Activity Report 2018

Project-Team MOKAPLAN

Advances in Numerical Calculus of Variations

IN COLLABORATION WITH: CEREMADE

RESEARCH CENTER
Paris

THEME
Numerical schemes and simulations
# Table of contents

1. Team, Visitors, External Collaborators .................................................... 1
   
2. Overall Objectives .................................................................................. 2
   
   2.1. Introduction ................................................................................. 2
   
   2.2. Static Optimal Transport and Generalizations ................................. 2
       2.2.1. Optimal Transport, Old and New. .................................. 2
       2.2.2. Monge-Ampère Methods. ............................................. 3
       2.2.3. Generalizations of OT. ................................................ 3
       2.2.4. Numerical Applications of Optimal Transportation. .......... 3
   
   2.3. Diffeomorphisms and Dynamical Transport .................................... 3
       2.3.1. Dynamical transport. .................................................... 3
       2.3.2. Gradient Flows for the Wasserstein Distance. .................. 4
       2.3.3. Geodesic on infinite dimensional Riemannian spaces. ......... 4
   
   2.4. Sparsity in Imaging ...................................................................... 5
       2.4.1. Sparse $\ell^1$ regularization. ......................................... 5
       2.4.2. Regularization over measure spaces. ............................... 6
       2.4.3. Low complexity regularization and partial smoothness. .......... 6
   
   2.5. Mokaplan unified point of view .................................................... 6
   
3. Research Program .................................................................................. 7
   
   3.1. Modeling and Analysis .................................................................... 7
       3.1.1. Static Optimal Transport and Generalizations ..................... 7
       3.1.1.1. Convexity constraint and Principal Agent problem in Economics. 7
       3.1.1.2. Optimal transport and conditional constraints in statistics and finance. 7
       3.1.1.3. JKO gradient flows. ............................................... 8
       3.1.1.4. From networks to continuum congestion models. ............. 8
       3.1.2. Diffeomorphisms and Dynamical Transport ......................... 10
       3.1.2.1. Growth Models for Dynamical Optimal Transport .......... 10
       3.1.2.2. Mean-field games. .................................................. 10
       3.1.2.3. Macroscopic Crowd motion, congestion and equilibria .......... 10
       3.1.2.4. Diffeomorphic image matching. .................................. 12
       3.1.2.5. Metric learning and parallel transport for statistical applications. 12
       3.1.3. Sparsity in Imaging ............................................................ 14
       3.1.3.1. Inverse problems over measures spaces. ......................... 14
       3.1.3.2. Sub-Riemannian diffusions. ...................................... 14
       3.1.3.3. Sparse reconstruction from scanner data. ...................... 14
       3.1.3.4. Tumor growth modeling in medical image analysis. ........... 15
   
   3.2. Numerical Tools ............................................................................. 15
       3.2.1. Geometric Discretization Schemes ..................................... 15
       3.2.1.1. Discretizing the cone of convex constraints. .................. 15
       3.2.1.2. Numerical JKO gradient flows. ................................ 16
       3.2.2. Sparse Discretization and Optimization ............................... 16
       3.2.2.1. From discrete to continuous sparse regularization and transport. 16
       3.2.2.2. Polynomial optimization for grid-free regularization. ....... 17
   
   3.2.3. First Order Proximal Schemes .................................................. 17
       3.2.3.1. $L^2$ proximal methods. .......................................... 17
       3.2.3.2. Bregman proximal methods. ..................................... 17
   
4. Highlights of the Year ............................................................................. 18
   
5. New Software and Platforms ................................................................... 18
   
   5.1. ALG2 ....................................................................................... 18
   
   5.2. Mokabajour .................................................................................. 19
6. **New Results** ................................................................. 19
   6.1. Rank optimality for the Burer-Monteiro factorization 19
   6.2. Representer theorems in variational problems 19
   6.3. The Sliding Frank-Wolfe algorithm for Super-resolution Microscopy Imaging 19
   6.4. Approximation of variational problems with a convexity constraint by PDEs of Abreu type 20
   6.5. Variational methods for tomographic reconstruction with few views 20
   6.6. A differential approach to the multi-marginal Schrödinger system 20
   6.7. Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem 20
   6.8. Second order models for optimal transport and cubic splines on the Wasserstein space 21
   6.9. An entropy minimization approach to second-order variational mean-field games 21
   6.10. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm 21
   6.11. Testing Gaussian Process with Applications to Super-Resolution 21
   6.13. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes 22
   6.15. Generalized compressible fluid flows and solutions of the Camassa-Holm variational model 22
   6.16. The Camassa-Holm equation as an incompressible Euler equation: a geometric point of view 22
   6.17. Variational Second-Order Interpolation on the Group of Diffeomorphisms with a Right-Invariant Metric 23
   6.18. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences 23
   7. **Partnerships and Cooperations** ......................................................... 23
   7.1. National Initiatives 23
   7.2. European Initiatives 24
   7.3. International Research Visitors 24
   7.3.1. Visits of International Scientists 24
   7.3.2. Visits to International Teams 24
   8. **Dissemination** ........................................................................... 24
   8.1. Promoting Scientific Activities 24
   8.1.1. Scientific Events Organisation 24
   8.1.2. Scientific Events Selection 24
   8.1.3. Journal 25
   8.1.3.1. Member of the Editorial Boards 25
   8.1.3.2. Reviewer - Reviewing Activities 25
   8.1.4. Invited Talks 25
   8.1.5. Research Administration 26
   8.2. Teaching - Supervision - Juries 26
   8.2.1. Teaching 26
   8.2.2. Supervision 26
   8.2.3. Juries 26
   8.3. Popularization 27
   8.3.1. Internal or external Inria responsibilities 27
   8.3.2. Education 27
   8.3.3. Interventions 27
   9. **Bibliography** ............................................................................ 27
Project-Team MOKAPLAN

Creation of the Team: 2013 January 01, updated into Project-Team: 2015 December 01

Keywords:

**Computer Science and Digital Science:**
- A5.3. - Image processing and analysis
- A5.9. - Signal processing
- A6.1.1. - Continuous Modeling (PDE, ODE)
- A6.2.1. - Numerical analysis of PDE and ODE
- A6.2.6. - Optimization

**Other Research Topics and Application Domains:**
- B1.2. - Neuroscience and cognitive science
- B9.5.2. - Mathematics
- B9.5.3. - Physics
- B9.5.4. - Chemistry
- B9.6.3. - Economy, Finance

1. Team, Visitors, External Collaborators

**Research Scientists**
- Jean-David Benamou [Team leader, Inria, Senior Researcher, HDR]
- Yohann de Castro [Univ Paris-Saclay and Ecole Ponts ParisTech, Researcher]
- Vincent Duval [Inria, Researcher]
- Thomas Gallouët [Inria, Researcher]
- Irene Waldspurger [Univ Paris Dauphine, Researcher]

**Faculty Members**
- François-Xavier Vialard [Univ Paris Dauphine, Associate Professor]
- Guillaume Carlier [Univ Paris Dauphine, Professor, HDR]
- Paul Pegon [Univ Paris Dauphine, Assistant Professor, from Oct 2018]

**Post-Doctoral Fellows**
- Jean Baptiste Courbot [Univ Paris Sciences et Lettres, until Oct 2018]
- Guillaume Garrigos [CNRS, until Sep 2018]
- Guillaume Mijoule [Inria, from Jul 2018]
- Andrea Natale [Inria]

**PhD Students**
- Paul Catala [Ecole Normale Supérieure Paris]
- Quentin Denoyelle [Univ Paris Dauphine]
- Aude Genevey [Univ Paris Dauphine]
- Lucas Martinet [Univ Paris Dauphine]
- Marco Masoero [Univ Paris Dauphine]
- Quentin Ismael Petit [Univ Paris Dauphine, from Sep 2018]
- Giorgi Rukhaia [Inria, from May 2018]
- Gabriele Todeschi [Univ Paris Dauphine, from Oct 2018]
- Miao Yu [Univ Denis Diderot]
- Gwendoline de Bie [Ecole Normale Supérieure Paris]
2. Overall Objectives

2.1. Introduction

The last decade has witnessed a remarkable convergence between several sub-domains of the calculus of variations, namely optimal transport (and its many generalizations), infinite dimensional geometry of diffeomorphisms groups and inverse problems in imaging (in particular sparsity-based regularization). This convergence is due to (i) the mathematical objects manipulated in these problems, namely sparse measures (e.g. coupling in transport, edge location in imaging, displacement fields for diffeomorphisms) and (ii) the use of similar numerical tools from non-smooth optimization and geometric discretization schemes. Optimal Transportation, diffeomorphisms and sparsity-based methods are powerful modeling tools, that impact a rapidly expanding list of scientific applications and call for efficient numerical strategies. Our research program shows the important part played by the team members in the development of these numerical methods and their application to challenging problems.

