

New estimates on the regularity of the pressure in density-constrained Mean Field Games

Hugo Lavenant, Filippo Santambrogio

▶ **To cite this version:**

Hugo Lavenant, Filippo Santambrogio. New estimates on the regularity of the pressure in density-constrained Mean Field Games. Journal of the London Mathematical Society, In press, 10.1112/jlms.12245 . hal-01930480v2

HAL Id: hal-01930480
https://hal.science/hal-01930480v2

Submitted on 13 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

New estimates on the regularity of the pressure in density-constrained Mean Field Games

Hugo Lavenant*, Filippo Santambrogio†

May 13, 2019

Abstract

We consider variational Mean Field Games endowed with a constraint on the maximal density of the distribution of players. Minimizers of the variational formulation are equilibria for a game where both the running cost and the final cost of each player is augmented by a pressure effect, i.e. a positive cost concentrated on the set where the density saturates the constraint. Yet, this pressure is a priori only a measure and regularity is needed to give a precise meaning to its integral on trajectories. We improve, in the limited case where the Hamiltonian is quadratic, which allows to use optimal transport techniques after time-discretization, the results obtained in a paper of the second author with Cardaliaguet and Mészáros. We prove H^1 and L^∞ regularity under very mild assumptions on the data, and explain the consequences for the MFG, in terms of the value function and of the Lagrangian equilibrium formulation.

1 Introduction

The main motivation of this paper is to provide improved regularity estimates about the pressure arising in a class of variational Mean Field Games (MFG) where the interaction between players is due to a density constraint $\rho \leqslant 1$ instead of arising from a penalization on the density itself. For the whole theory of Mean Field Games, a recent hot topic in applied mathematics introduced by Lasry and Lions in [17, 18] and, independently, by Caines, Huang and Malamé in [14], we refer to the lecture notes by Cardaliaguet [8] and to the video-recorded lectures by Lions, [19]. This theory is concerned with the behavior of a continuous family of rational agents, who need to choose a strategy on how to move in a domain where they meet other agents and their cost is affected by their presence. The goal is to study the Nash equilibria and characterize them in terms of PDEs. Most models assume stochastic effects on the trajectory of the agents, and the corresponding PDEs include diffusion terms which make the solution smooth and simplify the analysis, besides being reasonable from the modeling point of view. Analytically, the most difficult case consists in problems where the interaction between players is local (i.e. the cost at point x and time t depends on the value of the density $\rho_t(x)$, without averaging it in a neighborhood) and no diffusion is present. This case is essentially attacked when the game is of variational origin, i.e. it is a potential game, and the equilibrium condition arises as an optimality condition for an optimization problem in the class of density evolutions. For local potential MFG, we refer to [9, 10] and to the survey [6].

A particular potential MFG has been studied in [11], where the density ρ is constrained to be below a certain threshold, which represents a capacity constraint of the transportation network or of the medium where agents move. In this case a pressure appears: according to what we know from fluid mechanics, the pressure is a scalar field, vanishing where the density does not saturate the constraint, and its gradient affects the accelaration of the particles. In terms of the equilibrium problem, the pressure is a price to pay to pass through saturated regions. This means that agents compute their total cost by integrating the pressure along their trajectory, but this generates some issues about its regularity, since it is a priori not well-defined on negligible sets such as curves.

Inspired by the considerations in [2, 3], in [11] two facts were proven. First, there is a way to define a precise representative $\hat{p}$ of the pressure p such that almost every trajectory followed by the agents optimize a cost involving

*Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, cedex FRANCE, hugo.lavenant@math.u-psud.fr

†Institut Camille Jordan, Université Claude Bernard - Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex FRANCE, santambrogio@math.univ-lyon1.fr

$\omega \mapsto \int_0^1 \hat{p}(t,\omega(t))dt$ among curves ω such that $\int_0^1 (Mp)(t,\omega(t))dt < +\infty$ (Mp is the maximal function of p; such an integrability condition is required since $\hat{p}$ is defined as a limit of averages on balls and this is necessary to pass to the limit the averaged estimations). Second, the pressure p belongs, under some assumptions on the data, to $L^2_{t,loc}BV_x$ (same regularity as the one obtained in [2]), which guarantees $L^{1+\varepsilon}$ summability. This implies, by well-known harmonic analysis results, that Mp is also summable to the same power, and guarantees that the class of curves satisfying the integrability condition on Mp is large enough.

In the present paper this result is improved in several different ways, but we have to pay a price: we need to specialize to the case where the Hamiltonian is quadratic. This means that the cost payed by the agents following a trajectory ω is of the form

$$\int_0^1 \left(\frac{1}{2}|\dot{\omega}(t)|^2 + \tilde{V}(t,\omega(t))\right) dt + \tilde{\Psi}(\omega(1)),$$

where $\tilde{V} = V + p$ is given by a running cost plus the pressure, and $\tilde{\Psi}$ is given by the final cost, augmented by a final pressure P_1. Note that we cannot exclude the presence of a pressure effect concentrated at the final time (the pressure should be considered as a measure with a singular part concentrated on $t = 1$); we know that it is indeed possible to observe it in some particular cases, and this is indeed the role played by P_1 (see Sections 1.1 and 2.3).

In particular, the dependence in the velocity is quadratic, instead of using more general convex, and possibly space-dependent, functions $L(\omega(t),\dot{\omega}(t))$. This allows to use properties of optimal transport and of the Wasserstein distance W_2. In this precise setting, the results in [11] provided $L^2_{t,loc}BV_x$ regularity for the pressure if both the running cost V and the final cost Ψ were $C^{1,1}$; here (see Theorem 2.8 for the precise statement) we manage to obtain $L^\infty_t H^1_x$ under the only assumption $V \in H^1$. Moreover, we obtain an inequality on the Laplacian of the pressure which, thanks to quite standard Moser iterations, provides $p \in L^\infty$ as soon as $V \in W^{1,q}$ for $q > d$ (d being the space dimension). This boundedness is very important since it implies that the integrability condition on Mp is always satisfied, and the class of competitor curves in the equilibrium condition includes now all curves. Finally, similar result are obtained for the singular pressure P_1 at time $t = 1$, while [11] did not adress its behavior. As a last remark, the achievements of the present paper are global on a general bounded convex space domain, while the techniques in [11] could not easily be adapted to domains with boundary (they were presented in the torus; the technique can be adapted to other domains but only obtaining local results).

The main tool to obtain the desired estimate is an inequality on $\Delta(p + V)$, valid at any time t on the saturated region where $p > 0$, and all the estimates derive from this one (analogously, we also use a similar inequality on $\Delta(P_1 + \Psi)$). We first provide (in Section 1.1) an heuristic derivation of this inequality, based on the use of the convective derivative along the flow. In order to make the proof rigorous, the strategy is very much inspired from our previous work [20], based on time-discretization, even if the inequalities we use are not the same.

As we said at the beginning, MFG with density constraints are the motivation for this work and they are the setting where these estimates show better their potential of applications. However, we believe that the technique deserves attention both for its remarkable simplicity, and for the possibility of being applied to other settings.

A similar setting can be found in the variational formulation by Brenier of the incompressible Euler equation. Yet, the sharp regularity of the pressure is in such a setting an open problem, as semi-concavity is a reasonable conjecture advanced by Brenier, but the current achievements do not go beyond the $L^2_t BV_x$ result mentioned above. Yet, due to the multiphasic nature of the problem formulated by Brenier, it is in general not possible to translate all the available techniques into such a more complicated setting (see for instance [16] where the time-convexity of the entropy is proven, but, differently from [20], the same cannot be obtained for other internal energies; analogously, the same results of [16] are also recovered in [5], and the same algebraic obstruction prevents from generalizing the result to more general energies). On the other hand, the works on density-constrained MFG (including a first attempt, with a non-variational model, in [24]) were inspired by previous works of the second author on crowd motion formulated as a gradient flow with density constraints (see [22] and [25]), and the present technique seems possible to be applied to such a first-order (in time) setting. Indeed, as explained in the core of the article, the technique of proof for the regularity of the final pressure P_1 is performed exactly as if we had a JKO scheme for a gradient flow (see [15]).

From the point of view of generalizations, in particular to other Hamiltonians, or to quadratic Hamiltonians on manifold (which could require to use curvature assumptions on the manifold), it seems that the main point is the computation of the Laplacian of the Hamilton-Jacobi equation: once suitable inequalities are available on it, the approach could be generalized.

The paper is organized as follows. In the rest of Section 1 we present first the heuristic derivation of our estimates and then some consequences, in the framework of MFG, on the regularity of the value function. Section 2 presents

useful preliminaries, then the context and the precise results we will prove, and finally the time-discrete approximation we choose. Section 3 contains the main estimates, divided into interior regularity for p and then "boundary" regularity for P_1. Section 4 shows how to translate the discrete estimates into continuous ones, by providing limit results both on the primal and on the dual problem.

1.1 Heuristic derivation of the estimates

We start with the MFG system, which can be obtained as a consequence of the primal-dual optimality conditions of a variational problem, see [11]. These conditions read, for functions depending on time $t \in [0,1]$ and space $x \in \Omega$,

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla \phi) & = 0, \\ -\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 & \leqslant P + V \quad \text{(with equality on } \{\rho > 0\}), \\ \rho(0,\cdot) & = \bar{\rho}_0, \\ \phi(1,\cdot) & \leqslant \Psi \quad \text{(with equality on } \{\rho_1 > 0\}), \end{cases}$$

where $P \geqslant 0$ is a measure concentrated on the set $\{\rho = 1\}$. As far as this heuristic justification is concerned, we will just look at the conditions which are satisfied on the support of ρ, where the inequalities become equalities. As we are interested in estimates on the pressure P, i.e. on the set $\{\rho = 1\}$, this is a legit restriction. As we will see later, the pressure P is a measure which can be decomposed into two parts: its restriction to $[0,1) \times \Omega$ is absolutely continuous w.r.t. the Lebesgue measure on $[0,1) \times \Omega$, and its density is denoted by p; on the other hand, there is also a part on $\{1\} \times \Omega$ which is singular, but absolutely continuous w.r.t. the Lebesgue measure on Ω, and its density is denoted by P_1. This second part represents a jump of the function ϕ at $t = 1$, which allows to re-write the system as follows.

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla \phi) & = 0, \\ -\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 & = p + V, \\ \rho(0,\cdot) & = \bar{\rho}_0, \\ \phi(1,\cdot) & = \Psi + P_1. \end{cases} \tag{1.1}$$

where the density ρ satisfies $\rho \leqslant 1$ everywhere and $p, P_1 \geqslant 0$ are strictly positive only on the regions where the constraint involving ρ is saturated, i.e. where $\rho = 1$ ($\rho_1 = 1$ in the case of P_1).

