Optimal density evolution with congestion: L infinity bounds via flow interchange techniques and applications to variational Mean Field Games

Hugo Lavenant, Filippo Santambrogio

▶ **To cite this version:**

Hugo Lavenant, Filippo Santambrogio. Optimal density evolution with congestion: L infinity bounds via flow interchange techniques and applications to variational Mean Field Games. 2017. hal-01522084v1

HAL Id: hal-01522084

https://hal.science/hal-01522084v1

Preprint submitted on 16 May 2017 (v1), last revised 28 Sep 2017 (v2)

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.

Optimal density evolution with congestion: L^∞ bounds via flow interchange techniques and applications to variational Mean Field Games

Hugo Lavenant and Filippo Santambrogio

May 16, 2017

Abstract

We consider minimization problems for curves of measure, with kinetic and potential energy and a congestion penalization, as in the functionals that appear in Mean Field Games with a variational structure. We prove L^∞ regularity results for the optimal density, which can be applied to the rigorous derivations of equilibrium conditions at the level of each agent's trajectory, via time-discretization arguments, displacement convexity, and suitable Moser iterations. Similar L^∞ results have already been found by P.-L. Lions in his course on Mean Field Games, using a proof based on the use of a (very degenerate) elliptic equation on the dual potential (the value function) φ, in the case where the initial and final density were prescribed (planning problem). Here the strategy is highly different, and allows for instance to prove local-in-time estimates without assumptions on the initial and final data, and to insert a potential in the dynamics.

Contents

1 Introduction

The problem of optimal density evolution with congestion is a very natural question where an initial distribution ρ_0 of mass (particles, individuals...) is given, and it has to evolve from time $t = 0$ to time $t = T$ by minimizing an overall energy. This typically involves its kinetic energy expenditure and a cost depending on congestion effects, i.e. on how much it is concentrated along its trajectory. Then, at time $t = T$, either the final configuration ρ_T is prescribed, or a final cost depending on ρ_T is also considered.

When ρ_T is fixed and no congestion effect is present, the only quantity to be minimized is the kinetic energy and this boils down to what is usually known as the dynamic formulation of the optimal transport problem, studied by Benamou and Brenier in [7]. From the fluid mechanics point of view, this model corresponds to that of particles of a pressureless gas moving without acceleration in straight lines, and without interaction with each other. From the geometric point of view, this variational problem consists in looking for a geodesic in the Wasserstein space W_2 (for references on optimal transport and Wasserstein spaces, see [32, 30]). Inserting congestion effects corresponds to looking at deformed geodesics, i.e. curves which are optimal for other criteria which do not involve only their length (weighted lengths, length + penalizations...), and to add pressure terms in the corresponding gas equations. For instance, in [10, 4] geodesics in the Wasserstein space for different weights, minimizing energies of the form $\int E(\rho(t))|\dot\rho_t|\,\mathrm{d}t$ were considered, including cases where E penalized congestion. Yet, the case which is now the most studied is the one where a penalization on congestion is added to the kinetic energy, thus minimizing $\int(|\dot\rho_t|^2 + E(\rho(t)))\,\mathrm{d}t$, as it was done in [13]. Since this is the kind of problems this paper will be devoted to, it is important to clarify precisely its form. We can either look for a curve $\rho : [0,T] \to \mathcal{P}(\Omega)$ which minimizes

$$\rho \mapsto \int_0^T \frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t + \int_0^T E(\rho_t)\,\mathrm{d}t + \Psi(\rho_T) \tag{1.1}$$

with ρ_0 prescribed (here $|\dot\rho_t|$ is the metric derivative, i.e. the speed of this curve for the distance W_2, see Section 2.2), or look for a pair $(\rho, \mathbf{v})$ minimizing

$$\rho \mapsto \int_0^T \int_\Omega \frac{1}{2}\rho_t|\mathbf{v}_t|^2\,\mathrm{d}t + \int_0^T E(\rho_t)\,\mathrm{d}t + \Psi(\rho_T) \tag{1.2}$$

under the same constraint on ρ_0 and a differential constraint $\partial_t\rho_t + \nabla\cdot(\rho_t\mathbf{v}_t) = 0$. Here $\Omega \subset \mathbb{R}^d$ is a bounded and connected domain and the continuity equation is satisfied in the weak sense on $[0,T]\times\mathbb{R}^d$ (which corresponds to imposing no-flux boundary conditions on $\partial\Omega$), The equivalence between the two formulations essentially comes from [7] and from the characterization of absolutely continuous curves in the Wasserstein space studied in [3].The functional E usually takes the form of an integral functional such as

$$E(\rho) := \int_\Omega f(\rho(x))\,\mathrm{d}x + \int_\Omega V(x)\rho(x)\,\mathrm{d}x,$$

for a convex function f and a given potential V (where we identify measures with their densities; for the definition for measures which are not absolutely continuous, see Section 2.3). The final penalization Ψ can be either a functional of the same form of E, or a constraint which prescribes ρ_T.

The interest for this minimization problem, which is already very natural in itself, has increased a lot after the introduction in 2006 of the theory of Mean Field Games (MFG) (introduced essentially at the same time by Lasry and Lions, [20, 21, 22], and by Huang, Malhamé and Caines [18]). In the easiest version of these games, we consider a population of agents where everybody chooses its own trajectory, solving

$$\min \int_0^T \left(\frac{|x'(t)|^2}{2} + V(x(t)) + g(\rho_t(x(t)))\right)\,\mathrm{d}t + \Psi(x(T)), \tag{1.3}$$

with given initial point $x(0)$. Here g is a given increasing function of the density ρ_t at time t, i.e. an individual cost for each agent penalizing congested areas. The difficulty in the model is that every agent optimizes given the density of all agents ρ_t, but this density depends on turn on the choices of all the agents. An equilibrium problem arises, and we look for a Nash equilibrium in this continuum game (with infinitely many negligible players, who move continuously in time in a continuous space). This can be translated into a system of PDEs

$$\begin{cases} -\partial_t\varphi + \frac{|\nabla\varphi|^2}{2} = V(x) + g(\rho), \\ \partial_t\rho - \nabla\cdot(\rho\nabla\varphi) = 0, \\ \varphi(T,x) = \Psi(x), \quad \rho(0,x) = \rho_0(x). \end{cases}$$

This forward-backward system is composed of a Hamilton-Jacobi equation for the value function φ of the above optimization problem where the density ρ appears at the right-hand side and of a continuity equation for ρ which is advected by the vector field $\mathbf{v} = -\nabla\varphi$. Taking the gradient of the HJ equation gives a formula for the Lagrangian acceleration

$$\partial_t \mathbf{v}_t + (\mathbf{v}_t \cdot \nabla)\mathbf{v}_t = \nabla(V + g(\rho_t)),$$

where $g(\rho_t)$ plays the role of a pressure to be added to the potential V, as it is typical in compressible fluid mechanics. Alternatively, the same equilibrium problem can be formulated in terms of a probability measure Q on the set $H^1([0,T];\Omega)$ of paths valued in Ω, defining $\rho_t = (e_t)_\# Q$ (where $e_t : H^1([0,T];\Omega) \to \Omega$ is the evaluation map at time t), and requiring $(e_0)_\# Q = \rho_0$ and that Q-a.e. curve is optimal for (1.3) with this choice of ρ_t. For a general introduction to mean field games, other than the papers by Lasry-Lions and Huang-Malhamé-Caines, the reader can consult the lecture notes by Cardaliaguet [15], based on the lectures given by P.-L. Lions at Collège de France between 2006 and 2012 ([24]). In particular, for the simple model which is the object of this paper, and which is deterministic and first-order (no random effect in the motion of the agents, and no diffusion in the equations), we also refer to [16].

The remarkable fact is that this class of equilibrium problem has a variational origin, and one can find an equilibrium by minimizing (1.2) choosing $E(\rho) := \int f(\rho) + V(x)\rho$ with $f' = g$ (for a review on variational mean field games and on these questions, we refer to [8]). The optimality condition on the optimal $(\rho, \mathbf{v})$ will indeed show that we have $\mathbf{v} = -\nabla\varphi$ where φ solves the HJ part of (1), thus getting a solution of the system. The same can also be formally formulated in terms of probabilities Q on the set of path.

Yet, these considerations are essentially formal and not rigorous, so far. Indeed, the difficulty is the following: the function $h(t,x) := V(x) + g(\rho_t(x))$ is obtained from the density of a measure, and hence it is only defined a.e. Integrating it on a curve, as we do when we consider the action $\int_0^T h(t,x(t))\,\mathrm{d}t$ in (1.3) has absolutely no meaning! Of course, it would be different if we could prove some regularity (for instance, continuity) on ρ_t. The question of the regularity in mean field games is a very challenging one and deserves high attention. In [17] a stategy to overcome this difficulty, taken from [2], is used: indeed, it is sufficient to chose a suitable representative of h to give a precise meaning to the integral of h on a curve, and the correct choice is

$$\hat{h}(t,x) := \limsup_{r\to 0} h_r(t,x) := \fint_{B(x,r)} h(t,y)\,\mathrm{d}y;$$

to prove that Q is concentrated on optimal curves for $\hat{h}$ it is then enough to write estimates with h_r and then pass to the limit as $r \to 0$. This requires an upper bound on h_r, and the natural assumption is to require that the maximal function $Mh := \sup_r h_r$ is L^1 in space and time. Thanks to well-known results in harmonic analysis, $h \in L^1$ is not enough for this but $h \in L^m$ for $m > 1$ is instead enough. Once integrability of Mh is obtained, then one can say that the optimal measure Q is concentrated on curves which minimize in (1.3) in the class of curves $x(\cdot)$ such that $\int_0^T Mh(t,x(t))\,\mathrm{d}t < +\infty$. These curves are almost all curves in a suitable sense, thanks to the integrability properties of Mh in space-time, but they are in general *not all* curves.

It is interesting to observe that the strategy of [17] and [2] was first used in the framework of variational models for the incompressible Euler equation, in the sense of Brenier [11, 12]. Indeed, the problem of incompressible evolution has many similarities with the one of evolution with congestion effects, with the only difference that instead of penalizing high densities there is a constraint $\rho = 1$. Also, the precise mean field game studied in [17] is of very similar nature, since it included the constraint $\rho \leqslant 1$. Moreover, the techniques used in [17] to prove this extra summability of h come from the incompressible Euler framework: they are techniques based on convex duality taken from [12] and later improved in [1], which allow, in this case, to prove $h \in L^2_{loc}((0,T); BV_{loc}(\Omega))$. In the framework of more standard mean field games (i.e. with density penalization instead of constraints), the same technique (presented in more generality on some simpler examples in [31]) has been used in [29] to prove H^1 regularity results on the density ρ.

In the present paper, we present L^∞ bounds on the optimal ρ. For applications to MFG, whenever L^∞ results are available, it is possible to avoid all the assumptions on the maximal function Mh and obtain optimality in the larger class of all competing curves. This explains the interest of these results for MFG, but of course the reader can easily guess that they are interesting in themselves for the variational problem.

