New Monge Problems and Applications

14th-15th of September 2023

	Thursday 14th	**Friday 15th**
9:00-9:40	Welcome of participants	
9:40-10:30	Nathaël Gozlan	Gudmund Pammer
10:30-11:20	Nicolas Juillet	Beatrice Acciaio
11:20-11:40	Coffee Break	Coffee Break
11:40-12:30	Armand Ley Martin Rapaport	Dinh-Toan Nguyen Kexin Shao
12:30-14:00	Lunch (buffet)	Lunch (buffet)
14:00-14:50	Mathias Dus	Matthias Erbar
14:50-15:40	Pierre Ablin	Rémi Flamary
15:40-16:00	Coffee Break	Conclusion
16:00-16:50	Johannes Hertrich	
16:50-17:40	Maxime Laborde	

Pierre Ablin, Apple

Title: Monge, Bregman and Occam: Interpretable Optimal Transport in High-Dimensions with Feature-Sparse Maps.

Abstract: Optimal transport (OT) theory focuses, among all maps that can morph a probability measure onto another, on those that are the “thriftiest”, i.e. such that the averaged cost between a point and its image be as small as possible. Many computational approaches have been proposed to estimate such Monge maps when the cost is the squared L2 distance, e.g., using entropic maps or neural networks. We propose a new model for transport maps, built on a family of translation invariant costs which are the sum of the squared L2 distance and a regularizer. We propose a generalization of the entropic map in this context, and highlight a surprising link tying it with the Bregman centroids of the divergence generated by the cost, and the proximal operator of the regularizer. We show that choosing a sparsity-inducing norm for the regularizer results in maps that apply Occam's razor to transport, in the sense that the displacement vectors they induce are sparse, with a sparsity pattern that varies depending on the input. We showcase the ability of our method to estimate meaningful OT maps for high-dimensional single-cell transcription data, in the 34000-d space of gene counts for cells, without using dimensionality reduction, thus retaining the ability to interpret all displacements at the gene level.

Beatrice Acciaio, ETH Zurich

Title: Non-linear filtering via optimal transport.

Abstract: In this talk I will present an approach to non-linear non-Gaussian filtering based on a variational representation of Bayes' rule obtained through optimal transport theory. Based on ongoing work with T. Schmidt.

Mathias Dus, Ecole des Ponts ParisTech

Title: Numerical solution of Poisson-Neumann equation in the high dimensional regime using two-layer neural networks.

Abstract: The aim of this work is to analyze numerical schemes using two-layer neural networks with infinite width for the resolution of the high-dimensional Poisson partial differential equation (PDE) with Neumann boundary condition. Using Barron's representation of the solution with a probability measure defined on the set of parameter values, the energy is minimized thanks to a gradient curve dynamic on the 2-Wasserstein space of the set of parameter values defining the neural network. Inspired by the work from Bach and Chizat, we prove that if the gradient curve converges, then the represented function is the solution of the elliptic equation considered. In contrast to the works, the activation function we use here is not assumed to be homogeneous to obtain global convergence of the flow. Numerical experiments are given to show the potential of the method.

Matthias Erbar, University of Bielefeld

Title: Covariance-modulated optimal transport and gradient flows.

Abstract: We study a variant of the dynamical optimal transport problem in which the energy to be minimised is modulated by the covariance matrix of the current distribution. Such transport metrics arise naturally in mean field limits of recent particle filtering methods for inverse problems. We show that the transport problem splits into two coupled minimisation problems: one for the evolution of mean and covariance of the interpolating curve and one for its shape. The latter consists in minimising the usual Wasserstein length under the constraint of maintaining fixed mean and covariance along the interpolation. We analyse the geometry induced by this modulated transport distance on the space of probabilities as well as the dynamics of the associated gradient flows. Those show better convergence properties in comparison to the classical Wasserstein metric in terms of exponential convergence rates independent of the Gaussian target. This is joint work with Martin Burger, Franca Hoffmann, Daniel Matthes and André Schlichting.

Rémi Flamary, École Polytechnique

Title: SNEkhorn: Dimension Reduction with Symmetric Entropic Affinities.