We denote by $D_t := \partial_t - \nabla\phi \cdot \nabla$ the convective derivative. The idea is to look at the quantity $-D_{tt}(\ln \rho)$. Indeed, the first equation of (1.1) can be rewritten $D_t(\ln \rho) = \Delta \phi$. On the other hand, taking the Laplacian of the second equation in (1.1), it is easy to get, dropping a positive term, $-D_t(\Delta \phi) \leqslant \Delta(p + V)$. Hence,

$$-D_{tt}(\ln \rho) \leqslant \Delta(p + V). \tag{1.2}$$

Notice that if $\rho(\bar{t}, \bar{x}) = 1$, then ρ is maximal at $(\bar{t}, \bar{x})$ hence $-D_{tt}(\ln \rho)(\bar{t}, \bar{x}) \geqslant 0$. On the other hand, if $\rho(\bar{t}, \bar{x}) < 1$ then $p(\bar{t}, \bar{x}) = 0$. In other words, p satisfies $\Delta p \geqslant -\Delta V$ on $\{p > 0\}$, which looks like an obstacle problem. Multiplying (1.2) by p, integrating w.r.t. space at a given instant in time and doing an integration by parts, for all t

$$\int_\Omega \nabla p(t,\cdot) \cdot \nabla (p(t,\cdot) + V) \leqslant \int_{\partial\Omega} p(t,\cdot)[\nabla (p(t,\cdot) + V) \cdot \mathbf{n}_\Omega], \tag{1.3}$$

where $\mathbf{n}_\Omega$ is the outward normal to Ω. As $\nabla(p + V)$ is the acceleration of the agents, who are constrained to stay in Ω, under the assumption that Ω is convex we get $\nabla(p(t,\cdot) + V) \cdot \mathbf{n}_\Omega \leqslant 0$, hence the l.h.s. of (1.3) is negative. From this we immediately see that $\|\nabla p(t,\cdot)\|_{L^2(\Omega)} \leqslant \|\nabla V\|_{L^2(\Omega)}$, i.e. that $p \in L^\infty((0,1); H^1(\Omega))$. Moreover, taking $m > 1$ and mutliplying by p^m, and provided that $\nabla V \in L^q(\Omega)$ with $q > d$, we can use Moser iterations (i.e. iterating the inequality we obtain for different values of m) to prove that $p(t,\cdot) \in L^\infty(\Omega)$ with a norm depending only on V and Ω. For the final pressure P_1, we only look at $D_t(\ln \rho) = \Delta \phi$. Using the equation for the terminal value of ϕ,

$$D_t(\ln \rho)(1,\cdot) = \Delta(P_1 + \Psi). \tag{1.4}$$

The l.h.s. is positive at every point $\bar{x}$ such that $\rho(1, \bar{x}) = 1$, hence we get $\Delta(\Psi + P_1) \geqslant 0$ on $\{P_1 > 0\}$. From exactly the same computations, we deduce $\|\nabla P_1\|_{L^2(\Omega)} \leqslant \|\nabla \Psi\|_{L^2(\Omega)}$ and the $L^\infty(\Omega)$ norm of P_1 depends only on Ω and Ψ provided that $\nabla \Psi \in L^q(\Omega)$ with $q > d$.

Let us say that this strategy, namely looking at the convective derivative of quantities such as $\ln\rho$ was in fact already used by Loeper [21] to study a problem similar to ours (related to the reconstruction of the early universe), but in a case without potential and where $\Delta p := \rho - 1$. In his case, (1.2) leads to a differential inequality involving only ρ from which a L^∞ bound on ρ was deduced.

The assumption of convexity of Ω can be surprising, but is crucial for our method, as it was already the case in our previous work [20]. Roughly speaking, it prevents the interaction with the boundary from causing congestion: only the potentials V and Ψ are a source of congestion. As far as we can see, we do not know how to relax the assumption of convexity of Ω and still be able to control the r.h.s. of (1.3).

It is also possible to adapt these heuristic computations to the case of more general Hamiltonians. For instance, if we replace Ω by a Riemannian manifold (i.e. if we stick to a quadratic Lagrangian and Hamiltonian but insert a specific x-dependency), it is clear that the heuristic computation can be performed exactly in the same way provided that the manifold has a positive Ricci curvature, as the inequality involving the Laplacian of the Hamilton-Jacobi equation can be deduced from Bochner's formula. On the other hand, if we take a Lagrangian which is not quadratic, for instance of the form $L(\dot\omega)$ (we omit explicit dependence on the point, not to overburden the computations), the mean field game system now reads (we have not included the temporal boundary conditions)

$$\begin{cases} \partial_t\rho - \nabla\cdot(\rho\nabla H(\nabla\phi)) &= 0,\\ -\partial_t\phi + H(\nabla\phi) &= p+V, \end{cases}$$

where H is the Legendre transform of L. Compared to the previous system (1.1), the velocity of the agents appearing in the continuity equation is given by $-\nabla H(\nabla\phi)$ and the convective derivative is now $D_t = \partial_t - \nabla H(\nabla\phi)\cdot\nabla$. Expanding the continuity equation leads to

$$D_t(\ln\rho) = \nabla\cdot(\nabla H(\nabla\phi)).$$

For instance, for $L(\dot\omega) = |\dot\omega|^r/r$, we have $H(z) = |z|^q/q$, where q is the conjugate exponent to r, and we have $D_t(\ln\rho) = \Delta_q(\phi)$. Because of the non-linearity of this term, as the reader can see, computations become much more involved than the quadratic case, but it is still possible to obtain an inequality of the form

$$\nabla\cdot(D^2H(\nabla\phi)\nabla(p+V)) \geqslant 0,$$

which proves again that $p+V$ is a subsolution of a suitable elliptic equation. Unfortunately this equation, though being linear in $p+V$, can be highly degenerate if H is not uniformly elliptic, and the analysis of the resulting estimates is far from the objectives of this paper.

1.2 On the consequences for the value function and the Lagrangian point of view

Though they can be seen as interesting in themselves, the estimates obtained on the pressure have consequences for the interpretation of the MFG system.

The first one has to do with the regularity of the value function ϕ. Indeed, as understood in the works by Cardaliaguet and collaborators [9, 10, 12], a solution of a Hamilton-Jacobi equation exhibits regularity as soon as the r.h.s. is bounded from below and its positive part lies in $L^{1+d/2+\varepsilon}$. In the aforementioned articles, such an assumption on the r.h.s. was obtained by assuming a moderate growth on the penalization of congestion in the primal problem. In the case studied in this article, where the density is forced to stay below a given threshold, the naive estimate on the r.h.s. of the Hamilton-Jacobi equation leads to a L^1 bound. The estimates obtained previously in [9] do not allow to deduce regularity of the value function (except if $d=1$). However, with what we prove in the present paper, one can deduce the following (see Section 2.3 below for the definition of the dual and primal problem).

Proposition 1.1. *Assume that either $V\in H^1(\Omega)$ and $d\leqslant 4$ or that $V\in W^{1,q}(\Omega)$ with $q>d$. Then there exists $(\bar\phi,\bar P)$ a solution to the dual problem such that, for any $[0,T]\times\tilde\Omega$ compactly embedded in $[0,1)\times\mathring{\tilde\Omega}$, the function $\bar\phi$ is Hölder-continuous on $[0,T]\times\tilde\Omega$ and $\partial_t\bar\phi\in L^{1+\varepsilon}([0,T]\times\tilde\Omega)$, $\nabla\bar\phi\in L^{2+\varepsilon}([0,T]\times\tilde\Omega)$ for some $\varepsilon>0$.*

Sketch of the proof. The proof of this result can be obtained thanks to Theorem 2.8 proved below. Indeed, if either one of these assumption is true, it implies (thanks to Sobolev injections) that the density $\bar p$ of $\bar P$ is such that $\bar p+V\in L^r([0,T]\times\tilde\Omega)$ for some $r>1+d/2$. Combined with [9, Lemma 3.5] (for the Hölder regularity) and [12, Theorem 1.1] (for the Sobolev regularity), we deduce the result. □

On the other hand, we can also deduce some information about the Lagrangian point of view, for instance when we have L^∞ bounds on the pressure. The necessity for such bounds is explained in details in the introduction of our previous article [20], let us just say that it has to do with how one chooses the correct representative for the pressure. We do not reproduce the derivation here, but just state the result. For that, we need to choose $\bar{P}$ a solution of the dual problem. We know from our theorem 2.8 that it has a density that we will call $\bar{p}$ on $[0,1)\times\Omega$, and we select a precise representative of it, according to the formula in [3], in the following way:

$$\hat{p}(t,x) = \limsup_{\varepsilon\to 0} \frac{1}{|B(x,\varepsilon)|}\int_{B(x,\varepsilon)} \bar{p}(t,y)\,\mathrm{d}y. \tag{1.5}$$

Notice that if $V,\Psi\in W^{1,q}(\Omega)$, we know that V,Ψ are continuous and (thanks to Theorem 2.8), that $\bar{p}$ is in $L^\infty((0,1)\times\Omega)$. Analogously, we also set

$$\hat{P}_1(x) = \limsup_{\varepsilon\to 0} \frac{1}{|B(x,\varepsilon)|}\int_{B(x,\varepsilon)} \bar{P}_1(y)\,\mathrm{d}y. \tag{1.6}$$

We will need to use the notation $e_t : \omega\in H^1([0,1],\Omega)\to\omega(t)\in\Omega$ which denotes the evaluation operator.

Proposition 1.2. *Assume that $V,\Psi\in W^{1,q}(\Omega)$, that $(\bar{\phi},\bar{P})$ is the solution of the dual problem given in Theorem 2.8, and that $\hat{p},\hat{P}_1$ are defined everywhere by* (1.5) *and* (1.6)*. Then there exists $Q\in\mathcal{P}(H^1([0,1],\Omega))$ a measure on the set of H^1 curves valued in Ω such that $\rho : t\to e_t\#Q$ is a solution of the primal problem and, for Q-a.e. γ, the curve γ minimizes*

$$\omega\to\frac{1}{2}\int_0^1|\dot{\omega}(t)|^2\,\mathrm{d}t+\int_0^1\left(V(\omega(t))+\hat{p}(t,\omega(t))\right)\,\mathrm{d}t+\left(\Psi(\omega(1))+\hat{P}_1(\omega(1))\right)$$

among all curves $\omega\in H^1([0,1],\Omega)$ such that $\omega(0)=\gamma(0)$.

If we come back to the interpretation of the MFG system, Q exactly describes the strategy of all the agents: it is the ditribution of mass on the possible strategies, and Proposition 1.2 reads as the fact that Q is a Nash-equilibrium, provided that the agents pay the prices $\bar{p}$ and $\bar{P}_1$ in the regions where the constraint $\rho\leqslant 1$ is saturated.

2 Notations, optimal transport and the variational problems

In all the sequel, Ω will denote the closure of an open bounded convex domain of $\mathbb{R}^d$ with smooth boundary. We assume that the Lebesgue measure of Ω, denoted by $|\Omega|$ is strictly larger than 1: it will be necessary to get existence of probability measures on Ω with density bounded by 1. The generalization to the case where Ω is the d-dimensional torus is straightforward and we do not address it explicitly. The space of probability measures on Ω will be denoted by $\mathcal{P}(\Omega)$. This space $\mathcal{P}(\Omega)$ is endowed with the weak-* topology, i.e. the topology coming from the duality with $C(\Omega)$, the continuous functions from Ω valued in $\mathbb{R}$.

We will also make use of the space of positive measures on the product $[0,1]\times\Omega$ which we denote by $\mathcal{M}_+([0,1]\times\Omega)$.

In all the sequel, when the element of integration is not specified, it is assumed tacitly that integration is performed w.r.t. the Lebesgue measure (either the d-dimensional Lebesgue measure on Ω or the $d+1$-dimensional Lebesgue measure on $[0,1]\times\Omega$). Similarly, a measure is said to be absolutely continuous if it is the case w.r.t. the Lebesgue measure.

