# ERC Consolidator Grant 2018
# Research proposal, Part $\mathbb{B}2$

Filippo Santambrogio

Université Claude Bernard – Lyon 1, Villeurbanne, France

## Everything You Always Wanted to Know About the JKO Scheme
## –
## EYAWKAJKOS

## A State of the art and objectives

This proposal is concerned with the study of a wide class of evolutionary partial differential equations, their discetization and various properties of their solutions, in connection with the theory of optimal transportation (OT). The main object of the proposal are those PDEs which have a variational structure in the space of probability measures endowed with the Wasserstein distance (a distance induced by the optimal value of optimal transport problems), in the sense that they are gradient flows of suitable functionals in such a space.

The notion of *gradient flow*, or *steepest descent curve*, is very classical in evolution equations: given a functional $F$ defined on a Hilbert space $X$, instead of looking at points $x$ minizing $F$ (which solve $\nabla F(x) = 0$), we consider an initial point $x_0$ and look for a curve starting at $x_0$ which tries to minimize $F$ as fast as possible, solving an equation of the form $x'(t) = -\nabla F(x(t))$. In the finite-dimensional case, this equation is very easy to deal with, while the infinite-dimensional case can lead to PDEs. In order to detail some examples, we first need to (roughly) define the notion of first variation of a functional $F$ defined on a suitable functional space $X$: given $u_0 \in X$, if there exists a function $g$ (belonging to $X$ or to another functional space, we will not be precise now) such that $\lim_{\varepsilon\to 0} \frac{F(u_0+\varepsilon\chi)-F(u_0)}{\varepsilon} = \int g\chi$ for every $\chi$ belonging to a large enough class of smooth test functions, we say that $F$ admits a first variation at $u_0$ and we write $g = \frac{\delta F}{\delta u}(u_0)$. In the case $X = L^2$ this notion exactly coincides with that of Gateaux differential, so that $\frac{\delta F}{\delta u} = \nabla F$ whenever $F$ is differentiable. Hence, gradient flows in $L^2$ are easy to consider, as they correspond to $\partial_t u = -\frac{\delta F}{\delta u}$. For instance, the evolution equation $\partial_t u = \Delta u$, is the gradient flow, in the $L^2$ Hilbert space, of the Dirichlet energy $F(u) = \frac{1}{2}\int |\nabla u|^2$, since $\frac{\delta F}{\delta u} = -\Delta u$.

The renovated interest in last years for the notion of gradient flow arrived, at the end of last century, with the work of Jordan, Kinderleherer and Otto ([43]) and then of Otto [66], who saw a gradient flow structure in some equations of the form $\partial_t\rho - \nabla\cdot(\rho\mathbf{v}) = 0$, where the vector field $\mathbf{v}$ is given by $\mathbf{v} = \nabla[\delta F/\delta\rho]$ for a certain functional $F$ defined on the space of probability measures $\rho$. This requires to use the space of such probabilities on a given domain $\Omega$, and to endow it with a non-linear metric structure, derived from the theory of optimal transport. This theory, initiated by Monge in the 18th century ([64]), then developed by Kantorovich in the '40s ([44]), is now well-established (many texts present it, such as [76, 77, 70]) and is intimately connected with PDEs of the form of the *continuity equation* $\partial_t\rho - \nabla\cdot(\rho\mathbf{v}) = 0$.

In order to see the role of the metric structure in gradient flows, let us look at a natural time-discretization of the equation $x' = -\nabla F(x)$: fix a small time step parameter $\tau > 0$ and look for a sequence of points $(x^\tau_k)_k$ defined through the iterated scheme, called *Minimizing Movement Scheme* (see [31, 2]), based on the proximal operator: first, for every point $y$ and every value of $\tau > 0$ define $\text{Prox}_\tau^F(y) := \text{argmin}_x F(x) + \frac{|x-y|^2}{2\tau}$; then, define a sequence obtained in the following way: take $x_0^\tau := x_0$, then

$$x_{k+1}^\tau \in \text{Prox}_\tau^F(x_k^\tau). \tag{1}$$

Under suitable differentiability assumptions on $F$, optimal points satisfy the condition $\frac{x_{k+1}^\tau - x_k^\tau}{\tau} = -\nabla F(x_{k+1}^\tau)$, which is an implicit Euler discretization of the desired evolution equation.

Yet, it is important to notice that the equation $x' = -\nabla F(x)$ requires, to be meaningful, the space $X$ to be endowed with a linear structure and a scalar product (which explains why Hilbert spaces were considered at the beginning), so that we can define derivatives and gradients. On the other hand, the minimization probem in (1) can be easily be considered in general metric spaces: once a l.s.c. function $F : X \to \mathbb{R} \cup \{+\infty\}$ is given on a metric space $(X, d)$, we can interatively minimize $F(x) + \frac{d(x, x_k^\tau)^2}{2\tau}$ and consider the limit (should it exist) that we obtain after suitably interpolating the points $x_k^\tau$ and sending $\tau$ to 0.

This metric approach has been exploited in [43, 66] and in the whole theory developed in [3] using the space $\mathcal{P}(\Omega)$ of probability measures on a (bounded) domain $\Omega$, endowed with the Wasserstein distance $W_2$ induced by the optimal transport with quadratic cost:

$$W_2(\mu, \nu) := \sqrt{\min\left\{ \int |x-y|^2 d\gamma \ : \ \gamma \in \mathcal{P}(\Omega \times \Omega),\ (\pi_x)_\# \gamma = \mu, (\pi_y)_\# \gamma = \nu \right\}}.$$

Given a functional $F : \mathcal{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ we iteratively solve

$$\rho_{k+1}^\tau \in \text{argmin}_\rho F(\rho) + \frac{W_2^2(\rho, \rho_k^\tau)}{2\tau}, \tag{2}$$

and this minimization scheme is nowadays called JKO scheme. Generalizing the previous notation, we will also denote by $\text{Prox}_\tau^F(\nu)$ the set of minimizers of $F(\rho) + \frac{W_2^2(\rho,\nu)}{2\tau}$, so that the JKO scheme can also be defined with the same laguage as in (1). In this case, it is in general possible to prove, under suitable conditions, that the obtained sequence of minimizers converges as $\tau \to 0$ to a solution of the PDE

$$\partial_t \rho - \nabla \cdot (\rho \nabla \left( \frac{\delta F}{\delta \rho} \right)) = 0$$

with no-flux boundary conditions on $\partial\Omega$ that we will omit in the sequel of this text.

Some particular cases are of interest

- when $F(\rho) = \int \rho \log \rho + \rho V$ we obtain the Fokker-Plack equation $\partial_t \rho - \Delta \rho - \nabla \cdot (\rho \nabla V) = 0$ where $-\nabla V$ acts as a drift and is associated with linear diffusion;

- when $F(\rho) = \int \frac{\rho^m}{m-1} + \rho V$ we get the non-linear diffusion equation $\partial_t \rho - \Delta(\rho^m) - \nabla \cdot (\rho \nabla V) = 0$ (called *porous medium* equation for $m > 1$, see [75], or *fast diffusion* for $m < 1$);

- when $F(\rho) = \int \rho V$ for $\rho \le 1$ and $F(\rho) = +\infty$ if $\rho$ is not bounded by 1 (i.e., the limit $m \to \infty$ of the previous case), we obtain the Hele-Shaw type system

$$\partial_t \rho - \Delta p - \nabla \cdot (\rho \nabla V) = 0,\ p \ge 0, \rho \le 1,\ p(1-\rho) = 0, \tag{3}$$

  first seen as a gradient flow in [60] for applications to crowd motion models.

- aggregation-diffusion equations can be considered ([22, 23]) and the general framework is obtained choosing the functional $F(\rho) = \int f(\rho) + \rho V + \frac{1}{2} \int\int W(x-y) d\rho(x) d\rho(y)$ for an even interaction kernel $W$, which gives

$$\partial_t \rho - \nabla \cdot (\rho \nabla (f'(\rho) + V + W * \rho)) = 0.$$

Many other diffusion PDEs can be obtained in this way, including the Keller-Segel model for chemotaxis ([45, 46, 14]) and higher order equations ([58, 21]). Several cross-diffusion models can be dealt with in product spaces where several populations $\rho_i$ evolve ([51, 24]): if we consider a functional $F$ defined on $\mathcal{P}(\Omega)\times\mathcal{P}(\Omega)$ the JKO scheme becomes

$$(\rho^\tau_{k+1},\mu^\tau_{k+1}) \in \operatorname{argmin}_{\rho,\mu} F(\rho,\mu) + \frac{W_2^2(\rho,\rho^\tau_k) + W_2^2(\mu,\mu^\tau_k)}{2\tau}$$

and the equation becomes a system

$$\begin{cases} \partial_t\rho - \nabla\cdot(\rho\nabla\left(\frac{\delta F}{\delta\rho}\right)) = 0, \\ \partial_t\mu - \nabla\cdot(\mu\nabla\left(\frac{\delta F}{\delta\mu}\right)) = 0. \end{cases}$$

The JKO scheme, which provides a natural approximation, is a powerful tool to obtain existence results, as well as a theoretical numerical method, discrete in time but still continuous in space. If for well-known PDEs (such as the linear Fokker-Planck equation) there are many easier techniques for both existence and numerics, for other cases the JKO scheme could be crucial. Of course, other approximations could exist (adding vanishing viscosity, non-local terms…) but in many cases the advantage of the JKO scheme is that it preserves many features (including some powerful algebraic miracles in the computations) of the structure of the desired continuous-time equations.