2.2. Static Optimal Transport and Generalizations

2.2.1. Optimal Transport, Old and New.

Optimal Mass Transportation is a mathematical research topic which started two centuries ago with Monge’s work on the “Théorie des déblais et des remblais” (see [121]). This engineering problem consists in minimizing the transport cost between two given mass densities. In the 40’s, Kantorovich [128] introduced a powerful linear relaxation and introduced its dual formulation. The Monge-Kantorovich problem became a specialized research topic in optimization and Kantorovich obtained the 1975 Nobel prize in economics for his contributions to resource allocations problems. Since the seminal discoveries of Brenier in the 90’s [73], Optimal Transportation has received renewed attention from mathematical analysts and the Fields Medal awarded in 2010 to C. Villani, who gave important contributions to Optimal Transportation and wrote the modern reference monographs [168], [167], arrived at a culminating moment for this theory. Optimal Mass Transportation is today a mature area of mathematical analysis with a constantly growing range of applications. Optimal Transportation has also received a lot of attention from probabilists (see for instance the recent survey [138] for an overview of the Schrödinger problem which is a stochastic variant of the Benamou-Brenier dynamical formulation of optimal transport). The development of numerical methods for Optimal Transportation and Optimal Transportation related problems is a difficult topic and comparatively underdeveloped. This research field has experienced a surge of activity in the last five years, with important contributions of the MOKAPLAN group (see the list of important publications of the team). We describe below a few of recent and less recent Optimal Transportation concepts and methods which are connected to the future activities of MOKAPLAN.
Brenier’s theorem [74] characterizes the unique optimal map as the gradient of a convex potential. As such, Optimal Transportation may be interpreted as an infinite dimensional optimisation problem under “convexity constraint”: i.e. the solution of this infinite dimensional optimisation problem is a convex potential. This connects Optimal Transportation to “convexity constrained” non-linear variational problems such as, for instance, Newton’s problem of the body of minimal resistance. The value function of the optimal transport problem is also known to define a distance between source and target densities called the Wasserstein distance which plays a key role in many applications such as image processing.

### 2.2.2. Monge-Ampère Methods.

A formal substitution of the optimal transport map as the gradient of a convex potential in the mass conservation constraint (a Jacobian equation) gives a non-linear Monge-Ampère equation. Caffarelli [83] used this result to extend the regularity theory for the Monge-Ampère equation. In the last ten years, it also motivated new research on numerical solvers for non-linear degenerate Elliptic equations [109] [136] [58] [59] and the references therein. Geometric approaches based on Laguerre diagrams and discrete data [145] have also been developed. Monge-Ampère based Optimal Transportation solvers have recently given the first linear cost computations of Optimal Transportation (smooth) maps.

### 2.2.3. Generalizations of OT.

In recent years, the classical Optimal Transportation problem has been extended in several directions. First, different ground costs measuring the “physical” displacement have been considered. In particular, well posedness for a large class of convex and concave costs has been established by McCann and Gangbo [120]. Optimal Transportation techniques have been applied for example to a Coulomb ground cost in Quantum chemistry in relation with Density Functional theory [105]. Given the densities of electrons Optimal Transportation models the potential energy and their relative positions. For more than more than 2 electrons (and therefore more than 2 densities) the natural extension of Optimal Transportation is the so called Multi-marginal Optimal Transport (see [149] and the references therein). Another instance of multi-marginal Optimal Transportation arises in the so-called Wasserstein barycenter problem between an arbitrary number of densities [42]. An interesting overview of this emerging new field of optimal transport and its applications can be found in the recent survey of Ghoussoub and Pass [148].

### 2.2.4. Numerical Applications of Optimal Transportation.

Optimal transport has found many applications, starting from its relation with several physical models such as the semi-geostrophic equations in meteorology [125], [107], [106], [51], [135], mesh adaptation [134], the reconstruction of the early mass distribution of the Universe [117], [75] in Astrophysics, and the numerical optimisation of reflectors following the Optimal Transportation interpretation of Oliker [84] and Wang [169]. Extensions of OT such as multi-marginal transport has potential applications in Density Functional Theory , Generalized solution of Euler equations [72] (DFT) and in statistics and finance [48], [119] .... Recently, there has been a spread of interest in applications of OT methods in imaging sciences [66], statistics [63] and machine learning [108]. This is largely due to the emergence of fast numerical schemes to approximate the transportation distance and its generalizations, see for instance [54]. Figure 1 shows an example of application of OT to color transfer. Figure 9 shows an example of application in computer graphics to interpolate between input shapes.

### 2.3. Diffeomorphisms and Dynamical Transport

#### 2.3.1. Dynamical transport.

While the optimal transport problem, in its original formulation, is a static problem (no time evolution is considered), it makes sense in many applications to rather consider time evolution. This is relevant for instance in applications to fluid dynamics or in medical images to perform registration of organs and model tumor growth.
In this perspective, the optimal transport in Euclidean space corresponds to an evolution where each particule of mass evolves in straight line. This interpretation corresponds to the Computational Fluid Dynamic (CFD) formulation proposed by Brenier and Benamou in [50]. These solutions are time curves in the space of densities and geodesics for the Wasserstein distance. The CFD formulation relaxes the non-linear mass conservation constraint into a time dependent continuity equation, the cost function remains convex but is highly non smooth. A remarkable feature of this dynamical formulation is that it can be re-cast as a convex but non smooth optimization problem. This convex dynamical formulation finds many non-trivial extensions and applications, see for instance [52]. The CFD formulation also appears to be a limit case of Mean Fields games (MFGs), a large class of economic models introduced by Lasry and Lions [130] leading to a system coupling an Hamilton-Jacobi with a Fokker-Planck equation. In contrast, the Monge case where the ground cost is the euclidan distance leads to a static system of PDEs [68].

2.3.2. Gradient Flows for the Wasserstein Distance.

Another extension is, instead of considering geodesic for transportation metric (i.e. minimizing the Wasserstein distance to a target measure), to make the density evolve in order to minimize some functional. Computing the steepest descent direction with respect to the Wasserstein distance defines a so-called Wasserstein gradient flow, also known as JKO gradient flows after its authors [126]. This is a popular tool to study a large class of non-linear diffusion equations. Two interesting examples are the Keller-Segel system for chemotaxis [127], [100] and a model of congested crowd motion proposed by Maury, Santambrogio and Roudneff-Chupin [141]. From the numerical point of view, these schemes are understood to be the natural analogue of implicit scheme for linear parabolic equations. The resolution is however costly as it involves taking the derivative in the Wasserstein sense of the relevant energy, which in turn requires the resolution of a large scale convex but non-smooth minimization.

2.3.3. Geodesic on infinite dimensional Riemannian spaces.

To tackle more complicated warping problems, such as those encountered in medical image analysis, one unfortunately has to drop the convexity of the functional involved in defining the gradient flow. This gradient
Figure 2. Examples of displacement interpolation (geodesic for optimal transport) according to a non-Euclidean Riemannian metric (the mass is constrained to move inside a maze) between two input Gaussian distributions. Note that the maze is dynamic: its topology change over time, the mass being “trapped” at time $t = 1/3$.

Flow can either be understood as defining a geodesic on the (infinite dimensional) group of diffeomorphisms [47], or on a (infinite dimensional) space of curves or surfaces [170]. The de-facto standard to define, analyze and compute these geodesics is the “Large Deformation Diffeomorphic Metric Mapping” (LDDMM) framework of Trouvé, Younes, Holm and co-authors [47], [124]. While in the CFD formulation of optimal transport, the metric on infinitesimal deformations is just the $L^2$ norm (measure according to the density being transported), in LDDMM, one needs to use a stronger regularizing metric, such as Sobolev-like norms or reproducing kernel Hilbert spaces (RKHS). This enables a control over the smoothness of the deformation which is crucial for many applications. The price to pay is the need to solve a non-convex optimization problem through geodesic shooting method [142], which requires to integrate backward and forward the geodesic ODE. The resulting strong Riemannian geodesic structure on spaces of diffeomorphisms or shapes is also pivotal to allow us to perform statistical analysis on the tangent space, to define mean shapes and perform dimensionality reduction when analyzing large collection of input shapes (e.g. to study evolution of a disease in time or the variation across patients) [99].

2.4. Sparsity in Imaging

2.4.1. Sparse $\ell^1$ regularization.

Beside image warping and registration in medical image analysis, a key problem in nearly all imaging applications is the reconstruction of high quality data from low resolution observations. This field, commonly referred to as “inverse problems”, is very often concerned with the precise location of features such as point sources (modeled as Dirac masses) or sharp contours of objects (modeled as gradients being Dirac masses along curves). The underlying intuition behind these ideas is the so-called sparsity model (either of the data itself, its gradient, or other more complicated representations such as wavelets, curvelets, bandlets [140] and learned representation [171]).
The huge interest in these ideas started mostly from the introduction of convex methods to serve as proxy for these sparse regularizations. The most well known is the $\ell^1$ norm introduced independently in imaging by Donoho and co-workers under the name “Basis Pursuit” [103] and in statistics by Tibshirani [162] under the name “Lasso”. A more recent resurgence of this interest dates back to 10 years ago with the introduction of the so-called “compressed sensing” acquisition techniques [88], which make use of randomized forward operators and $\ell^1$-type reconstruction.