2.1 The Wasserstein space

The space $\mathcal{P}(\Omega)$ of probability measures on Ω is endowed with the Wasserstein distance: if μ and ν are two elements of $\mathcal{P}(\Omega)$, the 2-Wasserstein distance $W_2(\mu,\nu)$ between μ and ν is defined via

$$W_2(\mu,\nu) := \sqrt{\min\left\{\int_{\Omega\times\Omega}|x-y|^2\,\mathrm{d}\gamma(x,y)\ :\ \gamma\in\mathcal{P}(\Omega\times\Omega)\text{ and }\pi_0\#\gamma=\mu,\ \pi_1\#\gamma=\nu\right\}}. \tag{2.1}$$

In the formula above, π_0 and $\pi_1 : \Omega\times\Omega\to\Omega$ stand for the projections on respectively the first and second component of $\Omega\times\Omega$. If $T : X\to Y$ is a measurable application and μ is a measure on X, then the image measure of μ by T, denoted by $T\#\mu$, is the measure defined on Y by $(T\#\mu)(B)=\mu(T^{-1}(B))$ for any measurable set $B\subset Y$. For general results about optimal transport, the reader might refer to [27] or [26].

In all the sequel, we identify a measure with its density w.r.t. the Lebesgue measure on Ω. Moreover, if $\mu \in \mathcal{P}(\Omega)$, we write $\mu \leqslant 1$ if the measure μ is absolutely continuous and its density is a.e. bounded by 1.

The Wasserstein distance admits a dual formulation, the dual variables being the so-called Kantorovich potentials. The main properties of these potentials, in the case which is of interest to us, are summarized in the proposition below. We restrict to the cases where the measures have a strictly positive density a.e., as in this particular case the potentials are unique (up to a global additive constant). The proof of these results can be found in (but not exclusively) [26, Chapters 1 and 7].

Proposition 2.1. *Let $\mu, \nu \in \mathcal{P}(\Omega)$ be two absolutely continuous probability measures with strictly positive density. Then there exists a unique (up to adding a constant to φ and subtracting it from ψ) pair (φ, ψ) of Kantorovich potentials satisfying the following properties.*

1. *The squared Wasserstein distance $W_2^2(\mu, \nu)$ can be expressed as*

$$\frac{W_2^2(\mu,\nu)}{2} = \int_\Omega \varphi\mu + \int_\Omega \psi\nu.$$

2. *The "vertical" derivative of $W_2^2(\cdot, \nu)$ at μ is φ: if $\tilde{\rho} \in \mathcal{P}(\Omega)$ is any probability measure, then*

$$\lim_{\varepsilon \to 0} \frac{W_2^2((1-\varepsilon)\mu + \varepsilon\tilde{\rho}, \nu) - W_2^2(\mu,\nu)}{2} = \int_\Omega \varphi(\tilde{\rho} - \mu).$$

3. *The potentials φ and ψ are one the c-transform of the other, meaning that we have*

$$\begin{cases} \varphi(x) &= \inf\limits_{y\in\Omega} \left(\dfrac{|x-y|^2}{2} - \psi(y) \right) \\ \psi(y) &= \inf\limits_{x\in\Omega} \left(\dfrac{|x-y|^2}{2} - \varphi(x) \right). \end{cases}$$

4. *There holds $(\mathrm{Id} - \nabla\varphi)\#\mu = \nu$ and the transport plan $\gamma := (\mathrm{Id}, \mathrm{Id} - \nabla\varphi)\#\mu$ is optimal in the problem* (2.1).

The function φ (resp. ψ) is called the Kantorovitch potential from μ to ν (resp. from ν to μ).

2.2 Absolutely continuous curves in the Wasserstein space

We will denote by Γ the space of continuous curves from $[0,1]$ to $\mathcal{P}(\Omega)$. This space will be equipped with the distance d_Γ of the uniform convergence, i.e.

$$d_\Gamma(\rho^1, \rho^2) := \max_{t\in[0,1]} W_2(\rho_t^1, \rho_t^2).$$

Following [4, Definition 1.1.1], we will introduce the following subset of Γ.

Definition 2.2. *We say that a curve $\rho \in \Gamma$ is 2-absolutely continuous if there exists a function $a \in L^2([0,1])$ such that, for every $0 \leqslant t \leqslant s \leqslant 1$,*

$$W_2(\rho_t, \rho_s) \leqslant \int_t^s a(r)\, \mathrm{d}r.$$

The main interest of this notion lies in the following theorem, which we recall.

Theorem 2.3. *If $\rho \in \Gamma$ is a 2-absolutely continuous curve, then the quantity*

$$|\dot{\rho}_t| := \lim_{h\to 0} \frac{W_2(\rho_{t+h}, \rho_t)}{h}$$

exists and is finite for a.e. t. Moreover,

$$\int_0^1 |\dot{\rho}_t|^2 \,\mathrm{d}t = \sup_{N\geqslant 2} \sup_{0\leqslant t_1 < t_2 < \ldots < t_N \leqslant 1} \sum_{k=2}^{N} \frac{W_2^2(\rho_{t_{k-1}}, \rho_{t_k})}{t_k - t_{k-1}}. \tag{2.2}$$

Proof. The first part is just [4, Theorem 1.1.2]. The proof of the representation formula (2.2) can easily be obtained by adapting the proof of [1, Theorem 4.1.6]. □

The quantity $|\dot\rho_t|$ is called the metric derivative of the curve ρ and heuristically corresponds to the norm of the derivative of ρ at time t in the metric space $(\mathcal{P}(\Omega), W_2)$. The quantity $\int_0^1 |\dot\rho_t|^2 \,\mathrm{d}t$ behaves like a H^1 norm, see [20, Proposition 2.7], but we will not make a use of this.

2.3 Primal and dual problem

To state our main theorem, we do the following assumptions, which will hold throughout the whole article.

Assumptions.

(A1) The domain Ω is the closure of an open bounded convex subset of $\mathbb{R}^d$ with Lebesgue measure $|\Omega| > 1$.

(A2) We fix $V \in H^1(\Omega)$ (the "running" cost) and assume that it is positive.

(A3) We fix $\Psi \in H^1(\Omega)$ (the "final" cost) and assume that it is positive.

(A4) We take $\bar\rho_0 \in \mathcal{P}(\Omega)$ (the initial probability measure) such that $\bar\rho_0 \leqslant 1$.

We denote by $\Gamma_0 \subset \Gamma$ the set of curves such $\rho \in \Gamma$ that $\rho_0 = \bar\rho_0$. As we will see below in the definition of the primal problem, it does not change anything to add a constant to V or Ψ, hence (A2) and (A3) are equivalent to ask that V and Ψ are bounded from below. Note that the important assumption on Ω is its convexity, as already indicated by the formal computation in the introduction.

The primal objective functional reads

$$\mathcal{A}(\rho) := \begin{cases} \displaystyle\int_0^1 \frac{1}{2}|\dot\rho_t|^2 \,\mathrm{d}t + \int_0^1 \left(\int_\Omega V\rho_t\right) \,\mathrm{d}t + \int_\Omega \Psi\rho_1 & \text{if } \rho_t \leqslant 1 \text{ for all } t \in [0,1], \\ +\infty & \text{else.} \end{cases}$$

Definition 2.4. *The primal problem is*

$$\min\{\mathcal{A}(\rho) \;:\; \rho \in \Gamma_0\} \tag{2.3}$$

We will need to consider the dual of this problem. Let $\phi \in C^1([0,1]\times\Omega)$ and $P \in C([0,1]\times\Omega)$ be smooth functions with p positive and in such a way that the Hamilton Jacobi equation is satisfied as an inequality

$$-\partial_t\phi + \frac{1}{2}|\nabla\phi|^2 \leqslant V + P \tag{2.4}$$

and the final value of ϕ is constrained by

$$\phi(1,\cdot) \leqslant \Psi. \tag{2.5}$$

The dual functional is defined as follows:

$$\mathcal{B}(\phi, P) := \int_\Omega \phi(0,\cdot)\bar\rho_0 - \iint_{[0,1]\times\Omega} P.$$

and there is no duality gap between the primal and the dual problem. However, to get existence of a solution of the dual problem, it is too restrictive to look only at smooth functions. As understood in [11], the right functional space is the following.

Definition 2.5. *Let $\mathcal{K}$ be the set of pairs (ϕ, P) where $\phi \in \mathrm{BV}([0,1]\times\Omega) \cap L^2([0,1], H^1(\Omega))$ and $P \in \mathcal{M}_+([0,1]\times\Omega)$ is a positive measure, and the Hamilton Jacobi equation* (2.4) *is understood in the distributional sense, provided we set $\phi(1^+,\cdot) = \Psi$ and that we take in account the possible jump from $\phi(1^-,\cdot)$ to $\phi(1^+,\cdot)$ in the temporal distributional derivative. Specifically, we impose that for any smooth and non-negative function f defined on $[0,1]\times\Omega$,*

$$\iint_{[0,1]\times\Omega} \partial_t f\phi - \int_\Omega f(1,x)\Psi(x)\,\mathrm{d}x + \int_\Omega f(0,x)\phi(0^+,x)\,\mathrm{d}x + \iint_{[0,1]\times\Omega} \frac{1}{2}f|\nabla\phi|^2 \leqslant \iint_{[0,1]\times\Omega} fV + \iint_{[0,1]\times\Omega} f\,\mathrm{d}P.$$

For $(\phi, P) \in \mathcal{K}$, the dual functional is understood in the following sense:

$$\mathcal{B}(\phi, P) := \int_\Omega \phi(0^+, \cdot)\bar{\rho}_0 - P([0,1] \times \Omega).$$

Notice, in the definition of solutions of the Hamilton Jacobi equation we assume $\phi(0^-, \cdot) = \phi(0^+, \cdot)$ (no jump for $t = 0$) but we set $\phi(1^+, \cdot) = \Psi$. The measure P can have a part concentrated on $t = 1$, which may lead to $\phi(1^-, \cdot) > \phi(1^+, \cdot) = \Psi$, provided the jump is compensated by the part of P on $\{1\} \times \Omega$.

Definition 2.6. *The dual problem is*

$$\max \left\{ \mathcal{B}(\phi, P) \ : \ (\phi, P) \in \mathcal{K} \right\}. \tag{2.6}$$

These two problems are in duality in the following sense [11, Propositions 3.3 and 3.8].

