Meta Optimal Transport

Brandon Amos
Meta AI (FAIR) NYC

Giulia Luise
Microsoft Research

Samuel Cohen
University College London
Meta AI
Fairgen

Ievgen Redko
Noah's Ark Lab, Huawei
Aalto University

http://github.com/bamos/presentations
http://github.com/facebookresearch/meta-ot

ICML
International Conference On Machine Learning

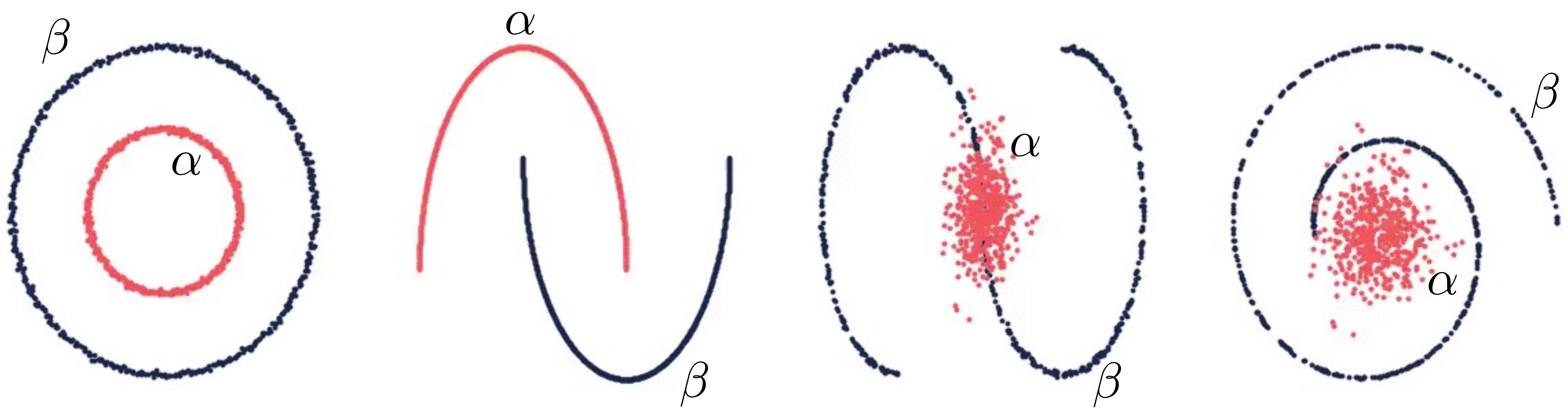
Optimal transport connects spaces

Monge (primal, Wasserstein-2)

$$T^\star(\alpha,\beta) \in \underset{T\in\mathcal{C}(\alpha,\beta)}{\operatorname{argmin}}\ \mathbb{E}_{x\sim\alpha}\|x - T(x)\|_2^2$$

α, β are **measures** $\mathcal{C}(\alpha,\beta)$ is the set of valid **couplings** T is a **transport map** from α to β

we also consider other/discrete OT formulations



Challenge: computing OT maps

Monge (primal, Wasserstein-2)

$$T^\star(\alpha,\beta) \in \underset{T\in\mathcal{C}(\alpha,\beta)}{\operatorname{argmin}}\ \mathbb{E}_{x\sim\alpha}\|x-T(x)\|_2^2$$

we also consider other/discrete OT formulations

Many OT problems are **numerically solved**

Improving OT solvers is active research

Solving multiple OT problems: even harder

Standard solution: independently solve

Meta Optimal Transport

Tutorial on amortized optimization. Brandon Amos. Foundations and Trends in ML, 2023.

Idea: predict the solution to OT problems with amortized optimization

Simultaneously solve many OT problems, sharing info between instances

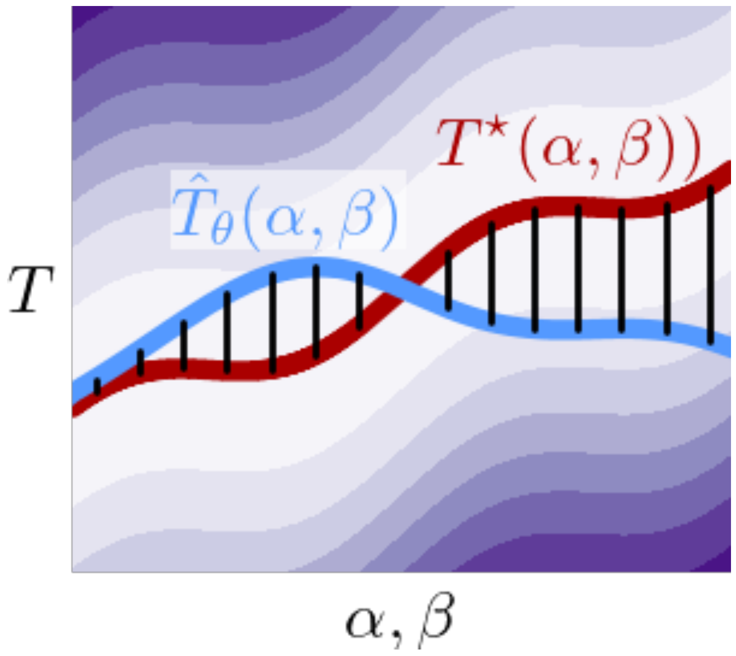
Monge (primal, Wasserstein-2)

