Statistical Bounds for Entropic Optimal Transport: Sample Complexity and the Central Limit Theorem

Gonzalo Mena & Jonathan Niles-Weed

A Spotlight presentation at NeurIPS 2019.

Optimal transport

- A method for comparing distributions featuring a tremendous richness of mathematical concepts.
- Successfully applied in Statistics and Machine Learning.
- This work: theory (why the success?)

Poster #220

	Optimal Transport	Entropic Optimal Transport
Definition	$W_2^2(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 \, d\pi(x,y)$	$S(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 d\pi(x,y) + I(\pi)$

POSTER #220

	Optimal Transport	Entropic Optimal Transport
Definition	$W_2^2(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 \, d\pi(x,y)$	$S(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 d\pi(x,y) + I(\pi)$
Computational Complexity	$W(P_n, Q_n)$ $\tilde{O}(n^3)$	$S(P_n, Q_n)$ $\tilde{O}(n^2)$

Poster #220

	Optimal Transport	Entropic Optimal Transport
Definition	$W_2^2(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 \, d\pi(x,y)$	$S(P,Q)$ $\inf_{\pi \in \Pi(P,Q)} \int \frac{1}{2}\|x-y\|^2 d\pi(x,y) + I(\pi)$
Computational Complexity	$W(P_n, Q_n)$ $\tilde{O}(n^3)$	$S(P_n, Q_n)$ $\tilde{O}(n^2)$
Statistical Complexity	$E\lvert W_2(P,Q) - W_2(P_n,Q_n)\rvert$ $O(n^{-1/d})$ **Curse of dimension!**	$E\lvert S(P,Q) - S(P_n,Q_n)\rvert$ **?** **Our work**