2.4.2. Regularization over measure spaces.

However, the theoretical analysis of sparse reconstructions involving real-life acquisition operators (such as those found in seismic imaging, neuro-imaging, astro-physical imaging, etc.) is still mostly an open problem. A recent research direction, triggered by a paper of Candès and Fernandez-Granda [87], is to study directly the infinite dimensional problem of reconstruction of sparse measures (i.e. sum of Dirac masses) using the total variation of measures (not to be mistaken for the total variation of 2-D functions). Several works [86], [113], [110] have used this framework to provide theoretical performance guarantees by basically studying how the distance between neighboring spikes impacts noise stability.

\[ \text{Segmentation input} \quad \text{output} \quad \text{Zooming input} \quad \text{output} \]

*Figure 3. Two example of application of the total variation regularization of functions. Left: image segmentation into homogeneous color regions. Right: image zooming (increasing the number of pixels while keeping the edges sharp).*

2.4.3. Low complexity regularization and partial smoothness.

In image processing, one of the most popular methods is the total variation regularization [157], [80]. It favors low-complexity images that are piecewise constant, see Figure 3 for some examples on how to solve some image processing problems. Beside applications in image processing, sparsity-related ideas also had a deep impact in statistics [162] and machine learning [44]. As a typical example, for applications to recommendation systems, it makes sense to consider sparsity of the singular values of matrices, which can be relaxed using the so-called nuclear norm (a.k.a. trace norm) [45]. The underlying methodology is to make use of low-complexity regularization models, which turns out to be equivalent to the use of partly-smooth regularization functionals [133], [164] enforcing the solution to belong to a low-dimensional manifold.

2.5. Mokaplan unified point of view

The dynamical formulation of optimal transport creates a link between optimal transport and geodesics on diffeomorphisms groups. This formal link has at least two strong implications that MOKAPLAN will elaborate on: (i) the development of novel models that bridge the gap between these two fields; (ii) the introduction of novel fast numerical solvers based on ideas from both non-smooth optimization techniques and Bregman metrics, as highlighted in Section 3.2.3.
In a similar line of ideas, we believe a unified approach is needed to tackle both sparse regularization in imaging and various generalized OT problems. Both require to solve related non-smooth and large scale optimization problems. Ideas from proximal optimization has proved crucial to address problems in both fields (see for instance [50], [155]). Transportation metrics are also the correct way to compare and regularize variational problems that arise in image processing (see for instance the Radon inversion method proposed in [54]) and machine learning (see [108]). This unity in term of numerical methods is once again at the core of Section 3.2.3.

3. Research Program

3.1. Modeling and Analysis

The first layer of methodological tools developed by our team is a set of theoretical continuous models that aim at formalizing the problems studied in the applications. These theoretical findings will also pave the way to efficient numerical solvers that are detailed in Section 3.2.

3.1.1. Static Optimal Transport and Generalizations

3.1.1.1. Convexity constraint and Principal Agent problem in Economics.

(Participants: G. Carlier, J-D. Benamou, V. Duval, Xavier Dupuis (LUSS Guido Carli University, Roma))

The principal agent problem plays a distinguished role in the literature on asymmetric information and contract theory (with important contributions from several Nobel prizes such as Mirrlees, Myerson or Spence) and it has many important applications in optimal taxation, insurance, nonlinear pricing. The typical problem consists in finding a cost minimizing strategy for a monopolist facing a population of agents who have an unobservable characteristic, the principal therefore has to take into account the so-called incentive compatibility constraint which is very similar to the cyclical monotonicity condition which characterizes optimal transport plans. In a special case, Rochet and Choné [156] reformulated the problem as a variational problem subject to a convexity constraint. For more general models, and using ideas from Optimal Transportation, Carlier [90] considered the more general $c$-convexity constraint and proved a general existence result. Using the formulation of [90] McCann, Figalli and Kim [114] gave conditions under which the principal agent problem can be written as an infinite dimensional convex variational problem. The important results of [114] are intimately connected to the regularity theory for optimal transport and showed that there is some hope to numerically solve the principal-agent problem for general utility functions.

Our expertise: We have already contributed to the numerical resolution of the Principal Agent problem in the case of the convexity constraint, see [95], [146], [143].

Goals: So far, the mathematical PA model can be numerically solved for simple utility functions. A Bregman approach inspired by [54] is currently being developed [93] for more general functions. It would be extremely useful as a complement to the theoretical analysis. A new semi-Discrete Geometric approach is also investigated where the method reduces to non-convex polynomial optimization.

3.1.1.2. Optimal transport and conditional constraints in statistics and finance.

(Participants: G. Carlier, J-D. Benamou, G. Peyré) A challenging branch of emerging generalizations of Optimal Transportation arising in economics, statistics and finance concerns Optimal Transportation with conditional constraints. The martingale optimal transport [48], [119] which appears naturally in mathematical finance aims at computing robust bounds on option prices as the value of an optimal transport problem where not only the marginals are fixed but the coupling should be the law of a martingale, since it represents the prices of the underlying asset under the risk-neutral probability at the different dates. Note that as soon as more than two dates are involved, we are facing a multimarginal problem.

Our expertise: Our team has a deep expertise on the topic of OT and its generalization, including many already existing collaboration between its members, see for instance [54], [60], [52] for some representative recent collaborative publications.
Goals: This is a non trivial extension of Optimal Transportation theory and MOKAPLAN will develop numerical methods (in the spirit of entropic regularization) to address it. A popular problem in statistics is the so-called quantile regression problem, recently Carlier, Chernozhukov and Galichon [91] used an Optimal Transportation approach to extend quantile regression to several dimensions. In this approach again, not only fixed marginals constraints are present but also constraints on conditional means. As in the martingale Optimal Transportation problem, one has to deal with an extra conditional constraint. The duality approach usually breaks down under such constraints and characterization of optimal couplings is a challenging task both from a theoretical and numerical viewpoint.

3.1.1.3. JKO gradient flows.

(Participants: G. Carlier, J-D. Benamou, M. Laborde, Q. Mérigot, V. Duval) The connection between the static and dynamic transportation problems (see Section 2.3) opens the door to many extensions, most notably by leveraging the use of gradient flows in metric spaces. The flow with respect to the transportation distance has been introduced by Jordan-Kindelherer-Otto (JKO) [126] and provides a variational formulation of many linear and non-linear diffusion equations. The prototypical example is the Fokker Planck equation. We will explore this formalism to study new variational problems over probability spaces, and also to derive innovative numerical solvers. The JKO scheme has been very successfully used to study evolution equations that have the structure of a gradient flow in the Wasserstein space. Indeed many important PDEs have this structure: the Fokker-Planck equation (as was first considered by [126]), the porous medium equations, the granular media equation, just to give a few examples. It also finds application in image processing [79]. Figure 4 shows examples of gradient flows.

Our expertise: There is an ongoing collaboration between the team members on the theoretical and numerical analysis of gradient flows.

Goals: We apply and extend our research on JKO numerical methods to treat various extensions:

- Wasserstein gradient flows with a non displacement convex energy (as in the parabolic-elliptic Keller-Segel chemotaxis model [97])
- systems of evolution equations which can be written as gradient flows of some energy on a product space (possibly mixing the Wasserstein and $L^2$ structures): multi-species models or the parabolic-parabolic Keller-Segel model [64]
- perturbation of gradient flows: multi-species or kinetic models are not gradient flows, but may be viewed as a perturbation of Wasserstein gradient flows, we shall therefore investigate convergence of splitting methods for such equations or systems.

3.1.1.4. From networks to continuum congestion models.

(Participants: G. Carlier, J-D. Benamou, G. Peyré) Congested transport theory in the discrete framework of networks has received a lot of attention since the 50’s starting with the seminal work of Wardrop. A few years later, Beckmann proved that equilibria are characterized as solution of a convex minimization problem. However, this minimization problem involves one flow variable per path on the network, its dimension thus quickly becomes too large in practice. An alternative, is to consider continuous in space models of congested optimal transport as was done in [94] which leads to very degenerate PDEs [70].

Our expertise: MOKAPLAN members have contributed a lot to the analysis of congested transport problems and to optimization problems with respect to a metric which can be attacked numerically by fast marching methods [60].

Goals: The case of general networks/anisotropies is still not well understood, general Γ'-convergence results will be investigated as well as a detailed analysis of the corresponding PDEs and numerical methods to solve them. Benamou and Carlier already studied numerically some of these PDEs by an augmented Lagrangian method see figure 5. Note that these class of problems share important similarities with metric learning problem in machine learning, detailed below.
Figure 4. Example of non-linear diffusion equations solved with a JKO flow [56]. The horizontal axis shows the time evolution minimizing the functional $\int \frac{\rho^\alpha}{\alpha-1}$ on the density $\rho$ (discretized here using point clouds, i.e. sum of Diracs' with equal mass). Each row shows a different value of $\alpha = (0.6, 2, 3)$.

Figure 5. Monge and Wardrop flows of mass around an obstacle [52]. the source/target mass is represented by the level curves. Left : no congestion, Right : congestion.
3.1.2. Diffeomorphisms and Dynamical Transport

3.1.2.1. Growth Models for Dynamical Optimal Transport.

(Participants: F-X. Vialard, J-D. Benamou, G. Peyré, L. Chizat) A major issue with the standard dynamical formulation of OT is that it does not allow for variation of mass during the evolution, which is required when tackling medical imaging applications such as tumor growth modeling [82] or tracking elastic organ movements [160]. Previous attempts [137], [153] to introduce a source term in the evolution typically lead to mass teleportation (propagation of mass with infinite speed), which is not always satisfactory.