The question of the L^∞ regularity of ρ was already studied, in the MFG framework, by P.-L. Lions (see the second hour of the video of the lecture of November 27, 2011, in [24]), but the analysis was limited to global results when both ρ_0 and ρ_T are fixed and L^∞, and no potential V is considered. The technique was essentially

taken from degenerate elliptic PDEs (note that adapting from global to local results would be very difficult, without strong assumptions on the degeneracy and growth of the corresponding equation). Here what we do is different. The technique is based on the time-discretization of (1.1) in the form

$$\min\left\{\sum_{k=1}^{N}\frac{W_2^2(\rho_{(k-1)\tau},\rho_{k\tau})}{2\tau}+\sum_{k=1}^{N-1}\tau E(\rho_{k\tau})+\Psi(\rho_{N\tau})\right\}, \tag{1.4}$$

where $\tau = T/N$. The interesting fact is that, as a necessary optimality condition, each measure $\rho_{k\tau}$ with $0 < k < N$ minimizes

$$\rho \;\mapsto\; \frac{W_2^2(\rho_{(k-1)\tau},\rho)}{2\tau}+\frac{W_2^2(\rho_{(k+1)\tau},\rho)}{2\tau}+\tau E(\rho),$$

which is very similar to what we see in the so-called Jordan-Kinderlehrer-Otto scheme for the gradient flow of the functional E (see [19] and [3]). The main difference is that we have now two Wasserstein terms, one referring to the distance to the previous measure and one to the next one. Techniques from the JKO scheme can be used, and in particular the so-called *flow-interchange* technique (introduced in [25]). Essentially, this technique consists in evaluating how much decreases another energy U along the gradient flow of E. In the JKO framework, it is usually used to obtain estimates of the form

$$U(\rho_{k\tau})-U(\rho_{(k+1)\tau})\geqslant\tau\int(\text{something positive}),$$

which allows to say that U is decreasing and to obtain integral estimates on the right hand side (r.h.s.) above. In this variational framework, which corresponds to a second-order-in-time equation, instead of monotonicity we obtain convexity:

$$\frac{U(\rho_{(k-1)\tau})+U(\rho_{(k+1)\tau})-2U(\rho_{k\tau})}{\tau^2}\geqslant\int(\text{something positive})$$

(more precisely: the integral term in the r.h.s. is nonnegative if $V = 0$ and extra lower-order terms appear in presence of a potential V). This allows for instance to obtain convexity in time of all the quantities of the form $\int\rho_t^m(x)\,\mathrm{d}x$ when $V = 0$ (a similar technique was used with similar results by the first author in [23]). A global L^∞ result if $\rho_0,\rho_T\in L^\infty$ are fixed is then easy to deduce in this case (actually, we will not even state it explicitly in this paper). Moreover, using the structure of the right-hand side and with tedious iterations inspired by Moser [27], we are also able to provide interior L^∞ regularity independent of the boundary data, and regularity on intervals of the form $[t_1,T]$ under some assumptions on the penalization Ψ. This very result is, by the way, the natural one for MFG applications, and improves upon the results announced in [24].

The paper is organized as follows: after this brief introduction Section 1 also contains a short summary of the main ideas of the proof, so that the reader does not get lost in the technical details. Then, in Section 2 we summarize the preliminaries about curves and functionals on the Wasserstein space, and give a precise statement for the variational problem we consider and the results we prove, distinguishing into two cases depending on the convexity of the congestion penalization f (in terms of lower bounds on f''). In Section 3 we present and prove the estimates that are obtained in this framework via the flow interchange technique. These estimates allow to bound increasing L^m norms of the solution, and in Section 4 we explain how to iterate in order to transform them into L^∞ estimates on the limit of the discretized problems. This involves a technical difficulty, as one needs a reverse Jensen inequality in time (passing from $\int||\rho_t||_{L^{\beta m}}^m\,\mathrm{d}t$ to $\left(\int||\rho_t||_{L^{\beta m}}^{\beta m}\,\mathrm{d}t\right)^{1/\beta}$); this can be fixed because we already proved a convexity-like property for $t\mapsto||\rho_t||_{L^{\beta m}}^{\beta m}$ but is quite technical. In Section 5 we detail how to pass to the limit from the time-discretization to the continuous problem, and in the Appendix we give a proof of the reverse Jensen inequality.

1.1 Structure of the proof

As the structure of the proof of L^∞ bounds may be hidden behind the technical details, we sketch in this subsection the formal computations on which our main results rely. Let us consider the simplest case, the one where there is no interior potential, and let us not worry about the temporal boundary terms for the moment. The variational problem reads

$$\min\left\{\int_0^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t+\int_0^T\int_\Omega f(\rho_t(x))\,\mathrm{d}x\,\mathrm{d}t+\Psi(\rho_T)\;:\;\rho:[0\,,T]\to\mathcal{P}(\Omega),\;\rho_0=\overline{\rho_0}\right\}.$$

First, we consider the time-discretization in (1.4) and, we apply the flow-interchange technique that we mentioned before (and that will be detailed later, see Section 3) to the functional $U = U_m$, where

$$U_m(\rho) := \frac{1}{m(m-1)} \int_\Omega \rho(x)^m \, \mathrm{d}x,$$

when $m > 1$ ($U_1(\rho)$ can be defined as the Boltzmann entropy of ρ, and the normalization constants are chosen for coherence with this case). The flow interchange technique will allow to obtain an estimate of the form

$$\frac{U_m(\rho_{(k-1)\tau}) + U_m(\rho_{(k+1)\tau}) - 2U_m(\rho_{k\tau})}{\tau^2} \geqslant C(m) \int f''(\rho_{k\tau}) \rho_k^{m-1} |\nabla \rho_{k\tau}|^2 \geqslant 0. \tag{1.5}$$

This gives a discrete-in-time convex behavior for the quantity $U_m(\rho_t)$. In the case $f = 0$, this is basically a restatement of McCann convexity principle [26]. It is easy to see that, if $\rho_0, \rho_T \in L^m$ (which is an assumption on ρ_0 and on Ψ), then automatically the same L^m bound is satisfied by all measures ρ_t. If $\rho_0, \rho_T \in L^\infty$, the same easily passes to the limit as $m \to \infty$ thus providing L^∞ bounds.

Moreover, the flow interchange applied to the last time step $k = N$ (with $N\tau = T$), gives a result of the form

$$\frac{U_m(\rho_{N\tau}) - U_m(\rho_{(N-1)\tau})}{\tau} \leqslant b(m) U_m(\rho_{N\tau}), \tag{1.6}$$

where the constant $b(m)$ depends on the penalization Ψ. This acts as a sort of Neumann condition for the function $t \mapsto U_m(\rho_t)$ and allows to obtain the bound on $U_m(\rho_t)$ with the only assumption $\rho_0 \in L^m$, with no need to assume the same for ρ_T. However, this can only be adapted to the limit $m \to \infty$ in the case where Ψ has the form $\Psi(\rho) := \int g(\rho(x)) \, \mathrm{d}x$ for a convex g, without potential terms, so that $b(m) = 0$. Otherwise, the dependence of $b(m)$ upon m prevents from letting $m \to \infty$.

Our paper includes L^m and L^∞ results which do not require assumptions on ρ_0, and which will be, of course, only of local nature on $(0, T]$. Our proof will look like Moser's proof of regularity for elliptic equations [27], as it will rely on a fine analysis of the growth (when $m \to +\infty$) of quantities of the form $\int_{T_1}^{T_2} \rho_t^m$. Indeed, one can guess from (1.5) that we may write (at the limit when $\tau \to 0$)

$$\frac{d^2}{dt^2} U_m(\rho_t) \geqslant \int_\Omega |\nabla \rho_t|^2 \rho_t^{m-1} f''(\rho_t). \tag{1.7}$$

To estimate more precisely the r.h.s. of (1.7), a natural assumption is $f''(s) \geqslant s^\alpha$ (with α which could be negative, of course): if this is the case, one can check that the integrand of the r.h.s. is larger than $|\nabla(\rho_t^{(m+1+\alpha)/2})|^2$ (up to a constant depending polynomially in m). Using the Sobolev injection $H^1 \hookrightarrow L^{2d/(d-2)}$, one can conclude (neglecting the 0-order term of the H^1 norm of $\rho_t^{(m+1+\alpha)/2}$), with $1 < \beta < d/(d-2)$, that

$$C(m) \frac{d^2}{dt^2} U_m(\rho_t) \geqslant \left(\int_\Omega \rho_t^{\beta(m+1+\alpha)} \right)^{1/\beta}$$

In the case $\alpha \geqslant -1$, we see that the r.h.s. is larger than $U_{\beta m}(\rho_t)^{1/\beta}$. In other words, we have obtained a control of $U_{\beta m}(\rho)$ in terms of $U_m(\rho)$. Such a control can be iterated. If we take a positive cutoff function χ which is equal to 1 on $[T_1 - \varepsilon, T_2 + \varepsilon]$ and which is null outside $[T_1 - 2\varepsilon, T_2 + 2\varepsilon]$, multiply (1.7) by χ and integrate the left hand side (l.h.s.) by parts twice, we can say that

$$\int_{T_1-\varepsilon}^{T_2+\varepsilon} U_{\beta m}(\rho_t)^{1/\beta} \, \mathrm{d}t \leqslant C(m, \varepsilon) \int_{T_1-2\varepsilon}^{T_2+2\varepsilon} U_m(\rho_t) \, \mathrm{d}t,$$

where the constant $C(m, \varepsilon)$ grows not faster than a polynomial function of m and ε^{-1}. We have to work a little bit more on the l.h.s. because we want to exchange the power $1/\beta$ and the integral sign, and unfortunately Jensen's inequality gives it the other way around. To this extent, we rely on the following observation: as the function $U_{\beta m}$ is convex (this can be seen in (1.7)) and positive, it is bounded on $[T_1, T_2]$ either by its values on $[T_1, T_1 - \varepsilon]$ or on $[T_2, T_2 + \varepsilon]$, thus we have a "reverse Jensen's inequality"

$$\left(\int_{T_1}^{T_2} U_{\beta m}(\rho_t) \, \mathrm{d}t \right)^{1/\beta} \leqslant \frac{(T_2 - T_1)^{1/\beta}}{\varepsilon} \left(\int_{T_1-\varepsilon}^{T_1} U_{\beta m}(\rho_t)^{1/\beta} \, \mathrm{d}t + \int_{T_2}^{T_2+\varepsilon} U_{\beta m}(\rho_t)^{1/\beta} \right).$$

Combining this inequality with the estimation we have on the r.h.s., we deduce that

$$\left(\int_{T_1}^{T_2} U_{\beta m}(\rho_t)\,\mathrm{d}t\right)^{1/\beta} \leqslant C(m,\varepsilon)\int_{T_1-2\varepsilon}^{T_2+2\varepsilon} U_m(\rho_t)\,\mathrm{d}t,$$

where the new constant $C(m,\varepsilon)$ has also a polynomial behavior in m and ε^{-1}. This estimation is ready to be iterated. Indeed, setting $m_n := \beta^n m_0$ and $\varepsilon_n = 2^{-n}\varepsilon_0$, given the moderate growth of $C(m,\varepsilon)$, it is not difficult to conclude that

$$\limsup_{n\to+\infty}\left(\int_{T_1-\varepsilon_n}^{T_2+\varepsilon_n} U_{m_n}(\rho_t)\,\mathrm{d}t\right)^{1/m_n} < +\infty.$$

As the l.h.s. controls the L^∞ norm of ρ on $[T_1\,,T_2]\times\Omega$, this is enough to conclude that ρ is bounded locally in time and globally in space.