Poster #220

Entropic optimal transport has better sample complexity

Genevay, Chizat, Bach, Cuturi , Peyré, AISTATS, 2018

For P, Q defined on a bounded domain with diameter D

$$E|S(P,Q) - S(P_n, Q_n)| \leq K_{d,D} \frac{e^{D^2}}{\sqrt{n}}$$

POSTER #220

Entropic optimal transport has better sample complexity

Genevay, Chizat, Bach, Cuturi , Peyré, AISTATS, 2018

For P, Q defined on a bounded domain with diameter D

$$E|S(P,Q) - S(P_n, Q_n)| \leq K_{d,D} \frac{e^{D^2}}{\sqrt{n}}$$

Problem #1 exponential blowup in D

POSTER #220

Entropic optimal transport has better sample complexity

Genevay, Chizat, Bach, Cuturi , Peyré, AISTATS, 2018

For P, Q defined on a bounded domain with diameter D

$$E|S(P,Q) - S(P_n, Q_n)| \leq K_{d,D} \frac{e^{D^2}}{\sqrt{n}}$$

Problem #1 exponential blowup in D

Problem #2: boundedness is a restrictive condition

POSTER #220

Entropic optimal transport has better sample complexity

Genevay, Chizat, Bach, Cuturi , Peyré, AISTATS, 2018

For P, Q defined on a bounded domain with diameter D

$$E|S(P,Q) - S(P_n, Q_n)| \leq K_{d,D} \frac{e^{D^2}}{\sqrt{n}}$$

Our result: a substantial improvement

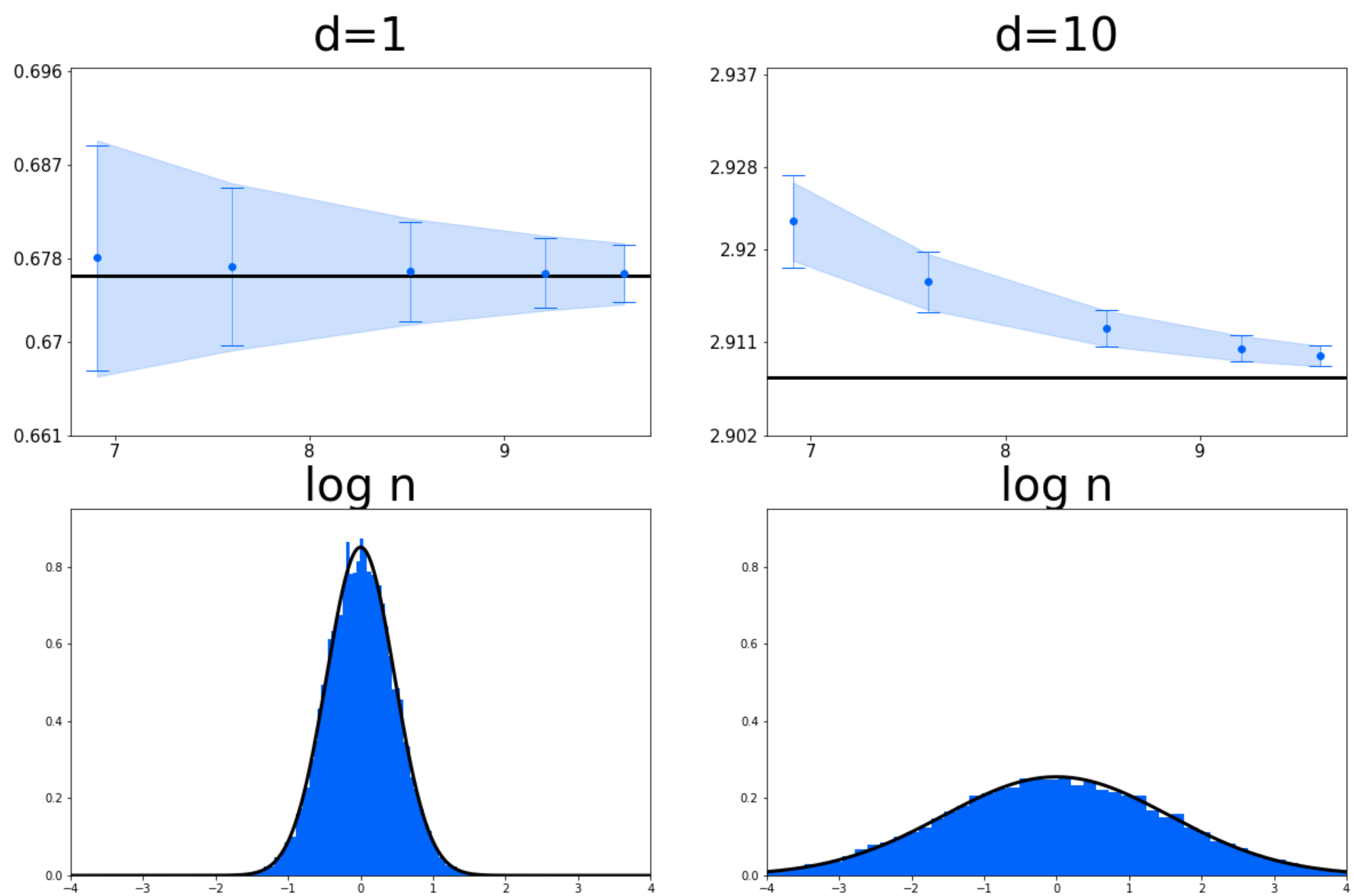
If P, Q are σ^2-subgaussian

$$E|S(P,Q) - S(P_n, Q_n)| \leq K_d \frac{1}{\sqrt{n}} \left(1 + \sigma^{O(d)}\right)$$

POSTER #220

Central Limit Theorem

$$\sqrt{n}\left(S(P_n,Q) - E(S(P_n,Q))\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma^2_{P,Q}\right)$$



- Explicit formula for $\sigma^2_{P,Q}$ in terms of so-called dual potentials
- Results and technique extend from Del Barrio and Loubes, The Annals of Probability, 2019
- A two-sample version is also available.