One of the goal of the present project is exactly to show to which extent many features of the corresponding PDEs can be observed in the JKO scheme, and we mainly focus on those properties of the solutions which can be inferred from this scheme. For instance, in [32] the PI and collaborators introduced a technique, based on an inequality nowadays called *five-gradients-inequality*, which allowed to prove that the BV norm of $\rho^\tau_k$ decreases with $k$ on the JKO scheme for the porous-medium equation: given an arbitrary probability density $\nu \in \mathcal{P}(\Omega)$ on a convex domain $\Omega$, taking $F(\rho) := \int f(\rho(x))dx$ for a convex and superlinear function $f$, then we have $\|\rho\|_{BV} \le \|\nu\|_{BV}$ for $\rho = \mathrm{Prox}^F_\tau(\nu)$. Moreover, considering $f(\rho) = \rho^m$ and sending $m \to \infty$, we also obtain the same BV estimate when $\rho$ solves

$$\min\left\{W_2^2(\rho,\nu) \;:\; \rho \le 1\right\},$$

i.e. when $\rho$ is the projection of $\nu$ onto the set of densities bounded by 1, a projection operator very useful in density-constrained models such as those for crowd motion (see [60]).

This kind of result, well-known for the continuous-in-time setting, was not known for the discrete-in-time case, and is useful to obtain compactness properties, asymptotic behavior, consistence of numerical schemes… For the linear Fokker-Planck equation, the spectrum of available results is much wider, starting from some very interesting works by Lee ([54] and [55]), which respectively prove Lipschitz and semiconcavity bounds in the case where $\Omega$ is the torus. The Lipschitz bound has been recently extended by the PI with a young collaborator in [37] to the case of a convex domain, with a sharp rate which exactly reproduces what happens in the continuous-in-time case. Similarly, some Sobolev bounds proven in [34] are sharp in what concerns Fokker-Planck, but have also been proven in a non-sharp form (the asymptotical behavior is non-optimal) for other aggregation-diffusion equations, including Keller-Segel. Among the other classes of bounds which can be proven on the JKO scheme, we cite $0-$order bounds, such as estimates on the $L^p$ norm of $\rho$, or on $L^\infty$ bounds from above or below. Other available bounds, instead of being uniform along the iterations (i.e., in time) are integrated in time. For instance, on the heat equation it is easy to see that if one differentiates in time the quantity $t \mapsto \int \rho_t^2$ we obtain at the same time a uniform $L^2$ bound on $\rho$ and an integral bound on the $L^2$ norm of $\nabla\rho$, since $\frac{d}{dt}\int\rho_t^2 = -2\int|\nabla\rho_t|^2$; a discrete counterpart can be proven using the so-called *flow interchange* technique (see [58]). Similarly, it is possible to differentiate 1st order quantities and obtain a bound on the integral of a 2nd order one: for instance, in the discrete-in-time setting it is possible to obtain bounds on quantities such as $\sum_k \tau \int \rho^\tau_k |D^2\log(\rho^\tau_k)|^2$ and hence $L^2_t H^2_x$ bounds.

The research project will be structured along three main themes, together with their connections.

### A.1 $(\tau, x)$ – Discrete-in-time regularity estimates and applications

This part of the project is the main core of it. It aims at developing a general theory for the JKO scheme for various equations and

- reproduce in discrete time integrability, decay, and uniform estimates which are known in the continuous case,
- find new ones which could be easier to observe in the discrete case and/or to prove rigorously.

The second above point (i.e. that it could be easier in some cases to act on the discrete setting than in the continuous one) could seem surprising but one has to think that in this case there is no regularity issue to justify the computations (as the time is discrete and functions can be assumed to be essentially as smooth as we want for fixed time step $\tau > 0$); moreover, it could happen that some involved error terms could not be immediately seen to have a sign, while this could be made apparent by using some black-box inequality coming either from the flow interchange (based on displacement convexity, see [61]) or the five-gradients-inequality, as it happened for the case of variational Mean Field Games in [53], where the desired result was indeed proven by discretization in time and flow interchange.

In order to exemplify the kind of properties we could consider we list here below the main results which are known in the easiest case, the Fokker-Planck equation, that we write as $\partial_t\rho - \nabla\cdot(\rho\nabla u) = 0$, with $u = \log\rho + V$. Note that we use $F(\rho) = \int \rho\log\rho + \rho V = \int \rho u$. We consider a datum $\rho_0$ and set $\rho_1 := \mathrm{Prox}^F_\tau(\rho_0)$, i.e. the new density produced by the JKO scheme. We also set, for $k = 0, 1$, $u_k = \log\rho_k + V$. Here is what is known, so far, when the domain $\Omega$ is convex:

- If $u_0$ is bounded from above (resp., below) then $u_1$ is bounded from above (resp., below) by the same constant, whatever is $V$ (this is proven in a more general setting in [41]).
- If $V$ is Lipschitz continuous then we have a bound on the $L^p$ norm: $\int \rho_1^p \le (1 + \tau C(p)\mathrm{Lip}(V)^2) \int \rho_0^p$. This provides uniform bounds on the $L^p$ norm of $\rho_k^\tau$ as soon as $\rho_0 \in L^p$ and $\tau k \le T$. Moreover, it is possible to improve the above estimate into

$$\int \rho_1^p \le (1 + \tau C(p)\mathrm{Lip}(V)^2) \int \rho_0^p - \tau C(p) \int \rho_1^{p-2}|\nabla\rho_1|^2,$$

  which allows to deduce integral bounds in time and space on the $H^1$ norm of $\rho^{p/2}$, i.e. on $\tau\sum_k \int_\Omega (\rho_k^\tau)^{p-2}|\nabla\rho_k^\tau|^2$.
- When $V$ is semiconvex, i.e. it satisfies $D^2V \ge \alpha I$ for some $\alpha \in \mathbb{R}$, if $u_0$ is Lipschitz continuous so is $u_1$, and $\mathrm{Lip}(u_1) \le (1 + \alpha\tau)^{-1}\mathrm{Lip}(u_0)$ (this is proven in [54] for the torus, and in [37] for the case of convex domains).
- Again for $D^2V \ge \alpha I$, setting $J_p(\rho) := \int |\nabla u|^p\rho$, we get $J_p(\rho_1) \le (1 + \alpha p\tau)^{-1}J_p(\rho_0)$ (see [34]). This estimate can also be improved with a higher-order remainder term depending on $p$, $\nabla\rho_1$ and $D^2\rho_1$. Using $p = 2$ we obtain in particular integral bounds in time and space on the Hessian of $u$, and more precisely a bound on $\tau\sum_k \int_\Omega \rho_k^\tau|D^2u_k^\tau|^2$.

The recent paper [34] partially generalizes some of the above results to the JKO scheme corresponding to other equations of the form $\partial_t\rho - \Delta\rho - \nabla\cdot(\rho\nabla V[\rho]) = 0$ where the potential $V$ is allowed to depend on $\rho$ (by convolution, or by solving a PDE as in the Keller-Segel case $-\Delta V[\rho] = \rho$). However, in this case the results are far from being sharp and the spectrum of available results becomes much more restricted when passing to non-linear diffusion such as the porous medium equation $\partial_t\rho - \Delta(\rho^m) = 0$.

The objective of this part is to generalize as much as possible (other equations, general domains, mild assumptions on the data…) the estimates on the solutions of the JKO scheme and apply these results in different directions. We list here below the main tasks to be accomplished. Note that the generalization to other JKO-like schemes will also be matter of study if the results on the "standard" JKO scheme are satisfactory enough. This includes splitting schemes with a JKO step coupled with explicit steps, minimizing movements in other distances involving $W_2$ (see $(t,x)$ for cross-diffusion in the product space $W_2 \times W_2$ or [15] for Hybrid variational schemes in $W_2 \times L^2$), higher-order (in time) JKO-like schemes as in [56, 59], or even the equations which can be obtained by replacing $W_2^2/\tau$ with $W_p^p/\tau^{p-1}$ which gives (see [1])

$$\partial_t\rho + \nabla\cdot(\rho\mathbf{v}) = 0, \quad |\mathbf{v}|^{p-2}\mathbf{v} = -\nabla\frac{\delta F}{\delta\rho}.$$

### A.1.1 Task $(\tau, x) - 1$: strong convergence for linear diffusion

Let us consider the decay estimate of the $L^2$ norm for the JKO scheme of the Heat equation (Fokker-Planck with $V = 0$, for simplicity): it is of the form

$$\int \rho_1^2 \leq \int \rho_0^2 - 2\tau \int |\nabla \rho_1|^2,$$

which, iterated along the JKO scheme, provides

$$\int_0^T \int_\Omega |\nabla \rho^\tau|^2 + \int_\Omega (\rho_T^\tau)^2 \leq \int_\Omega (\rho_0)^2,$$

where $\rho^\tau$ denotes the piecewise constant interpolation of the sequence $\rho_0, \rho_1, \dots$. When sending $\tau \to 0$ this allows to estimate from above $\limsup_\tau \int_0^T \int_\Omega |\nabla \rho^\tau|^2$ and, using the characterization of strong $L^2$ convergence as weak convergence completed by the convergence of the $L^2$ norm, proves strong $L^2$ convergence of $\nabla \rho^\tau$ to $\nabla \rho$, the gradient of the continuous-in-time solution. This exploits the fact that the estimate obtained in the discrete setting is exactly the translation of the continuous one. The extenstion to the case $V \neq 0$ is easy, and the extension to higher-order convergence (using the functional $J_2$ and estimating the $L^2$ norm of $D^2 u^\tau$) is a work-in-progress of the PI with G. Toshpulatov (current master student of the PI). The core of this task, after finalizing the study of the case of the Fokker-Planck equation, will consist in generalizing it to aggregation-diffusion equations, with linear diffusion but density-dependent potential (for instance with potential of the form $V[\rho] := W * \rho$), starting from the most regular cases and possibly arriving to more singular convolution kernels $W$, up to the Keller-Segel case.