Theorem 2.7. *There holds*

$$\min \left\{ \mathcal{A}(\rho) \ : \ \rho \in \Gamma_0 \right\} = \max \left\{ \mathcal{B}(\phi, P) \ : \ (\phi, P) \in \mathcal{K} \right\}.$$

Notice that the existence of a solution to both the primal and the dual problem are included in this statement. The main result of this paper is the following:

Theorem 2.8. *There exists a solution $(\bar{\phi}, \bar{P})$ of the dual problem such that:*

- *The restriction of $\bar{P}$ to $[0,1) \times \Omega$ has a density w.r.t. to the $d+1$-dimensional Lebesgue measure and this density, denoted by $\bar{p}$; satisfies $\|\nabla \bar{p}(t, \cdot)\|_{L^2(\Omega)} \leqslant \|\nabla V\|_{L^2(\Omega)}$ for a.e. $t \in [0,1]$. Moreover, if $V \in W^{1,q}(\Omega)$ with $q > d$, then $\|\bar{p}\|_{L^\infty([0,1)\times\Omega)} \leqslant C < +\infty$ with C depending only on $\|\nabla V\|_{L^q(\Omega)}$ and Ω.*
- *The restriction of $\bar{P}$ to $\{1\} \times \Omega$ has a density w.r.t. to the d-dimensional Lebesgue measure and this density (denoted by $\bar{P}_1$) satisfies $\|\nabla \bar{P}_1\|_{L^2(\Omega)} \leqslant \|\nabla \Psi\|_{L^2(\Omega)}$. Moreover, if $\Psi \in W^{1,q}(\Omega)$ with $q > d$, then $\|\bar{P}_1\|_{L^\infty(\Omega)} \leqslant C < +\infty$ with C depending only on $\|\nabla \Psi\|_{L^q(\Omega)}$ and Ω.*

As already understood in [11, Section 5], there are situations where the pressure is concentrated on $\{1\} \times \Omega$: one cannot expect $\bar{P}$ to have a density with respect to the Lebesgue measure on the closed interval $[0,1]$. Nevertheless we prove in our theorem that the part of the pressure concentrated on $\{1\} \times \Omega$ has spatial regularity, namely $H^1(\Omega)$ and even $L^\infty(\Omega)$ if $\Psi \in W^{1,q}(\Omega)$ with $q > d$. The rest of this paper is devoted to the proof of this theorem, the wrapping of the arguments being located at page 18.

2.4 The discrete problem

To tackle this problem and make rigorous the estimate given in the introduction, we will approximate the primal problem in the following way:

- We introduce a time-discretization. The integer $N+1$ denotes the number of time steps. The time step will be denoted by τ and we use the approximation

$$\int_0^1 \frac{1}{2}|\dot{\rho}_t|^2 \, \mathrm{d}t \simeq \sum_{k=0}^{N-1} \frac{W_2^2(\rho_{k\tau}, \rho_{(k+1)\tau})}{2\tau}.$$

- We add an infinitesimal entropic penalization. The goal is to make sure that the density of the minimizers of the discrete problem will be bounded from below, which is necessary when we want to write the optimality conditions.
- For technical reasons, we need to regularize V and Ψ. We take $(V_N)_{N\in\mathbb{N}}$ a sequence which converges to V in $H^1(\Omega)$ and such that V_N is Lipschitz for any $N \geqslant 1$. We can assume moreover that $\|\nabla V_N\|_{L^2(\Omega)} \leqslant \|\nabla V\|_{L^2(\Omega)}$ and V_N is positive. Similarly, we take a sequence Ψ_N going to Ψ in $H^1(\Omega)$ satisfying analogous properties.

The entropic penalization will be realized with the help of the following functional: for any $\rho \in \mathcal{P}(\Omega)$, we set

$$H(\rho) := \begin{cases} \displaystyle\int_\Omega \ln(\rho)\rho & \text{if } \rho \text{ is absolutely continuous,} \\ +\infty & \text{else.} \end{cases}$$

It is known [26, Chapter 7] that H is lower semi-continuous on $\mathcal{P}(\Omega)$. Moreover, a simple application of Jensen's inequality yields

$$-\ln(|\Omega|) \leqslant H(\rho) \leqslant 0$$

as soon as $\rho \leqslant 1$.

To define the discrete problem, we take $N \geqslant 1$ and denote by

$$\Gamma_0^N := \{(\rho_{k\tau})_{k\in\{0,1,\dots,N\}} \ : \ \rho_{k\tau} \in \mathcal{P}(\Omega) \text{ and } \rho_0 = \bar{\rho}_0\} \subset (\mathcal{P}(\Omega))^{N+1}$$

the set of discrete curves starting from $\bar{\rho}_0$. We denote by $\tau := 1/N$ the time step. We choose $(\lambda_N)_{N\in\mathbb{N}}$ which goes to 0 while being strictly positive, it will account for the scale of the entropic penalization. The speed at which $\lambda_N \to 0$ is irrelevant for the analysis, hence we do not need to specify it. The discrete functional $\mathcal{A}^N$ is defined on Γ_0^N as

$$\mathcal{A}^N(\rho) := \begin{cases} \displaystyle\sum_{k=0}^{N-1} \frac{W_2^2(\rho_{k\tau},\rho_{(k+1)\tau})}{2\tau} + \sum_{k=1}^{N-1} \tau\left(\int_\Omega V_N\rho_{k\tau} + \lambda_N H(\rho_{k\tau})\right) + \int_\Omega \Psi_N\rho_1 + \lambda_N H(\rho_1) & \text{if } \rho_{k\tau} \leqslant 1 \text{ for all } k \in \{0,1,\dots,N\}, \\ +\infty & \text{else.} \end{cases}$$

The discrete problem reads as

$$\min\left\{\mathcal{A}^N(\rho) \ : \ \rho \in \Gamma_0^N\right\}. \tag{2.7}$$

Proposition 2.9. *For any $N \geqslant 1$, there exists a unique solution to the discrete problem.*

Proof. The functional $\mathcal{A}^N$ is l.s.c. on Γ_0^N. Moreover, the curve ρ which is constant and equal to $\bar{\rho}_0$ belongs to Γ_0^N and is such that $\mathcal{A}^N(\rho) < +\infty$. As Γ_0^N is compact (for the topology of the weak convergence of measures), the direct method of calculus of variations ensures the existence of a minimizer.

Uniqueness clearly holds as $\lambda_N > 0$ and the entropy is a strictly convex function on $\mathcal{P}(\Omega)$. □

From now on, for any $N \geqslant 1$, we fix $\bar{\rho}^N$ the unique solution of the discrete problem

3 Estimates on the discrete problem

Let us comment on a technical refinement: for some computations to be valid, we will need to assume that $\bar{\rho}_0$ is smooth is strictly positive. If it is not the case, it is easy to approximate (for fixed N) the measure $\bar{\rho}_0$ with a sequence $\bar{\rho}_0^{(n)}$ of smooth densities. For such a $\bar{\rho}_0^{(n)}$, the estimates obtained below for a given N (Corollary 3.6) do not depend on n. Hence it is easy to send n to $+\infty$, using the stability of the Kantorovich potentials [26, Theorem 1.52] to see that these estimates are still satisfied for the solution of the discrete problem with initial condition $\bar{\rho}_0$. In short: we will do as if our initial condition $\bar{\rho}_0$ were smooth, and as long as the final estimates do not depend on the smoothness of $\bar{\rho}_0$ this will be legitimate.

3.1 Interior regularity

We begin with the interior regularity. In this subsection, we fix $N \geqslant 1$ and $k \in \{1,2,\dots,N-1\}$ a given instant in time. We use the shortcut $\bar{\rho} := \bar{\rho}^N_{k\tau}$ and we also denote $\mu := \bar{\rho}^N_{(k-1)\tau}$ and $\nu := \bar{\rho}^N_{(k+1)\tau}$. As $\bar{\rho}^N$ is a solution of the discrete problem, we know that $\bar{\rho}$ is a minimizer, among all probability measures with density bounded by 1, of

$$\rho \mapsto \frac{W_2^2(\mu,\rho) + W_2^2(\rho,\nu)}{2\tau} + \tau\left(\int_\Omega V_N\rho + \lambda_N H(\rho)\right).$$

Lemma 3.1. *The density $\bar{\rho}$ is strictly positive a.e.*

Proof. This is exactly the same proof as [20, Lemma 3.1], as the construction done in this proof preserves the constraint of having a density smaller than 1. □

Proposition 3.2. *Let us denote by φ_μ and φ_ν the Kantorovich potentials for the transport from $\bar\rho$ to μ and ν respectively. There exists $p \in L^1(\Omega)$, positive, such that $\{p > 0\} \subset \{\bar\rho = 1\}$ and a constant C such that*

$$\frac{\varphi_\mu + \varphi_\nu}{\tau^2} + V_N + p + \lambda_N \ln(\bar\rho) = C \quad a.e. \tag{3.1}$$

Moreover p and $\ln(\bar\rho)$ are Lipschitz and $\nabla p \cdot \nabla \ln(\bar\rho) = 0$ a.e.

Proof. Let $\tilde\rho \in \mathcal{P}(\Omega)$ such that $\tilde\rho \leqslant 1$. We define $\rho_\varepsilon := (1-\varepsilon)\bar\rho + \varepsilon\tilde\rho$ and use it as a competitor. Clearly $\rho_\varepsilon \leqslant 1$, i.e. it is an admissible competitor. Comparing $\mathcal{A}^N(\rho_\varepsilon)$ to $\mathcal{A}^N(\rho)$, we extract the following information. Using Proposition 2.1, as $\bar\rho > 0$, the Kantorovich potentials φ_μ and φ_ν are unique (up to a constant) and

$$\lim_{\varepsilon\to 0} \frac{W_2^2(\mu,\rho_\varepsilon) - W_2^2(\mu,\bar\rho) + W_2^2(\rho_\varepsilon,\nu) - W_2^2(\bar\rho,\nu)}{2\tau^2} = \int_\Omega \frac{\varphi_\mu + \varphi_\nu}{\tau}(\tilde\rho - \bar\rho).$$

The term involving V_N is straightforward to handle as it is linear. The only remaining term is the one involving the entropy. But here, using the same reasoning as in [20, Proposition 3.2], we can say that

$$\limsup_{\varepsilon\to 0} \frac{H(\rho_\varepsilon) - H(\bar\rho)}{\varepsilon} \leqslant \int_\Omega \ln(\bar\rho)(\tilde\rho - \bar\rho).$$

Putting the pieces together, we see that $\int_\Omega h\,(\tilde\rho - \bar\rho) \geqslant 0$ for any $\tilde\rho \in \mathcal{P}(\Omega)$ with $\tilde\rho \leqslant 1$, provided that h is defined by

$$h := \frac{\varphi_\mu + \varphi_\nu}{\tau^2} + V_N + \lambda_N \ln(\bar\rho)$$

It is known, analogously to [22, Lemma 3.3], that this leads to the existence of a constant C such that

$$\begin{cases} \bar\rho = 1 & \text{on } \{h < C\} \\ \bar\rho \leqslant 1 & \text{on } \{h = C\} \\ \bar\rho = 0 & \text{on } \{h > C\} \end{cases} \tag{3.2}$$

Specifically, C is defined as the smallest real $\tilde C$ such that $\mathcal{L}(\{h \leqslant \tilde C\}) \geqslant 1$ ($\mathcal{L}$ being the Lebesgue measure on Ω), and it is quite straightforward to check that this choice works. Note that the case $\{h > C\}$ can be excluded by Lemma 3.1. The pressure p is defined as $p = (C-h)_+$, thus (3.1) holds. It satisfies $p \geqslant 0$, and $\bar\rho < 1$ implies $p = 0$.

It remains to answer the question of the integrability properties of p and $\ln(\bar\rho)$. Notice that p is positive, and non zero only on $\{\bar\rho = 1\}$. On the other hand, $\ln(\bar\rho) \leqslant 0$ and it is non zero only on $\{\bar\rho < 1\}$. Hence, one can write

$$p = \left(C - \frac{\varphi_\mu + \varphi_\nu}{\tau^2} + V_N\right)_+ \text{ and } \ln(\bar\rho) = -\frac{1}{\lambda_N}\left(C - \frac{\varphi_\mu + \varphi_\nu}{\tau^2} + V_N\right)_- . \tag{3.3}$$

Given that the Kantorovich potentials and V_N are Lipschitz, it implies the Lipschitz regularity for p and $\ln(\bar\rho)$. Moreover, the identity $\nabla p \cdot \nabla \ln(\bar\rho) = 0$ is straightforward using $\nabla f_+ = \nabla f \mathbb{1}_{f>0}$ a.e., which is valid for any $f \in H^1(\Omega)$. □

Let us note that φ_μ and φ_ν have additional regularity properties, even though they depend heavily on N.