Our expertise: Our team has already established key contributions both to connect OT to fluid dynamics [50] and to define geodesic metrics on the space of shapes and diffeomorphisms [102].

Goals: Lenaic Chizat’s PhD thesis aims at bridging the gap between dynamical OT formulation, and LDDDM diffeomorphisms models (see Section 2.3). This will lead to biologically-plausible evolution models that are both more tractable numerically than LDDDM competitors, and benefit from strong theoretical guarantees associated to properties of OT.

3.1.2.2. Mean-field games.

(Participants: G. Carlier, J-D. Benamou) The Optimal Transportation Computational Fluid Dynamics (CFD) formulation is a limit case of variational Mean-Field Games (MFGs), a new branch of game theory recently developed by J-M. Lasry and P-L. Lions [130] with an extremely wide range of potential applications [122]. Non-smooth proximal optimization methods used successfully for the Optimal Transportation can be used in the case of deterministic MFGs with singular data and/or potentials [53]. They provide a robust treatment of the positivity constraint on the density of players.

Our expertise: J.-D. Benamou has pioneered with Brenier the CFD approach to Optimal Transportation. Regarding MFGs, on the numerical side, our team has already worked on the use of augmented Lagrangian methods in MFGs [52] and on the analytical side [89] has explored rigorously the optimality system for a singular CFD problem similar to the MFG system.

Goals: We will work on the extension to stochastic MFGs. It leads to non-trivial numerical difficulties already pointed out in [41].

3.1.2.3. Macroscopic Crowd motion, congestion and equilibria.

(Participants: G. Carlier, J-D. Benamou, Q. Mérigot, F. Santambrogio (U. Paris-Sud), Y. Achdou (Univ. Paris 7), R. Andreev (Univ. Paris 7)) Many models from PDEs and fluid mechanics have been used to give a description of people or vehicles moving in a congested environment. These models have to be classified according to the dimension (1D model are mostly used for cars on traffic networks, while 2-D models are most suitable for pedestrians), to the congestion effects (“soft” congestion standing for the phenomenon where high densities slow down the movement, “hard” congestion for the sudden effects when contacts occur, or a certain threshold is attained), and to the possible rationality of the agents Maury et al [141] recently developed a theory for 2D hard congestion models without rationality, first in a discrete and then in a continuous framework. This model produces a PDE that is difficult to attack with usual PDE methods, but has been successfully studied via Optimal Transportation techniques again related to the JKO gradient flow paradigm. Another possibility to model crowd motion is to use the mean field game approach of Lions and Lasry which limits of Nash equilibria when the number of players is large. This also gives macroscopic models where congestion may appear but this time a global equilibrium strategy is modelled rather than local optimisation by players like in the JKO approach. Numerical methods are starting to be available, see for instance [41], [78].

Our expertise: We have developed numerical methods to tackle both the JKO approach and the MFG approach. The Augmented Lagrangian (proximal) numerical method can actually be applied to both models [52], JKO and deterministic MFGs.

Goals: We want to extend our numerical approach to more realistic congestion model where the speed of agents depends on the density, see Figure 6 for preliminary results. Comparison with different numerical approaches will also be performed inside the ANR ISOTACE. Extension of the Augmented Lagrangian approach to Stochastic MFG will be studied.
Figure 6. Example of crowd congestion with density dependent speed. The macroscopic density, at 4 different times, of people forced to exit from one room towards a meeting point in a second room.
3.1.2.4. Diffeomorphic image matching.

(Participants: F-X. Vialard, G. Peyré, B. Schmitzer, L. Chizat) Diffeomorphic image registration is widely used in medical image analysis. This class of problems can be seen as the computation of a generalized optimal transport, where the optimal path is a geodesic on a group of diffeomorphisms. The major difference between the two approaches being that optimal transport leads to non smooth optimal maps in general, which is however compulsory in diffeomorphic image matching. In contrast, optimal transport enjoys a convex variational formulation whereas in LDDMM the minimization problem is non convex.

Our expertise: F-X. Vialard is an expert of diffeomorphic image matching (LDDMM) [165], [76], [163]. Our team has already studied flows and geodesics over non-Riemannian shape spaces, which allows for piecewise smooth deformations [102].

Goals: Our aim consists in bridging the gap between standard optimal transport and diffeomorphic methods by building new diffeomorphic matching variational formulations that are convex (geometric obstructions might however appear). A related perspective is the development of new registration/transport models in a Lagrangian framework, in the spirit of [159], [160] to obtain more meaningful statistics on longitudinal studies.

Diffeomorphic matching consists in the minimization of a functional that is a sum of a deformation cost and a similarity measure. The choice of the similarity measure is as important as the deformation cost. It is often chosen as a norm on a Hilbert space such as functions, currents or varifolds. From a Bayesian perspective, these similarity measures are related to the noise model on the observed data which is of geometric nature and it is not taken into account when using Hilbert norms. Optimal transport fidelity have been used in the context of signal and image denoising [132], and it is an important question to extends these approach to registration problems. Therefore, we propose to develop similarity measures that are geometric and computationally very efficient using entropic regularization of optimal transport.

Our approach is to use a regularized optimal transport to design new similarity measures on all of those Hilbert spaces. Understanding the precise connections between the evolution of shapes and probability distributions will be investigated to cross-fertilize both fields by developing novel transportation metrics and diffeomorphic shape flows.

The corresponding numerical schemes are however computationally very costly. Leveraging our understanding of the dynamic optimal transport problem and its numerical resolution, we propose to develop new algorithms. These algorithms will use the smoothness of the Riemannian metric to improve both accuracy and speed, using for instance higher order minimization algorithm on (infinite dimensional) manifolds.

3.1.2.5. Metric learning and parallel transport for statistical applications.

(Participants: F-X. Vialard, G. Peyré, B. Schmitzer, L. Chizat) The LDDMM framework has been advocated to enable statistics on the space of shapes or images that benefit from the estimation of the deformation. The statistical results of it strongly depend on the choice of the Riemannian metric. A possible direction consists in learning the right invariant Riemannian metric as done in [166] where a correlation matrix (Figure 7) is learnt which represents the covariance matrix of the deformation fields for a given population of shapes. In the same direction, a question of emerging interest in medical imaging is the analysis of time sequence of shapes (called longitudinal analysis) for early diagnosis of disease, for instance [115]. A key question is the inter subject comparison of the organ evolution which is usually done by transport of the time evolution in a common coordinate system via parallel transport or other more basic methods. Once again, the statistical results (Figure 8) strongly depend on the choice of the metric or more generally on the connection that defines parallel transport.

Our expertise: Our team has already studied statistics on longitudinal evolutions in [115], [116].

Goals: Developing higher order numerical schemes for parallel transport (only low order schemes are available at the moment) and developing variational models to learn the metric or the connections for improving statistical results.
Figure 7. Learning Riemannian metrics in diffeomorphic image matching to capture the brain variability: a diagonal operator that encodes the Riemannian metric is learnt on a template brain out of a collection of brain images. The values of the diagonal operator are shown in greyscale. The red curves represent the boundary between white and grey matter. For more details, we refer the reader to [166], which was a first step towards designing effective and robust metric learning algorithms.

Figure 8. Statistics on initial momenta: In [115], we compared several intersubject transport methodologies to perform statistics on longitudinal evolutions. These longitudinal evolutions are represented by an initial velocity field on the shapes boundaries and these velocity fields are then compared using logistic regression methods that are regularized. The four pictures represent different regularization methods such as $L^2$, $H^1$ and regularization including a sparsity prior such as Lasso, Fused Lasso and TV.
3.1.3. Sparsity in Imaging

3.1.3.1. Inverse problems over measures spaces.

(Participants: G. Peyré, V. Duval, C. Poon, Q. Denoyelle) As detailed in Section 2.4, popular methods for regularizing inverse problems in imaging make use of variational analysis over infinite-dimensional (typically non-reflexive) Banach spaces, such as Radon measures or bounded variation functions.

**Our expertise:** We have recently shown in [164] how – in the finite dimensional case – the non-smoothness of the functionals at stake is crucial to enforce the emergence of geometrical structures (edges in images or fractures in physical materials [65]) for discrete (finite dimensional) problems. We extended this result in a simple infinite dimensional setting, namely sparse regularization of Radon measures for deconvolution [110]. A deep understanding of those continuous inverse problems is crucial to analyze the behavior of their discrete counterparts, and in [111] we have taken advantage of this understanding to develop a fine analysis of the artifacts induced by discrete (i.e. which involve grids) deconvolution models. These works are also closely related to the problem of limit analysis and yield design in mechanical plasticity, see [92], [65] for an existing collaboration between MOKAPLAN’s team members.

**Goals:** A current major front of research in the mathematical analysis of inverse problems is to extend these results for more complicated infinite dimensional signal and image models, such as for instance the set of piecewise regular functions. The key bottleneck is that, contrary to sparse measures (which are finite sums of Dirac masses), here the objects to recover (smooth edge curves) are not parameterized by a finite number of degrees of freedom. The relevant previous work in this direction are the fundamental results of Chambolle, Caselles and co-workers [49], [43], [98]. They however only deal with the specific case where there is no degradation operator and no noise in the observations. We believe that adapting these approaches using our construction of vanishing derivative pre-certificate [110] could lead to a solution to these theoretical questions.