Once these strong convergence results are proven, they allow to use the densities obtained through the JKO scheme as a reasonable proxy for the solution of the continuous equation, including in what concerns the pointwise (a.e.) values of their derivatives, thus improving the role of numerical simulations for visualizing the solutions.

### A.1.2 Task $(\tau, x) - 2$: BV, sobolev or Lipschitz bounds for non-linear diffusion

First-order estimates on the JKO scheme besides the case of linear diffusion (considered in [34]) are so far only available for the gradient flow of functionals of the form $F(\rho) = \int f(\rho(x))dx$, i.e. for equations of the form $\partial_t \rho = \nabla \cdot (\rho \nabla (f'(\rho)))$ (which can also be written as $\partial_t \rho = \Delta(h(\rho))$ for a suitable function $h$ depending on $f$). In this case the results in [32] prove that the BV norm decreases in time. Yet, the very same proof cannot even be used in the case where a potential is added ($F(\rho) = \int f(\rho(x))dx + \int V d\rho$), and the only known way to tackle the case of a potential is by splitting: instead of a standard JKO scheme we first follow the drift $-\nabla V$ for a time $\tau$, which allows for BV estimate if $V$ is very smooth ($C^3$, actually), and then use a step of the JKO scheme for the functional without $V$. Not only this is not very satisfactory, but another natural question also arises, i.e. obtaining estimates which go beyond the BV case, such as the $W^{1,p}$ estimates obtained in [34] for linear diffusion. Of course, non-sharp estimates would already be a great success.

This task, even if it stems from previous results of the PI ([32, 41, 34, 37]), is one of the more difficult challenges of the project, as the currently used techniques do not guarantee obtaining such a result. The new techniques which will be developed in such a task will for sure help in the study of other equations, including cross-diffusion systems.

### A.1.3 Task $(\tau, x) - 3$: functional inequalities using discrete flows

If $V$ is a convex potential with $D^2 V \geq \alpha I$, $\alpha > 0$, and $\int e^{-V} = 1$, the well-known logarithmic Sobolev inequality states that for every $g$ we have $\alpha \int (g^2 \log g) e^{-V} \leq \int |\nabla g|^2 e^{-V}$. This can be formulated, setting $g^2 = \rho e^V$, $F(\rho) = \int \rho \log \rho + \rho V$, $u = \log(g^2) = \log \rho + V$ and $J_2(\rho) = \int |\nabla u|^2 d\rho$, as $2\alpha F \leq J_2$ and can be proven, according to the Bakry-Emery theory (see for instance [5]), starting from an arbitrary initial datum $\rho_0$ in the Fokker-Planck equation $\partial_t \rho = \Delta \rho + \nabla \cdot (\rho \nabla V)$ and differentiating both $F$ and $J_2$. We obtain $\frac{d}{dt} F(\rho_t) = -J_2(\rho_t)$ and $\frac{d}{dt} J_2(\rho_t) \leq -\alpha J_2(\rho_t)$. Using the fact that both $F$ and $J_2$ tend to 0 as $t \to \infty$ and integrating in time, this provides the desired inequality for the initial datum.

The same technique has now been popularized for other functional inequalities by differentiating twice suitable energies. Yet, this requires the study of the corresponding continous-in-time PDE, and in particular to prove existence and regularity of the solution, which may be hard in some cases. We cite, for instance [78], together as the PhD thesis of the same author, where the study of a non-linear flow required an impressing number of approximations. Hence, a natural idea is to replace continuous flows with the JKO scheme, as a consequence of the 0 and 1st order inequalities which can sometimes be proven on the solutions of such a scheme. The advantage would be to completely skip the difficulties related to regulairty and existence. A proof of the logarithmic Sobolev inequality using the estimates on the Fokker-Planck JKO scheme from [34] seems doable, even if an extra difficulty arises from the fact that the decrease rate of $F$ is not expressed in terms of an equality with $J_2$, but only of an inequality (one should prove that this is asymptotically an equality as $\tau \to 0$). The core of this task, after detailing this proof about the Fokker-Planck case and the log-Sobolev inequality, will be the generalization to other flows and the application of this technique to other functional inequalities. This part will be mainly developed in collaboration with I. Gentil (see the Methodology Section B2 for the list of collaborators and their roles).

### A.2 $(\tau, h)$ – Estimates for fully discrete schemes

This part of the project aims at exploiting some of the results of the first part, together with extra ideas from the algorithmic and computational side, for efficient numerical methods. Due to the variational structures of the equations we consider, the resulting algorithm will be a procedure based on optimization tools (convex and linear optimization) to approximate the solution of an evolution PDE.

We start this presentation from a connection with the estimates from the $(\tau, x)$ part, and in particular the Lipschitz bound on $u$ for Fokker-Planck. Indeed, since the drift in the equation is $-\nabla u$, Lipschitz bounds on $u$ translate into $L^\infty$ bounds on $\nabla u$, hence on the speed of the particles. More precisely, at the discrete level with time step $\tau$, every particle moves along the optimal transport between $\rho_k^\tau$ and $\rho_{k+1}^\tau$ of a distance being at most $C\tau$. This can be used for numerical purposes, as it simplifies the computational cost of some algorithms for optimal transport.

In order to present in few words the existing numerical methods for optimal transport, the most efficient ones can be classified in three classes: purely continuous, PDE-based (Benamou-Brenier, [8]); semi-discrete (one measure has a density but the other is atomic, thus transforming the problem into a partition problem into Laguerre cells, [62, 52]); purely discrete, exploiting the linear programming formulation of the Kantorovich problem. All these algorithms admit suitable variants where one of the marginal measures is not prescribed but is part of the optimization, as it happens for the JKO scheme (see [11, 12]). We are interested here in the last class, where the most used algorithm is the Sinkhorn one, popularized first by [28] and then by [10]. Sinkhorn provides a “lightspeed” approximation of the optimizer. In order to have the exact optimizer other, slower, algorithms exist, based for instance on the network simplex (see for instance [16] for an efficient application of this method to optimal transport). Yet, Sinkhorn's strength is to be able to deal with arbitrary costs and/or non-structured sets of points, wich is not crucial when dealing with the discretization of a PDE. Indeed, in this case the best is to have points on a uniform or structured grid. For this latter situation, a recent variant of the network simplex, exploiting the separability of the cost according to the different coordinate variables, strongly reduces the number of arcs in the network to consider ([4], see also the methodology section B1), and is a reasonable choice which will be mainly followed in this project. Also note other reasons to prefer the use of the network simplex rather than Sinkhorn: first, for iterating the minimization problem, it is crucial to have a very precise estimate of the minimizer ; then, we would like to exploit the information on the size of the displacement of each particle, which allows for an extra redution of the number of arcs, an advantage which does not seem to be exploited by Sinkhorn (knowing that some entries of the transport plan will necessarily be 0 goes in a direction incompatible with the explicit lightspeed structure of such an algorithm).

We then see that the Lipschitz estimates on the JKO scheme for Fokker-Planck allow to reduce, essentially by a factor $\tau$, the number of arcs to be considered in a network simplex approach to the optimal transport problems in the JKO scheme (where the number of OT problems to be solved is of order $\tau^{-1}$). Similar results for other equations would be extremely useful for the numerical approximation of the solutions.

This part of the project will focus on the convergence when $\tau, h \to 0$, on the computational complexity, on the numerical implementation, and on the search for new efficient bounds on the displacement in terms of $\tau$ in both a continuous and a discrete setting in $x$.

### A.2.1 Task $(\tau, h) - 1$: displacement estimates for nonlinear diffusion

This task starts with a question which could also fit the part $(\tau, x)$: for a given functional $F$, for instance of the form $F(\rho) = \int f(\rho(x))dx + \int Vd\rho$, can we bound in $L^\infty$ norm the displacement $T(x) - x$ for the optimal transport map $T$ from the optimal density $\rho^\tau_{k+1}$ to the previous one $\rho^\tau_k$? the goal would be to bound it in terms of a power $\tau^\alpha$ of the time-step $\tau$ (if possible, $\alpha = 1$, as it is the case for the Fokker-Planck case when $\log\rho_0$ is Lipschitz continuous).

A case on which we can say something is the crowd motion case where $f$ is the constraint $\rho \le 1$, (3). In this case the choice $\alpha = 1$ cannot be met, but we can quite easily obtain $\alpha = 2/(d+2)$. The impossibility of proving a bound with $\alpha = 1$ comes from [29], which proves an $L^4$ bound on the gradient of the pressure of the incompressible limit and shows that it is sharp, while the displacement being $O(\tau)$ would imply an $L^\infty$ bound on $\nabla p$. On the other hand, the techniques in [18] can be adapted to the crowd motion case (3), where the important region is the saturated region $\{\rho = 1\}$, in order to obtain $||T - id||_{L^\infty} \le CW_2^{2/(d+2)}$ and then estimate this in terms of $\tau$ (but we do not know whether this exponent is sharp for the crowd motion case). Since this case corresponds to $m = \infty$ in the porous medium equation $\partial_t\rho = \Delta(\rho^m)$, it would be natural to look for similar displacement estimates for other exponents $m \in (1, \infty)$ and It would be interesting to obtain an exponent $\alpha$ depending on $m$. This is the core of the first part of this task.