$$T^\star(\alpha,\beta) \in \underset{T\in\mathcal{C}(\alpha,\beta)}{\operatorname{argmin}} \mathbb{E}_{x\sim\alpha}\|x - T(x)\|_2^2$$

$$\wr\wr$$

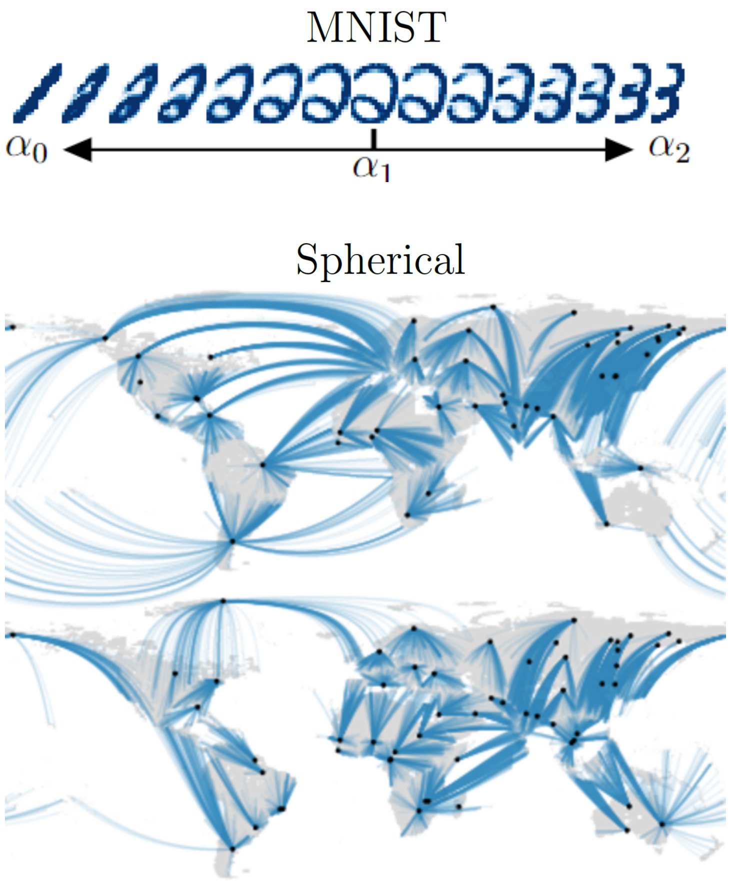
$\hat{T}_\theta(\alpha,\beta)$ (parameterize dual potential via an MLP)

we also consider other/discrete OT formulations



Meta OT for Discrete OT (Sinkhorn)

Sinkhorn Distances: Lightspeed Computation of Optimal Transport. Marco Cuturi. NeurIPS 2013.