Abstract: Many approaches in machine learning rely on a weighted graph to encode the similarities between samples in a dataset. Entropic affinities (EAs), which are notably used in the popular Dimensionality Reduction (DR) algorithm t-SNE, are particular instances of such graphs. To ensure robustness to heterogeneous sampling densities, EAs assign a kernel bandwidth parameter to every sample in such a way that the entropy of each row in the affinity matrix is kept constant at a specific value, whose exponential is known as perplexity. EAs are inherently asymmetric and row-wise stochastic, but they are used in DR approaches after undergoing heuristic symmetrization methods that violate both the row-wise constant entropy and stochasticity properties. In this work, we uncover a novel characterization of EA as an optimal transport problem, allowing a natural symmetrization that can be computed efficiently using dual ascent. The corresponding novel affinity matrix derives advantages from symmetric doubly stochastic normalization in terms of clustering performance, while also effectively controlling the entropy of each row thus making it particularly robust to varying noise levels. Following, we present a new DR algorithm, SNEkhorn, that leverages this new affinity matrix. We show its clear superiority to state-of-the-art approaches with several indicators on both synthetic and real-world datasets. This is a collaborative work with Hugues Van Assel, Titouan Vayer, Nicolas Courty.

Nathaël Gozlan, Université Paris Cité

Title: Weak Optimal Transport with Unnormalized Kernels.

Abstract: This talk will present a new variant of the weak optimal transport problem where mass is distributed from one space to the other through unnormalized kernels. This problem is motivated by economic questions. We give sufficient conditions for primal attainment and prove a dual formula for this transport problem. We also obtain dual attainment conditions

for some specific cost functions. As a byproduct we obtain a transport characterization of the stochastic order defined by convex positively 1-homogenous functions, in the spirit of Strassen theorem for convex domination. Joint work with P. Choné and F. Kramarz.

Johannes Hertrich, TU Berlin

Title: Neural Wasserstein Gradient Flows for Maximum Mean Discrepancies with Riesz Kernels.

Abstract: We analyze Wasserstein gradient flows of maximum mean discrepancy (MMD) functionals with non-smooth Riesz kernels. For lambda-convex functionals, we characterize Wasserstein gradient flows via steepest descent directions, which allows analytic computations in certain cases. For numerical simulations, we propose to approximate the backward scheme of Jordan, Kinderlehrer and Otto for computing such Wasserstein gradient flows as well as a forward scheme for Wasserstein steepest descent flows by neural networks. To reduce the computational effort of these simulations, we prove that the MMD with Riesz kernels coincides with its sliced version. Therefore, all computations can be done in a one-dimensional setting. Here, a simple sorting algorithm can be applied to reduce the complexity for computing (gradients of) the MMD from $O((M+N)^2)$ to $O((M+N)log(M+N))$ for two measures with M and N support points. Finally, we apply this efficient simulation of MMD flows for generative modelling and image generation on MNIST, FashionMNIST and CIFAR10.

Nicolas Juillet, Université de Haute-Alsace

Title: Monge problem on the real line approached from concave cost problems.

Abstract: One of the most basic instance of a Monge-Kantorovich problem is the one given by the distance on the real line for which it is known that the monotone rearrangement provides a solution. However, it is far from being the unique one (think of the books shelve example). In this talk we present the solution by approximation through solutions of the problem for the distance cost to the power $1-h$ when $h>0$ goes to zero —the monotone rearrangement corresponds to the power $1+h$ when h goes to 0+.

Maxime Laborde, Université Paris Cité

Title: Stability of Schrödinger potentials and application to PDEs.

Abstract: The function that maps a family of probability measures to the solution of the dual entropic optimal transport problem is known as the Schrödinger map. We prove that when the cost function is $\mathcal{C}^{k+1}$ with $k \in \mathbb{N}^*$ then this map is Lipschitz continuous from the L^2-Wasserstein space to the space of $\mathcal{C}^k$ functions. Our result holds on compact domains and covers the multi-marginal case. As applications, we prove displacement smoothness of the entropic optimal transport cost and the well-posedness of certain Wasserstein gradient flows involving this functional, including multi-species system, and exponential convergence to the equilibrium. This work is in collaboration with G. Carlier and L. Chizat.

Armand Ley, Université de Haute-Alsace

Title: A structure result on the solutions of the Monge-Kantorovich problem on the real line.