However, the linear programming approach that we propose for numerical purposes requires space discretization, and the same happens for these $L^\infty$ bounds on the displacement. Indeed, it is possible to impose as an extra constraint on the (fully-discrete) linear programming problem the $L^\infty$ displacement bound obtained from the JKO scheme (discrete in time but not in space), but we have to prove that this extra constraint does not affect the convergence to the desired limit; or we can follow another, more natural path: re-prove similar bounds, adapted to the case where also space is discretized. The technique for the $L^\infty$ bounds on the displacement in the continuous (in space) or discrete cases are necessarily very different from each other, since in the continous case we can exploit the fact that we have $T - id = -\nabla\varphi$ and use the Monge-Ampère equation to find the maximum point for $|\nabla\varphi|$ (see the Methodology Section B1).

### A.2.2 Task $(\tau, h) - 2$: efficient JKO algorithms via network simplex

*First part: efficient LP algorithm for crowd motion*

This part of the project will require computer implementation and optimization of the network simplex algorithm for the JKO scheme. The most natural case where to apply the approach we described above is the crowd motion case (3). The reason is the fact that the functional $F$ is linear, up to a part which imposes a constraint ($\rho \le 1$), which is also linear. Hence, the whole JKO scheme can be written as a sequence of Linear Programming (LP) prolems on the unknown $\gamma$, which is the transport plan between the new and the old measure: the functional to be minimized is, in continuous language, of the form

$$\gamma \mapsto \int \left( \frac{|x - y|^2}{2\tau} + V(x) \right) d\gamma$$

under the constraints $(\pi_x)_\#\gamma \le 1$ and $(\pi_y)_\#\gamma = \nu$ (the fixed measure $\nu$ being the optimizer of the previous step). This can be discretized, and transformed into a network problem with a reduced number of arcs using at the same time the separability trick of [4] and the estimates on the displacement of the previous task. Its complexity and comparison with other previously used algorithms for the same equation will be studied.

*Second part: efficient algorithm for other equations*

After studying the case where the optimization problems are linear, we will pass on to the case of convex optimization, in particular for functionals of the form $F(\rho) = \int f(\rho(x))dx + \int Vd\rho$. Note that the Fokker-Planck case is peculiar in this sense: the equation is linear, but the optimization problem is not, because of the logarithmic entropy. Using general $f$ arises two difficulties: the first consists in the displacement bounds, which should come from the results of Task $(\tau, h) - 1$; the second is the fact that the discretized optimization problem is a network problem where the flow on some arcs is not bounded in capacity but penalized via a convex function, and this requires adapations of the algorithm.

Both parts of this task will be mainly considered in collaboration with N. Bonneel.

### A.2.3 Task $(\tau, h) - 3$: convergence proofs

This task concerns the rigorous proofs of convergence of the above fully-discrete schemes, consisting in a sequence of discrete optimization problems approximating the solution of the desired PDE, for which we will be able to provide explicit bounds on the complexity of their solution. Yet, for the convergence when $\tau, h \to 0$ some conditions will for sure be needed, since it is well-known that $\tau$ cannot be too small compared to $h$: otherwise, the cost for moving a unit mass would be at least $O(|h|^2/\tau)$ which could result in a completely frozen evolution if $|h|^2 \gg \tau$.

Of course, the rate of convergence as a function of $\tau$ and $h$ will also be investigated and, thanks to the possible results in $(\tau, x) - 1$, they could be quantified either in $W_2$ distance or in stronger norms.

This last question will require involving specialists from numerical analysis and in particular F. Lagoutière.

## A.3 $(t, x)$ – Analysis of continuous-in-time models

This part of the project is more focused on PDE questions, but of course will exploit many of the results from the two previous parts. In what concerns existence of a solution, the JKO scheme will be a tool but not the only one, even if many other ideas are anyway inspired from the variational structure of the problem, such as the use of the Energy Dissipation Inequality (EDI, see [3, 72]).

For some PDEs for which existence has already been proven the question is to study sharp uniqueness results and/or regularity and qualitiative or asymptotic properties of the solutions. For some equations another approach is to consider the JKO scheme as the main object, and then study the PDE properties of the limits when $\tau \to 0$ of its solutions: if the class of weak solutions is large, find which one is selected by the JKO scheme; if the equation has a formal gradient-flow structure, but associated with a functional which is not lsc findd the connections between the gradient flow of the lsc envelop and the initial equation (more generally, study the limits of the JKO scheme when such a scheme is well-posed an the continuous equation is not, or not known to be well-posed).

### A.3.1 Task $(t, x) - 1$: improved estimates on crowd motion models and applications

One of the main contribution of the PI to the theory of Wasserstein gradient flows was the study of the crowd motion model started in [60]. Such a paper mainly proved existence (and provided numerical simulations), while uniqueness results are contained in [33] and in the recent preprint [42]. We do not know much about the regularity, in particular of the pressure term, but the techniques in [30] should provide $\nabla p \in L^4_{t,x}$ and some weak second-order estimates (this is a work in progress of the PI with the authors of [30]).

Improving the regularity, and the stability, of the solutions of such a gradient flow would be, besides an interest *per se*, important for applications to systems where this PDE appears but is not the only ingredient, in order for instance to apply fixed point theorems. This is the case of the Mean Field Game (MFG) model under density constraints proposed in [71], which consists in a (time-dependent) gradient evolution driven by a value function solving a Hamilton-Jacobi equation where the gradient of the pressure appears. Existence could be proven via a fixed point procedure on the pressure, but requires strong stability and uniqueness results in order to apply Schauder's fixed point theorem.

For simplicity reasons, since it is known that this makes uniqueness easier (see [33]), the same MFG has been considered adding diffusion to the players, which means adding a Laplacian to both equations. This leads to the variant of the crowd motion model involving diffusion studied in [63], which also arises its number of questions. In particular, even if we expect to have better estimates, some of the algebraic tricks used in [30] should be adapted, and reproducing the same results is not trivial in this case.

In this task, mainly involving the PI and exterior collaborators, the goal will be to find sharp regularity results on the pressure and study their applications to other models in density-constrained fluid problems, including MFG.

### A.3.2 Task $(t, x) - 2$: existence results for the total variation flow and higher-order equations

Some higher-order equations can be considered in the framework of this project. We can obtain fourth-order equations as soon as the functional $F$ involves derivatives of $\rho$. Note that in general these functionals (except for few examples in dimension 1, see [25]) are not geodesically convex, which prevents from using the general theory of [3] and requires ad-hoc estimates, as iti is intensively done in [58]. Among the equations that are of interest for EYAWKAJKOS, we mention the gradient flow of the total Variation functional $F(\rho) = \int |\nabla\rho|$ (to be intended in the BV sense). Here the non-linear equation can be written as

$$\partial_t\rho + \nabla\cdot(\rho\nabla(\nabla\cdot z)) = 0, \quad z\cdot\nabla\rho = |\nabla\rho|, \quad |z| \le 1.$$

This PDE has been studied in [21] but the existence result was only presented in 1D, the reason being the fact that some estimates to guarantee compactness and pass to the limit $\tau \to 0$ required lower bounds on the density $\rho$. If upper bounds can be obtained by estimating integrals of the form $\int \rho^q$ for $q \to \infty$, and these quantities can be tackled via the usual flow interchange technique, lower bounds would require $q \to -\infty$, which does not allow to use the flow interchange technique since for $q < 0$ we do not have a displacement convex functional (except in dimension 1, which explains the partial result in [21]). Yet, we already saw that lower bounds can be preserved for other equations, such as in the Fokker-Planck case, via a different technique introduced in [41], and we could hope for a similar result in this setting, thus allowing to generalize the existence result of [21] to arbitrary dimensions.
Besides existence, the asymptotic behavior of the solution and the rate of convergence to the stable configuration (a constant value equal to $1/|\Omega|$ could also be investigated. Note that this stable solution cannot be attained if a constraint $\rho \in \{0, 1\}$ is added, and the variant with this extra constraint (for which the JKO scheme is perfectly well-posed) gives an interesting geometric evolution flow (see [26]).

Among other fourth-order equations that can be considered, we cite thin-film equations of the form $\partial_t\rho + \nabla\cdot(\rho\nabla\Delta\rho) = 0$, (see [7] as well as [58]) or their variants $\partial_t\rho + \nabla\cdot(M(\rho)\nabla\Delta\rho) = 0$, coming from modified Wasserstein distances with a different mobility function, in the spirit of [35]. Note that the theory in [35] requires $M$ to be concave in order to have lsc action functionals, but here the regularizing effect coming from the higher-order energy $F(\rho) := \int |\nabla\rho|^2$ should allow more general $M$.

Finally, we cite the possibility of applying five-gradient inequalities to this higher-order setting, which could come up into uniform-in-time second-order bounds (since we estimate quantities of the form $-\int \nabla\rho\cdot\nabla\varphi$, where $\varphi = -\frac{\delta F}{\delta\rho}\dots$ considering for instance $\varphi = -\Delta\rho$, integrating by parts,we obtain a bound on $\int |\Delta\rho|^2$).