Lemma 3.3. *The Kantorovich potentials φ_μ and φ_ν belong to $C^{2,\alpha}(\mathring\Omega) \cap C^{1,\alpha}(\Omega)$.*

Proof. If $k \in \{2,\dots,N\}$, thanks to Proposition 3.2 (applied in $k-1$ and $k+1$), we know that μ and ν have a Lipschitz density and are bounded from below. Using the regularity theory for the Monge Ampère-equation [27, Theorem 4.14], we can conclude that φ_μ and φ_ν belong to $C^{2,\alpha}(\mathring\Omega) \cap C^{1,\alpha}(\Omega)$. □

Theorem 3.4. *For any $m \geqslant 1$, the following inequality holds:*

$$\int_\Omega \nabla(p^m) \cdot \nabla(p + V_N) \leqslant 0. \tag{3.4}$$

Proof. The (optimal) transport map from $\bar\rho$ to μ is given by $\mathrm{Id} - \nabla\varphi_\mu$, and similarly for ν. We consider the following quantity, (defined on the whole Ω given the regularity of μ, ν, φ_μ and φ_ν), which is a discrete analogue of the l.h.s. of (1.2):

$$D(x) := -\frac{\ln(\mu(x - \nabla\varphi_\mu(x))) + \ln(\nu(x - \nabla\varphi_\nu(x))) - 2\ln(\bar\rho(x))}{\tau^2}.$$

Notice that if $\bar\rho(x) = 1$, then by the constraint $\mu(x - \nabla\varphi_\mu(x)) \leqslant 1$ and $\nu(x - \nabla\varphi_\nu(x)) \leqslant 1$ the quantity $D(x)$ is positive. On the other hand, using $(\mathrm{Id} - \nabla\varphi_\mu)\#\bar\rho = \mu$ and the Monge-Ampère equation, for all $x \in \mathring{\Omega}$ there holds

$$\mu(x - \nabla_\mu\varphi_\mu(x)) = \frac{\bar\rho(x)}{\det(\mathrm{Id} - D^2\varphi_\mu(x))},$$

and a similar identity holds for φ_ν. Hence the quantity $D(x)$ is equal, for all $x \in \mathring{\Omega}$, to

$$D(x) = \frac{\ln(\det(\mathrm{Id} - D^2\varphi_\mu(x))) + \ln(\det(\mathrm{Id} - D^2\varphi_\nu(x)))}{\tau^2}.$$

Diagonalizing the matrices $D^2\varphi_\mu, D^2\varphi_\nu$ and using the convexity inequality $\ln(1-y) \leqslant -y$, we end up with

$$D(x) \leqslant -\frac{\Delta(\varphi_\mu(x) + \varphi_\nu(x))}{\tau^2}.$$

We multiply this identity by p^m and integrate. Thanks to the fact that D is positive on $\{\bar\rho = 1\}$, as p is positive and does not vanish only on $\{\bar\rho = 1\}$, the quantity $p^m D$ is positive on $\mathring{\Omega}$. As the latter coincides, up to a Lebesgue negligible set, with Ω, we get

$$\int_\Omega p^m \frac{\Delta(\varphi_\mu + \varphi_\nu)}{\tau^2} \leqslant 0. \tag{3.5}$$

We do an integration by parts, which reads

$$\int_\Omega p^m \frac{\Delta(\varphi_\mu + \varphi_\nu)}{\tau^2} = \int_{\partial\Omega} p^m \frac{\nabla(\varphi_\mu + \varphi_\nu)}{\tau^2} \cdot \mathbf{n}_\Omega - \int_\Omega \nabla(p^m) \cdot \frac{\nabla(\varphi_\mu + \varphi_\nu)}{\tau^2} \tag{3.6}$$

To handle the boundary term, recall that $\nabla\varphi_\mu$ is continuous up to the boundary and that $x - \nabla\varphi_\mu(x) \in \Omega$ for every $x \in \Omega$ as $(\mathrm{Id} - \nabla\varphi_\mu)\#\bar\rho = \mu$. Given the convexity of Ω, it implies $\nabla\varphi_\mu(x) \cdot \mathbf{n}_\Omega(x) \geqslant 0$ for every point $x \in \partial\Omega$ for which the outward normal $\mathbf{n}_\Omega(x)$ is defined. As it is the case for a.e. point of the boundary, as a similar inequality holds for φ_ν, and given that p^m is positive, we can drop the boundary term in (3.6) and get

$$\int_\Omega p^m \frac{\Delta(\varphi_\mu + \varphi_\nu)}{\tau^2} \geqslant -\int_\Omega \nabla(p^m) \cdot \frac{\nabla(\varphi_\mu + \varphi_\nu)}{\tau^2}.$$

We emphasize that dropping this boundary term corresponds exactly to the heuristic inequality $\nabla(p(t,\cdot) + V) \cdot \mathbf{n}_\Omega \leqslant 0$ evoked in the introduction below equation (1.3). Using the optimality conditions (3.1), we see that

$$0 \geqslant \int_\Omega p^m \frac{\Delta(\varphi_\mu + \varphi_\nu)}{\tau^2} \geqslant \int_\Omega \nabla(p^m) \cdot \nabla(p + V_N + \lambda_N \ln(\bar\rho))$$

Now remember that in Proposition 3.2 we have proved that $\nabla p \cdot \nabla \ln(\bar\rho) = 0$ a.e., which is sufficient to drop the term involving $\nabla\ln(\bar\rho)$ and get (3.4). □

The inequation (3.4) implies the $H^1(\Omega)$ and $L^\infty(\Omega)$ regularity for the pressure: this can be seen as a consequence of Moser's regularity for elliptic equations [23]. We still give the proof for the sake of completeness, and also because in the inequality (3.4), the boundary terms have already been taken in account, which enables to get regularity up to the spatial boundary in a single set of iterations.

Lemma 3.5. *Let f, W be Lipschitz functions defined on Ω such that f vanishes on a set of measure at least $|\Omega| - 1 > 0$ and such that, for any $m \geqslant 1$,*

$$\int_\Omega \nabla(f^m) \cdot \nabla(f + W) \leqslant 0.$$

Then there holds $\|\nabla f\|_{L^2(\Omega)} \leqslant \|\nabla W\|_{L^2(\Omega)}$. Moreover, if $\nabla W \in L^q(\Omega)$ with $q > d$, then $f \in L^\infty(\Omega)$ and $\|f\|_{L^\infty(\Omega)}$ is bounded by a constant which depends only on Ω and $\|\nabla W\|_{L^q(\Omega)}$.

Proof. With $m = 1$ we immediately get

$$\|\nabla f\|_{L^2(\Omega)} \leqslant \|\nabla W\|_{L^2(\Omega)},$$

In particular, using the Poincaré inequality and the fact that $|\{f = 0\}| \geqslant |\Omega| - 1$, we see that $\|f\|_{L^1(\Omega)}$ is bounded by a constant depending only on Ω and V.

In the rest of the proof, we denote by C a constant which depends only on Ω and $\|\nabla W\|_{L^q(\Omega)}$, and can change from line to line. We write the estimate, for any $m \geqslant 1$, as

$$\int_\Omega |\nabla f|^2 f^{m-1} \leqslant -\int_\Omega (\nabla f \cdot \nabla W) f^{m-1}$$

Using Young's inequality, it is clear that

$$\frac{2}{(m+1)^2} \int_\Omega \left|\nabla \left(f^{(m+1)/2}\right)\right|^2 = \frac{1}{2}\int_\Omega |\nabla f|^2 f^{m-1} \leqslant \frac{1}{2}\int_\Omega |\nabla W|^2 f^{m-1}.$$

Take $\tilde\beta < \beta < \frac{d}{d-2}$ sufficiently close to $\frac{d}{d-2}$ in such a way that $2\tilde\beta/(\tilde\beta - 1) \leqslant q$. In particular, the $L^{2\tilde\beta/(\tilde\beta-1)}(\Omega)$ norm of ∇W is bounded by $C\|\nabla W\|_{L^q(\Omega)}$. Moreover, we know that $H^1(\Omega) \hookrightarrow L^{2\beta}(\Omega)$. Considering the fact that $f^{(m+1)/2}$ vanishes on a subset of measure at least $|\Omega| - 1$, it enables us to write [23, Lemma 2]

$$\begin{aligned}\left(\int_\Omega f^{(m+1)\beta}\right)^{1/\beta} \leqslant C\int_\Omega \left|\nabla\left(f^{(m+1)/2}\right)\right|^2 &\leqslant C(m+1)^2 \int_\Omega |\nabla W|^2 f^{m-1} \\ &\leqslant C(m+1)^2 \left(\int_\Omega |\nabla W|^{2\tilde\beta/(\tilde\beta-1)}\right)^{(\tilde\beta-1)/\tilde\beta} \left(\int_\Omega f^{(m-1)\tilde\beta}\right)^{1/\tilde\beta},\end{aligned}$$

where the last inequality is Hölder's inequality with an exponent $\tilde\beta$. Thanks to this choice, taking the power $1/(m+1)$ on both sides,

$$\|f\|_{L^{(m+1)\beta}(\Omega)} \leqslant \left[C(m+1)^2\right]^{1/(m+1)} \|f\|_{L^{(m-1)\tilde\beta}}^{(m-1)/(m+1)}.$$

It is easy to iterate this inequation. With $r = (m-1)\tilde\beta$, as $(m+1)\beta \geqslant \beta r/\tilde\beta$, one can write that

$$\|f\|_{L^{\beta/\tilde\beta r}(\Omega)} \leqslant [C(r+1)]^{C/r} \max\left(\|f\|_{L^r(\Omega)}, 1\right).$$

An easy induction (recall that we already know that f is bounded in $L^1(\Omega)$ by a constant depending only on Ω and W) with $r_n = \left(\beta/\tilde\beta\right)^n$ shows that $\|f\|_{L^{r_n}(\Omega)}$ is bounded by a constant which depends only on $\|\nabla W\|_{L^q(\Omega)}$ and Ω, which implies the claimed $L^\infty(\Omega)$ bound. □

Corollary 3.6. *There holds $\|\nabla p\|_{L^2(\Omega)} \leqslant \|\nabla V\|_{L^2(\Omega)}$. Moreover, if $V \in W^{1,q}(\Omega)$ with $q > d$, then $p \in L^\infty(\Omega)$ and $\|p\|_{L^\infty(\Omega)}$ is bounded by a constant which depends only on Ω and $\|\nabla V\|_{L^q(\Omega)}$.*

Proof. It is enough to combine Lemma 3.5 and (3.4): one has to remember that p vanishes where $\bar\rho = 1$, which is of measure at least $|\Omega| - 1$, that $\|\nabla V_N\|_{L^2(\Omega)} \leqslant \|\nabla V\|_{L^2(\Omega)}$, and that $\|\nabla V_N\|_{L^q(\Omega)}$ is bounded independently on N if $V \in W^{1,q}(\Omega)$. □

3.2 Boundary regularity

As we said, we will see that the pressure has a part which is concentrated on the time-boundary $t = 1$. The regularity of this part is proved exactly by the same technique than in the interior, hence we will only sketch it. In this subsection, we fix $N \geqslant 1$. We use the shortcut $\bar\rho := \bar\rho^N_{N\tau} = \bar\rho^N_1$ for the final measure and we also denote $\mu := \bar\rho^N_{(N-1)\tau}$. As $\bar\rho^N$ is a solution of the discrete problem, we know that $\bar\rho$ is a minimizer, among all probability measures with density bounded by 1, of

$$\rho \mapsto \frac{W_2^2(\mu,\rho)}{2\tau} + \left(\int_\Omega \Psi_N \rho + \lambda_N H(\rho)\right).$$

This variational problem would be exactly the one obtained if we were to discretize the (Wasserstein) gradient flow of the functional $\Psi_N + \lambda_N H$ (with the constraint that the density does not exceed 1) using the *minimizing movement scheme* (known in this context as the JKO scheme [15]). In particular, all the computations of this paragraph could be translated in the framework of gradient flows, i.e. first order evolutions in time.