Abstract: In this talk we will consider the transport problem on the real line for the distance cost function $c : (x, y) \in \mathbb{R}^2 \mapsto |x - y|$. We will see that both marginals are decomposable in such a way that each optimal transport plan can be uniquely expressed as a sum of optimal transport plans relatively to this decompositions. We will give some properties of this decomposition and show how it allows to distinguish a particular optimal transport plan, which appears naturally in the context of convergence in entropic optimal transport.

Dinh-Toan Nguyen, Université Gustave Eiffel

Title: Limit Theorems in Wasserstein Distance for Invariant Measure Estimator of Diffusion Processes on Manifolds.

Abstract: Let $(\mathcal{M}, d)$ be a connected compact Riemannian manifold without boundary, let $\mu(\mathrm{d}x) = p(x)\mathrm{d}x$ be a probability measure on $\mathcal{M}$, where $\mathrm{d}x$ is the volume measure and d denotes the geodesic distance. Consider a diffusion process $(X_t)_{t\geq 0}$ generated by an operator $\mathcal{L} := \Delta + \dfrac{\nabla p}{p} \cdot \nabla$.

In this talk, we give limit theorems for the convergence speed in Wasserstein distance of the invariant density estimator $p_{T,h}$

$$p_{T,h}(y) := \frac{1}{T} \int_0^T K_h(X_t, y)\mathrm{d}t,$$

with $K_h(x, y) := \rho_h(x)^{-1} K\left(\frac{d(x,y)}{h}\right)$ and $\rho_h(x) = \int_{\mathcal{M}} K\left(\frac{d(x,y)}{h}\right)\mathrm{d}y$, where $K : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a kernel function.

We also discuss the dependence of the convergence speed on the order of K and the regularity of p and $\mathcal{M}$.

This work is jointly researched with Hélène Guérin, Viet-Chi Tran, and Vincent Divol.

Gudmund Pammer, ETH Zurich

Title: Stretched Brownian Motion: Analysis of a Fixed-Point Scheme.

Abstract: The fitting problem is a classical challenge in mathematical finance about finding martingales that satisfy specific marginal constraints. Building on the Bass solution to the Skorokhod embedding problem and optimal transport, Backhoff, Beiglböck, Huesmann, and Källblad propose a solution for the two-marginal problem: the stretched Brownian motion. Notably rich in structure, this process is an Ito diffusion and a continuous, strong Markov martingale.

Following a similar approach, Conze and Henry-Larbordère recently introduced a novel local volatility model. This model, rooted in an extension of the Bass construction, is efficiently computable through a fixed-point scheme. In our presentation, we reveal the fixed-point

scheme's intricate connection to the stretched Brownian motion and analyse its convergence. This presentation is based on joint work with Beatrice Acciaio and Antonio Marini.

Martin Rapaport, Université Gustave Eiffel

Title: Entropic curvature: local criteria and motivating examples.

Abstract: There are numerous proposals for discrete curvatures in the literature (such as Ollivier curvature, discrete calculus Γ_2 ...). Paul-Marie Samson introduced the notion of discrete entropic curvature based on Schrödinger bridges at zero temperature. We will briefly describe this curvature and give local criteria or geometrical conditions relative to this curvature proposal, we will evoke some basic properties as well as some functional consequences of this curvature. This will be exemplified with tree models, the hypercube and interaction models (Ising models) as a current work in progress. Based on joint work with P.M Samson: Criteria for entropic curvature on graph spaces (arXiv).

Kexin Shao, Ecole des Ponts ParisTech

Title: Non-decreasing martingale couplings

Abstract: For many examples of couples (μ, ν) of probability measures on the real line in the convex order, we observe numerically that the Hobson and Neuberger martingale coupling, which maximizes for $\rho = 1$ the integral of $|y - x|^\rho$ with respect to any martingale coupling between μ and ν, is still a maximizer for $\rho \in (0, 2)$ and a minimizer for $\rho > 2$. We investigate the theoretical validity of this numerical observation and give rather restrictive sufficient conditions for the property to hold. We also exhibit couples (μ, ν) such that it does not hold. The support of the Hobson and Neuberger coupling is known to satisfy some monotonicity property which we call non-decreasing. We check that the non-decreasing property is preserved for maximizers when $\rho \in (0, 1]$. In general, there exist distinct non-decreasing martingale couplings, and we find some decomposition of ν which is in one-to-one correspondence with martingale couplings non-decreasing in a generalized sense.