### A.3.3 Task $(t, x) - 3$: new results for cross-diffusion models and for the Muskat problem

Among other systems deserving their study, cross-diffusion problems are a natural application of the gradient flow theory. The simplest example of a gradient-flow cross-diffusion problem for which there is so far, in general, no existence result is the following system

$$\begin{cases} \partial_t\rho - \nabla\cdot(\rho\nabla(f'(\rho+\mu))) - \nabla\cdot(\rho\nabla V) = 0, \\ \partial_t\mu - \nabla\cdot(\mu\nabla(f'(\rho+\mu))) - \nabla\cdot(\mu\nabla W) = 0. \end{cases} \tag{4}$$

This is the gradient flow, in $W_2(\Omega)\times W_2(\Omega)$, of the energy $(\rho,\mu) \mapsto \int f(\rho+\mu) + \int V d\rho + \int W d\mu$. It has been studied under a very restrictive assumption in dimension 1 in [51], but the higher-dimensional problem is open. The difficulty is to pass to the limit the product of $\rho$ or $\mu$ times $\nabla(f'(\rho+\mu))$, two terms on which it is easy to obtain weak convergence but not strong. A similar system, with reaction terms instead of different potentials, was studied in [40, 13] and then with optimal transport methods in [24], but, again, only in the 1D case. Note that the presence of reaction terms made the equation non-conservative and required a slight modification of the JKO scheme, i.e. a splitting technique (first we apply a reaction phenomen for a duration $\tau$, then we apply a step of a JKO scheme; this is not a major change and this part of the project will also consider problems of this form, adapting the main techniques).

Treating the above cross-diffusion system will require new higher-order estimates on $\rho$ and $\mu$ (or on $\nabla(\rho+\mu)$) which are not available so far and is a challenging part of the ERC project EYAWKAJKOS.

Recently, the PI, in collaboration with R. Ducasse and H. Yoldas, also studied a different cross-diffusion problem which is motivated by [6], a paper devoted to the very natural system

$$\begin{cases} \partial_t\rho - \Delta\rho - \nabla\cdot(\rho\nabla\mu) = 0, \\ \partial_t\mu - \Delta\mu - \nabla\cdot(\mu\nabla\rho) = 0, \end{cases}$$

which is formally the Wasserstein gradient flow of the functional $F(\rho,\mu) = \int \rho\log\rho + \mu\log\mu + \rho\mu$. Unfortunately, this functional is not l.s.c. since the function $(a,b) \mapsto f(a,b) := a\log a + b\log b + ab$ fails to be convex on $\{ab > 1\}$. Hence, in such a work in progress with Ducasse and Yoldas, the PI is considering the gradient flow of the lower semi-continuous envelop of $F$ (i.e. the integral functional defined by convexifying $f$ on $\mathbb{R}^2_+$), and an existence result is available using a technique based on the Energy Dissiparion Inequality (EDI, see [3]), proving the lower semicontinuity of the slope of the new functional (note that the same technique cannot be applied to (4) as the corresponding slope fails to be lsc for the weak convergence of probability measures, which confirms that stronger bounds and better convergence need to be enforced). An extra question which has not been considered yet is to find the relation between the two PDEs, the one which is the gradient flow of $F$ and the one which is the gradient flow of its lsc envelop. Indeed, if in the JKO scheme, which is a minimization problem, the lsc envelop of $F$ automatically appears (minimizing a functional which is not lsc, considering that the extra terms are Wasserstein distances which, on a compact domain, are continuous for the weak convergence, produces a minimizing sequence, with possible oscillating behavior, which weakly concerges to the minimizer of the relaxed functional), it is not clear what happens in the PDE setting.

Some of the features of the two above cross-diffusion problems meet when considering another, independent, PDE model, which is part of the research project of another collaborator of EYAWKAJKOS, Aymeric Baradat, with whom this task will be mainly considered and whose competences will be useful in most of this part of the project. This model is the so-called Muskat problem, see [65], which is a model for the evolution of two incompressible and immiscible fluids subjects to different gravity effects because of their different densities. The same problem can be reformulated in terms of a gradient flow for two species (see also [50]), $\rho$ and $\mu$, subject to two different potentials acting in the same direction (consider $\nabla V$ and $\nabla W$ to be two opposite constant multiples of the vertical vector), so that we can write

$$F(\rho,\mu) = \int z d\rho - \int z d\mu \quad \text{if } \rho + \mu = 1,$$

where $z$ denotes the vertical coordinate function. The $W_2 \times W_2$ gradient flow of $F$ gives a formulation for the Muskat problem, if we forget about the fact that we would also like to impose the immiscibility (segregation) constraint $\rho\mu = 0$ (equivalently, this means that $\rho$ and $\mu$ are indicator functions). Note that in general the solutions of the JKO scheme for this functional, which is a sequence of linear minimization problems, automatically satisfy this segregation constraint (since $\rho$ and $\mu$ will be the concentrated on the upper and lower level sets of some function), but this condition is not stable for weak convergence.

This underlines many questions and similarities between this problem and the two cross-diffusion problems previously exposed. First, even ignoring the immiscibility constraint, the Muskat problem formally fits the framework of (4) using $f = I_{\{1\}}$ (or, equivalently, $f = I_{[0,1]}$ as imposing $\rho + \mu = 1$ everywhere will be the same as $\rho + \mu \le 1$ if the domain $\Omega$ has suitable volume). Morever, the constraint $\rho\mu = 0$ is a non-convex constraint which prevents lower semicontinuity, and its relaxation exactly consists in removing this constraint and admitting mixing of the two fluids. This procedure (studying the JKO scheme of the relaxed functional) is exactly what is done for the second above problem. For the Muskat problem, general notions of subsolution have been introduced in various works [27, 74, 38], and taking the limit of the JKO scheme is indeed a way of selecting a particular subsolution in this class.

As a consequence, this task will on the one hand look for existence results for general cross-diffusion models which have a variational structure in the (squared) Wasserstein space, but also investigate the connections between the limits of the solutions of the JKO scheme (which automatically applies a relaxation procedure, the class of functionals being lsc when involving two densities instead of one being much more restricted than in the classical case) and the solutions or subsolutions of continuous PDEs, in collaboration with A. Baradat and with the current collaborators of the PI.

#### A.3.4 Task $(t,x)-4$: the sliced Wasserstein flow

Before the rise of strong numerical methods for OT and in connections with some procedures already applied in image processing [67], in the years 2000 proposed a flow whose goal was to moe the particles from an initial configuration $\rho_0$ to a target $\nu$. In this flow every particle followed the vector field $\mathbf{v}$ obtained in the following way: given the two measures $\rho, \nu \in \mathcal{P}(\mathbb{R}^d)$, we project them onto any one-dimensional direction $e \in \mathbb{S}^{d-1}$ via the map $\pi_e : \mathbb{R}^d \to \mathbb{R}$ given by $\pi_e(x) = x \cdot e$ and call $T_e : \mathbb{R} \to \mathbb{R}$ the monotone optimal transport between the two image measures $(\pi_e)_\# \rho$ and $(\pi_e)_\# \nu$. Then we define $\mathbf{v}_e(x) := (T_e(\pi_e(x)) - \pi_e(x))e$ and $\mathbf{v}(x) = ⨍_{S^{d-1}} \mathbf{v}_e(x)\, d\mathcal{H}^{d-1}(e)$, where $\mathcal{H}^{d-1}$ is the uniform measure on the sphere.

An old idea by Bernot, motivated by the approximation of well-behaved transport maps from $\rho_0$ to $\nu$, was to iterate a construction with a time step $\tau > 0$, but a natural continuous counterpart exists: we consider $\nu$ as a fixed target and we define, for every $\rho \in \mathcal{P}(\mathbb{R}^d)$, the vector field $\mathbf{v}[\rho]$ as the one computed above, and we solve the equation

$$\partial_t \rho_t + \nabla \cdot (\rho_t \mathbf{v}[\rho_t]) = 0.$$

It happens that this equation has a gradient flow structure: it is indeed the gradient flow of the functional $F(\rho) := \frac{1}{2} \mathrm{S}W_2^2(\rho, \nu)$, where $\mathrm{S}W_2$ is the so-called Sliced Wasserstein distance (see [69, 17]): given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, set

$$\mathrm{S}W_2(\mu, \nu) := \left( ⨍_{S^{d-1}} W_2^2((\pi_e)_\# \mu, (\pi_e)_\# \nu)\, d\mathcal{H}^{d-1}(e) \right)^{1/2}.$$

Existence and estimates on the solution of this equation are proven in [17], and the nonlinearity of $\mathbf{v}_{(\rho)}$ is quite easy to deal with. On the other hand, many, natural and useful, questions are still open, starting from the most crucial one: is it true that $\rho_t \rightharpoonup \nu$ as $t \to \infty$? This is the first and main question, and a partial answer is contained in [17], but requires to prove that the limit measure $\rho_\infty$ has strictly positive density a.e. If on the one hand this requires an assumption on $\nu$ (which is natural, and some counter-examples are known in the discrete case), it also requires to prove lower bounds on the densities, which could be attacked via the JKO scheme. Another question is whether we can define (at least under regularity assumptions on the initial data) the flow of the vector field $\mathbf{v}[\rho_t]$, and what is the limit of this flow as $t \to \infty$ (it will be a transport map between $\rho_0$ and $\nu$, which are its properties?).