Lemma 3.7. *The density $\bar\rho$ is strictly positive a.e.*

Proof. This property holds for exactly the same reason as in Lemma 3.1. □

Proposition 3.8. *Let us denote by φ_μ the Kantorovich potential for the transport from $\bar\rho$ to μ. There exists $p \in L^1(\Omega)$, positive, such that $\{p > 0\} \subset \{\bar\rho = 1\}$ and a constant C such that*

$$\frac{\varphi_\mu}{\tau} + \Psi_N + p + \lambda_N \ln(\bar\rho) = C. \tag{3.7}$$

Moreover p and $\ln(\bar\rho)$ are Lipschitz and $\nabla p \cdot \nabla \ln(\bar\rho) = 0$ a.e.

Proof. We use exactly the same competitor as in the proof of Proposition 3.2. It leads to the conclusion that $\int_\Omega h(\tilde\rho - \bar\rho) \geqslant 0$ for any $\tilde\rho \in \mathcal{P}(\Omega)$ with $\tilde\rho \leqslant 1$ where h is defined as

$$h := \frac{\varphi_\mu}{\tau} + \Psi_N + \lambda_N \ln(\bar\rho).$$

It implies the existence of a constant C such that (3.2) holds, and we define p exactly in the same way, as $p := (C-h)_+$. The integrability properties of p and $\ln(\bar\rho)$ are derived in the same way as in the proof of Proposition 3.2. □

The additional regularity for φ_μ is exactly the same than for the interior case (this is why we have also used an entropic penalization at the boundary).

Lemma 3.9. *The Kantorovich potential φ_μ belongs to $C^{2,\alpha}(\mathring\Omega) \cap C^{1,\alpha}(\Omega)$.*

Theorem 3.10. *For any $m \geqslant 1$, the following inequality holds:*

$$\int_\Omega \nabla(p^m) \cdot \nabla(p + \Psi_N) \leqslant 0. \tag{3.8}$$

Proof. On the set $\mathring\Omega$ we consider the following quantity, which is the analogue of the l.h.s. of (1.4):

$$D(x) := \frac{\ln(\bar\rho(x)) - \ln(\mu(x - \nabla\varphi_\mu(x)))}{\tau}.$$

If $\bar\rho(x) = 1$, then by the constraint $\mu(x - \nabla\varphi_\mu(x)) \leqslant 1$ the quantity $D(x)$ is positive. On the other hand, exactly by the same estimate than in the proof of Theorem 3.4,

$$D(x) \leqslant -\frac{(\Delta\varphi_\mu)(x)}{\tau}.$$

We multiply this inequality by p^m, do an integration by parts (the boundary term is handled exactly as in Theorem 3.4), and we end up with (3.8). □

Corollary 3.11. *There holds $\|\nabla p\|_{L^2(\Omega)} \leqslant \|\nabla\Psi\|_{L^2(\Omega)}$. Moreover, if $\Psi \in W^{1,q}(\Omega)$ with $q > d$, then $p \in L^\infty(\Omega)$ and $\|p\|_{L^\infty(\Omega)}$ is bounded by a constant which depends only on Ω and $\|\nabla\Psi\|_{L^q(\Omega)}$.*

Proof. Exactly as in the proof of Corollary 3.6, it is enough to combine Lemma 3.5 and the estimate (3.8). □

4 Convergence to the continuous problem

Recall that for any $N \geqslant 1$, $\bar\rho^N$ denotes the solution of the discrete problem.

4.1 Convergence of the primal problem

This convergence is very similar to the one performed in [20] hence we will not really reproduce it. Furthermore, as we are ultimately interested in the dual problem, we need only the convergence of the value of the primal problem, not of the minimizers.

Define $\tilde{\mathcal{A}}^N$ on Γ_0^N exactly as the discrete primal functional $\mathcal{A}^N$, but where the regularized potentials V_N and Ψ_N are replaced by the true potentials V and Ψ. Given the L^∞ bound on ρ (which holds if $\mathcal{A}^N$ or $\tilde{\mathcal{A}}^N$ are finite), one can see that for any $\rho \in \Gamma_0^N$ with density bounded by 1,

$$\left|\mathcal{A}^N(\rho) - \tilde{\mathcal{A}}^N(\rho)\right| \leqslant \|V - V_N\|_{L^1(\Omega)} + \|\Psi - \Psi_N\|_{L^1(\Omega)}, \tag{4.1}$$

and the r.h.s. goes to 0 uniformly in ρ as $N \to +\infty$.

On the other hand, using exactly the same proofs as in [20], Section 5.1 and 5.2, one can easily check (the only thing to check is that all the constructions are compatible with the constraint of having a density bounded by 1 but it is straightforward) that the value of the discrete problem

$$\min\left\{\tilde{A}^N(\rho) \ : \ \rho \in \Gamma_0^N\right\}$$

converges to the minimal value of the primal problem (notice that it is for this result that we need the scale λ_N of the entropic penalization to go to 0). Combined with (4.1), one can conclude the following.

Proposition 4.1. *The value of the discrete problem converges to the one of the continuous one in the sense that*

$$\lim_{N\to+\infty} \mathcal{A}^N(\bar{\rho}^N) = \min\left\{\mathcal{A}(\rho) \ : \ \rho \in \Gamma_0\right\}.$$

4.2 Convergence to the dual problem

In this subsection, we want to build a value function ϕ^N which will go, as $N \to +\infty$, to a solution of the (continuous) dual problem. Notice that the discrete functional $\mathcal{A}^N$ is convex, hence we could consider discrete dual problem but we will not do it explicitly: indeed, the approximate value function ϕ^N will not be a solution of the discrete dual problem and we will not prove a duality result at the discrete level.

On the contrary, we will just guess the expression of ϕ^N (we have to say to the inspiration for this kind of construction was found in the work of Loeper [21]) and use the explicit expression to prove that the value of some quantity which looks like the continuous dual objective, evaluated at ϕ^N, is closed to the value of the discrete primal problem. Then, sending N to $+\infty$, we recover an admissible $(\bar{\phi}, \bar{P})$ for the continuous dual problem such that $\mathcal{B}(\bar{\phi}, \bar{P})$ is larger than the optimal value of the continuous primal problem (and this comes from estimates proved at the discrete level). It will allow us to conclude that $(\bar{\phi}, \bar{P})$ is a solution of the dual problem thanks to the absence of duality gap at the continuous level. Eventually, we pass to the limit the discrete estimates in Corollary 3.6 and Corollary 3.11 to get the ones for $\bar{p}$ and $\bar{P}_1$.

Let us recall that $\bar{\rho}^N$ is the solution of the discrete problem. For any $k \in \{0, 1, \dots, N-1\}$, we choose $(\varphi^N_{k\tau}, \psi^N_{k\tau})$ a pair of Kantorovich potential between $\bar{\rho}^N_{k\tau}$ and $\bar{\rho}^N_{(k+1)\tau}$, such choice is unique up to an additive constant. According to Proposition 3.2 and Proposition 3.8, making the dependence on N and k explicit, for any $k \in \{1, 2, \dots, N\}$, there exists a pressure $p^N_{k\tau}$ positive and Lipschitz, and a constant $C^N_{k\tau}$ such that

$$\begin{cases} \dfrac{\psi^N_{(k-1)\tau} + \varphi^N_{k\tau}}{\tau^2} + V_N + p^N_{k\tau} + \lambda_N \ln(\bar{\rho}^N_{k\tau}) = C^N_{k\tau} & k \in \{1, 2, \dots, N-1\}, \\ \dfrac{\psi^N_{(k-1)\tau}}{\tau} + \Psi_N + p^N_1 + \lambda_N \ln(\bar{\rho}^N_{k\tau}) = C^N_1 & k = N. \end{cases} \tag{4.2}$$

We define the following value function, defined on the whole interval $[0, 1]$ which can be thought as a function which looks like a solution of what could be called a discrete dual problem.

Definition 4.2. *Let $\phi^N \in \mathrm{BV}([0,1] \times \Omega) \cap L^2([0,1] \times H^1(\Omega))$ the function defined as follows. The "final" value is given by*

$$\phi^N(1^-, \cdot) := \Psi_N + p^N_1. \tag{4.3}$$

Provided that the value $\phi^N((k\tau)^-,\cdot)$ is defined for some $k \in \{1,2,\dots,N\}$, the value of ϕ^N on $((k-1)\tau, k\tau)\times\Omega$ is defined by

$$\phi^N(t,x) := \inf_{y\in\Omega}\left(\frac{|x-y|^2}{2(k\tau-t)} + \phi^N((k\tau)^-,y)\right). \tag{4.4}$$

If $k \in \{1,2,\dots,N-1\}$, the function ϕ^N has a temporal jump at $t = k\tau$ defined by

$$\phi^N((k\tau)^-,x) := \phi^N((k\tau)^+,x) + \tau\left(V_N + p^N_{k\tau}\right)(x) \tag{4.5}$$

Notice that we have not included the entropic term: its only effect would have been to decrease ϕ^N (which in the end decreases the value of the dual functional) and it would have prevented us from getting compactness on the sequence ϕ^N. The link between this value function and the Kantorovich potentials is the following.

Lemma 4.3. *For any $k \in \{1,2,\dots,N\}$, one has*

$$\phi^N((k\tau)^-,\cdot) \geqslant C^N_1 + \tau\sum_{j=k}^{N-1} C^N_{j\tau} - \frac{\psi^N_{(k-1)\tau}}{\tau}. \tag{4.6}$$

For any $k \in \{0,1,\dots,N-1\}$, one has

$$\phi^N((k\tau)^+,\cdot) \geqslant C^N_1 + \tau\sum_{j=k+1}^{N-1} C^N_{j\tau} + \frac{\varphi^N_{k\tau}}{\tau}. \tag{4.7}$$

Proof. We will prove it by (decreasing) induction on $k \in \{0,1,\dots,N\}$. For $k = N$, by the optimality conditions (4.2) and the fact that $\ln(\bar\rho^N_1) \leqslant 0$, it is clear that (4.6) holds.