Let us comment the technical refinements and generalization of the above argument that are used in the present article:

- As we do not have enough time regularity to differentiate twice w.r.t. time, we decided to work with a time discretization of the problem. Hence, instead of(1.7) we use (1.5).
- If we add an interior potential, the r.h.s. of (1.7) contains lower order terms that are controlled by the term involving f''. However, the sign of the l.h.s. is no longer known and the function U_m is no longer convex but rather satisfies
$$\frac{d^2}{dt^2}U_m(\rho_t) + \omega^2 U_m(\rho_t) \geqslant 0,$$
where ω grows linearly with m. In particular, the "reverse Jensen inequality" becomes more difficult to prove, but it is still doable.
- With assumptions on the final penalization, the regularity can be extended to the final time. More precisely, if we assume that the final penalization is given by the sum of a potential term and a congestion term, then formally (and this can be proven by taking the limit $\tau\to 0$ of (1.6)),
$$\left.\frac{d}{dt}U_m(\rho_t)\right|_{t=T} \leqslant b(m)U_m(\rho_T), \tag{1.8}$$
where the constant $b(m)$ depends on the potential and can be taken equal to 0 if there is no potential. This inequality enables to control the value of U_m at the boundary $t=T$ by its values in the interior. Thus the same kind of iterations can be performed and gives L^∞ regularity up to the boundary.
- If $\alpha < -1$, we only have a control of U_m by $U_{\beta(m+1+\alpha)}$. Thus we must start the iterative procedure with a value m such that $m < \beta(m+1+\alpha)$, i.e. we must impose *a priori* some L^m regularity on ρ (with a m which depends on α and β, the latter depending itself only on the dimension of the ambient space). Such a regularity is imposed by assuming that ρ_0 (which is fixed) is in $L^m(\Omega)$ and that the boundary penalization in $t=T$ is the sum of a potential and a congestion term. Indeed, if this is the case, the boundary condition (1.8) combined with the interior estimate (1.7) shows that if T is small enough (given the potentials and the congestion function f), the L^m norm of ρ on $[0\,,T]\times\Omega$ must be bounded.

2 Notation and presentation of the optimal density evolution problem

In all the sequel, Ω will denote the closure of an open bounded convex domain of $\mathbb{R}^d$ with smooth boundary. To avoid normalization constants, we will assume that its Lebesgue measure is 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 $\mathcal{P}(\Omega)$. The Lebesgue measure restricted to Ω, which is therefore a probability measure, will be denoted by $\mathcal{L}$. The 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 $\mathbb{R}$).

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 by

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

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 $B\subset Y$. It can also be defined by

$$\int_Y a(y)\,\mathrm{d}(T\#\mu)(y) := \int_X a(T(x))\,\mathrm{d}\mu(x),$$

this identity being valid as soon as $a : Y\to\mathbb{R}$ is any integrable function. For general results about optimal transport, the reader might refer to [32] or [30]. We recall that W_2 admits a dual formulation: for any $\mu,\nu\in\mathcal{P}(\Omega)$,

$$W_2(\mu,\nu) = \sqrt{\max\left\{\int\varphi\,\mathrm{d}\mu+\int\varphi^c\,\mathrm{d}\nu \;:\; \varphi\in C(\Omega)\right\}}, \tag{2.2}$$

where $\varphi^c(y) := \inf_{x\in\Omega}(|x-y|^2-\varphi(x))$ for any $y\in\Omega$. A function $\varphi\in C(\Omega)$ which is optimal in (2.2) is called a Kantorovitch potential for the transport from μ to ν. The following result, giving the derivative of the Wasserstein distance, can be found in [30, Propositions 7.18 and 7.19].

Proposition 2.1. *Let $\mu,\nu\in\mathcal{P}(\Omega)$ and assume that μ is absolutely continuous w.r.t. $\mathcal{L}$ and that its density is strictly positive a.e. Then there exists a unique Kantorovitch potential φ for the transport from μ to ν. Moreover, φ is Lipschitz and if $\tilde\mu\in\mathcal{P}(\Omega)\cap L^\infty(\Omega)$, then*

$$\lim_{\varepsilon\to0}\frac{W_2^2((1-\varepsilon)\mu+\varepsilon\tilde\mu,\nu)-W_2^2(\mu,\nu)}{\varepsilon} = \int_\Omega\varphi\,\mathrm{d}(\tilde\mu-\mu).$$

We recall that W_2 defines a metric on $\mathcal{P}(\Omega)$ that metrizes the weak-* topology. Therefore, thanks to Prokhorov Theorem, the space $(\mathcal{P}(\Omega),W_2)$ is a compact metric space. We also recall that $(\mathcal{P}(\Omega),W_2)$ is a geodesic space. If μ and ν are probability measures such that μ admits a strictly positive density w.r.t. $\mathcal{L}$, then there exists a unique constant-speed geodesic $\rho : [0\,,1]\to\mathcal{P}(\Omega)$ joining μ to ν and it is given by

$$\rho(t) = (\mathrm{Id}-t\nabla\varphi)\#\mu,$$

where φ is the unique Kantorovitch potential for the transport from μ to ν.

We will need to define functionals of the form $\mu\in\mathcal{P}(\Omega)\mapsto\int_\Omega h(\mu)\,\mathrm{d}\mathcal{L}$. To this extent, we rely on the following proposition (see [30, Chapter 7]; see also [9] for the most advanced results on the semicontinuity of this kind of functionals on measures)

Proposition 2.2. *Let $h : [0\,,+\infty)\to\mathbb{R}$ be a convex function bounded from below. Let $h'(+\infty)\in(-\infty\,,+\infty]$ be the limit of $h'(t)$ as $t\to+\infty$. Then, the functional*

$$\rho\in\mathcal{P}(\Omega)\mapsto\int_\Omega h(\rho^{ac})+h'(+\infty)\rho^{sing}(\Omega), \tag{2.3}$$

(where $\rho =: \rho^{ac}\mathcal{L}+\rho^{sing}$ is the decomposition of ρ as an absolutely continuous part $\rho^{ac}\mathcal{L}$ and a singular part ρ^{sing} w.r.t. $\mathcal{L}$) is convex and l.s.c.

In particular, we will make a strong use of the following functionals.

Definition 2.3. *For any $m\geqslant1$, we define $u_m : [0\,,+\infty)\to\mathbb{R}$ for any $t\geqslant0$ through*

$$u_m(t) := \begin{cases} t\ln t+1 & \text{if } m=1\\ \dfrac{t^m}{m(m-1)} & \text{if } m>1\end{cases}.$$

For any $m \geqslant 1$, the functional $U_m : \mathcal{P}(\Omega) \to \mathbb{R}$ is defined, for $\rho \in \mathcal{P}(\Omega)$, via

$$U_m(\rho) := \begin{cases} \int_\Omega u_m(\rho) & \text{if } \rho \text{ is absolutely continuous w.r.t. } \mathcal{L} \\ +\infty & \text{else} \end{cases}.$$

One can notice that $u_m''(t) = t^{m-2}$ for any $m \geqslant 1$ and any $t > 0$, hence the functions u_m are convex for all m. One can also notice that U_1 is (up to an additive constant) the entropy w.r.t. $\mathcal{L}$. Moreover, some useful properties of U_m are summarized below.

Proposition 2.4. *For any $m \geqslant 1$,*

1. *One has $m^2 U_m \geqslant 1$.*
2. *The functional U_m is convex and l.s.c.*
3. *The functional U_m is geodesically convex: it is convex along every constant-speed geodesic of $(\mathcal{P}(\Omega), W_2)$.*

Proof. The first point derives from Jensen's inequality. The second point is an application of Proposition 2.2. To prove the third point, recall that Ω is convex: thus it is enough to check that the functions u_m satisfy McCann's conditions (see [26] or [32, Theorem 5.15]), which is the case. □

2.2 Absolutely continuous curves in the Wasserstein space

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

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

Following [3, Definition 1.1.1], we will use the following definition.

Definition 2.5. *We say that a curve $\rho \in \Gamma$ is 2-absolutely continuous if there exists a function $\lambda \in L^2([0,T])$ such that, for every $0 \leqslant t \leqslant s \leqslant T$,*

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

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

Theorem 2.6. *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,

$$\int_0^T |\dot\rho_t|^2 \,\mathrm{d}t = \sup_{N \geqslant 2} \sup_{0 \leqslant t_1 < t_2 < \ldots < t_N \leqslant T} \sum_{k=2}^N \frac{W_2^2(\rho_{t_{k-1}}, \rho_{t_k})}{t_k - t_{k-1}}. \tag{2.4}$$

Proof. The first part is just [3, Theorem 1.1.2]. The proof of the representation formula (2.4) can easily be obtained by adapting the proof of [5, 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 $(\mathcal{P}(\Omega), W_2)$. Thus, the quantity $\int_0^T |\dot\rho_t|^2 \,\mathrm{d}t$ behaves like a H^1 norm. In particular, we have the following.

Proposition 2.7. *The function $\rho \in \Gamma \mapsto \int_0^T |\dot\rho_t|^2 \,\mathrm{d}t$ is l.s.c., convex, and its sublevel sets are compact.*

Proof. The lower semi-continuity and convexity are a consequence of the representation formula (2.4) (because the square of the Wasserstein distance is a continuous convex function of its two arguments, see [30, Chapter 7]).