POSTER #220

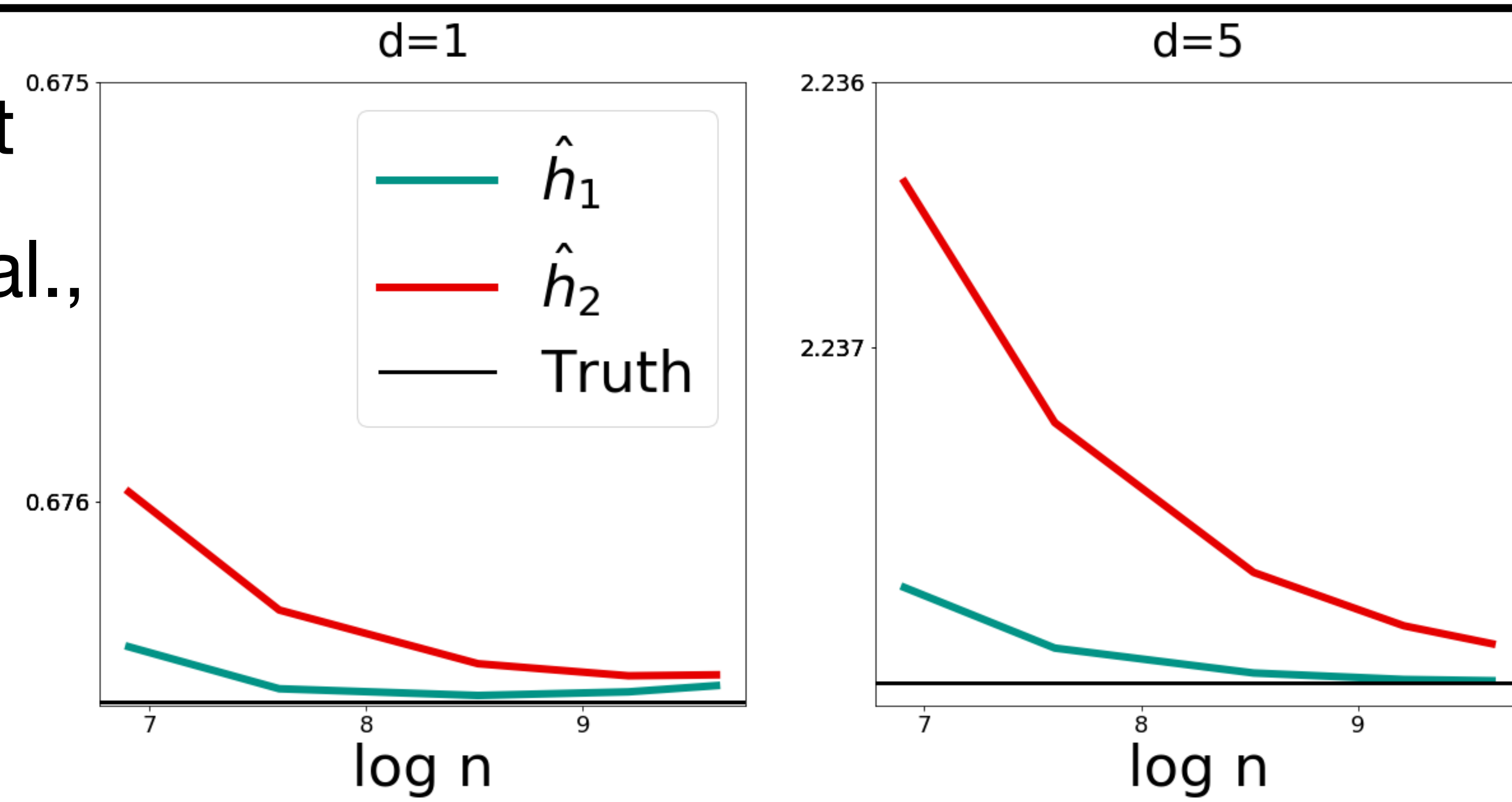
Entropy estimation

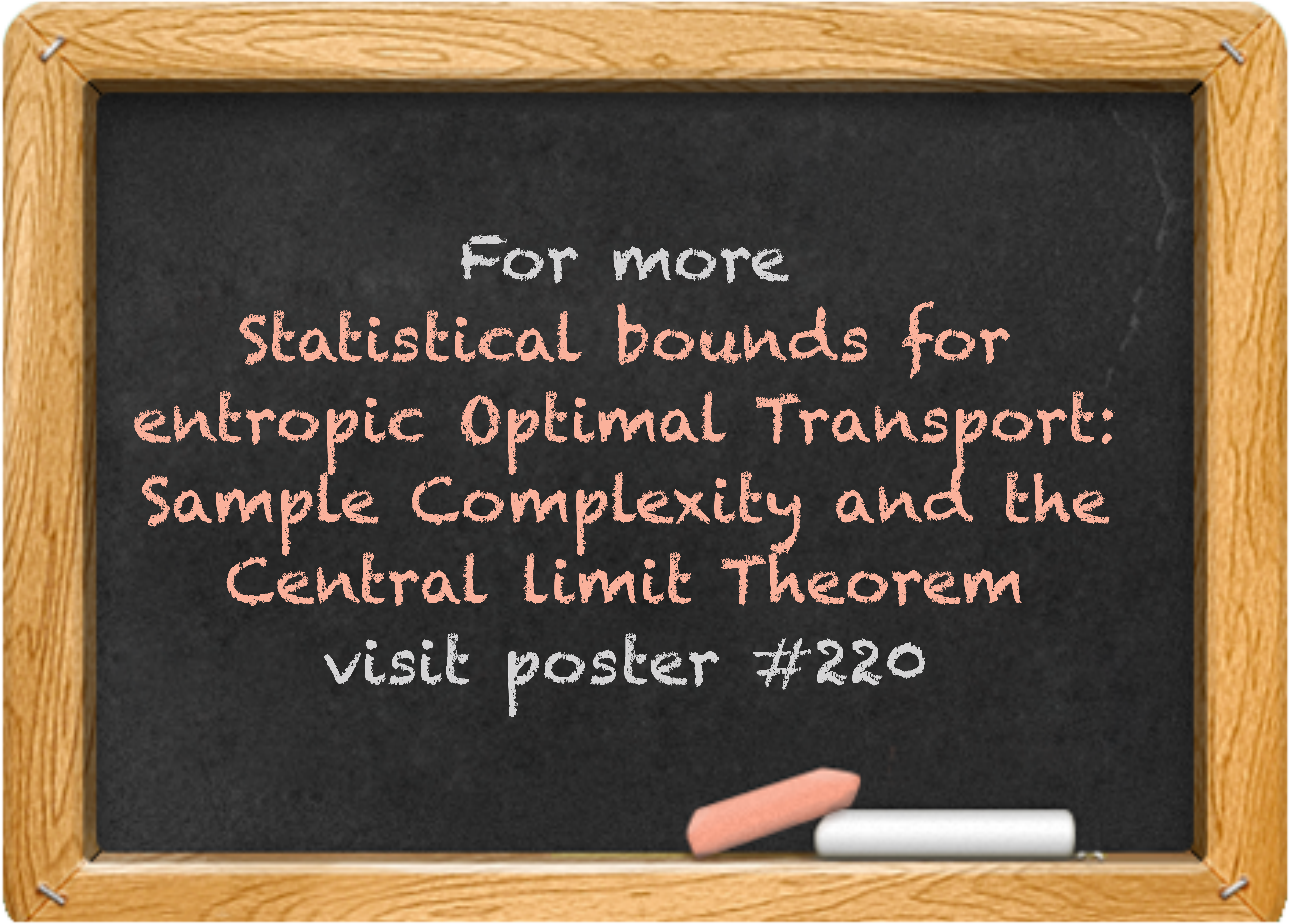
Recent work (Goldfeld et al. 2019, Berret et al., 2019) shows that the differential entropy $h(P * \mathcal{N}(0,I))$ of random variables corrupted by gaussian noise can be estimated at the rate $1/\sqrt{n}$

Our approach

1. Prove that $h\left(P * \mathcal{N}(0,I)\right) = S(P, P * \mathcal{N}(0,I)) + \frac{d}{2}\log\left(2\pi\right)$
2. Estimate $\hat{h}_1(P * \mathcal{N}(0,I)) := S(P_n, (P * \mathcal{N}(0,I)_n) + \frac{d}{2}\log\left(2\pi\right)$

- $\hat{h}_1$ better than $\hat{h}_2$, the best available (from Goldfeld et al., 2019)
- Has distributional limits
- New perspective




For more
Statistical bounds for
entropic Optimal Transport:
Sample Complexity and the
Central limit Theorem
visit poster #220