Note that a recent variant has been introduced in [57] adding an entropy term to $F$ (i.e. a Laplacian in the equation). This has a meaning in data-generation models: we observe some sample points (providing an atomic measure $\nu$) drawn from a distribution $\nu_0$ that we do not know, but which we expect to be close to $\nu$ and smoother (even if possibly concentrated very close to a submanifold of the ambient space, standing for a feasible subset); we want an evolution model which at the same time finds an approximation of $\nu_0$, converges as $t \to \infty$ to it, and provides, via a flow, a sort of projection map onto the support of $\nu_0$ or the submanifold which it represents. For this modified equation the convergence in long time to the minimizer of the modified functional is easier to prove, but many questions on the reconstruction of individual trajectories in the diffusive case have still to be well-understood.

The work on this task will be mainly done in collaboration with N. Bonneel and J. Digne.

# B Methodology

## B.1 Mathematical tools

As for many mathematical projects, the exact methods to use on these problems will be adapted to partial results and are difficult to know in advance. It is anyhow possible to detail some tools, and in particular the most recent ones, which will for sure be used

**Flow interchange** The flow interchange technique, introduced in [58] allows to consider the rate of dissipation of a functional $G$ along the iterations of the JKO scheme for a functional $F$. To make an example (computations can be adapted to more general situations), if $G(\rho) = \int g(\rho(x))dx$ and $F(\rho) = \int f(\rho(x))dx$ and $\rho \in \mathrm{Prox}_\tau^F(\nu)$ and $f$ is convex and $g$ satisfies McCann's conditions for displacement convexity ([61]) then we have

$$G(\rho) \leq G(\nu) - \tau \int f''(\rho) g''(\rho) |\nabla \rho|^2 \leq G(\nu).$$

This allows at the same time to prove that $G$ decreases along iterations and to obtain an integral bound (in space-time) on a higher-order quantity, a squared $H^1$ norm. The same computation is easy to obtain in the continuous setting: if $\rho$ solves $\partial_t\rho = \nabla\cdot(\rho\nabla(f'(\rho)))$, then an integration by parts provides the equality

$$\frac{d}{dt}\int g(\rho_t) = -\int g''(\rho)f''(\rho)|\nabla\rho|^2.$$

By the way, this continuous computation does not require at all the displacement convexity of $G$, while this is required in the discrete setting as one needs to replace a derivative along the geodesic with the increment $G(\nu) - G(\rho)$. This is just a technical point, and similar results could be obtained for $\lambda-$convex functionals, with an extra error term of the order of $|\lambda|\tau^2$, thus disappearing at the limit $\tau \to 0$. An interesting question which has not been developed enough yet is to exploit this fact by approximating general non-displacement convex functionals with functionals which are $\lambda-$convex for $\lambda = -O(\tau^{-1/2})$, which could be an extra tool in the analysis.

**Monge-Ampère equation** The monge-Ampère equation $\det(I - D^2\varphi) = \rho/\nu(x - \nabla\varphi(x))$ solved by the Kantorovich potential $\varphi$ in the transport from $\rho = \mathrm{Prox}^F_\tau(\nu)$ to $\nu$, which is at the same time related to $\rho$ via the equality $\frac{\varphi}{\tau} + \frac{\delta F}{\delta\rho} = const$, has been a useful tool to establish some $L^\infty$ bounds on $\rho$ or $\nabla\rho$, by looking at the maximum/minimum points for functions of $\varphi$ or $\nabla\varphi$ (see Chapter 7 in [70] or [37]). We refer for instance to [39] for the theory about this equation which is crucail in optimal transportation: since the general theory for this kind of fully nonlinear PDE goes much beyond the simple application of maximum principle and Bernstein estimates, we can expect to obtain stronger results if employeing stronger tools from this theory.

**Five-gradients inequality** The so-called *Five Gradients Inequality* has been developed by the PI and some coauthors in [32] to prove BV bounds for the solutions of variational problems involving the Wasserstein distance. This inequality states that, whenever $\rho$ and $\nu$ are smooth enough, if $\phi$ and $\psi$ are the corresponding Kantorovich potentials and $H$ is a convex function, then

$$\int \nabla H(\nabla\phi)\cdot\nabla\rho + \int \nabla H(\nabla\psi)\cdot\nabla\nu \;\geq\; 0.$$

This is of much interest when $\rho \in \mathrm{Prox}^F_\tau(\nu)$ for some functional $F$, since in this case $\nabla\varphi$ and $\nabla\rho$ are strongly related to each other, and allows to obtain bounds on $|\nabla\rho|$. In particular, $H(z) = |z|$ allows to obtain BV bounds when $F$ is of the form $\int f(\rho(x))dx$ for $f$ convex, and a suitable use of $H(z) = |z|^p$ helps in some sort of $W^{1,p}$ bounds, as explained in [34]. It is also possible to be explicit about the remainder term, thus coupling a uniform (in time) first-order bound (in space) with an integral bound on second-order quantities.

**Regularity by duality** Another tool for regularity is the method called *regularity via duality*, first found in [19], then used in [20, 68] for MFG, and presented in its full generality in [73]. This strategy allows to obtain Sobolev (or, in some cases, BV or fractional Sobolev) regularity for the solutions of convex variational problems, by estimating quantities such as $\|u_h - u\|$ (where $u_h$ is a translation of $u$ and the norm is often an $L^2$ norm) in terms of the duality gap and of how much $u_h$ is not optimal. This has given original results in the framework of MFG, but has never been explicitly used in gradient flows. Yet, many among the optimazion problems used in the JKO scheme are convex optimization probelms, and this strategy could also be used, both to obtain regularity and quantitative results. By the way, even the use of the dual problem of a JKO step is not so common, but starts to be taken into consideration, see for instance [49].

$L^\infty$ **bounds in OT** A paper by Bouchitté, Jimenez and Rajesh ([18]) proved the following useful inequality: given two probability densities $\rho, \nu$ on a convex set, denoting $T$ the optimal transport map from $\rho$ to $\nu$, if $\rho \geq \alpha$, then we have $\alpha\|T - id\|^{d+2}_{L^\infty} \leq cW_2^2(\rho,\nu)$ for a constant $c = c(d)$ only depending on the dimension $d$. This is extremely useful when estimating, for numerical purposes, the maximal displacement in a JKO step, and perfectly fits the framework of crowd motion modelling, where we can assume $\rho = 1$ on the region of interest. On the other hand, variant of this inequality need to be proven in order to deal with other cases (in particular if we cannot guarantee uniform lower bounds on the densities), in order to apply it to other JKO schemes, such as the porous-medium case. Let us mention that more refined local inequalities in the same spirit have been used by Goldman-Otto in [47] and then generalized to other costs in [48]

**Network simplex for separable costs** The optimal transport problem between two discrete measures with $N$ atoms each can be considered as a network linear programming problem, with a network containing $N \times N$ edges (one per each possible origin-destination pair). Yet, when the cost is the quadratic cost, which can be decomposed into the sum of $d$ costs depending each on one variable $x_1, \dots x_d$ only, if the $N$ points are on a regular grid, with $N = n^d$, then an idea presented in [4] allows to strongly reduce the number of edges: one has to insert $d-1$ intermediate layers of $N$ points, and connect each layer to the next one only accepting movement on a given direction. In this way the total number of edges to be used is $dn^{d+1}$, which is in general much smaller than $n^{2d}$. For applications to PDEs (and to image processing, as it is done in [4]) the grid structure is a natural assumption and allows to reduce the computational cost of the linear programming problem. When passing to a JKO scheme, this improvement can be easily applied to the crowd motion case, where the functional $F$ is linear with a linear costraint, while it requires some more adaptations for general $F$. Moreover, if it is possible to provide $L^\infty$ bounds on the displacement, it is possible to reduce even further the number of edges, since if the displacement is bounded by $\varepsilon(\tau)$, then the nuber of edges in each transport problem becomes of the order of $\varepsilon(\tau)n^{d+1}$, which is useful considering that this optimization has to be iterated $\tau^{-1}$ times.

**Recent results on the mesa problem** When considering porous-medium–type reaction diffusion equations, with an exponent $m$, in the limit $m \to \infty$ many authors, starting from [36], have observed the formation of a plateau-like region, which they refer to as "mesa", of nearly constant density $\rho = 1$. This limit equation is, up to the choice of the lower-order terms, the same as in Hele-Shaw models or crowd motion models. Different community have studied these equations and putting together the different technique is something which is starting to be done in a very fruitful way. For instance, [29] and [30], motivated by tumor-growth applications, prove improved limit therorems and uniform estimates independent of $m$ and new properties on the solution such as an $L^4$ bound on the gradient of the pressure. In what concerns uniqueness, the classical theory from the Hele-Shaw community ignored possible advection by a drift, which is on the contrary the most interesting case when the mass is preserved, as it happens in the JKO scheme; the first results in this direction were in [33], but only covered the case where either the drift is smooth or diffusion is added but, very recently, the preprint [42] came out with a different technique to prove uniqueness for drifts which are only Sobolev, an assumption which is reasonably sharp in view of the Di-Perna Lions theory in the absence of density constraints.

## B.2 Management of the team, of the work, and of the competences

In order to attack all the taks included in the research project a team composed of mathematicians and computer scientists will be created. This team will be fully based in the Lyon 1 campus and will include permanent staff (researchers who will be officially associated with the project, and accepted to devote a part of their working time to EYAWKAJKOS), local collaborators (other colleagues or students already present in Lyon working on related topics) and young members to be hired by the project.

