Master 2 Internship - Université de Bordeaux

# Semi-discrete Optimal Transport for Large-Scale problems

| | |
|---|---|
| **When:** | 5 to 6 months Internship starting between January and April 2020 |
| **Where:** | Institut de Mathématiques de Bordeaux, Talence, France |
| **Salary:** | ≈540 €/month |
| **Expected skills:** | Applied mathematics<br>Image processing and analysis, machine learning, Matlab, python. |
| **Application:** | Send by email, **before December 31th**, a CV and a statement of interest to<br>Arthur.Leclaire@math.u-bordeaux.fr and Nicolas.Papadakis@math.u-bordeaux.fr |

**Context** Optimal transportation [12] is a very active research topic with applications in various fields, like economics, machine learning, or image processing. Given two probability measures $\mu, \nu$ on $\mathbb{R}^d$ and a convex cost function $c(x, y)$ on $\mathbb{R}^d$, the Monge's formulation of optimal transport (OT) consists in solving

$$\inf_T \int c(x, T(x)) d\mu(x), \tag{1}$$

where the infimum is taken on all measurable maps $T : \mathbb{R}^d \to \mathbb{R}^d$ whose image measure $T_\sharp\mu$ equals $\nu$. Kantorovich proposed the convex relaxation

$$W(\mu, \nu) = \inf_\pi \int c(x, y) d\pi(x, y) \tag{2}$$

where the infimum is taken on all couplings $\pi$ of $(\mu, \nu)$. Using convex duality [10], this problem is equivalent to solving

$$\max_{\varphi,\psi} \int \varphi d\mu + \int \psi d\nu \tag{3}$$

on all continuous bounded functions $\varphi, \psi$ satisfying $\varphi(x) + \psi(y) \leqslant c(x, y)$ almost everywhere. One can even reduce this problem to a single variable $\psi$ since the corresponding optimal $\varphi$ is the so-called $c$-transform of $\psi$ [10]

$$\psi^c(x) = \min_y c(x, y) - \psi(y).$$

**Objectives** The main goal of this internship is to investigate regularity properties of the optimal $c$-transform, depending on the regularity of the measures $\mu, \nu$.

We will focus primarily on the semi-discrete case of optimal transportation [7] meaning that $\mu$ is absolutely continuous whereas $\nu$ is supported on a finite set $Y$. In this setting, computing the $c$-transform is essentially a nearest neighbor search. Besides, this setting leads to a finite-dimensional concave problem that can be solved with deterministic [7] and stochastic [6, 2] solvers. However, the stochastic gradient descent proposed in [6] does not scale well when the cardinal of $Y$ is large. In contrast, the stochastic approach of [11] based on a parameterization with neural networks should scale better, but remains to be analyzed precisely.

The regularity properties of the optimal dual variables will allow to propose new approximation classes for $\varphi, \psi$, and thus design new scalable numerical OT solvers that rely on a principled parameterization of the dual variables $\varphi, \psi$. The developed numerical solutions could be integrated to several applications that require large-scale optimal transport, for example domain adaptation [3], generative networks [1], texture synthesis [8], or shape analysis [5]. The project thus consists of the following tasks:

- Analyze the regularity of optimal dual variables $\varphi, \psi$.
- Quantize the impact of restricting to a sub-class of dual variables.
- Propose a stochastic solver based on a new parameterization of the dual problem.
- Evaluate the performance of this numerical scheme on synthetic examples.
- Compare with other techniques based on entropic regularization [4] and regularization of Brenier potentials [9].
- Incorporate the proposed numerical solver to address OT problems in machine learning and image processing.

# References

[1] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan. *preprint arXiv:1701.07875*, 2017.

[2] B. Bercu and J. Bigot. Asymptotic distribution and convergence rates of stochastic algorithms for entropic optimal transportation between probability measures. *preprint arXiv:1812.09150*, 2018.

[3] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy. Optimal transport for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1853–1865, 2015.

[4] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Conference on Neural Information Processing Systems (NIPS'13)*, pages 2292–2300, 2013.

[5] J. Feydy, P. Roussillon, A. Trouvé, and P. Gori. Fast and scalable optimal transport for brain tractograms. In *Proceedings of MICCAI*, pages 636–644. Springer, 2019.

[6] A. Genevay, M. Cuturi, G. Peyré, and F. Bach. Stochastic optimization for large-scale optimal transport. In *Advances in neural information processing systems*, pages 3440–3448, 2016.

[7] J. Kitagawa, Q. Mérigot, and B. Thibert. A Newton algorithm for semi-discrete optimal transport. *Journal of the European Math Society*, 2017.

[8] A. Leclaire and J. Rabin. A fast multi-layer approximation to semi-discrete optimal transport. In *Proceedings of SSVM*, pages 341–353. Springer, 2019.

[9] F.-P. Paty, A. d'Aspremont, and M. Cuturi. Regularity as regularization: Smooth and strongly convex brenier potentials in optimal transport. *preprint arXiv:1905.10812*, 2019.

[10] F. Santambrogio. Optimal transport for applied mathematicians. *Progress in Nonlinear Differential Equations and their applications*, 87, 2015.

[11] V. Seguy, B. Bhushan Damodaran, R. Flamary, N. Courty, A. Rolet, and M. Blondel. Large-Scale Optimal Transport and Mapping Estimation. *preprint arXiv:1711.02283*, 2017.

[12] C. Villani. *Topics in Optimal Transportation*. American Math. Society, 2003.