Now assume that (4.6) holds for some k. Using (4.4), one has

$$\begin{aligned}
\phi^N(((k-1)\tau)^+,x) &= \inf_{y\in\Omega}\left(\frac{|x-y|^2}{2\tau} + \phi^N((k\tau)^-,y)\right)\\
&\geqslant C^N_1 + \tau\sum_{j=k}^{N-1} C^N_{j\tau} + \inf_{y\in\Omega}\left(\frac{|x-y|^2}{2\tau} - \frac{\psi^N_{(k-1)\tau}(y)}{\tau}\right)\\
&= C^N_1 + \tau\sum_{j=k}^{N-1} C^N_{j\tau} + \frac{1}{\tau}\inf_{y\in\Omega}\left(\frac{|x-y|^2}{2} - \psi^N_{(k-1)\tau}(y)\right) = C^N_1 + \tau\sum_{j=k}^{N-1} C^N_{j\tau} + \frac{\varphi^N_{(k-1)\tau}(x)}{\tau},
\end{aligned}$$

where the last inequality comes from the fact that $\varphi^N_{(k-1)\tau}$ is the c-transform of $\psi^N_{(k-1)\tau}$. This gives us (4.7) for $k-1$. On the other hand, assume that (4.7) holds for some k. Using (4.5) and the optimality conditions (4.2) ,

$$\begin{aligned}
\phi^N((k\tau)^-,x) &= \phi^N((k\tau)^+,x) + \tau\left(V_N + p^N_{k\tau}\right)\\
&\geqslant C^N_1 + \tau\sum_{j=k+1}^{N-1} C^N_{j\tau} + \frac{\varphi^N_{k\tau}}{\tau} + \tau\left(V_N + p^N_{k\tau}\right)\\
&= C^N_1 + \tau\sum_{j=k+1}^{N-1} C^N_{j\tau} + C^N_{k\tau} - \frac{\psi^N_{k\tau}}{\tau} - \lambda_N\tau\ln(\bar\rho^N_{k\tau}) \geqslant C^N_1 + \tau\sum_{j=k}^{N-1} C^N_{j\tau} - \frac{\psi^N_{k\tau}}{\tau},
\end{aligned}$$

which means that (4.6) holds for k. □

From this identity, we can express some kind of duality result at the discrete level, which reads as follows.

Proposition 4.4. *For $N \geqslant 1$, the following inequality holds:*

$$\mathcal{A}^N(\bar\rho^N) \leqslant \int_\Omega \phi^N(0^+,\cdot)\bar\rho_0 - \tau\sum_{k=1}^{N-1}\int_\Omega p^N_{k\tau} - \int_\Omega p^N_1. \tag{4.8}$$

We have an inequality and not an equality because we have not included the entropic terms in the value function.

Proof. The idea is to evaluate $\mathcal{A}^N(\bar\rho^N)$ by expressing the Wasserstein distances with the help of the Kantorovich potentials.

$$\begin{aligned}\mathcal{A}^N(\bar\rho^N) &= \sum_{k=0}^{N-1}\frac{W_2^2(\bar\rho^N_{k\tau},\bar\rho^N_{(k+1)\tau})}{\tau}+\sum_{k=1}^{N-1}\tau\left(\int_\Omega V_N\bar\rho^N_{k\tau}+\lambda_N H(\bar\rho^N_{k\tau})\right)+\int_\Omega \Psi_N\bar\rho^N_1+\lambda_N H(\bar\rho^N_1)\\ &= \frac{1}{\tau}\sum_{k=0}^{N-1}\left(\int_\Omega\varphi^N_{k\tau}\bar\rho^N_{k\tau}+\int_\Omega\psi^N_{k\tau}\bar\rho^N_{(k+1)\tau}\right)+\sum_{k=1}^{N-1}\tau\left(\int_\Omega V_N\bar\rho^N_{k\tau}+\lambda_N H(\bar\rho^N_{k\tau})\right)+\int_\Omega\Psi_N\bar\rho^N_1+\lambda_N H(\bar\rho^N_1)\\ &= \frac{1}{\tau}\int_\Omega\varphi^N_0\bar\rho_0+\sum_{k=1}^{N-1}\int_\Omega\left(\frac{\varphi^N_{k\tau}+\psi^N_{(k-1)\tau}}{2\tau}+\tau(V_N+\lambda_N\ln(\bar\rho^N_{k\tau}))\right)\bar\rho^N_{k\tau}+\int_\Omega\left(\frac{\psi^N_{(N-1)\tau}}{2\tau}+\Psi_N+\lambda_N\ln(\bar\rho^N_1)\right)\bar\rho^N_1,\end{aligned}$$

where the last equality comes from a reindexing of the sums. Now we use the optimality conditions (4.2) to handle the second and third term. Notice that, as $p^N_{k\tau}$ lives only where $\bar\rho^N_{k\tau}=1$, that we can replace $\bar\rho^N_{k\tau}$ by 1 when it is multiplied by the pressure. Recall also that the probability distributions, when integrated against a constant, are equal to this constant. We are left with

$$\begin{aligned}\mathcal{A}^N(\bar\rho^N) &= \frac{1}{\tau}\int_\Omega\varphi^N_0\bar\rho_0+\sum_{k=1}^{N-1}\left(C^N_{k\tau}-\tau\int_\Omega p^N_{k\tau}\right)+\left(C^N_1-\int_\Omega p^N_1\right)\\ &\leqslant \int_\Omega\phi^N(0^+,\cdot)\bar\rho_0-\tau\sum_{k=1}^{N-1}\int_\Omega p^N_{k\tau}-\int_\Omega p^N_1,\end{aligned}$$

where the last equality comes from Lemma 4.3 which allows to make the link between the Kantorovich potential φ^N_0 and $\phi^N(0^+,\cdot)$. □

We want to pass to the limit $N\to+\infty$. To this extent, we rely on the fact that ϕ^N satisfies an explicit equation in the sense of distributions. We start to define the distribution which will be the r.h.s. of the Hamilton Jacobi equation.

Definition 4.5. *Let α^N and P^N the positive measures on $[0,1]\times\Omega$ defined as*

$$\begin{cases}\alpha^N &:= \tau\displaystyle\sum_{k=1}^{N-1}\delta_{t=k\tau}(p^N_{k\tau}+V_N)+\delta_{t=1}p^N_1,\\ P^N &:= \tau\displaystyle\sum_{k=1}^{N-1}\delta_{t=k\tau}p^N_{k\tau}+\delta_{t=1}p^N_1.\end{cases}$$

More precisely, for any test function $a\in C([0,1]\times\Omega)$,

$$\int_{[0,1]\times\Omega}a\,\mathrm{d}\alpha^N:=\tau\sum_{k=1}^{N-1}\int_\Omega a(k\tau,\cdot)\left(V_N+p^N_{k\tau}\right)+\int_\Omega a(1,\cdot)p^N_1,$$

and similarly for P^N.

In other words, α^N is, from the temporal point of view, a sum of delta function, each of them corresponding to the jump of the value function ϕ^N.

Proposition 4.6. *Provided that we set $\phi^N(0^-,\cdot)=\phi^N(0^+,\cdot)$ and $\phi^N(1^+,\cdot)=\Psi_N$, the following equation holds in the sense of distributions on $[0,1]\times\Omega$:*

$$-\partial_t\phi^N+\frac{1}{2}|\nabla\phi^N|^2\leqslant\alpha^N. \tag{4.9}$$

Proof. As the pressures and the potentials V_N,Ψ_N are Lipschitz, for any $t\in[0,1]$, the value function $\phi^N(t^+,\cdot)$ and $\phi^N(t^-,\cdot)$ are Lipschitz (but with a Lipschitz constant which may diverge as $N\to+\infty$).

Notice that on each interval $(((k-1)\tau)^+,(k\tau)^-)$, the function ϕ^N is defined by the Hopf-Lax formula, hence solves the Hamilton-Jacobi equation $-\partial_t\phi^N+\frac{1}{2}|\nabla\phi^N|^2=0$ a.e. [13, Section 3.3]. It implies that the inequality $-\partial_t\phi^N+\frac{1}{2}|\nabla\phi^N|^2\leqslant 0$

is also satisfied in the sense of distributions, as $\nabla\phi^N$ is bounded and $\partial_t\phi$ may have some singular parts, but they are positive.

Provided that we set $\phi^N(0^-,\cdot) = \phi^N(0^+,\cdot)$ and $\phi^N(1^+,\cdot) = \Psi_N$, the measure $\partial_t\phi^N$ has a singular negative part at $\{\tau, 2\tau, \dots, 1\}$ corresponding to the jumps of the function ϕ^N; but, given (4.3) and (4.5), the negative part of $\partial_t\phi^N$ is exactly $-\alpha^N$. □

The next step is to pass to the limit $N \to +\infty$. To this extent, we need uniform bounds on α^N, which derive easily from the bounds that we have on the pressure.

Lemma 4.7. *There exists a constant C, independent of N, such that $\alpha^N([0,1]\times\Omega) \leqslant C$ and $P^N([0,1]\times\Omega) \leqslant C$.*

Recall that both α^N and P^N are positive measures as we have chosen V_N in such a way that it is positive.

Proof. We know that the $p^N_{k\tau}$, for $k \in \{1,2,\dots,N\}$ have a gradient which is bounded uniformly in $L^2(\Omega)$. As moreover they all vanish on a set of measure at least $|\Omega| - 1$, they are bounded uniformly (w.r.t. N) in $L^1(\Omega)$. This is enough, in order to get the uniform bound on P^N. Given the way V_N is built, the one for α^N is a straightforward consequence of the one on P^N. □

Now that we have a bound on α^N, to get compactness on the sequence ϕ^N, we use the same kind of estimates used to prove existence of a solution in the dual at the continuous level, see for instance [11, Section 3]. We recall that $\mathcal{K}$, the set of admissible competitors for the dual problem, was defined in Definition 2.5.

Proposition 4.8. *There exists $(\bar\phi, \bar P) \in \mathcal{K}$ admissible for the dual problem such that*

$$\begin{cases} \lim\limits_{N\to+\infty} \phi^N = \bar\phi & \text{weakly in } \mathrm{BV}([0,1]\times\Omega) \cap L^2([0,1], H^1(\Omega)), \\ \lim\limits_{N\to+\infty} P^N = \bar P & \text{in } \mathcal{M}_+([0,1]\times\Omega). \end{cases}$$

Proof. Given Lemma 4.7, we know that P^N is bounded in $\mathcal{M}_+([0,1]\times\Omega)$ independently of N. Up to the extraction of a subsequence, it converges weakly as a measure to some $\bar P$. On the other hand, once we know this convergence, it is easy to see that α^N converges as a measure on $\mathcal{M}_+([0,1]\times\Omega)$ to $\bar P + V$.

We have assumed that V and Ψ are positive, and so are V_N and Ψ_N, independently of N. Using the definition of ϕ^N and the positivity of the pressures, it is not hard to see that ϕ^N is positive $[0,1]\times\Omega$. Integrating the Hamilton Jacobi equation w.r.t. space and time and using the bound on α^N (Lemma 4.7), we see that

$$\int_\Omega \phi^N(0^+,\cdot) - \int_\Omega \Psi_N + \frac{1}{2} \iint\limits_{[0,1]\times\Omega} |\nabla\phi^N|^2 \leqslant C. \tag{4.10}$$

Combined with the positivity of $\phi^N(0^+,\cdot)$ and a $L^1(\Omega)$ bound on Ψ_N, we see that $\nabla\phi^N$ is uniformly bounded in $L^2([0,1]\times\Omega)$.