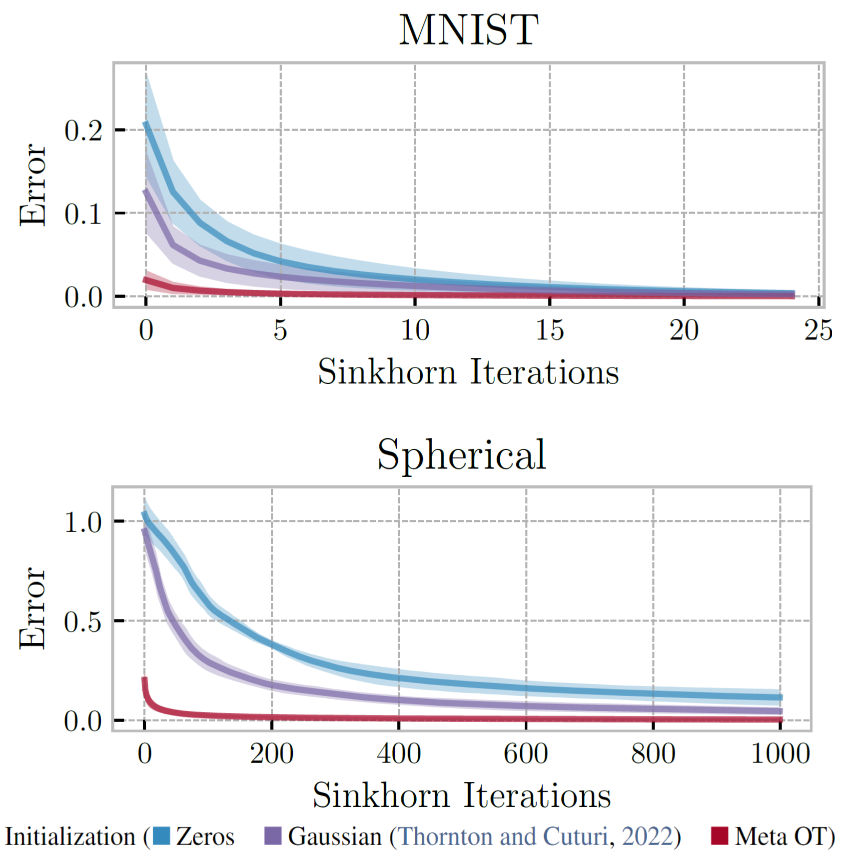


Table 1. Sinkhorn runtime (seconds) to reach a marginal error of 10^{-2}. Meta OT's initial prediction takes $\approx 5 \cdot 10^{-5}$ seconds. We report the mean and std across 10 test instances.

Initialization	MNIST	Spherical
Zeros (t_{zeros})	$4.5 \cdot 10^{-3}$ $\pm 1.5 \cdot 10^{-3}$	0.88 ± 0.13
Gaussian	$4.1 \cdot 10^{-3}$ $\pm 1.2 \cdot 10^{-3}$	0.56 $\pm 9.9 \cdot 10^{-2}$
Meta OT (t_{Meta})	$2.3 \cdot 10^{-3}$ $\pm 9.2 \cdot 10^{-6}$	$7.8 \cdot 10^{-2}$ $\pm 3.4 \cdot 10^{-2}$
Improvement ($t_{\text{zeros}}/t_{\text{Meta}}$)	1.96	11.3