Moreover if $\rho \in \Gamma$ is a curve with finite action and $s < t$, then, again with (2.4), one can see that $W_2(\rho_s, \rho_t) \leqslant \sqrt{\int_0^T |\dot\rho_t|^2 \,\mathrm{d}t}\, \sqrt{t-s}$. This shows that the sublevel sets of $\int_0^T |\dot\rho_t|^2 \,\mathrm{d}t$ are uniformly equicontinuous, therefore they are relatively compact thanks to Ascoli-Arzela's theorem. As we know moreover that the sublevel sets are closed (by the lower semi-continuity we just proved), we can conclude that they are compact. □

2.3 Continuous and discrete problems

In all the sequel, we will make the following assumptions:

1. Recall that Ω is the closure of an open convex bounded domain with smooth boundary.
2. We assume that $f : [0, +\infty) \to \mathbb{R}$ is a strictly convex function, bounded from below and C^2 on $(0, +\infty)$. We define the congestion penalization F by, for any $\rho \in \mathcal{P}(\Omega)$,

$$F(\rho) := \int_\Omega f(\rho^{ac}) + f'(+\infty)\rho^{sing}(\Omega),$$

   where $\rho =: \rho^{ac}\mathcal{L} + \rho^{sing}$ is the decomposition of ρ as an absolutely continuous part ρ^{ac} (identified with its density) and a singular part ρ^{sing} w.r.t. $\mathcal{L}$. Thanks to Proposition 2.2, we know that F is a convex l.s.c. functional on $\mathcal{P}(\Omega)$.
3. We assume that $V : \Omega \to \mathbb{R}$ is a Lipschitz function.
4. We assume that $\Psi : \mathcal{P}(\Omega) \to \mathbb{R}$ is a l.s.c. and convex functional, bounded from below.

We will consider variational problems with a running cost of the form $\rho \mapsto E(\rho) := F(\rho) + \int_\Omega V \, \mathrm{d}\rho$, while Ψ will penalize the final density, and the initial one will be prescribed.

Definition 2.8. *We define the the functional $\mathcal{A} : \Gamma \to \mathbb{R}$ by*

$$\mathcal{A}(\rho) := \int_0^T \frac{1}{2}|\dot\rho_t|^2 \, \mathrm{d}t + \int_0^T E(\rho_t) \, \mathrm{d}t + \Psi(\rho_T).$$

We state the continuous problem as

$$\min\{\mathcal{A}(\rho) \ : \ \rho \in \Gamma, \ \rho_0 = \overline{\rho_0}\}. \tag{ContPb}$$

A curve ρ that minimizes $\mathcal{A}$ will be called a solution of the continuous problem.

Proposition 2.9. *Let us assume that there exists $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$ such that $\mathcal{A}(\rho) < +\infty$. Then the problem* (ContPb) *admits a unique solution.*

Proof. The functional $\mathcal{A}$ is the sum of l.s.c., convex and bounded functionals. Moreover, as $\mathcal{A}(\rho) \geqslant \int_0^T \frac{1}{2}|\dot\rho_t|^2 \, \mathrm{d}t - C$ (where C depends on the lower bounds of f, V and Ψ), we see (thanks to Proposition 2.7) that the sublevel sets of $\mathcal{A}$ are compact. The existence of a solution to (ContPb) follows from the direct method of calculus of variations.

To prove uniqueness, we need to prove that $\mathcal{A}$ is strictly convex. If ρ^1 and ρ^2 are two distinct minimizers of $\mathcal{A}$, we define $\rho := (\rho^1 + \rho^2)/2$. As ρ^1 and ρ^2 are distinct, by continuity there exists $T_1 < T_2$ such that ρ^1_t and ρ^2_t differ for every $t \in [T_1, T_2]$. In particular, for any $t \in [T_1, T_2]$, by strict convexity of F, $F(\rho) < (F(\rho^1) + F(\rho^2))/2$. Thus,

$$\int_0^T F(\rho_t) \, \mathrm{d}t < \frac{1}{2}\int_0^T F(\rho^1_t) \, \mathrm{d}t + \frac{1}{2}\int_0^T F(\rho^2_t) \, \mathrm{d}t.$$

As all the other terms appearing in $\mathcal{A}$ are convex, one concludes that $\mathcal{A}(\rho) < (\mathcal{A}(\rho^1) + \mathcal{A}(\rho^2))/2$, which contradicts the optimality of ρ^1 and ρ^2. □

In order to get the L^∞ bounds, we will consider two different cases (strong and weak congestion), depending on the second derivative of f. This allows to quantify how much F penalizes concentrated measures.

Assumption 1 (strong congestion)**.** *There exists $\alpha \geqslant -1$ and $C_f > 0$ such that $f''(t) \geqslant C_f t^\alpha$ for any $t > 0$.*

Assumption 2 (strong congestion-variant)**.** *There exist $\alpha \geqslant -1$, $t_0 > 0$ and $C_f > 0$ such that $f''(t) \geqslant C_f t^\alpha$ for any $t \geqslant t_0$.*

In particular, integrating twice, we see that under either of the above assumptions, for $\rho \in \mathcal{P}(\Omega)$ we have $U_{\alpha+2}(\rho) \leqslant C_f F(\rho) + C$, where C is a constant that depends on f (but not on ρ). One can also see that $f'(+\infty) = +\infty$. The function u_m is the typical example of a function satisfying Assumption 1 with $\alpha = m - 2$. To produce functions satisfying Assumption 2 but not Assumption 1, think at $f(t) = \sqrt{1 + t^4}$ (if we try to satisfy Assumption 1 we need $\alpha \leqslant 0$ for large t, and $\alpha \geqslant 2$ for small t) or at $f(t) = (t-1)_+^2$ (the difference between these two examples is that in the first case on could choose an aribtrary $t_0 > 0$, while in the second it is necessary to use $t_0 \geqslant 1$).

Assumption 3 (weak congestion)**.** *There exist $\alpha < -1$, $t_0 > 0$ and $C_f > 0$ such that $f''(t) \geqslant C_f t^\alpha$ for any $t \geqslant t_0$.*

For example, $f(t) := \sqrt{1+t^2}$ satisfies $f''(t) \geqslant C_f t^\alpha$ for $t \geqslant 1$ with $\alpha = -3$.

Assumption 4 (higher regularity of the potential)**.** *The potential V is of class $C^{1,1}$ (it is C^1 and its gradient is Lipschitz) and $\nabla V \cdot \mathbf{n} \geqslant 0$ on $\partial\Omega$, where $\mathbf{n}$ is the outward normal to Ω.*

We will see that only Assumption 1, where we require a control of f'' everywhere, allows to deal with Lipschitz potentials, while in general we will need the use of Assumption 4. The condition $\nabla V \cdot \mathbf{n} \geqslant 0$ on $\partial\Omega$ can be interpreted by the fact that the minimum of V is reached in the interior of Ω: it prevents the mass of ρ to concentrate on the boundaries.

Assumption 5 (final penalization)**.** *The penalization Ψ is of the following form*

$$\Psi(\rho_T) = \begin{cases} \displaystyle\int_\Omega g(\rho_T) + \int_\Omega W \,\mathrm{d}\rho_T & \text{if } \rho_T \text{ is absolutely continuous w.r.t. } \mathcal{L} \\ +\infty & \text{if } \rho_T \text{ is singular w.r.t. } \mathcal{L}, \end{cases}$$

where $g : [0,+\infty) \to \mathbb{R}$ is a convex and superlinear (i.e. $g'(+\infty) = +\infty$) function, bounded from below, and $W : \Omega \to \mathbb{R}$ is a potential of class $C^{1,1}$ satisfying $\nabla W \cdot \mathbf{n} \geqslant 0$ on $\partial\Omega$.

The mains results of this paper can be stated as follows.

Theorem 2.10 (strong congestion, interior regularity)**.** *Suppose that either Assumption 1 holds or Assumption 2 and 4 hold, and that $\mathcal{A}(\rho) < +\infty$ for some $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$. Let ρ be the unique solution to* (ContPb)*. Then for any $0 < T_1 < T_2 < T$, the restriction of ρ to $[T_1, T_2]$ belongs to $L^\infty([T_1, T_2] \times \Omega)$.*

Theorem 2.11 (strong congestion, boundary regularity)**.** *Suppose that either Assumption 1 holds or Assumption 2 and 4 hold, and that Assumption 5 holds as well, and that $\mathcal{A}(\rho) < +\infty$ for some $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$. Let ρ be the unique solution to* (ContPb)*. Then, for any $0 < T_1 < T$, the restriction of ρ to $[T_1, T]$ belongs to $L^\infty([T_1, T] \times \Omega)$.*

Theorem 2.12 (weak congestion case)**.** *Suppose Assumptions 5, 3 and 4 hold and that $\mathcal{A}(\rho) < +\infty$ for some $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$. We assume that the prescribed initial measure $\overline{\rho_0}$ satisfies $\overline{\rho_0} \in L^{m_0}$ with $m_0 > d|\alpha+1|/2$ and $F(\overline{\rho_0}) < +\infty$, and that T is small enough (smaller than a constant that depends on f, g, V, W and $\overline{\rho_0}$). Let ρ be the unique solution to* (ContPb)*. Then $\rho \in L^{m_0}([0,T] \times \Omega)$ and for any $0 < T_1 < T$, the restriction of ρ to $[T_1, T]$ belongs to $L^\infty([T_1, T] \times \Omega)$.*

The rest of the paper is devoted to the proof of these theorems. In particular, we will always assume in the sequel that there exists $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$ such that $\mathcal{A}(\rho) < +\infty$. In order to prove these theorems, we will introduce a discrete (in time) variational problem that will approximate the continuous one. For this problem, we will be able to show the existence of a unique smooth (in space) solution and write down the optimality conditions. From these optimality conditions, we will be able to derive a *flow interchange* estimate whose iteration will give uniform (in the approximation parameters, and in p) L^p estimates.

Let us introduce the discrete problem here. We will use two approximations parameters:

- $N+1 \geqslant 2$ will denote the number of time steps. We will write $\tau := T/N$ for the distance between two time steps. The set T^N will stand for the set of all time steps, namely

  $$T^N := \{k\tau;\ k = 0, 1, \dots, N\}.$$

  We set $\Gamma_N := \mathcal{P}(\Omega)^{T^N} \simeq \mathcal{P}(\Omega)^{N+1}$ (i.e. an element $\rho \in \Gamma_N$ is a $N+1$-uplet $(\rho_0, \rho_\tau, \dots, \rho_{k\tau}, \dots, \rho_T)$ of probability measures indexed by T^N). A natural discretization of the action of a curve is

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

- We will also add a (vanishing) entropic penalization (recall that U_1 denotes the entropy w.r.t. $\mathcal{L}$). It will ensure that the solution of the discrete problem is smooth. The penalization will be a discretized version of

  $$\lambda \int_0^T U_1(\rho_t) \,\mathrm{d}t,$$

  where λ is a parameter that will be sent 0.