It remains to get a bound on $\partial_t\phi^N$. Of course, as a measure, it can be decomposed as a positive and a negative part. The negative part is concentrated on the instants $\{\tau, 2\tau, \dots, 1\}$ as $\partial_t\phi^N \geqslant 0$ on the intervals $(((k-1)\tau)^+, (k\tau)^-)$. On the other hand, on $\{\tau, 2\tau, \dots, 1\}$, the temporal derivative $\partial_t\phi^N$ coincides with $-\alpha^N$, hence the negative part is bounded as a measure. On the other hand, given that

$$\iint\limits_{[0,1]\times\Omega} \partial_t\phi^N = \int_\Omega \Psi_N - \int_\Omega \phi^N(0^+,\cdot) \leqslant \int_\Omega \Psi_N$$

is bounded independently of N, we see that $(\partial_t\phi^N)_+ = \partial_t\phi^N + (\partial_t\phi^N)_-$ is also bounded as a measure.

As a consequence, up to the extraction of a subsequence we know that ϕ^N converges weakly in $\mathrm{BV}([0,1]\times\Omega) \cap L^2([0,1], H^1(\Omega))$ to some $\bar\phi$. This convergence allows easily to pass to the limit in the Hamilton-Jacobi equation satisfied (in the sense of distributions) by ϕ^N, hence $(\bar\phi, \bar P)$ is admissible in the dual problem. □

The last step, to show the optimality of the limit $(\bar\phi, \bar P)$, is to pass to the limit in (4.8).

Proposition 4.9. *The pair $(\bar\phi, \bar P) \in \mathcal{K}$ is a solution of the dual problem.*

Proof. We have already proved in Proposition 4.1 that

$$\lim_{N\to+\infty} \mathcal{A}^N(\bar\rho^N) = \min\left\{\mathcal{A}(\rho) \ : \ \rho\in\Gamma_0\right\}.$$

Given (4.8) and the duality result which holds for the continuous problem (Theorem 2.7), it is enough to show that

$$\limsup_{N\to+\infty}\left(\int_\Omega \phi^N(0^+,\cdot)\bar\rho_0 - \tau\sum_{k=1}^{N-1}\int_\Omega p^N_{k\tau} - \int_\Omega p^N_1\right) \leqslant \mathcal{B}(\bar\phi,\bar P) = \int_\Omega \bar\phi(0^+,\cdot)\bar\rho_0 - \bar p([0,1]\times\Omega).$$

The convergence of the term involving the pressure is quite easy to show. Indeed, given the positivity of the pressures,

$$\tau\sum_{k=1}^{N-1}\int_\Omega p^N_{k\tau} + \int_\Omega p^N_1 = P^N([0,1]\times\Omega) \to \bar p([0,1]\times\Omega)$$

by weak convergence. On the other hand, using the definition of the trace,

$$\int_\Omega \bar\phi(0^+,x)\bar\rho_0 = \lim_{t\to0}\frac{1}{t}\iint\limits_{[0,t]\times\Omega}\bar\phi(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x.$$

We fix some $t>0$. Due to the convergence of ϕ^N to $\bar\phi$, it clearly holds

$$\lim_{N\to+\infty}\frac{1}{t}\iint\limits_{[0,t]\times\Omega}\phi^N(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x = \frac{1}{t}\iint\limits_{[0,t]\times\Omega}\bar\phi(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x.$$

For the value function ϕ^N, we can use the information that we have on the temporal derivative, namely $\partial_t\phi^N \geqslant -\alpha^N$. It allows us to write, given the positivity of $\bar\rho_0$,

$$\begin{aligned}
\frac{1}{t}\iint\limits_{[0,t]\times\Omega}\phi^N(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x &= \frac{1}{t}\iint\limits_{[0,t]\times\Omega}\left(\phi^N(0^+,x)+\int_0^s\partial_t\phi^N(r,x)\,\mathrm{d}r\right)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x\\
&\geqslant \int_\Omega\phi^N(0^+,\cdot)\bar\rho_0 - \frac{1}{t}\iint\limits_{[0,t]\times\Omega}s\alpha^N(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x\\
&\geqslant \int_\Omega\phi^N(0^+,\cdot)\bar\rho_0(x) - \iint\limits_{[0,t]\times\Omega}\alpha^N(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x.
\end{aligned}$$

Now, recall that $\bar\rho_0\leqslant 1$ and α^N converges as a measure to $\bar p + V$, hence

$$\frac{1}{t}\iint\limits_{[0,t]\times\Omega}\bar\phi(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x = \lim_{N\to+\infty}\frac{1}{t}\iint\limits_{[0,t]\times\Omega}\phi^N(s,x)\bar\rho_0(x)\,\mathrm{d}s\,\mathrm{d}x \geqslant \limsup_{N\to+\infty}\left(\int_\Omega\phi^N(0^+,\cdot)\bar\rho_0\right) - (\bar p+V)([0,t]\times\Omega).$$

Now we send t to 0, and use the fact that $(\bar p+V)(\{0\}\times\Omega)=0$ (this can be seen as a consequence of Corollary 3.6) to conclude that

$$\limsup_{N\to+\infty}\int_\Omega\phi^N(0^+,\cdot)\bar\rho_0 \leqslant \int_\Omega\bar\phi(0^+,x)\bar\rho_0,$$

which gives us the announced result. □

To reach the conclusion of our main theorem, it is enough to show that $\bar P$ has the regularity we announced. But this easily derives from the weak convergence of P^N to $\bar P$ and the estimates of Corollary 3.6 and Corollary 3.11.

Proof of Theorem 2.8. For any smooth test functions a, b (with a being real-valued and b being vector valued), given the convergence of P^N to $\bar{p}$ it holds

$$\begin{cases} \displaystyle\iint_{[0,1]\times\Omega} a\bar{P} & = \lim_{N\to+\infty} \left(\sum_{k=1}^{N-1} \tau \int_\Omega a(k\tau,\cdot) p^N_{k\tau} + \int_\Omega a(1,\cdot) p^N_1 \right), \\ \displaystyle\iint_{[0,1]\times\Omega} (\nabla\cdot b)\bar{P} & = -\lim_{N\to+\infty} \left(\sum_{k=1}^{N-1} \tau \int_\Omega g(k\tau,\cdot)\cdot\nabla(p^N_{k\tau}) + \int_\Omega b(1,\cdot)\cdot\nabla p^N_1 \right), \end{cases}$$

where on the second line we have done an integration by parts in the r.h.s. Using the estimates given by Corollary 3.6 and Corollary 3.11, it is clear that

$$\iint_{[0,1]\times\Omega} (\nabla\cdot b)\bar{P} \leqslant \int_0^1 \left(\|b(t,\cdot)\|_{L^2(\Omega)} \|\nabla V\|_{L^2(\Omega)} \right) \mathrm{d}t + \|b(1,\cdot)\|_{L^2(\Omega)} \|\nabla\Psi\|_{L^2(\Omega)}.$$

On the other hand, if $V, \Psi \in W^{1,q}$ with $q > d$ then, using the same propositions,

$$\iint_{[0,1]\times\Omega} a\bar{P} \leqslant C \int_0^1 \left(\|a(t,\cdot)\|_{L^1(\Omega)} \right) \mathrm{d}t + \|a(1,\cdot)\|_{L^\infty(\Omega)},$$

where C depends only on $\|\nabla V\|_{L^q(\Omega)}$, $\|\nabla\Psi\|_{L^q(\Omega)}$ and Ω. Standard functional analysis manipulations provide the conclusion of Theorem 2.8. □

Acknowledgments

Both authors acknowledge the support of the French ANR via the contract MFG (ANR- 16-CE40-0015-01) and benefited from the support of the FMJH "Program Gaspard Monge for Optimization and operations research and their interactions with data science" and EDF via the PGMO project VarPDEMFG.

References

[1] L. Ambrosio, P. Tilli. *Topics on analysis in metric spaces*. Vol. 25. Oxford University Press on Demand, 2004.

[2] L. Ambrosio, A. Figalli, On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations, *Calc. Var. PDE*, 31 (2008) No. 4, 497-509.

[3] L. Ambrosio, A. Figalli, Geodesics in the space of measure-preserving maps and plans, *Arch. Rational Mech. Anal.*, 194 (2009), 421-462.

[4] L. Ambrosio, N. Gigli, G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Springer Science , 2008.

[5] A. Baradat, L. Monsaingeon Small noise limit and convexity for generalized incompressible flows, Schrödinger problems, and optimal transport, preprint 2018, arXiv:1810.12036.

[6] J. D. Benamou, G. Carlier, F. Santambrogio, Variational Mean Field Games, in *Active Particles, Volume 1: Theory, Models, Applications*, edited by N. Bellomo, P. Degond, E. Tadmor, 141–171, 2017.

[7] Y. Brenier, Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations, *Comm. Pure Appl. Math.*, 52 (1999) 4, 411–452.

[8] P. Cardaliaguet, Notes on Mean Field Games, unublished, available online at `https://www.ceremade.dauphine.fr/~cardalia/MFG20130420.pdf`

[9] P. Cardaliaguet, Weak solutions for first order mean field games with local coupling. *Analysis and geometry in control theory and its applications*. Springer, Cham, 2015. 111-158.

[10] P. Cardaliaguet, P. J. Graber. Mean field games systems of first order. *ESAIM: Control, Optimisation and Calculus of Variations* 21.3 (2015): 690-722.

[11] P. Cardaliaguet, A. R. Mészáros, F. Santambrogio First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization* 54.5 (2016): 2672-2709.

[12] P. Cardaliaguet, A. Porretta, D. Tonon. Sobolev regularity for the first order Hamilton–Jacobi equation. *Calculus of Variations and Partial Differential Equations* 54.3 (2015): 3037-3065.

[13] L. C. Evans *Partial differential equations.* AMS Graduate Studies in Mathematics, 2010 (2nd edition).

[14] M. Huang, R.P. Malhamé, P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle, *Communication in information and systems*, 6 (2006), No. 3, 221-252.

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

[16] H. Lavenant, Time-convexity of the entropy in the multiphasic formulation of the incompressible Euler equation, *Calculus of Variations and PDEs*, Vol 56 no 6 (2017), 170.

[17] J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal, *C. R. Math. Acad. Sci. Paris*, 343 (2006), No. 10, 679–684.

[18] J.-M. Lasry, P.-L. Lions, Mean field games, *Jpn. J. Math.*, 2 (2007), no. 1, 229–260.

[19] P.-L. Lions, *Cours au Collège de France*, `www.college-de-france.fr`.

[20] H. Lavenant, F. Santambrogio. Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational Mean Field Games. *Communications in Partial Differential Equations*, Vol 43 no 12 (2018), p. 1761–1802.

[21] G. Loeper The reconstruction problem for the Euler-Poisson system in cosmology. *Archive for rational mechanics and analysis* 179.2 (2006): 153-216.

[22] B. Maury, A. Roudneff-Chupin, F. Santambrogio, A macroscopic crowd motion model of gradient flow type. *Mathematical Models and Methods in Applied Sciences* 20.10 (2010): 1787-1821.

[23] J. Moser A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations.*Communications on Pure and Applied Mathematics* 13.3 (1960): 457-468.

[24] F. Santambrogio, A modest proposal for MFG with density constraints, *Netw. Heterog. Media*, 7 (2012) No. 2, 337–347.

[25] F. Santambrogio, Crowd motion and population dynamics under density constraints, to appear in the Proceedings of the SMAI 2017 conference, in (*ESAIM Proceedings and Surveys*).

[26] F. Santambrogio, *Optimal transport for applied mathematicians: calculus of variations, PDEs, and modeling*. Vol. 87. Birkhäuser, 2015.

[27] C. Villani *Topics in optimal transportation*. No. 58. American Mathematical Soc., 2003.