3.1.3.2. Sub-Riemannian diffusions.

(Participants: G. Peyré, J-M. Mirebeau, D. Prandi) Modeling and processing natural images require to take into account their geometry through anisotropic diffusion operators, in order to denoise and enhance directional features such as edges and textures [152], [112]. This requirement is also at the heart of recently proposed models of cortical processing [151]. A mathematical model for these processing is diffusion on sub-Riemannian manifold. These methods assume a fixed, usually linear, mapping from the 2-D image to a lifted function defined on the product of space and orientation (which in turn is equipped with a sub-Riemannian manifold structure).

**Our expertise:** J-M. Mirebeau is an expert in the discretization of highly anisotropic diffusions through the use of locally adaptive computational stencils [144], [112]. G. Peyré has done several contributions on the definition of geometric wavelets transform and directional texture models, see for instance [152]. Dario Prandi has recently applied methods from sub-Riemannian geometry to image restoration [67].

**Goals:** A first aspect of this work is to study non-linear, data-adaptive, lifting from the image to the space/orientation domain. This mapping will be implicitly defined as the solution of a convex variational problem. This will open both theoretical questions (existence of a solution and its geometrical properties, when the image to recover is piecewise regular) and numerical ones (how to provide a faithful discretization and fast second order Newton-like solvers). A second aspect of this task is to study the implication of these models for biological vision, in a collaboration with the UNIC Laboratory (directed by Yves Fregnac), located in Gif-sur-Yvette. In particular, the study of the geometry of singular vectors (or “ground states” using the terminology of [61]) of the non-linear sub-Riemannian diffusion operators is highly relevant from a biological modeling point of view.

3.1.3.3. Sparse reconstruction from scanner data.

(Participants: G. Peyré, V. Duval, C. Poon) Scanner data acquisition is mathematically modeled as a (sub-sampled) Radon transform [123]. It is a difficult inverse problem because the Radon transform is ill-posed and the set of observations is often aggressively sub-sampled and noisy [158]. Typical approaches [129] try to recover piecewise smooth solutions in order to recover precisely the position of the organ being imaged. There is however a very poor understanding of the actual performance of these methods, and little is known on how to enhance the recovery.
**Our expertise:** We have obtained a good understanding of the performance of inverse problem regularization on compact domains for pointwise sources localization [110].

**Goals:** We aim at extending the theoretical performance analysis obtained for sparse measures [110] to the set of piecewise regular 2-D and 3-D functions. Some interesting previous work of C. Poon et al [154] (C. Poon is currently a postdoc in MOKAPLAN) have tackled related questions in the field of variable Fourier sampling for compressed sensing application (which is a toy model for fMRI imaging). These approaches are however not directly applicable to Radon sampling, and require some non-trivial adaptations. We also aim at better exploring the connection of these methods with optimal-transport based fidelity terms such as those introduced in [40].

3.1.3.4. Tumor growth modeling in medical image analysis.

*(Participants: G. Peyré, F-X. Vialard, J-D. Benamou, L. Chizat)* Some applications in medical image analysis require to track shapes whose evolution is governed by a growth process. A typical example is tumor growth, where the evolution depends on some typically unknown but meaningful parameters that need to be estimated. There exist well-established mathematical models [82], [150] of non-linear diffusions that take into account recently biologically observed property of tumors. Some related optimal transport models with mass variations have also recently been proposed [139], which are connected to so-called metamorphoses models in the LDDMM framework [62].

**Our expertise:** Our team has a strong experience on both dynamical optimal transport models and diffeomorphic matching methods (see Section 3.1.2).

**Goals:** The close connection between tumor growth models [82], [150] and gradient flows for (possibly non-Euclidean) Wasserstein metrics (see Section 3.1.2) makes the application of the numerical methods we develop particularly appealing to tackle large scale forward tumor evolution simulation. A significant departure from the classical OT-based convex models is however required. The final problem we wish to solve is the backward (inverse) problem of estimating tumor parameters from noisy and partial observations. This also requires to set-up a meaningful and robust data fidelity term, which can be for instance a generalized optimal transport metric.

3.2. Numerical Tools

The above continuous models require a careful discretization, so that the fundamental properties of the models are transferred to the discrete setting. Our team aims at developing innovative discretization schemes as well as associated fast numerical solvers, that can deal with the geometric complexity of the variational problems studied in the applications. This will ensure that the discrete solution is correct and converges to the solution of the continuous model within a guaranteed precision. We give below examples for which a careful mathematical analysis of the continuous to discrete model is essential, and where dedicated non-smooth optimization solvers are required.

3.2.1. Geometric Discretization Schemes

3.2.1.1. Discretizing the cone of convex constraints.

*(Participants: J-D. Benamou, G. Carlier, J-M. Mirebeau, Q. Mérigot)* Optimal transportation models as well as continuous models in economics can be formulated as infinite dimensional convex variational problems with the constraint that the solution belongs to the cone of convex functions. Discretizing this constraint is however a tricky problem, and usual finite element discretizations fail to converge.
Our expertise: Our team is currently investigating new discretizations, see in particular the recent proposal [59] for the Monge-Ampère equation and [143] for general non-linear variational problems. Both offer convergence guarantees and are amenable to fast numerical resolution techniques such as Newton solvers. Since [59] explaining how to treat efficiently and in full generality Transport Boundary Conditions for Monge-Ampère, this is a promising fast and new approach to compute Optimal Transportation viscosity solutions. A monotone scheme is needed. One is based on Froese Oberman work [118], a new different and more accurate approach has been proposed by Mirebeau, Benamou and Collino [57]. As shown in [104], discretizing the constraint for a continuous function to be convex is not trivial. Our group has largely contributed to solve this problem with G. Carlier [95], Quentin Mérigot [146] and J-M. Mirebeau [143]. This problem is connected to the construction of monotone schemes for the Monge-Ampère equation.

Goals: The current available methods are 2-D. They need to be optimized and parallelized. A non-trivial extension to 3-D is necessary for many applications. The notion of c-convexity appears in optimal transport for generalized displacement costs. How to construct an adapted discretization with “good” numerical properties is however an open problem.

3.2.1.2. Numerical JKO gradient flows.

(Participants: J-D. Benamou, G. Carlier, J-M. Mirebeau, G. Peyré, Q. Mérigot) As detailed in Section 2.3, gradient Flows for the Wasserstein metric (aka JKO gradient flows [126]) provides a variational formulation of many non-linear diffusion equations. They also open the way to novel discretization schemes. From a computational point, although the JKO scheme is constructive (it is based on the implicit Euler scheme), it has not been very much used in practice numerically because the Wasserstein term is difficult to handle (except in dimension one).

Our expertise:

Solving one step of a JKO gradient flow is similar to solving an Optimal transport problem. A geometrical discretization of the Monge-Ampère operator approach has been proposed by Mérigot, Carlier, Oudet and Benamou in [56] see Figure 4. The Gamma convergence of the discretisation (in space) has been proved.

Goals: We are also investigating the application of other numerical approaches to Optimal Transport to JKO gradient flows either based on the CFD formulation or on the entropic regularization of the Monge-Kantorovich problem (see section 3.2.3). An in-depth study and comparison of all these methods will be necessary.

3.2.2. Sparse Discretization and Optimization

3.2.2.1. From discrete to continuous sparse regularization and transport.

(Participants: V. Duval, G. Peyré, G. Carlier, Jalal Fadili (ENSICAEN), Jérôme Malick (CNRS, Univ. Grenoble)) While pervasive in the numerical analysis community, the problem of discretization and Γ-convergence from discrete to continuous is surprisingly over-looked in imaging sciences. To the best of our knowledge, our recent work [110], [111] is the first to give a rigorous answer to the transition from discrete to continuous in the case of the spike deconvolution problem. Similar problems of Γ-convergence are progressively being investigated in the optimal transport community, see in particular [96].

Our expertise: We have provided the first results on the discrete-to-continuous convergence in both sparse regularization variational problems [110], [111] and the static formulation of OT and Wasserstein barycenters [96].

Goals: In a collaboration with Jérôme Malick (Inria Grenoble), our first goal is to generalize the result of [110] to generic partly-smooth convex regularizers routinely used in imaging science and machine learning, a prototypical example being the nuclear norm (see [164] for a review of this class of functionals). Our second goal is to extend the results of [96] to the novel class of entropic discretization schemes we have proposed [54], to lay out the theoretical foundation of these ground-breaking numerical schemes.
3.2.2.2. Polynomial optimization for grid-free regularization.

(Participants: G. Peyré, V. Duval, I. Waldspurger) There has been a recent spark of attention of the imaging community on so-called “grid free” methods, where one tries to directly tackle the infinite dimensional recovery problem over the space of measures, see for instance [87], [110]. The general idea is that if the range of the imaging operator is finite dimensional, the associated dual optimization problem is also finite dimensional (for deconvolution, it corresponds to optimization over the set of trigonometric polynomials).

Our expertise: We have provided in [110] a sharp analysis of the support recovery property of this class of methods for the case of sparse spikes deconvolution.