Let us state formally our problem. We fix $N \geqslant 1$ ($\tau := T/N$) and $\lambda > 0$, and we set $\lambda_N = \lambda$ if Assumption 5 is satisfied, $\lambda_N = 0$ otherwise. We define $\mathcal{A}^{N,\lambda} : \Gamma_N \to \mathbb{R}$ by

$$\mathcal{A}^{N,\lambda}(\rho) := \sum_{k=1}^{N} \frac{W_2^2(\rho_{(k-1)\tau}, \rho_{k\tau})}{2\tau} + \sum_{k=1}^{N-1} \tau \left(E(\rho_{k\tau}) + \lambda U_1(\rho_{k\tau})\right) + \Psi(\rho_T) + \lambda_N U_1(\rho_T).$$

This means that in the case of Assumption 5 we penalize ρ_T by $\int_\Omega g(\rho_T) + \lambda U_1(\rho_T) + \int_\Omega W \, \mathrm{d}\rho_T$, while we do not modify the boundary condition otherwise (the reason for not always adding $\lambda U_1(\rho_T)$ lies in the possibility of having a prescribed value for ρ_T with infinite entropy). In all the cases, we enforce strictly $\rho_0 = \overline{\rho_0}$. The discrete minimization problem reads

$$\min\{\mathcal{A}^{N,\lambda}(\rho) \ : \ \rho \in \Gamma_N, \ \rho_0 = \overline{\rho_0}\}, \tag{DiscrPb}$$

and a $\rho \in \Gamma_N$ which minimizes $\mathcal{A}^{N,\lambda}$ will be called a solution of (DiscrPb).

Theorem 2.13. *For any $N \geqslant 1$ and any $\lambda > 0$, the discrete problem* (DiscrPb) *admits a solution.*

Proof. The functional $\mathcal{A}^{N,\lambda}$ is a sum of convex and l.s.c. functionals, bounded from below, hence it is itself convex, l.s.c. and bounded from below. Moreover, the space $\Gamma_N = \mathcal{P}(\Omega)^{N+1}$ is compact (for the weak-* convergence). Thus, to use the direct method of calculus of variations, it is enough to show that $A^{N,\lambda}(\rho) < +\infty$ for some $\rho \in \Gamma_N$.

This is easy in this discrete framework: just take $\rho_{k\tau} = \mathcal{L}$ if $k \in \{1, 2, \dots, N-1\}$, $\rho_0 = \overline{\rho_0}$ and $\rho_{N\tau}$ equal to an arbitrary measure ρ such that $\Psi(\rho) + \lambda_N U_1(\rho) < +\infty$. □

Remark 2.14. *We did not adress the uniqueness of the minimizer in the above problem since we do not really care about it, but indeed it also holds. Indeed, the strict convexity of F (or the term λU_1 that we added) guarantees uniqueness of $\rho_{k\tau}$ for all $k \leqslant N-1$. The uniqueness of the last measure (which cannot be deducted from strict convexity for an arbitrary functional Ψ, as we do not always add a term of the form $\lambda U_1(\rho_T)$) can be obtained from the strict convexity of the last Wasserstein distance term $\rho \mapsto W_2^2(\rho, \rho_{(N-1)\tau})$, as $\rho_{(N-1)\tau}$ is absolutely continuous (see [30, Proposition 7.19]).*

In all the following, for any $N \geqslant 1$ and $\lambda > 0$, we denote by $\bar{\rho}^{N,\lambda} \in \Gamma_N$ the unique solution of (DiscrPb) with parameters N and λ. Moreover, In all the sequel, we fix $1 < \beta < d/(d-2)$. It is well known that the space $H^1(\Omega)$ is continuously embedded into $L^{2\beta}(\Omega)$. Moreover, in the case where the assumptions of Theorem 2.12 are satisfied, we choose β in such a way that

$$\frac{\beta}{\beta - 1} m_0 > |\alpha + 1|. \tag{2.5}$$

3 Flow interchange estimate

3.1 Interior flow interchange

In this subsection, we study the optimality conditions of (DiscrPb) away from the temporal boundaries. We fix for the rest of the subsection $N \geqslant 1$, $0 < \lambda \leqslant 1$ and $0 < k < N$, and we use the shortcut $\bar{\rho} := \bar{\rho}^{N,\lambda}_{k\tau}$. Let us also denote $\mu := \bar{\rho}^{N,\lambda}_{(k-1)\tau}$ and $\nu := \bar{\rho}^{N,\lambda}_{(k+1)\tau}$. As $\bar{\rho}^{N,\lambda}$ is a solution of the discrete problem, we know that $\bar{\rho}$ is a minimizer (among all probability measures) of

$$\rho \mapsto \frac{W_2^2(\mu, \rho) + W_2^2(\rho, \nu)}{2\tau} + \tau \left(F(\rho) + \lambda U_1(\rho) + \int_\Omega V \, \mathrm{d}\rho \right).$$

In particular, we know that $U_1(\bar{\rho}) < +\infty$, thus $\bar{\rho}$ is absolutely continuous w.r.t. $\mathcal{L}$.

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

Proof. For $0 < \varepsilon < 1$, we define $\rho_\varepsilon := (1-\varepsilon)\bar{\rho} + \varepsilon\mathcal{L}$. As $\mathcal{L}$ is a probability measure, we know that ρ_ε is a probability measure too. Thus, using ρ_ε as a competitor, we get

$$\lambda(U_1(\bar{\rho}) - U_1(\rho_\varepsilon)) \leqslant \frac{W_2^2(\mu, \rho_\varepsilon) + W_2^2(\rho_\varepsilon, \nu)}{2\tau} + \tau E(\rho_\varepsilon) - \frac{W_2^2(\mu, \bar{\rho}) + W_2^2(\bar{\rho}, \nu)}{2\tau} - \tau E(\bar{\rho}).$$

We estimate the r.h.s. by convexity (as W_2^2 and F are convex) to see that

$$U_1(\bar\rho) - U_1(\rho_\varepsilon) \leqslant \frac{\varepsilon}{\lambda}\left(\frac{W_2^2(\mu,\mathcal{L}) + W_2^2(\mathcal{L},\nu)}{2\tau} + \tau E(\mathcal{L}) - \frac{W_2^2(\mu,\bar\rho) + W_2^2(\bar\rho,\nu)}{2\tau} - \tau E(\bar\rho)\right).$$

Thus, there exists a constant C, independent of ε, such that $U_1(\bar\rho) - U_1(\rho_\varepsilon) \leqslant C\varepsilon$. This can be easily seen to imply (see for instance the proof of [30, Lemma 8.6]) that $\bar\rho$ is strictly positive a.e. □

We can then write the first-order optimality conditions.

Proposition 3.2. *The measure $\bar\rho$ (or more precisely its density w.r.t. $\mathcal{L}$) is Lipschitz and bounded away from 0 and ∞. Moreover, let us denote by φ_μ and φ_ν the Kantorovitch potentials for the transport from $\bar\rho$ to respectively μ and ν. Then the following identity holds a.e.:*

$$\frac{\nabla\varphi_\mu + \nabla\varphi_\nu}{\tau^2} + \left(f''(\bar\rho) + \frac{\lambda}{\bar\rho}\right)\nabla\bar\rho + \nabla V = 0. \tag{3.1}$$

Proof. Let $\tilde\rho \in \mathcal{P}(\Omega)\cap L^\infty(\Omega)$ and for $0<\varepsilon<1$ define $\rho_\varepsilon = (1-\varepsilon)\bar\rho + \varepsilon\tilde\rho$. We use ρ_ε as a competitor. We use Proposition 2.1: as $\bar\rho > 0$ a.e., the Kantorovitch potentials φ_μ and φ_ν for the transport from $\bar\rho$ to respectively μ and ν are unique and

$$\lim_{\varepsilon\to0}\frac{W_2^2(\mu,\bar\rho) - W_2^2(\mu,\rho_\varepsilon) + W_2^2(\bar\rho,\nu) - W_2^2(\rho_\varepsilon,\nu)}{2\tau^2} = \int_\Omega \frac{\varphi_\mu+\varphi_\nu}{\tau}(\bar\rho-\tilde\rho).$$

The term involving V is straightforward to handle as it is linear. Hence, by optimality of $\bar\rho$ we get

$$\int_\Omega\left(\frac{\varphi_\mu+\varphi_\nu}{\tau^2} + V\right)(\bar\rho-\tilde\rho) \leqslant \liminf_{\varepsilon\to0}\frac{F(\rho_\varepsilon)+\lambda U_1(\rho_\varepsilon) - F(\bar\rho) - \lambda U_1(\bar\rho)}{\varepsilon}. \tag{3.2}$$

$$\frac{F(\rho_\varepsilon)+\lambda U_1(\rho_\varepsilon) - F(\bar\rho) - \lambda U_1(\bar\rho)}{\varepsilon} = \int_\Omega \frac{f_\lambda[(1-\varepsilon)\bar\rho+\varepsilon\tilde\rho] - f_\lambda[\bar\rho]}{\varepsilon}.$$

The integrand of the integral of the r.h.s. converges pointewisely, as $\varepsilon\to0$, to $(f'(\bar\rho)+\lambda\ln\bar\rho)(\tilde\rho-\bar\rho)$. Moreover, as the function f_λ is convex, we see that for $0<\varepsilon<1$,

$$\frac{f_\lambda[(1-\varepsilon)\tilde\rho+\varepsilon\bar\rho] - f_\lambda[\bar\rho]}{\varepsilon} \leqslant f_\lambda(\tilde\rho) - f_\lambda(\bar\rho).$$

As $\tilde\rho\in L^\infty(\Omega)$ and $F(\bar\rho)+\lambda U_1(\bar\rho) < +\infty$, the r.h.s. of the equation is integrable on Ω. Thus, by a reverse Fatou's lemma,

$$\limsup_{\varepsilon\to0}\int_\Omega \frac{F(\rho_\varepsilon)+\lambda U_1(\rho_\varepsilon) - F(\bar\rho) - \lambda U_1(\bar\rho)}{\varepsilon} \leqslant \int_\Omega (f'(\bar\rho)+\lambda\ln\bar\rho)\,(\tilde\rho-\bar\rho).$$

Combing this equation with (3.2), we see that $\int_\Omega h\,\mathrm{d}(\tilde\rho-\bar\rho) \geqslant 0$ with

$$h := \frac{\varphi_\mu+\varphi_\nu}{\tau^2} + f'(\bar\rho) + \lambda\ln\bar\rho + V.$$

We know that h is finite a.e., thus its essential infimum cannot be $+\infty$. Moreover, starting from $\bar\rho f'(\bar\rho) \geqslant f(\bar\rho) - f(0)$, we see that $\int_\Omega h\bar\rho > -\infty$. Taking probability measures $\tilde\rho$ concentrated on sets where h is close to its essential infimum, we see that the essential infimum of h cannot be $+\infty$ and that h coincides with its essential infimum $\bar\rho$-a.e. As $\bar\rho>0$ a.e., there exists C such that we have a.e. on Ω

$$f'(\bar\rho) + \lambda\ln\bar\rho = C - \frac{\varphi_\mu+\varphi_\nu}{\tau^2} - V. \tag{3.3}$$

As f' is C^1 and increasing, it is easy to see that $f'+\lambda\ln$ is an homeomorphism of $(0\,,+\infty)$ on $(-\infty\,,+\infty)$ which is bilipschitz on compact sets. As the function $C-(\varphi_\mu+\varphi_\nu)/\tau^2 - V$ takes its values in a compact set and is Lispchitz, we see that $\bar\rho$ is bounded away from 0 and ∞ and is Lipschitz. With all this regularity (recall that f is assumed to be C^2 on $(0\,,+\infty)$), we can take the gradient of (3.3) to obtain (3.1). □

