gamma \)
- Comparable result for general problem? Local exponential rate?
Large Deviations Principle as $\varepsilon \to 0$

**Theorem**

Let $\pi_* = \lim_{\varepsilon \to 0} \pi_{\varepsilon}$ and $I(x, y) = c(x, y) - f_*(x) - g_*(y)$.

(a) For any compact set $C \subset X \times Y$,

$$\limsup_{\varepsilon \to 0} \varepsilon \log \pi_{\varepsilon}(C) \leq - \inf_{(x, y) \in C} I(x, y).$$

(b) For any open set $U \subset X_0 \times Y_0$,

$$\liminf_{\varepsilon \to 0} \varepsilon \log \pi_{\varepsilon}(U) \geq - \inf_{(x, y) \in U} I(x, y).$$

- $X_0 := \text{proj}_X \Gamma$ and $Y_0 := \text{proj}_Y \Gamma$, where $\Gamma = \text{spt} \pi_*$
Capturing the Rate for the Lower Bound

- Look for a good $k$-tuple including the point $(x, y)$!

**Lemma**

Fix $(x, y)$. Suppose there exist $(x_i, y_i)_{2 \leq i \leq k} \subset \text{spt} \pi_*$ such that

$$
\delta_0 := \sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{i+1}) > 0, \quad \text{where} \quad (x_1, y_1) := (x, y).
$$

Given $\delta < \delta_0$, there exist $\alpha, r, \varepsilon_0 > 0$ such that

$$
\pi_\varepsilon(B_r(x, y)) \leq \alpha e^{-\delta/\varepsilon} \quad \text{for} \quad \varepsilon \leq \varepsilon_0.
$$

- Optimizing over $\delta$ will give a good lower bound

**Definition of (rate) Function $I$**:

$$
I(x, y) = \sup_{k \geq 2} \sup_{(x_i, y_i)_{i=2}^k \subset \Gamma} \left( \sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{i+1}) \right),
$$
Dual Problem and Potentials

Let $c \in L^1(\mu \otimes \nu)$. The dual EOT problem is

$$S_\varepsilon := \sup_{f \in L^1(\mu), g \in L^1(\nu)} \left( \int f(x) \mu(dx) + \int g(y) \nu(dy) - \varepsilon \int e^{\frac{f(x)+g(y)-c(x,y)}{\varepsilon}} \mu(dx)\nu(dy) + \varepsilon \right).$$

- Unique (up to constant) solution $(f_\varepsilon, g_\varepsilon) \in L^1(\mu) \times L^1(\nu)$
- $f_\varepsilon, g_\varepsilon$ are called the Schrödinger potentials
- $(-f_\varepsilon, -g_\varepsilon)$ is the optimal (static) portfolio for an exponential utility maximizer with random endowment $c$ and risk aversion $\varepsilon^{-1}$
Dual Convergence

Theorem

Let \( c \in L^1(\mu \otimes \nu) \). Let \( f_\varepsilon, g_\varepsilon \) be the rescaled dual solution of EOT (= the Schrödinger potentials):

\[
\frac{d\pi_\varepsilon}{dP}(x, y) = \exp \left( \frac{1}{\varepsilon} \left[ f_\varepsilon(x) + g_\varepsilon(y) - c(x, y) \right] \right)
\]

normalized with \( \int f_\varepsilon \, d\mu = \int g_\varepsilon \, d\nu \). Then

\( f_\varepsilon \to f_\ast \) in \( L^1(\mu) \), \( g_\varepsilon \to g_\ast \) in \( L^1(\nu) \)

where \( (f_\ast, g_\ast) = \text{Kantorovich potentials with } \int f_\ast \, d\mu = \int g_\ast \, d\nu \).

- More generally (without uniqueness): compactness in \( L^1 \)
- One can also vary \( c_\varepsilon \to c \) with locally uniform convergence
- Compact case has uniform convergence and follows from Arzelà–Ascoli. Previously shown by Gigli–Tamanini
Thank you!