Goals: A key bottleneck of these approaches is that, while being finite dimensional, the dual problem necessitates to handle a constraint of polynomial positivity, which is notoriously difficult to manipulate (except in the very particular case of 1-D problems, which is the one exposed in [87]). A possible, but very costly, methodology is to ressort to Lasserre’s SDP representation hierarchy [131]. We will make use of these approaches and study how restricting the level of the hierarchy (to obtain fast algorithms) impacts the recovery performances (since this corresponds to only computing approximate solutions). We will pay a particular attention to the recovery of 2-D piecewise constant functions (the so-called total variation of functions regularization [157]), see Figure 3 for some illustrative applications of this method.

3.2.3. First Order Proximal Schemes

3.2.3.1. \(L^2\) proximal methods.

(Participants: G. Peyré, J-D. Benamou, G. Carlier, Jalal Fadili (ENSICAen)) Both sparse regularization problems in imaging (see Section 2.4) and dynamical optimal transport (see Section 2.3) are instances of large scale, highly structured, non-smooth convex optimization problems. First order proximal splitting optimization algorithms have recently gained lots of interest for these applications because they are the only ones capable of scaling to giga-pixel discretizations of images and volumes and at the same time handling non-smooth objective functions. They have been successfully applied to optimal transport [50], [147], congested optimal transport [81] and to sparse regularizations (see for instance [155] and the references therein).

Our expertise: The pioneering work of our team has shown how these proximal solvers can be used to tackle the dynamical optimal transport problem [50], see also [147]. We have also recently developed new proximal schemes that can cope with non-smooth composite objectives functions [155].

Goals: We aim at extending these solvers to a wider class of variational problems, most notably optimization under divergence constraints [52]. Another subject we are investigating is the extension of these solvers to both non-smooth and non-convex objective functionals, which are mandatory to handle more general transportation problems and novel imaging regularization penalties.

3.2.3.2. Bregman proximal methods.

(Participants: G. Peyré G. Carlier, L. Nenna, J-D. Benamou, L. Nenna, Marco Cuturi (Kyoto Univ.)) The entropic regularization of the Kantorovich linear program for OT has been shown to be surprisingly simple and efficient, in particular for applications in machine learning [108]. As shown in [54], this is a special instance of the general method of Bregman iterations, which is also a particular instance of first order proximal schemes according to the Kullback-Leibler divergence.

Our expertise: We have recently [54] shown how Bregman projections [71] and Dykstra algorithm [46] offer a generic optimization framework to solve a variety of generalized OT problems. Carlier and Dupuis [93] have designed a new method based on alternate Dykstra projections and applied it to the principal-agent problem in microeconomics. We have applied this method in computer graphics in a paper accepted in SIGGRAPH 2015 [161]. Figure 9 shows the potential of our approach to handle giga-voxel datasets: the input volumetric densities are discretized on a \(100^3\) computational grid.
Figure 9. Example of barycenter between shapes computed using optimal transport barycenters of the uniform densities inside the 3 extremal shapes, computed as detailed in [161]. Note that the barycenters are not in general uniform distributions, and we display them as the surface defined by a suitable level-set of the density.

**Goals:** Following some recent works (see in particular [101]) we first aim at studying primal-dual optimization schemes according to Bregman divergences (that would go much beyond gradient descent and iterative projections), in order to offer a versatile and very effective framework to solve variational problems involving OT terms. We then also aim at extending the scope of usage of this method to applications in quantum mechanics (Density Functional Theory, see [105]) and fluid dynamics (Brenier’s weak solutions of the incompressible Euler equation, see [72]). The computational challenge is that realistic physical examples are of a huge size not only because of the space discretization of one marginal but also because of the large number of marginals involved (for incompressible Euler the number of marginals equals the number of time steps).

**4. Highlights of the Year**

**4.1. Highlights of the Year**

**4.1.1. Awards**

G. Carlier was a John von Neumann invited Professor at TUM (Munich) in 2018.

**5. New Software and Platforms**

**5.1. ALG2**

**FUNCTIONAL DESCRIPTION:** ALG2 for Monge Mean-Field Games, Monge problem and Variational problems under divergence constraint. A generalisation of the ALG2 algorithm has been implemented in FreeFem++.

- Contact: Jean-David Benamou
- URL: https://team.inria.fr/mokaplan/augmented-lagrangian-simulations/
5.2. Mokabajour

FUNCTIONAL DESCRIPTION: We design a software resolving the following inverse problem: define the shape of a mirror which reflects the light from a source to a defined target, distribution and support of densities being prescribed. Classical applications include the conception of solar oven, public lightning, car headlights... Mathematical modeling of this problem, related to the optimal transport theory, takes the form of a nonlinear Monge-Ampere type PDE. The numerical resolution of these models remained until recently a largely open problem. MOKABAJOUR project aims to develop, using algorithms invented especially at Inria and LJK, a reflector design software more efficient than geometrical methods used so far. The final step is to realize and physically test prototype reflectors.

- Participants: Boris Thibert, Jean-David Benamou and Quentin Mérigot
- Contact: Jean-David Benamou
- URL: https://project.inria.fr/mokabajour/

6. New Results

6.1. Rank optimality for the Burer-Monteiro factorization

I. Waldspurger, A. Watersw

In [39], Numerically solving a large scale semidefinite program, in full generality, is a challenge: The complexity of generic algorithms blows up quickly with the size of the unknown matrix. Fortunately, in many situations, the solution of the program has low rank, and this can be exploited to achieve algorithmic speedups. The most classical way to do this is the Burer-Monteiro factorization, introduced in [77]. It consists in writing the unknown matrix as the product of low-rank factors, and optimizing the factors instead of the matrix itself. The first theoretical guarantees for this method appeared in [69], where it was shown that this strategy almost always succeeds when the size of the factors is of the order of the square root of the full matrix. In our article, we show that, up to a marginal improvement, this result is optimal: Contrarily to what numerical experiments might suggest, there exist situations where the method fails if the size of the factors is chosen smaller.

6.2. Representer theorems in variational problems

C. Boyer, A. Chambolle, Y. De Castro, V. Duval, F. De Gournay, P. Weiss

In [29], we have established a general principle which states that regularizing an inverse problem with a convex function yields solutions which are convex combinations of a small number of atoms. These atoms are identified with the extreme points and elements of the extreme rays of the regularizer level sets. An extension to a broader class of quasi-convex regularizers is also discussed. As a side result, we characterize the minimizers of the total gradient variation, describing the solutions of total variation problem as a superposition of indicator functions of simply connected sets. That result provides an explanation of the so-called staircasing phenomenon.

6.3. The Sliding Frank-Wolfe algorithm for Super-resolution Microscopy Imaging

Q. Denoyelle, V. Duval, G. Peyré, E. Soubies
In [32], we have studied the theoretical and numerical performance of the Sliding Frank-Wolfe, a novel optimization algorithm to solve the BLASSO sparse spikes super-resolution problem. The BLASSO is a continuous (i.e. off-the-grid or grid-less) counterpart to the well-known $\ell^1$ sparse regularisation method (also known as LASSO or Basis Pursuit). Our algorithm is a variation on the classical Frank-Wolfe (also known as conditional gradient) which follows a recent trend of interleaving convex optimization updates (corresponding to adding new spikes) with non-convex optimization steps (corresponding to moving the spikes). Our main theoretical result is that this algorithm terminates in a finite number of steps under a mild non-degeneracy hypothesis. We then target applications of this method to several instances of single molecule fluorescence imaging modalities, among which certain approaches rely heavily on the inversion of a Laplace transform. Our second theoretical contribution is the proof of the exact support recovery property of the BLASSO to invert the 1-D Laplace transform in the case of positive spikes. On the numerical side, we conclude this paper with an extensive study of the practical performance of the Sliding Frank-Wolfe on different instantiations of single molecule fluorescence imaging, including convolutive and non-convolutive (Laplace-like) operators. This shows the versatility and superiority of this method with respect to alternative sparse recovery techniques.

6.4. Approximation of variational problems with a convexity constraint by PDEs of Abreu type

G. Carlier, T. Radice
In [31], motivated by some variational problems subject to a convexity constraint, we consider an approximation using the logarithm of the Hessian determinant as a barrier for the constraint. We show that the minimizer of this penalization can be approached by solving a second boundary value problem for Abreu’s equation which is a well-posed nonlinear fourth-order elliptic problem. More interestingly, a similar approximation result holds for the initial constrained variational problem.

6.5. Variational methods for tomographic reconstruction with few views

M. Bergounioux, I. Abraham, R. Abraham, G. Carlier, E. Le Pennec, E. Trélat
In [16], we deal with a severe ill posed problem, namely the reconstruction process of an image during tomography acquisition with (very) few views. We present different methods that we investigated during the past decade. They are based on variational analysis. This is a survey paper and we refer to the quoted papers for more details.

6.6. A differential approach to the multi-marginal Schrödinger system

G. Carlier, M. Laborde
In [30], we develop an elementary and self-contained differential approach, in an $L^\infty$ setting, for well-posedness (existence, uniqueness and smooth dependence with respect to the data) for the multi-marginal Schrödinger system which arises in the entropic regularization of optimal transport problems.