Theorem 3.3 (Flow interchange inequality)**.** *For any $m \geqslant 1$, the following inequality holds:*

$$\int_\Omega |\nabla\bar\rho|^2 f''(\bar\rho)\bar\rho^{m-1} + \int_\Omega (\nabla\bar\rho\cdot\nabla V)\bar\rho^{m-1} \leqslant \frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2}.$$

Proof. We multiply pointewisely (3.1) by $\bar\rho^{m-1}\nabla\bar\rho$ and integrate over Ω. Dropping the entropic term, we easily get

$$\int_\Omega |\nabla\bar\rho|^2 f''(\bar\rho)\bar\rho^{m-1} + \int_\Omega (\nabla\bar\rho\cdot\nabla V)\bar\rho^{m-1} \leqslant -\frac{1}{\tau^2}\int_\Omega \left[\nabla\bar\rho\cdot(\nabla\varphi_\mu+\nabla\varphi_\nu)\right]\bar\rho^{m-1}.$$

To prove the flow interchange inequality, it is enough to show that

$$-\int_\Omega (\nabla\bar\rho\cdot\nabla\varphi_\mu)\bar\rho^{m-1} \leqslant U_m(\mu)-U_m(\bar\rho),$$

as a similar inequality will hold for the term involving φ_ν. To this purpose, we denote by $\rho:[0\,,1]\to\mathcal{P}(\Omega)$ the constant-speed geodesic joining $\bar\rho$ to μ. We know that it is given by

$$\rho(t) = (\mathrm{Id}-t\nabla\varphi_\mu)\#\bar\rho.$$

By geodesic convexity of U_m, the function $t\mapsto U_m(\rho(t))$ is convex. Hence,

$$\begin{aligned}
U_m(\mu)-U_m(\bar\rho) &= U_m(\rho(1))-U_m(\rho(0))\\
&\geqslant \limsup_{t\to0}\frac{U_m(\rho(t))-U_m(\rho(0))}{t}\\
&= \limsup_{t\to0}\int_\Omega \frac{u_m(\rho(t))-u_m(\bar\rho)}{t}\\
&\geqslant \limsup_{t\to0}\int_\Omega \frac{(\rho(t)-\bar\rho)u_m'(\bar\rho)}{t}\\
&= \limsup_{t\to0}\int_\Omega \frac{u_m'(\bar\rho[x-t\nabla\varphi_\mu(x)])-u_m'(\bar\rho[x])}{t}\bar\rho(x)\,\mathrm{d}x,
\end{aligned}$$

where we also have used that u_m is convex. It is clear that for a.e. $x\in\Omega$,

$$\lim_{t\to0}\frac{u_m'(\bar\rho[x-t\nabla\varphi_\mu(x)])-u_m'(\bar\rho[x])}{t} = -\left[(\nabla\bar\rho\cdot\nabla\varphi_\mu)u_m''(\bar\rho)\right](x).$$

Moreover, we have the uniform (in t) bound

$$\left|\frac{u_m'(\bar\rho[x-t\nabla\varphi_\mu(x)])-u_m'(\bar\rho[x])}{t}\right| \leqslant \|u_m''(\bar\rho)\|_\infty\|\nabla\bar\rho\|_\infty\|\nabla\varphi_\mu\|_\infty.$$

At this point, one can remember that $u_m''(x)=x^{m-2}$. Moreover, as $\bar\rho$ is bounded away from 0 and ∞ and Lipschitz, the r.h.s. of the equation above is finite. Thus, by dominated convergence,

$$\limsup_{t\to0}\int_\Omega \frac{u_m'(\bar\rho[x-t\nabla\varphi_\mu(x)])-u_m'(\bar\rho[x])}{t}\bar\rho(x)\,\mathrm{d}x = -\int_\Omega(\nabla\bar\rho\cdot\nabla\varphi_\mu)\bar\rho^{m-1}. \qquad \square$$

From the result of Theorem 3.3 we need to deduce estimates on improved L^m norms. To this aim, we treat in a slightly different way the cases of weak and strong congestion even if the result are similar. The main issue is to control the term involving ∇V.

Corollary 3.4 (Strong congestion case)**.** *Suppose that Assumption 1 holds. Then, for any $m\geqslant\alpha+2$ one has*

$$U_{\beta m}(\bar\rho)^{1/\beta} \leqslant Cm^2\left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2}+Cm^2U_m(\bar\rho)\right],$$

where $C>0$ depends only on f, V and Ω.

Proof. Let us start from the case of Assumption 1. In this case, we recall that C_f is the constant such that $f''(t) \geqslant C_f t^\alpha$ for any $t > 0$. We transform the term involving ∇V in the following way:

$$\begin{aligned}\int_\Omega (\nabla\bar\rho\cdot\nabla V)\bar\rho^{m-1} &= \int_\Omega (\bar\rho^{\alpha/2}\nabla\bar\rho)\cdot(\bar\rho^{-\alpha/2}\nabla V)\bar\rho^{m-1}\\ &\geqslant -\frac{C_f}{2}\int_\Omega |\bar\rho^{\alpha/2}\nabla\bar\rho|^2\bar\rho^{m-1} - \frac{1}{2C_f}\int_\Omega |\bar\rho^{-\alpha/2}\nabla V|^2\bar\rho^{m-1}\\ &= -\frac{C_f}{2}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} - \frac{1}{2C_f}\int_\Omega |\nabla V|^2\bar\rho^{m-1-\alpha}\\ &\geqslant -\frac{C_f}{2}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} - \frac{\|\nabla V\|_\infty^2 m^2}{2C_f}U_m(\bar\rho).\end{aligned}$$

For the last inequality, we have used the fact that

$$\int_\Omega \bar\rho^{m-1-\alpha} \leqslant \left(\int_\Omega \bar\rho^m\right)^{(m-1-\alpha)/m} \leqslant \int_\Omega \bar\rho^m \leqslant m^2 U_m(\bar\rho),$$

which is valid because $1 \leqslant m-1-\alpha \leqslant m$ and $\mathcal{L}(\Omega) = 1$. Thus, using Theorem 3.3, we get

$$\begin{aligned}\frac{C_f}{2}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} &\leqslant \int_\Omega |\nabla\bar\rho|^2 f''(\bar\rho)\bar\rho^{m-1} - \frac{C_f}{2}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha}\\ &\leqslant \left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + \frac{\|\nabla V\|_\infty^2}{2C_f}m^2U_m(\bar\rho)\right].\end{aligned}$$

We are interested only in the large values taken by $\bar\rho$. Let us introduce $\hat\rho := \max(1,\bar\rho)$. This function is larger than $\bar\rho$ and 1 and its gradient satisfies $|\nabla\hat\rho| = |\nabla\bar\rho| 1_{\bar\rho\geqslant 1}$. Thus,

$$\int_\Omega |\nabla\hat\rho^{m/2}|^2 = \frac{m^2}{4}\int_\Omega |\nabla\hat\rho|^2\hat\rho^{m-2} \leqslant \frac{m^2}{4}\int_\Omega |\nabla\hat\rho|^2\hat\rho^{m-1+\alpha} \leqslant \frac{m^2}{4}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha}.$$

(the last inequality is true since $\nabla\hat\rho = 0$ on the points where $\hat\rho > \bar\rho$, and the first inequality is exactly the point where we exploit the fact $\hat\rho \geqslant 1$, which explains the use of $\hat\rho$ instead of $\bar\rho$). On the other hand, if we use the injection of $H^1(\Omega)$ into $L^{2\beta}(\Omega)$ for the function $\hat\rho^{m/2}$, we get (with C_Ω a constant that depends only on Ω),

$$\left(\int_\Omega \hat\rho^{m\beta}\right)^{1/\beta} \leqslant C_\Omega\left(\int_\Omega |\nabla\hat\rho^{m/2}|^2 + \int_\Omega \hat\rho^m\right).$$

As $\bar\rho^{\beta m} \leqslant \hat\rho^{\beta m}$ and $\hat\rho^m \leqslant 1+\bar\rho^m$, we see that

$$\begin{aligned}\left(\int_\Omega \bar\rho^{m\beta}\right)^{1/\beta} &\leqslant \left(\int_\Omega \hat\rho^{m\beta}\right)^{1/\beta}\\ &\leqslant C_\Omega\left(\frac{m^2}{4}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} + \int_\Omega \bar\rho^m + 1\right)\\ &\leqslant C_\Omega m^2\left(\frac{1}{4}\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} + 2U_m(\bar\rho)\right)\\ &\leqslant Cm^2\left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + Cm^2U_m(\bar\rho)\right].\end{aligned}$$

Notice that to go from the second to the third line, we have used the fact that $1 \leqslant \int_\Omega \bar\rho^m \leqslant m^2 U_m(\bar\rho)$. To conclude, it remains to notice that, as $m\beta \geqslant \beta > 1$, that we can control (uniformly in m) $U_{m\beta}(\bar\rho)$ by $\int_\Omega \bar\rho^{m\beta}$. Indeed

$$\left(\int_\Omega \bar\rho^{m\beta}\right)^{1/\beta} \geqslant \frac{1}{(\beta(\beta-1))^{1/\beta}}U_{m\beta}(\bar\rho)^{1/\beta}.$$

Thus, up to a change in the constant C, we get the result we claimed. □

Corollary 3.5 (Weak congestion case)**.** *Suppose Assumption 3 and 4 both hold. Then, for any $m \geqslant 1$ such that $\beta(m+\alpha+1) \geqslant 1$ one has*

$$\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + CmU_m(\bar\rho) \geqslant 0$$

and

$$U_{\beta(m+1+\alpha)}(\bar\rho)^{1/\beta} \leqslant Cm^2\left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + CmU_m(\bar\rho)\right] + Ct_0^{m+1+\alpha},$$

where C depends only on f, V and Ω.