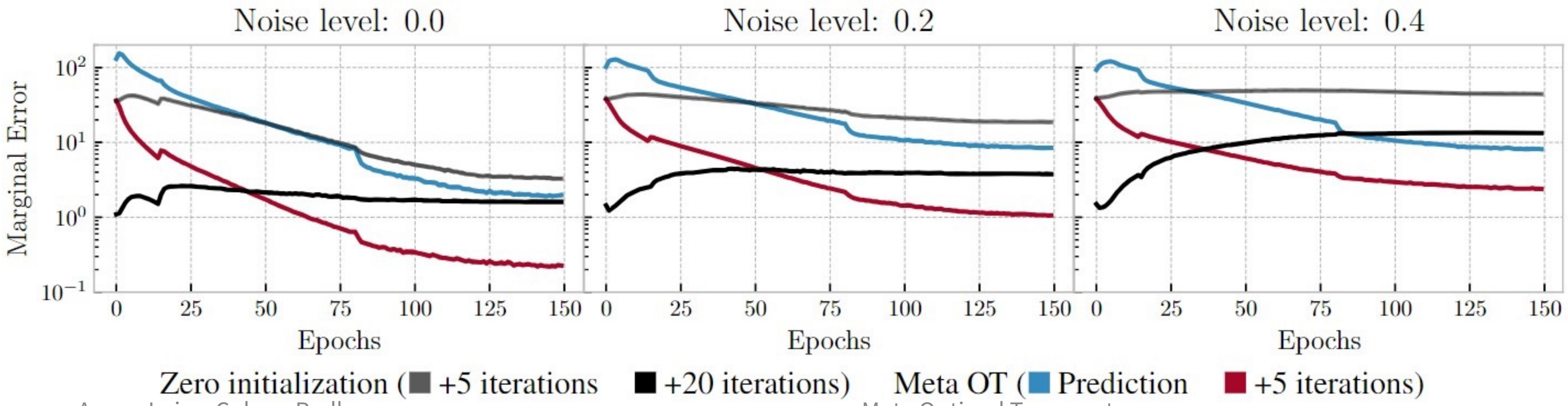
Wasserstein adversarial regularization

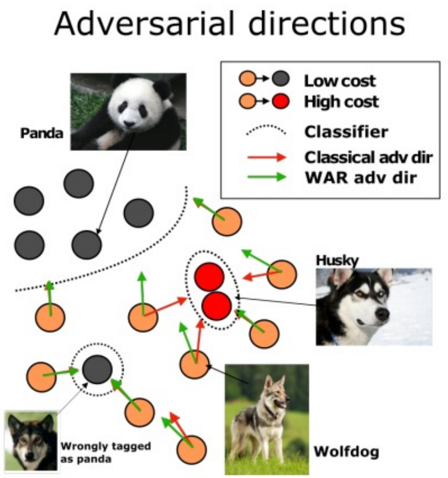
Wasserstein adversarial regularization for learning with label noise. Kilian Fatras et al., TPAMI 2021.

Setting: Discrete OT for classification with label noise

OT is repeatedly solved across minibatches
Use Meta OT to learn better solutions

Fig. 1: AR vs. WAR. Given a number of samples, both methods regularize along adversarial directions (arrows in the left panel), leading to updated decision functions (right panel). While both regularizations prevent the classifier to overfit on the noisy labelled sample, AR also tends over-smooth between similar classes (*wolfdog* and *husky*), while WAR preserves them by changing the adversarial direction.