The permament staff of EYAWKAJKOS will be composed of

- the PI, professor at the ICJ laboratory (mathematics), in the modeling and computing group. He will devote 55% of his time to the project.
- Ivan Gentil, also professor at ICJ, but in the PDE and Analysis group, who will be mainly involved in task $(\tau, x) - 3$ but whose competences will more generally help in both the $(\tau, x)$ and the $(t, x)$ parts. He will devote 15% of his time to the project.
- Aymeric Baradat, junior CNRS researcher at ICJ in the PDE and Analysis group, who will be mainly involved in task $(t, x) - 3$ but whose competences will also more generally help in both the $(\tau, x)$ and the $(t, x)$ parts. He will devote 20% of his time to the project.
- Nicolas Bonneel, junior CNRS researcher at the LIRIS lab (computer science), who will be mainly involved in tasks $(\tau, h) - 2$ and $(t, x) - 4$. He will devote 15% of his time to the project.

The local collaborators will be

- Frédéric Lagoutière, professor at ICJ in the modeling and computing group, who will help in the numerical analysis (task $(\tau,h)-3$) and in aggregation-diffusion equtions.
- Sébastien Tran Tien, PhD student of F. Lagoutière (PhD started in 2020), who could help in the same areas.
- Annette Dumas, PhD student of the PI (starting in 2021), who works on a different but related topic (Mean Field Games) and could be involved in task $(t,x)-1$ or in the use of time-discretization techniques common to MFG and gradient flows.
- Julie Digne, junior CNRS researcher at LIRIS, who will be involved in task $(t,x)-4$.

EYAWKAJKOS will also hire 6 young members throughout the duration fo the project, one PhD student and one postdoc per part. A completely tentative hiring plan would include

- a three-years postdoc working on the $(\tau,x)$ part and in particular attacking task $(\tau,x)-2$ (this task requires higher competences than what a PhD student could offer) and helping in task $(\tau,x)-1$
- a PhD student working on the $(\tau,x)$ part and in particular attacking task $(\tau,x)-3$, most likely co-supervised with I. Gentil.
- a three-years postdoc working on the $(\tau,h)$ part and in particular attacking tasks $(\tau,h)-1$ and $(\tau,h)-3$, with competences from numerical analysis and/or optimization.
- a PhD student working on the $(\tau,h)$ part and in particular attacking task $(\tau,h)-2$, most likely co-supervised with N. Bonneel. Note that this PhD thesis could start before the corresponding postdoc, even if it could profit of the results from the postdoc itself (bounds on the displacement, proof of convergence) beause of the time it will be needed to train the student in efficient coding. The consequences of the theoretical results coming from tasks $(\tau,h)-1$ and $(\tau,h)-3$ could be easily inserted in the code at a later stage.
- a three-years postdoc working on the $(t,x)$ part and in particular attacking tasks $(t,x)-2$ and $(t,x)-4$ with possible collaborations with N. Bonneel, J. Digne and other colleagues.
- a PhD student working on cross-diffusion in the $(t,x)$ part and in particular attacking task $(t,x)-3$, most likely co-supervised with A. Baradat.

Finally, as it was mentioned in the previous section, exterior collaborators will also be involved, such as G. Toshpulatov, N. David, M. Schmidtchen, R. Ducasse, H. Yoldaş… Previous collaborators of the PI on related topics (A. Mészáros, S. Di Marino, M. Iacobelli…) are also likely to be involved in joint researches.

A tentative schedule for the workplan is in the following chart: (of course the work by PhD students and postdocs is always accompanied by parallel, preliminary, and subsequent work by the PI and exterior collaborations; "Before EYAWKAJKOS" stands for preliminary results or work-in-progress, not funded by ERC).

<table>
<tr><th>Task</th><th>Before EYAWKAJKOS</th><th>Year 1</th><th>Year 2</th><th>Year 3</th><th>Year 4</th><th>Year 5</th></tr>
<tr><td rowspan="2">$(\tau,x)-1$</td><td colspan="2">Work by PI & collaborators</td><td colspan="4"></td></tr>
<tr><td></td><td colspan="3">Post-doc $(\tau,x)$</td><td colspan="2"></td></tr>
<tr><td>$(\tau,x)-2$</td><td></td><td colspan="3">Post-doc $(\tau,x)$</td><td colspan="2"></td></tr>
<tr><td>$(\tau,x)-3$</td><td colspan="3"></td><td colspan="3">PhD $(\tau,x)$</td></tr>
<tr><td>$(\tau,h)-1$</td><td colspan="3"></td><td colspan="3">Post-doc $(\tau,h)$</td></tr>
<tr><td>$(\tau,h)-2$</td><td></td><td colspan="3">PhD $(\tau,h)$</td><td colspan="2"></td></tr>
<tr><td>$(\tau,h)-3$</td><td colspan="3"></td><td colspan="3">Post-doc $(\tau,h)$</td></tr>
<tr><td>$(t,x)-1$</td><td colspan="2">Work by PI & collaborators</td><td colspan="4"></td></tr>
<tr><td>$(t,x)-2$</td><td></td><td colspan="3">Post-doc $(t,x)$</td><td colspan="2"></td></tr>
<tr><td>$(t,x)-3$</td><td colspan="2">Work by PI & collaborators</td><td colspan="3">PhD $(t,x)$</td><td></td></tr>
<tr><td rowspan="2">$(t,x)-4$</td><td colspan="2">Work by PI & collaborators</td><td colspan="4"></td></tr>
<tr><td></td><td colspan="3">Post-doc $(t,x)$</td><td colspan="2"></td></tr>
</table>

## References


[1] M. Agueh Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory *Adv. Differential Equations* 10(3): 309-360 (2005).

[2] L. Ambrosio, Movimenti minimizzanti, *Rend. Accad. Naz. Sci. XL Mem. Mat. Sci. Fis. Natur.* 113, 191–246, 1995.

[3] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the spaces of probability measures.* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.

[4] G.Auricchio, F. Bassetti, S. Gualandi, M.Veneroni Computing Kantorovich-Wasserstein Distances on $d-$dimensional histograms using $(d+1)-$partite graphs, *(NeurIPS 2018)*

[5] D. Bakry, I. Gentil and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, Springer, 2014

[6] A. B. T. Barbaro, N. Rodriguez, H. Yoldaş and N. Zamponi, Analysis of a cross-diffusion model for rival gangs interaction in a city, To appear in *Communications in Mathematical Sciences*.

[7] J. Becker, G. Grün, The thin-film equation: recent advances and some new perspectives. *J. Phys.: Condens. Matter* 17 (2005) 291–307.

[8] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84 (2000), 375–393.

[9] J.-D. Benamou and G. Carlier, Augmented Lagrangian Methods for Transport Optimization, Mean Field Games and Degenerate Elliptic Equations, *J. Opt. Theor. Appl.*, 167, pages1?26 (2015)

[10] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, G. Peyré Iterative Bregman projections for regularized transportation problems SIAM Journal on Scientific Computing 37 (2), 1111–1138, 2015.

[11] J.-D. Benamou, G. Carlier and M. Laborde An augmented Lagrangian approach to Wasserstein gradient flows and applications, *ESAIM: Proc.* 54, 1–17, 2016.

[12] J.-D. Benamou, G. Carlier, Q. Mérigot and É. Oudet Discretization of functionals involving the Monge-Ampère operator, *Numerische Mathematik*, 134(3), 611–636, 2016.

[13] M. Bertsch, R. Dal Passo, and M. Mimura. A free boundary problem arising in a simplified tumour growth model of contact inhibition. *Interfaces and Free Boundaries*, 12(2):235–250, 2010.

[14] A. Blanchet, V. Calvez and J.A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model, *SIAM J. Numer. Anal.* 46 (2), 691–721, 2008.

[15] A. Blanchet, J.-A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot and S. Lisini, A Hybrid Variational Principle for the Keller-Segel System in $\mathbb{R}^2$, *ESAIM M2AN*, Vol. 49, no 6, 1553–1576, 2015.

[16] N. Bonneel, M. van de Panne, S. Paris, W.Heidrich Displacement Interpolation Using Lagrangian Mass Transport, *Proceedings of ACM SIGGRAPH Asia* 2011, 30 (6), Article no 158.

[17] N. Bonnotte *Unidimensional and Evolution Methods for Optimal Transportation*, PhD thesis, U. Paris-Sud, 2013.

[18] G. Bouchitte, C. Jimenez, and R. Mahadevan, A new $L^\infty$ estimate in optimal mass transport. *Proc. Amer. Math. Soc.*, 135(11):3525?3535, 2007.

[19] Y. Brenier: The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Mat. Soc.*, 2, 225–255, 1989.

[20] P. Cardaliaguet, A. R. Mészáros, F. Santambrogio First order Mean Field Games with density constraints: pressure equals price., *SIAM J. Contr. Opt.*, 54(5), 2672-2709, 2016.

[21] G. Carlier, C. Poon, On the total variation Wasserstein gradient flow and the TV-JKO scheme, *Esaim-Control Optimisation and Calculus of Variations*, Vol 25, nr 42, 2019.

[22] J.-A. Carrillo, R.J. McCann and C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates, *Revista Matemática Iberoamericana* 19, 1–48, 2003.

[23] J.-A. Carrillo, R.J. McCann and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media, *Arch. Rat. Mech. An.* 179, 217–263, 2006.