Proof. We use an integration by parts to treat the term involving ∇V. Recall that $\mathbf{n}$ denotes the exterior normal to Ω.

$$\begin{aligned}\int_\Omega (\nabla\bar\rho\cdot\nabla V)\bar\rho^{m-1} &= \frac{1}{m}\int_\Omega \nabla(\bar\rho^m)\cdot\nabla V\\ &= \frac{1}{m}\int_{\partial\Omega}(\nabla V\cdot\mathbf{n})\bar\rho^m - \frac{1}{m}\int_\Omega \Delta V\bar\rho^m\\ &\geqslant -\|\Delta V\|_\infty m U_m(\bar\rho),\end{aligned}$$

where we have used the assumption $\nabla V\cdot\mathbf{n}\geqslant 0$ on $\partial\Omega$. Thus, using Theorem 3.3, we get (recall that $f''(t)\geqslant C_f t^\alpha$ but only for $t\geqslant t_0$)

$$C_f\int_{\{\bar\rho\geqslant t_0\}}|\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} \leqslant C_f\int_\Omega |\nabla\bar\rho|^2\bar\rho^{m-1}f''(\bar\rho) \tag{3.4}$$

$$\leqslant \left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + \|\Delta V\|_\infty m U_m(\bar\rho)\right]. \tag{3.5}$$

This gives us the first inequality of the corollary. In a similar manner to the strong congestion case, we introduce $\hat\rho := \max(t_0,\bar\rho)$. This time we notice that

$$\int_\Omega |\nabla\hat\rho^{(m+1+\alpha)/2}|^2 \leqslant \frac{m^2}{4}\int_{\{\bar\rho\geqslant t_0\}}|\nabla\bar\rho|^2\bar\rho^{m-1+\alpha}.$$

Thus, if we use the injection of $H^1(\Omega)$ into $L^{2\beta}(\Omega)$ with the function $\hat\rho^{(m+1+\alpha)/2}$,

$$\left(\int_\Omega \hat\rho^{\beta(m+1+\alpha)}\right)^{1/\beta} \leqslant C_\Omega\left(\int_\Omega|\nabla\hat\rho^{(m+1+\alpha)/2}|^2 + \int_\Omega\hat\rho^{m+1+\alpha}\right).$$

Then, we proceed as in the proof of the strong congestion case, but this time $m+1+\alpha\leqslant m$ and $\hat\rho^{m+1+\alpha}\leqslant \bar\rho^{m+1+\alpha}+t_0^{m+1+\alpha}$:

$$\begin{aligned}\left(\int_\Omega\bar\rho^{\beta(m+1+\alpha)}\right)^{1/\beta} &\leqslant \left(\int_\Omega\hat\rho^{\beta(m+1+\alpha)}\right)^{1/\beta}\\ &\leqslant C_\Omega\left(\frac{m^2}{4}\int_{\{\bar\rho\geqslant 1\}}|\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} + \int_\Omega\bar\rho^{m+1+\alpha} + t_0^{m+1+\alpha}\right)\\ &\leqslant C_\Omega\left(\frac{m^2}{4}\int_{\{\bar\rho\geqslant 1\}}|\nabla\bar\rho|^2\bar\rho^{m-1+\alpha} + \int_\Omega\bar\rho^{m} + t_0^{m+1+\alpha}\right)\\ &\leqslant Cm^2\left[\frac{U_m(\mu)+U_m(\nu)-2U_m(\bar\rho)}{\tau^2} + CmU_m(\bar\rho)\right] + Ct_0^{m+1+\alpha}.\end{aligned}$$

Notice that if $t_0\leqslant 1$, we can control t_0^m by $m^2U_m(\bar\rho)$ (as we did in the strong congestion case), but in the general case this is not possible and we have to keep an explicit dependence in t_0. To conclude, we notice that, thanks to (2.5), one has $\beta(m+1+\alpha)\geqslant m\geqslant m_0$ and thus

$$\left(\int_\Omega\bar\rho^{\beta(m+1+\alpha)}\right)^{1/\beta} \geqslant \frac{1}{(m_0(m_0-1))^{1/\beta}}U_{\beta(m+1+\alpha)}(\bar\rho)^{1/\beta}.$$ □

The last case is a combination of the previous two cases.

Corollary 3.6 (Strong congestion case-variant)**.** *Suppose that Assumption 2 and 4 both hold. Then, for any $m \geqslant m_0$ one has*

$$\frac{U_m(\mu) + U_m(\nu) - 2U_m(\bar{\rho})}{\tau^2} + CmU_m(\bar{\rho}) \geqslant 0 \tag{3.6}$$

and

$$U_{\beta m}(\bar{\rho})^{1/\beta} \leqslant Cm^2 \left[\frac{U_m(\mu) + U_m(\nu) - 2U_m(\bar{\rho})}{\tau^2} + CmU_m(\bar{\rho}) \right] + Ct_0^m, \tag{3.7}$$

where C depends only on f, V and Ω.

Proof. We begin with the same computations as in Corollary 3.5. We can obtain the same result as in (3.4), but on the set $\{\bar{\rho} \geqslant t_0\}$ we can use $\alpha \geqslant -1$ to write

$$\int_{\{\bar{\rho} \geqslant t_0\}} |\nabla \bar{\rho}|^2 \bar{\rho}^m \leqslant C \left[\frac{U_m(\mu) + U_m(\nu) - 2U_m(\bar{\rho})}{\tau^2} + \|\Delta V\|_\infty m U_m(\bar{\rho}) \right].$$

With $\hat{\rho} := \max(t_0, \bar{\rho})$ we get

$$\int_\Omega |\nabla \hat{\rho}^{m/2}|^2 \leqslant C \frac{m^2}{4} \int_{\{\bar{\rho} \geqslant t_0\}} |\nabla \bar{\rho}|^2 \bar{\rho}^m,$$

and the conclusion comes from the same Sobolev injection, with the function $\hat{\rho}^{m/2}$, and similar computations as in the previous cases. □

Remark 3.7. *For simplicity, inequality* (3.6) *and* (3.7) *will be used by replacing the term $mU_m(\bar{\rho})$ with $m^2 U_m(\bar{\rho})$, so as to allow a unified presentation with the inequality obtained in Corollary 3.4. Notice also that Corollary 3.4 is basically giving us the same inequality as* (3.7)*, as long as we set $t_0 = 0$.*

3.2 Boundary flow interchange

In the case of Assumption 5, we can derive some estimate right at the point T ($k = N$). We will only sketch the proof, at it mimicks the proof of the interior case and these computations are well-known in the case of the applications to the JKO scheme. We know that, with $\bar{\rho} = \bar{\rho}_T^{N,\lambda}$ and $\mu := \bar{\rho}_{T-\tau}^{N,\lambda}$, the measure $\bar{\rho}$ is a minimizer (among all probability measures) of

$$\rho \mapsto \frac{W_2^2(\mu, \rho)}{2\tau} + G(\rho) + \lambda U_1(\rho) + \int_\Omega W \, \mathrm{d}\rho.$$

Let us remark that it correspond to one step of the JKO scheme: it is in the context of such variational problems that the flow interchange was firstly used, see [25]. In any case, with these notation, we obtain:

Proposition 3.8. *Suppose Assumption 5 holds. Then, for any $m \geqslant 1$,*

$$\frac{U_m(\mu) - U_m(\bar{\rho})}{\tau} \geqslant -(m-1)\|\Delta W\|_\infty U_m(\bar{\rho}).$$

Proof. Following the same strategy than in Lemma 3.1 and Proposition 3.2, we know that $\bar{\rho}$ is bounded away from 0 and ∞, is a Lipschitz function, and that

$$\frac{\nabla \varphi_\mu}{\tau} + \left(g''(\bar{\rho}) + \frac{\lambda}{\bar{\rho}} \right) \nabla \bar{\rho} + \nabla W = 0$$

a.e. on Ω, where φ_μ is the unique Kantorovitch potential for the transport from $\bar{\rho}$ to μ. Thus, if we multiply by $\bar{\rho}^{m-1} \nabla \bar{\rho}$, we get, by the same estimation than in Theorem 3.3 (we drop both the entropic penalization and the congestion term),

$$\frac{U_m(\mu) - U_m(\bar{\rho})}{\tau} \geqslant \int_\Omega (\nabla W \cdot \nabla \bar{\rho}) \bar{\rho}^{m-1}.$$

It remains to perform an integration by parts, using the sign of $\nabla W \cdot \mathbf{n}$ on $\partial\Omega$, to conclude that

$$\frac{U_m(\mu) - U_m(\bar{\rho})}{\tau} \geqslant -\frac{1}{m} \int_\Omega \Delta W \bar{\rho}^m \geqslant -(m-1)\|\Delta W\|_\infty U_m(\bar{\rho}).$$ □

4 Moser-like iterations

Corollaries 3.4, 3.5 and 3.6 allow us to control the $L^{m\beta}$ or $L^{(m+1+\alpha)\beta}$ norm of $\bar\rho$ in terms of its L^m norm. The strategy will consist in integrating w.r.t. to time and iterating such a control in order to get a bound on the $L^m([T_1\,,T_2]\times\Omega)$ norm of $\bar\rho^{N,\lambda}$ that does not depend on λ and N and to control how this bounds grows in m. For any $N\geqslant 1$ and any $0<\lambda<1$, recall that $\bar\rho^{N,\lambda}$ is a solution of the discrete problem (DiscrPb).

Definition 4.1. *For any $m\geqslant 1$ and any $0\leqslant T_1\leqslant T_2\leqslant T$, we define $L^m_{T_1,T_2}$ as*

$$L^m_{T_1,T_2}:=\liminf_{N\to+\infty,\lambda\to 0}\left(\sum_{T_1\leqslant k\tau\leqslant T_2}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\right)^{1/m}.$$

The quantity $L^m_{T_1,T_2}$ can be seen as a discrete counter part of (up to a factor $1/(m(m-1))^{1/m}$) the L^m norm of the restriction to $[T_1\,,T_2]$ of the limit (whose existence will be proven in the next section) of $\bar\rho^{N,\lambda}$ when $N\to+\infty$ and $\lambda\to 0$.

4.1 The strong congestion case

First, we integrate w.r.t. time the estimate obtained in Corollary 3.4.

Proposition 4.2. *Suppose that either Assumption 1 holds or Assumptions 2 and 4 both hold. Then there exists two constants C_1 and C_2 (depending on f,V,T and Ω) such that, for any $0<\varepsilon\leqslant C_1/m$ and any $0<T_1<T_2<T$ such that $[T_1-\varepsilon\,,T_2+\varepsilon]\subset(0\,,T)$, and any $m\geqslant\alpha+2$,*

$$L^{\beta m}_{T_1,T_2}\leqslant\left[C_2\frac{m^3}{\varepsilon}\left(m^2+\frac{1}{\varepsilon^2}\right)\right]^{1/m}\max\left(L^m_{T_1-\varepsilon,T_2+\varepsilon},t_0\right).$$

As pointed out in Remark 3.7, in the case where Assumption 1 holds, we set $t_0=0$.