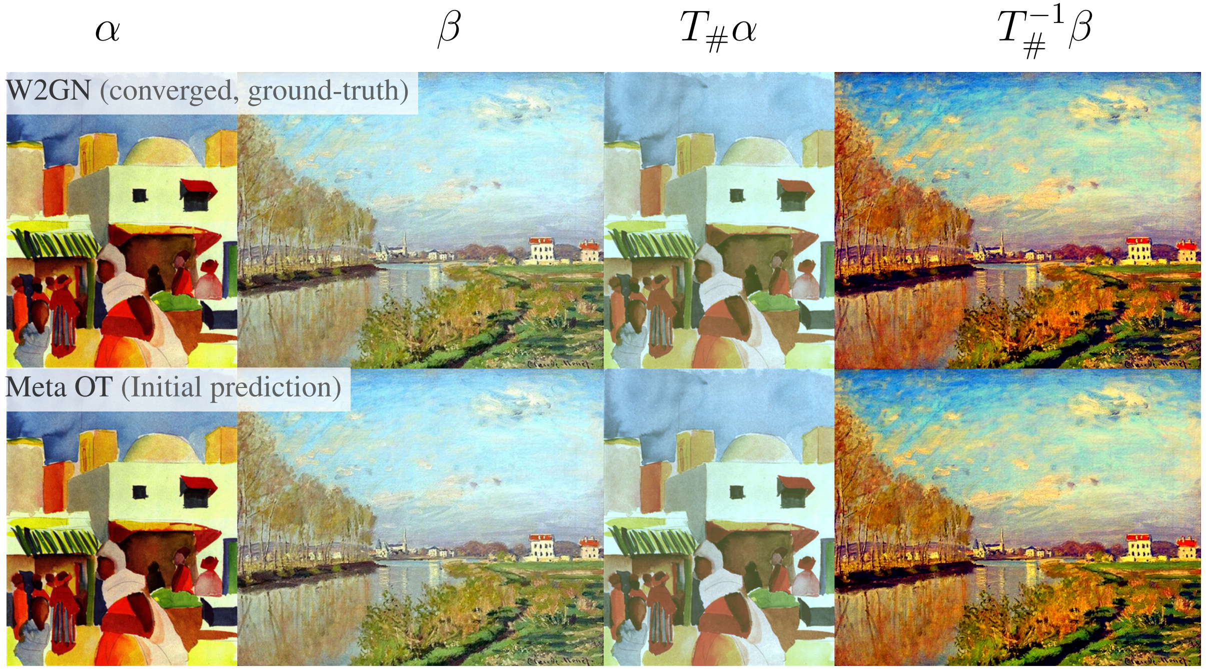


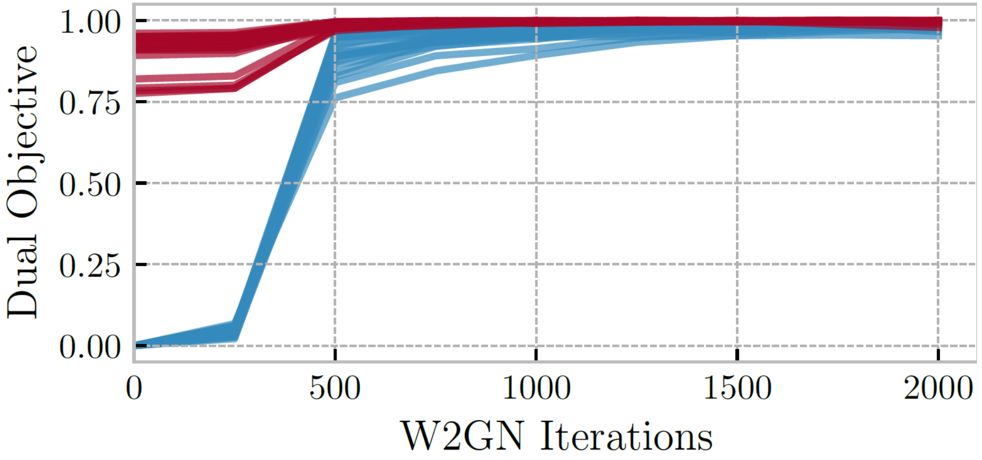


Meta OT in continuous settings (W2GN)

Wasserstein-2 Generative Networks. Alexander Korotin et al., ICLR 2021.

RGB color palette transport



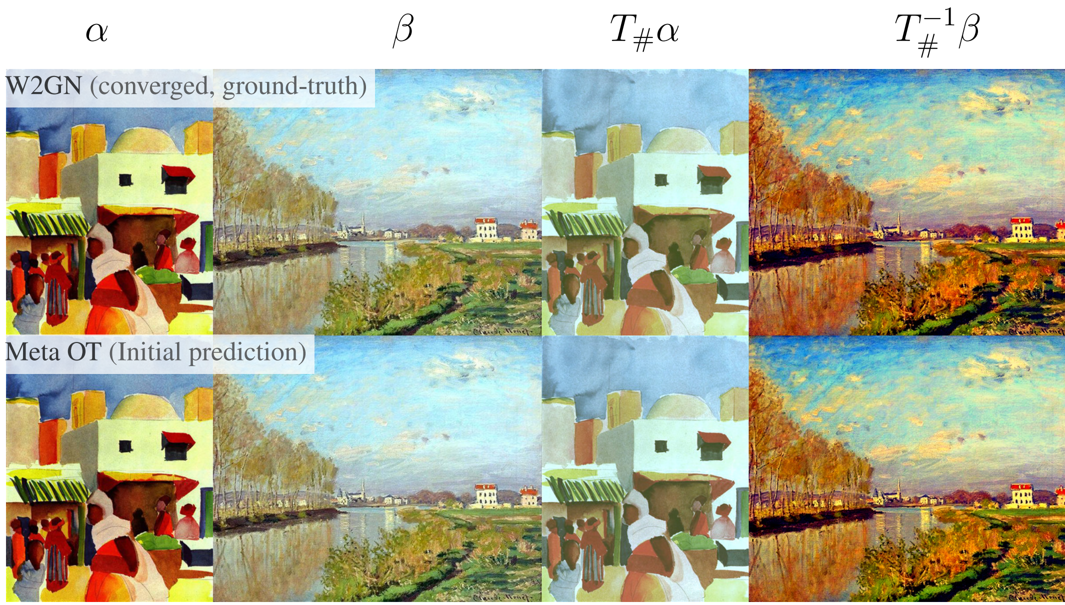


	Iter	Runtime (s)	Dual Value
Meta OT + W2GN	None	$3.5 \cdot 10^{-3} \pm 2.7 \cdot 10^{-4}$	$0.90 \pm 6.08 \cdot 10^{-2}$
	1k	$0.93 \pm 2.27 \cdot 10^{-2}$	$1.0 \pm 2.57 \cdot 10^{-3}$
	2k	$1.84 \pm 3.78 \cdot 10^{-2}$	$1.0 \pm 5.30 \cdot 10^{-3}$
W2GN	1k	$0.90 \pm 1.62 \cdot 10^{-2}$	$0.96 \pm 2.62 \cdot 10^{-2}$
	2k	$1.81 \pm 3.05 \cdot 10^{-2}$	$0.99 \pm 1.14 \cdot 10^{-2}$

More Meta OT color transfer predictions

Conclusions

Deploying OT in real applications will almost always result in **repeated solves**
Use **Meta OT** and amortized optimization to **learn a better solver**



Meta Optimal Transport

Brandon Amos
Meta AI (FAIR) NYC

Giulia Luise
Microsoft Research

Samuel Cohen
University College London
Meta AI
Fairgen

Ievgen Redko
Noah's Ark Lab, Huawei
Aalto University

http://github.com/bamos/presentations
http://github.com/facebookresearch/meta-ot

ICML
International Conference
On Machine Learning