6.7. Minimal convex extensions and finite difference discretization of the quadratic Monge-Kantorovich problem

J-D. Benamou, V. Duval
In [15] we present an adaptation of the MA-LBR scheme to the Monge-Ampré equation with second boundary value condition, provided the target is a convex set. This yields a fast adaptive method to numerically solve the Optimal Transport problem between two absolutely continuous measures, the second of which has convex support. The proposed numerical method actually captures a specific Brenier solution which is minimal in some sense. We prove the convergence of the method as the grid stepsize vanishes and we show with numerical experiments that it is able to reproduce subtle properties of the Optimal Transport problem.
6.8. Second order models for optimal transport and cubic splines on the Wasserstein space

J.-D. Benamou, T. O. Gallouët, F.-X. Vialard

On the space of probability densities, we extend in [28] the Wasserstein geodesics to the case of higher-order interpolation such as cubic spline interpolation. After presenting the natural extension of cubic splines to the Wasserstein space, we propose a simpler approach based on the relaxation of the variational problem on the path space. We explore two different numerical approaches, one based on multi-marginal optimal transport and entropic regularization and the other based on semi-discrete optimal transport.

6.9. An entropy minimization approach to second-order variational mean-field games

J.-D. Benamou, G. Carlier, S. Di Marino, L. Nenna

In [26] we propose a new viewpoint on variational mean-field games with diffusion and quadratic Hamiltonian. We show the equivalence of such mean-field games with a relative entropy minimization at the level of probabilities on curves. We also address the time-discretization of such problems, establish Gamma-Convergence results as the time step vanishes and propose an efficient algorithm relying on this entropic interpretation as well as on the Sinkhorn scaling algorithm.

6.10. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm

J.-D. Benamou, G. Carlier, L. Nenna

Starting from Brenier’s relaxed formulation of the incompressible Euler equation in terms of geodesics in the group of measure-preserving diffeomorphisms, we propose in [27] a numerical method based on Sinkhorn’s algorithm for the entropic regularization of optimal transport. We also make a detailed comparison of this entropic regularization with the so-called Bredinger entropic interpolation problem (see [1]). Numerical results in dimension one and two illustrate the feasibility of the method.

6.11. Testing Gaussian Process with Applications to Super-Resolution

J.-M. Azaïs, Y. De Castro, S. Mourareau

In [13], we introduce exact testing procedures on the mean of a Gaussian process \( X \) derived from the outcomes of \( \ell_1 \)-minimization over the space of complex valued measures. The process \( X \) can be thought as the sum of two terms: first, the convolution between some kernel and a target atomic measure (mean of the process); second, a random perturbation by an additive centered Gaussian process. The first testing procedure considered is based on a dense sequence of grids on the index set of \( X \) and we establish that it converges (as the grid step tends to zero) to a randomized testing procedure: the decision of the test depends on the observation \( X \) and also on an independent random variable. The second testing procedure is based on the maxima and the Hessian of \( X \) in a grid-less manner. We show that both testing procedures can be performed when the variance is unknown (and the correlation function of \( X \) is known). These testing procedures can be used for the problem of deconvolution over the space of complex valued measures, and applications in frame of the Super-Resolution theory are presented. As a byproduct, numerical investigations may demonstrate that our grid-less method is more powerful (it detects sparse alternatives) than tests based on very thin grids.


Y. De Castro, F. Gamboa, D. Henrion, R. Hess, J.-B Lasserre
In [19], we introduce a new approach aiming at computing approximate optimal designs for multivariate polynomial regressions on compact (semi-algebraic) design spaces. We use the moment-sum-of-squares hierarchy of semidefinite programming problems to solve numerically the approximate optimal design problem. The geometry of the design is recovered via semidefinite programming duality theory. This article shows that the hierarchy converges to the approximate optimal design as the order of the hierarchy increases. Furthermore, we provide a dual certificate ensuring finite convergence of the hierarchy and showing that the approximate optimal design can be computed numerically with our method. As a byproduct, we revisit the equivalence theorem of the experimental design theory: it is linked to the Christoffel polynomial and it characterizes finite convergence of the moment-sum-of-square hierarchies.

6.13. Simulation of multiphase porous media flows with minimizing movement and finite volume schemes

C. Cancès, T. O. Gallouët, M. Laborde, L. Monsaingeon

In [17]: the Wasserstein gradient flow structure of the PDE system governing multiphase flows in porous media was recently highlighted in [85]. The model can thus be approximated by means of the minimizing movement (or JKO) scheme. We solve the JKO scheme using the ALG2-JKO scheme proposed in [55]. The numerical results are compared to a classical upstream mobility Finite Volume scheme, for which strong stability properties can be established.


T. O. Gallouët, M. Laborde, L. Monsaingeon

In [21] the authors show that unbalanced optimal transport provides a convenient framework to handle reaction and diffusion processes in a unified metric framework. We use a constructive method, alternating minimizing movements for the Wasserstein distance and for the Fisher-Rao distance, and prove existence of weak solutions for general scalar reaction-diffusion-advection equations. We extend the approach to systems of multiple interacting species, and also consider an application to a very degenerate diffusion problem involving a Gamma-limit. Moreover, some numerical simulations are included.

6.15. Generalized compressible fluid flows and solutions of the Camassa-Holm variational model

T. O. Gallouët, A. Natale, F-X. Vialard

In [35]: The Camassa-Holm equation on a domain $M \subseteq \mathbb{R}^d$, in one of its possible multi-dimensional generalizations, describes geodesics on the group of diffeomorphisms with respect to the $H(\text{div})$ metric. It has been recently reformulated as a geodesic equation for the $L^2$ metric on a subgroup of the diffeomorphism group of the cone over $M$. We use such an interpretation to construct an analogue of Brenier’s generalized incompressible Euler flows for the Camassa-Holm equation. This involves describing the fluid motion using probability measures on the space of paths on the cone, so that particles are allowed to split and cross. Differently from Brenier’s model, however, we are also able to account for compressibility by employing an explicit probabilistic representation of the Jacobian of the flow map. We formulate the boundary value problem associated to the Camassa-Holm equation using such generalized flows. We prove existence of solutions and that, for short times, smooth solutions of the Camassa-Holm equations are the unique solutions of our model. We propose a numerical scheme to construct generalized solutions on the cone and present some numerical results illustrating the relation between the generalized Camassa-Holm and incompressible Euler solutions.

6.16. The Camassa-Holm equation as an incompressible Euler equation: a geometric point of view

T. O. Gallouët, F-X. Vialard
In [23]: The group of diffeomorphisms of a compact manifold endowed with the $L^2$ metric acting on the space of probability densities gives a unifying framework for the incompressible Euler equation and the theory of optimal mass transport. Recently, several authors have extended optimal transport to the space of positive Radon measures where the Wasserstein-Fisher-Rao distance is a natural extension of the classical $L^2$-Wasserstein distance. In this paper, we show a similar relation between this unbalanced optimal transport problem and the Hdiv right-invariant metric on the group of diffeomorphisms, which corresponds to the Camassa-Holm (CH) equation in one dimension. On the optimal transport side, we prove a polar factorization theorem on the automorphism group of half-densities. Geometrically, our point of view provides an isometric embedding of the group of diffeomorphisms endowed with this right-invariant metric in the automorphisms group of the fiber bundle of half densities endowed with an $L^2$ type of cone metric. This leads to a new formulation of the (generalized) CH equation as a geodesic equation on an isotropy subgroup of this automorphisms group: On $S_1$, solutions to the standard CH thus give particular solutions of the incompressible Euler equation on a group of homeomorphisms of $\mathbb{R}^2$ which preserve a radial density that has a singularity at 0. An other application consists in proving that smooth solutions of the Euler-Arnold equation for the Hdiv right-invariant metric are length minimizing geodesics for sufficiently short times.

6.17. Variational Second-Order Interpolation on the Group of Diffeomorphisms with a Right-Invariant Metric

F-X. Vialard

In [38] we propose a variational framework in which the minimization of the acceleration on the group of diffeomorphisms endowed with a right-invariant metric is well-posed. It relies on constraining the acceleration to belong to a Sobolev space of higher-order than the order of the metric in order to gain compactness. It provides the theoretical guarantee of existence of minimizers which is compulsory for numerical simulations.

6.18. Interpolating between Optimal Transport and MMD using Sinkhorn Divergences


In [33]: Comparing probability distributions is a fundamental problem in data sciences. Simple norms and divergences such as the total variation and the relative entropy only compare densities in a point-wise manner and fail to capture the geometric nature of the problem. In sharp contrast, Maximum Mean Discrepancies (MMD) and Optimal Transport distances (OT) are two classes of distances between measures that take into account the geometry of the underlying space and metrize the convergence in law. This paper studies the Sinkhorn divergences, a family of geometric divergences that interpolates between MMD and OT. Relying on a new notion of geometric entropy, we provide theoretical guarantees for these divergences: positivity, convexity and metrization of the convergence in law. On the practical side, we detail a numerical scheme that enables the large scale application of these divergences for machine learning: on the GPU, gradients of the Sinkhorn loss can be computed for batches of a million samples.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

J-D. Benamou and G. Carlier are members of the ANR MFG (ANR-16-CE40-0015-01). Scientific topics of the project: Mean field analysis Analysis of the MFG systems and of the Master equation Numerical analysis Models and applications
J-D. Benamou G. Carlier F-X. Vialard and T. O. Gallouët are members of ANR MAGA (ANR-13-JS01-0007-01). The Monge-Ampère equation is a fully nonlinear elliptic equation, which plays a central role in geometry and in the theory of optimal transport. However, the singular and non-linear nature of the equation is a serious obstruction to its efficient numerical resolution. The first aim of the MAGA project is to study and to implement discretizations of optimal transport and Monge-Ampère equations which rely on tools from computational geometry (Laguerre diagrams). In a second step, these solvers will be applied to concrete problems from various fields involving optimal transport or Monge-Ampère equations such as computational physics: early universe reconstruction problem, congestion/incompressibility constraints economics: principal agent problems, geometry: variational problems over convex bodies, reflector and refractor design for non-imaging optics.