[24] J.-A. Carrillo, S. Fagioli, F. Santambrogio, M. Schmidtchen Splitting schemes and segregation in reaction cross-diffusion systems *SIAM Journal on Mathematical Analysis* **50-5** (2018), 5695-5718

[25] J.-A. Carrillo, D. Slepcev, Example of a displacement convex functional of first order *Calculus of Variations and Partial Differential Equations* 36(4):547-564, 2008

[26] A. Chambolle and T. Laux Mullins-Sekerka as the Wasserstein flow of the perimeter *Proc. Amer. Math. Soc.* 149 (2021), 2943-2956

[27] D. Córdoba, D. Faraco and F. Gancedo, Lack of uniqueness for weak solutions of the incompressible porous media equation, *Arch. Ration. Mech. Anal.* 200 (2011), 725-746.

[28] M. Cuturi Sinkhorn Distances: Lightspeed Computation of Optimal Transport, *Proceedings of the 26th International Conference on Advances in Neural Information Processing Systems*, NIPS 26

[29] N. David, B. Perthame Free boundary limit of tumor growth model with nutrient. Preprint.

[30] N. David, M. Schmidtchen On the Incompressible Limit for a Tumour Growth Model incorporating Convective Effects. Preprint

[31] E. De Giorgi, New problems on minimizing movements, *Boundary Value Problems for PDE and Applications, C. Baiocchi and J. L. Lions eds.* (Masson, 1993) 81–98.

[32] G. De Philippis, A. R. Mészáros, F. Santambrogio and B. Velichkov BV estimates in optimal transportation and applications, *Arch. Rati. Mech. An.* 219 (2), 829-860, 2016.

[33] S. Di Marino and A. R. Mészáros Uniqueness issues for evolution equations with density constraints, *Math. Models Methods Appl. Sci.*, Vol. 26, No. 09, 1761–1783(2016)

[34] S. Di Marino F. Santambrogio JKO estimates in linear and non-linear Fokker-Planck equations, and Keller-Segel: $L^p$ and Sobolev bounds. Preprint.

[35] J. Dolbeault, B. Nazaret, G. Savaré A new class of transport distances between measures *Calculus of Variations and Partial Differential Equations* 34 (2), 193-231, 2009

[36] C.M. Elliott, M.A. Herrero, J.R. King, and J.R. Ockendon The mesa problem: Diffusion patterns for $u_t = \nabla\cdot(u^m\nabla u)$ as $m\to\infty$. *IMA journal of applied mathematics,* 37(2):147-154, 1986

[37] V. Ferrari, F. Santambrogio Lipschitz estimates on the JKO scheme for the Fokker–Planck equation on bounded convex domains, *Appli. Math. Letters*, Vol 112, 2021, 106806

[38] C. Förster and L. Székelyhidi. Piecewise constant subsolutions for the muskat problem. *Communications in Mathematical Physics*, 363(3) :1051–1080, 2018.

[39] A. Figalli, *The Monge-Ampère Equation and Its Applications*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017.

[40] M. E. Gurtin and A. Pipkin. A note on interacting populations that disperse to avoid crowding. *Quarterly of Applied Mathematics,* pages 87–94, 1984.

[41] M. Iacobelli, F. S. Patacchini, F. Santambrogio Weighted ultrafast diffusion equations: from well-posedness to long-time behaviour, *Arch. Rati. Mech. An.*, 232 (3), 1165-1206, 2019.

[42] N. Igbida $L^1$ Theory for reaction-diffusion Hele-Shaw flow with linear drift, preprint

[43] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker-Planck equation, *SIAM J. Math. Anal.* 29 (1), 1–17, 1998.

[44] L. Kantorovich, On the transfer of masses. *Dokl. Acad. Nauk. USSR*, 37, 7–8, 1942.

[45] E. F. Keller and L. A. Segel, Initiation of slide mold aggregation viewed as an instability, *J. Theor. Biol.,* 26, 399–415, 1970.

[46] E. F. Keller and L. A. Segel, Model for chemotaxis, *J. Theor. Biol.*, 30, 225–234, 1971.

[47] M. Goldman and F. Otto, A variational proof of partial regularity for optimal transportation maps. *Ann. Scient. Ec. Norm. Sup.*, 53:1209?1233, 2020

[48] C. Gutierrez, A. Montanari $L^\infty$-estimates in optimal transport for non quadratic costs, preprint, 2021

[49] M. Jacobs, I. Kim and J. Tong, $L^1$ contraction in Optimal Transport, to appear in *Ann. Sc. Norm. Super. Pisa Cl. Sci.*

[50] I. Kim, M. Jacobs and A. Meszaros, Weak solutions to the Muskat problem with surface tension via optimal transport *Arch. Ration. Mech. Anal.,* 239 (2021), 389?430

[51] I. Kim, A. R. Mészáros On nonlinear cross-diffusion systems: an optimal transport approach *Calculus of Variations and Partial Differential Equations* **57-3** (2018) 1-40

[52] J. Kitagawa, Q. Mérigot and B. Thibert, Convergence of a Newton algorithm for semi-discrete optimal transport, *J. Eur. Math. Soc.* 21 , no 9, pages 2603–2651, 2019.

[53] H. Lavenant, F. Santambrogio, Optimal density evolution with congestion: $L^\infty$ bounds via flow interchange techniques and applications to variational Mean Field Games, *Comm. Par. Diff. Eq.* Vol 43 (12), 1761-1802, 2018

[54] P. Lee, On the Jordan-Kinderlehrer-Otto scheme, *J. Math. An. Appl.* Vol 429, Issue 1, 131-142, 2015.

[55] P. Lee, A Harnack inequality for the Jordan-Kinderlehrer-Otto scheme, *Journal of Evolution Equations* volume 18, pages 143–152 (2018)

[56] G. Legendre and G. Turinici, Second order in time schemes for gradient flows in Wasserstein and geodesic metric spaces, *Compte Rendus Mathématiques*, Vol 355 no 3, pages 345-353, 2017

[57] A. Liutkus, U. Şimşekli, S. Majewski, A. Durmus, F.-R. Stöter Sliced-Wasserstein Flows: Nonparametric Generative Modeling via Optimal Transport and Diffusions, *International Conference on Machine Learning (ICML) 2019*

[58] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352– 1397, 2009.

[59] D. Matthes, S. Plazotta A variational formulation of the BDF2 method for metric gradient flows, *ESAIM: M2AN* 53 (2019) 145–172.

[60] B. Maury, A. Roudneff-Chupin and F. Santambrogio A macroscopic crowd motion model of gradient flow type, *Math. Models and Methods in Appl. Sciences* 20 (10), 1787–1821, 2010.

[61] R. J. McCann A convexity principle for interacting gases. *Adv. Math.* 128 (1), 153–159, 1997.

[62] Q. Mérigot A multiscale approach to optimal transport. *Computer Graphics Forum* 30 (5) 1583–1592, 2011.

[63] A. R. Mészáros and F. Santambrogio Advection-diffusion equations with density constraints, *Analysis & PDE* 9-3 (2016), 615–644.

[64] G. Monge, Mémoire sur la théorie des déblais et des remblais, *Histoire de l'Académie Royale des Sciences de Paris*, 666–704, 1781.

[65] M. Muskat. Two fluid systems in porous media. The encroachment of water into an oil sand. *Physics,* 5(9) :250–264, 1934.

[66] F. Otto, The geometry of dissipative evolution equations: The porous medium equation, *Comm. Partial Differential Equations*, 26, 101–174, 2011.

[67] F. Pitié, A. C. Kokaram and R. Dahyot. Automated colour grading using colour distribution transfer. *Comp. Vis. Image Understanding,* 107 (1–2), 123–137, 2007.

[68] A. Prosinski, F. Santambrogio Global-in-time regularity via duality for congestion-penalized Mean Field Games, *Stochastics*, Vol 89, 2017 - Issue 6-7 : Proceedings of the Hammamet Conference, 19-23 October 2015, pages 923–942.

[69] J. Rabin, G. Peyré, J. Delon and M. Bernot. Wasserstein Barycenter and Its Application to Texture Mixing. In *Scale Space and Variational Methods in Computer Vision*, edited by A. M. Bruckstein, B. M. Haar Romeny, A.M. Bronstein, and M. M. Bronstein, Lecture Notes in Computer Science, vol. 6667, 435–446, Springer Berlin Heidelberg, 2012.

[70] F. Santambrogio *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and Their Applications no 87, Birkhäuser Basel (2015).

[71] F. Santambrogio, A Modest Proposal for MFG with Density Constraints, proceedings of the conference *Mean Field Games and related Topics, Roma 1* (2011), published in *Net. Het. Media*, 7(2), 2012, 337-347.

[72] F. Santambrogio {Euclidean, Metric, and Wasserstein} Gradient Flows: an overview, *Bull. Math. Sci.*, Vol 7 (1), 87–154, 2017.

[73] F. Santambrogio Regularity via duality in calculus of variations and degenerate elliptic PDEs, *Journal of Mathematical Analysis and Applications* Vol 457 (2), 1649–1674, 2018

[74] L. Székelyhidi Relaxation of the incompressible porous media equation *Annales scientifiques de l'École Normale Supérieure*, Série 4, Tome 45 (2012) no. 3, pp. 491-509.

[75] J. L Vázquez, *The porous medium equation. Mathematical theory*, The Clarendon Press, Oxford University Press, (2007).

[76] C. Villani *Topics in Optimal Transportation*. Graduate Studies in Mathematics, AMS, 2003.

[77] C. Villani, *Optimal transport: Old and New*, Springer Verlag, 2008.

[78] S. Zugmeyer Entropy flows and functional inequalities in convex sets, preprint.