T. O. Gallouët is member of the ANR GEOPOR (JCJC of C. Cancès) Scientific topic: geometrical approach, based on Wasserstein gradient flow, for multiphase flows in porous media. Theory and Numerics.

T. O. Gallouët is member of the ANR MESA (JCJC of M. Fathi) Scientific topic: Stein methods.

7.2. European Initiatives

7.2.1. FP7 & H2020 Projects

J-D. Benamou and G. Rukhaia are members of the ROMSOC ITN.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

- Shuangjian Zhang, (PostDoc), Université de Toronto, June-August 2018.
- Clarice Poon, Imperial College London, January 2018
- Teresa Radice, Université de Naples, many short stays.

7.3.2. Visits to International Teams

7.3.2.1. Research Stays Abroad

- P. Pegon was invited for 10 days to Penn State College by Alberto Bressan in order to start a collaboration on the theory of ramified transport and applications to biology, and to give lectures (2) in the seminar series on Computational and Applied Mathematics.
- G. Carlier was a John von Neumann invited Professor at TUM (Munich) in 2018.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

- G. Carlier was in the organizing committee of the conference Des mathématiques de la décision aux jeux à champ moyen held in Honor of Jean-Michel Lasry at Dauphine in July 2018.
- T. O. Gallouët was in the organizing committee of the international conference Modern mathematical methods for Data analysis held at Liège (Belgium) in June 2018.
- F-X. Vialard was in the organizing committee of the December 2018 Banff, applied optimal transport, stochastic geometric mechanic and shapes conference.

8.1.2. Scientific Events Selection

8.1.2.1. Reviewer

V. Duval has reviewed contributions to the LVA/ICA conference.
8.1.3. Journal

8.1.3.1. Member of the Editorial Boards


8.1.3.2. Reviewer - Reviewing Activities

- J-D. Benamou has reviewed papers for the following journals: SINUM, Numerische Math., ...
- T. O. Gallouët has reviewed papers for ARMA, ...
- P. Pegon has reviewed articles for Journal de Mathématiques Pures et Appliquées and Journal of Functional Analysis.
- F-X. Vialard has reviewed papers for Numerische Math., Siam IS, IEEE TMI, Siaga, P AMS.

8.1.4. Invited Talks

- I. Waldspurger gave talks at the Séminaire de mathématiques appliquées du collège de France, at the Mathematical Image Analysis conference (Berlin), at the Colloquium in Applied and Computational Mathematics of ETH Zurich, at the 7th international conference on computational harmonic analysis (Nashville) and at the SIAM conference on imaging science (Bologne).
- V. Duval has given invited talks at the MAGA workshop (Jan.), Journée Parcimonie Bordeaux of the GdR MIA (May), GdT StatNum at CEREMADE (May), SIAM Imaging Conference (June), ISMP Conference (July), Journées MAS (Aug.).
- G. Carlier gave talks at TUM (Munich), workshop Optimisation and Machine Learning in Economics at UCL (Londres), workshop Entropies, the Geometry of Nonlinear Flows, and their Applications (Banff, Canada), SFB Colloquium, TUM, Munich, Workshop PDEs and optimal Transport (Essaouira), PDE seminar, Imperial College (Londres), Optimal Transport and Applications (Pise), ERC Readi closing conference (EHESS Paris).
- Y. De Castro has given invited talks at Société Mathématiques de France national conference, Ecole Normale Supérieure de Lyon, Oxford numerical analysis seminar (invited by Pr Tanner), Ecole des Ponts ParisTech, and Labex Bezout day.
- P. Pegon was invited for 10 days to Penn State College by Alberto Bressan in order to start a collaboration on the theory of ramified transport and applications to biology, and to give lectures (2) in the seminar series on Computational and Applied Mathematics.
• Andrea Natale gave the following talks:

8.1.5. Research Administration

J-D. Benamou is a member of the Commision de restauration locale.

J-D. Benamou is the Commission Bureau referent for the 4th floor of building A.

J-D. Benamou is a member of PSL Conseil Académique.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Licence: I. Waldspurger, Analyse 1, 92 heures d’équivalent TD, L1, Université Paris Dauphine, France
Master: V. Duval, project supervision (M2), INSA de Rouen, France
Licence: T. O. Gallouët, Optimisation (TD, L3, Orsay), EDP (TD, 2 ème année, ENS).
Master, Licence: G. Carlier taught algebra (Licence, Dauphine, 91h), transport and variational methods in economics (Dauphine, M2, 18h) and Convex duality and applications in mass transport and the calculus of variations in Munich (18h).
Licence: Y. De Castro taught optimization to first year engineering school student (10h).
Licence: A. Natale taught (TD) for the course “Calcul Différentiel et Optimisation” (Instructor: Emeric Bouin), 09-12/2018, Université Paris Dauphine, Paris (France).
Licence, Master: P. Pegon taught a crash course in analysis (M1, Dauphine, 12h), and gave exercise sessions in Measure theory and probability (L3, Dauphine, 39h) and on Functional Analysis and PDEs (M1, Dauphine, 19.5h).

8.2.2. Supervision

• PhD : Quentin Denoyelle, Theoretical and Numerical Analysis of Super-Resolution without Grid, defended on 09/07/2018, G. Peyré and V. Duval.
• PhD in progress : Gabriele Todeschi, Optimal transport and finite volumes, 01/10/2018, T. O. Gallouët.
• PhD in progress : Miao Yu , Optimal Transport distances for Full Waveform Inversion, 01/10/2016, J-D. Benamou.
• PhD in progress: Paul Catala, Low-rank Approaches for Off-the-grid Superresolution, 01/10/2016, G. Peyré and V. Duval.
• PhD in progress : Lucas Martinet , Decomposed and Parallel Sinkhorn Algorithm for Optimal Transport, 01/10/2017, J-D. Benamou.
• PhD in progress : Giorgi Rukhaia , On the application of Sinkhorn methods in Freeform Optics, 01/05/2018, J-D. Benamou.
• PhD in progress : Ernesto Araya , Random Graphs, Y. De Castro.

8.2.3. Juries

G. Carlier was in the HDR committee of Nicolas Juillet and in the Ph.D committees of Michael Orieux, Hadrien de March and Thomas Dumas
8.3. Popularization

8.3.1. Internal or external Inria responsibilities

V. Duval was a member of the Commission d’évaluation scientifique (CES) of the CRI Paris in 2018, and has been a member of the Comité de suivi doctoral since September 2018.

J-D. Benamou was a member of the CR2 Recruiting Commission of the CRI Paris in 2018.

8.3.2. Education

I. Waldspurger has given a talk at the Mathematic Park seminar (for L1 and L2 students), on the minimization of convex functions with Lipschitz gradient.

8.3.3. Interventions

Participation à l’accueil de 15 collégiens dans le cadre de leur stage d’observation (3 heures dans l’équipe MOKAPLAN supervisé par Lucas Martinet).

9. Bibliography

Major publications by the team in recent years


Publications of the year

Doctoral Dissertations and Habilitation Theses


Articles in International Peer-Reviewed Journals


Conferences without Proceedings


[25] J. M. Fadili, G. Garrigos, J. Malick, G. Peyré. *Model Consistency for Learning with Mirror-Stratifiable Regularizers*, in "International Conference on Artificial Intelligence and Statistics (AISTATS)", Naha, Japan, April 2019, [https://hal.archives-ouvertes.fr/hal-01988309](https://hal.archives-ouvertes.fr/hal-01988309)

Other Publications

[26] J.-D. Benamou, G. Carlier, S. Di Marino, L. Nenna. *An entropy minimization approach to second-order variational mean-field games*, August 2018, working paper or preprint, [https://hal.archives-ouvertes.fr/hal-01848370](https://hal.archives-ouvertes.fr/hal-01848370)

[27] J.-D. Benamou, G. Carlier, L. Nenna. *Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm*, March 2018, working paper or preprint, [https://hal.archives-ouvertes.fr/hal-01621311](https://hal.archives-ouvertes.fr/hal-01621311)


[31] G. Carlier, T. Radice. *Approximation of variational problems with a convexity constraint by PDEs of Abreu type*, May 2018, working paper or preprint, [https://hal.archives-ouvertes.fr/hal-01802925](https://hal.archives-ouvertes.fr/hal-01802925)


### References in notes


[137] D. Lombardi, E. Maître. Eulerian models and algorithms for unbalanced optimal transport, in "Preprint hal-00976501", 2013