Proof. Let us recall that in Corollary 3.4 and Corollary 3.6 with Remark 3.7, we have proved (if we explicit the dependence in N and λ) that for any $N\geqslant 1$, $\lambda>0$ and any $k\in\{1,2,\dots,N-1\}$, one has

$$U_{\beta m}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta}\leqslant Cm^2\left[\frac{U_m(\bar\rho^{N,\lambda}_{(k-1)\tau})+U_m(\bar\rho^{N,\lambda}_{(k+1)\tau})-2U_m(\bar\rho^{N,\lambda}_{k\tau})}{\tau^2}+Cm^2U_m(\bar\rho^{N,\lambda}_{k\tau})\right]+Ct_0^m. \tag{4.1}$$

Let us take $\chi:[0\,,T]\to[0\,,1]$ a positive $C^{1,1}$ cutoff function such that $\chi(t)=1$ if $t\in[T_1-\varepsilon/3\,,T_2+\varepsilon/3]$ and $\chi(t)=0$ if $t\notin[T_1-2\varepsilon/3\,,T_2+2\varepsilon/3]$. Such a function χ can be chosen with $\|\chi''\|_\infty\leqslant 54/\varepsilon^2$. We multiply (4.1) by $\tau\chi(k\tau)$ and sum over $k\in\{1,2,\dots,N-1\}$. After performing a discrete integration by parts, we are left with

$$\sum_{k=1}^{N-1}\tau\chi(k\tau)U_{m\beta}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta}\leqslant Ct_0^m+Cm^2\sum_{k=1}^{N-1}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\left[Cm^2+\frac{\chi((k+1)\tau)+\chi((k-1)\tau)-2\chi(k\tau)}{\tau^2}\right].$$

Given the bound on the second derivative of χ, and if $\tau\leqslant\varepsilon/3$, we get

$$\sum_{T_1-\varepsilon/3\leqslant k\tau\leqslant T_2+\varepsilon/3}\tau U_{m\beta}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta}\leqslant Ct_0^m+Cm^2\left(m^2+\frac{1}{\varepsilon^2}\right)\sum_{T_1-\varepsilon\leqslant k\tau\leqslant T_2-\varepsilon}\tau U_m(\bar\rho^{N,\lambda}_{k\tau}).$$

The l.h.s. is not exactly $\left(L^{m\beta}_{T_1,T_2}\right)^{1/m}$ as we would like to exchange the sum and the power $1/\beta$. Unfortunately, Jensen's inequality gives the inequality the other way around. To overcome this difficulty, we will use the fact that the function $k\mapsto U_{\beta m}(\bar\rho^{N,\lambda}_{k\tau})$ is almost a convex function of k. More precisely, we will use the "reverse Jensen inequality", whose proof is postponed in Appendix A.

Lemma 4.3. *Let $(u_k^\tau)_{k\in\mathbb{Z}}$ be a family of real sequences indexed by a parameter τ. We assume that there exists $\omega \geqslant 0$ such that for any $k \in \mathbb{Z}$ and any τ, one has $u_k^\tau > 0$ and*

$$\frac{u_{k+1}^\tau + u_{k-1}^\tau - 2u_k^\tau}{\tau^2} + \omega^2 u_k^\tau \geqslant 0. \tag{4.2}$$

Then, for any $T_1 < T_2$ and any $\eta < \pi/(8\omega)$, there exists τ_0 (which depends on ω), such that, if $\tau \leqslant \tau_0$, then

$$\left(\sum_{T_1 \leqslant k\tau \leqslant T_2} \tau u_k^\tau\right)^{1/\beta} \leqslant C\frac{(\omega+1)(T_2-T_1+1)^{1+1/\beta}}{\eta} \sum_{T_1-\eta\leqslant k\tau\leqslant T_2+\eta} \tau (u_k^\tau)^{1/\beta},$$

where C is a universal constant.

To use this lemma, we observe that $u_k^\tau := U_{m\beta}(\bar\rho_{k\tau}^{N,\lambda})$ satisfies (4.2) with $\omega^2 = Cm^2$ (thanks again to Corollary 3.4 and Remark 3.7). Thus, if we take C_1 small enough, we have $\varepsilon/3 < \pi/(8\omega)$ as soon as $\varepsilon \leqslant C_1/m$. If τ is small enough, we can exchange the sum and the power $1/\beta$ to get

$$\begin{aligned}\left(\sum_{T_1\leqslant k\tau\leqslant T_2} \tau U_{m\beta}(\bar\rho_{k\tau}^{N,\lambda})\right)^{1/\beta} &\leqslant C\frac{(Cm+1)(T+1)^{1+1/\beta}}{\varepsilon} \sum_{T_1-\varepsilon/3\leqslant k\tau\leqslant T_2+\varepsilon/3} \tau U_{m\beta}(\bar\rho_{k\tau}^{N,\lambda})^{1/\beta} \\ &\leqslant C_2\frac{m^3}{\varepsilon}\left(m^2+\frac{1}{\varepsilon^2}\right)\left(t_0 + \sum_{T_1-\varepsilon\leqslant k\tau\leqslant T_2+\varepsilon} \tau U_m(\bar\rho_{k\tau}^{N,\lambda})\right).\end{aligned}$$

Notice that we have put the constant $C_2 m^3 \varepsilon^{-1}(m^2+\varepsilon^{-2})$ also in factor of t_0^m, as it is anyway larger than 1 as soon as ε is small enough. Then we take the power $1/m$ on both sides, use the identity $(a+b)^{1/m} \leqslant C\max(a^{1/m}, b^{1/m})$ and send $N \to +\infty$ and $\lambda \to 0$ to get the result. □

In other words, on a slightly larger time interval, the $L^{\beta m}$ norm is control by the L^m norm. We just have to iterate this inequality.

Proposition 4.4. *Suppose that either Assumption 1 holds or Assumptions 2 and 4 both hold. For any $0 < T_1 < T_2 < T$, there exists C (that depends on T_1, T_2, T, f, V and Ω) such that*

$$\limsup_{m\to+\infty} L^m_{T_1,T_2} \leqslant C\max\left(L^{\alpha+2}_{0,T}, t_0\right).$$

Proof. Let $\varepsilon_0 > 0$ be small enough such that $0 < T_1 - \varepsilon_0\beta/(\beta-1) \leqslant T_2 + \varepsilon_0\beta/(\beta-1) < 1$ and $\varepsilon_0 \leqslant C_1/(\alpha+2)$ (where C_1 is the constant defined in Proposition 4.2). For any $n \in \mathbb{N}$, let us define

$$T_1^n := T_1 - \sum_{k=n}^{+\infty} \frac{\varepsilon_0}{\beta^n} \text{ and } T_2^n := T_2 + \sum_{k=n}^{+\infty} \frac{\varepsilon_0}{\beta^n},$$

and set $m_n := (\alpha+2)\beta^n$. Using Proposition 4.2, as we have $|T_i^{n+1} - T_i^n| = \varepsilon_0\beta^{-n} \leqslant C_1/m_n$ for $i \in \{1,2\}$, we can say that, with $l_n := \max\left(L^{m_n}_{T_1^n,T_2^n}, t_0\right)$

$$\begin{aligned} l_{n+1} &\leqslant \left[\max\left\{1, C_2\frac{m_n^3}{\varepsilon_0\beta^{-n}}\left(m_n^2 + \frac{1}{(\varepsilon_0\beta^{-n})^2}\right)\right\}\right]^{1/m_n} l_n \\ &\leqslant \left[C\beta^{6n}\right]^{\beta^{-n}/(\alpha+2)} l_n.\end{aligned}$$

One can easily check, as $\beta > 1$, that

$$\prod_{n=0}^{+\infty} \left[C\beta^{6n}\right]^{\beta^{-n}/(\alpha+2)} < +\infty,$$

thus we get that

$$\sup_{n\in\mathbb{N}} L^{m_n}_{T_1,T_2} \leqslant \sup_{n\in\mathbb{N}} L^{m_n}_{T_1^n,T_2^n} \leqslant \sup_{n\in\mathbb{N}} l_n \leqslant Cl_0 = C\max\left(L^{m_0}_{T_1^0,T_2^0}, t_0\right) \leqslant C\max\left(L^{\alpha+2}_{0,T}, t_0\right).$$

To conclude, we notice that, if $m > 1$ and $m_n \geqslant m$, one has (using Jensen's inequality)

$$L^m_{T_1,T_2} \leqslant \frac{(m_n(m_n-1))^{1/m_n}}{(m(m-1))^{1/m}} L^{m_n}_{T_1,T_2},$$

thus sending $m \to +\infty$ (hence $n \to +\infty$) we conclude that

$$\limsup_{m\to+\infty} L^m_{T_1,T_2} \leqslant \sup_{n\in\mathbb{N}} L^{m_n}_{T_1,T_2}. \tag{4.3}$$

□

As we will see later, the fact that $L^{\alpha+2}_{0,T}$ is finite is a consequence of the fact that the solution $\bar\rho$ of the continuous problem (ContPb) satisfies $\int_0^T F(\bar\rho_t)\,\mathrm{d}t < +\infty$.

4.2 Estimates up to the final time

In this subsection, still supposing that either Assumption 1 holds or Assumptions 2 and 4 both hold, we exploit Assumption 5 to extend the L^∞ bound up to the final time $t = T$. We will prove a result similar to Proposition 4.2, but this time up to the boundary.

Proposition 4.5. *Suppose that either Assumption 1 holds or Assumptions 2 and 4 both hold, and that Assumption 5 also holds. Then there exists two constants C_1 and C_2 (depending on f, T, V, g, W and Ω) such that for any $0 < \varepsilon < C_1/m$ and any $0 < T_1 < T$ with $0 < T_1 - \varepsilon$, then for any $m \geqslant \alpha + 2$,*

$$L^{\beta m}_{T_1,T} \leqslant \left[C_2 \frac{m^3}{\varepsilon}\left(\frac{m}{\varepsilon} + m^2 + \frac{1}{\varepsilon^2}\right)\right]^{1/m} \max\left(L^m_{T_1-\varepsilon,T}, t_0\right).$$

Again, we recall (Remark 3.7) that if we are under Assumption 1, we take $t_0 = 0$.

Proof. Let us recall that equation (4.1) holds for any $N \geqslant 1$, $\lambda > 0$ and $k \in \{1, 2, \dots, N-1\}$. We take $\chi : [0\,,T] \to [0\,,1]$ a positive $C^{1,1}$ cutoff function such that $\chi(t) = 1$ if $t \in [T_1 - \varepsilon/3\,, T]$ and $\chi(t) = 0$ if $t \in [0\,, T_1 - 2\varepsilon/3]$. Such a function χ can be chosen with $\|\chi''\|_\infty \leqslant 54/\varepsilon^2$. We multiply (4.1) by $\tau\chi(k\tau)$ and sum over $k \in \{1, 2, \dots, N-1\}$. After performing a discrete integration by parts, we are left with (now a boundary term is appearing):

$$\sum_{k=1}^{N-1} \tau\chi(k\tau) U_{m\beta}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta} \leqslant$$
$$Cm^2\left(\frac{U_m(\bar\rho^{N,\lambda}_T) - U_m(\bar\rho^{N,\lambda}_{T-\tau})}{\tau}\chi(T) + \sum_{k=1}^{N-1}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\left[Cm^2 + \frac{\chi(k\tau+\tau)+\chi(k\tau-\tau)-2\chi(k\tau)}{\tau^2}\right]\right) + Ct_0^m.$$

With the help of Proposition 3.8 and Corollary 3.4 or Corollary 3.6, and as $\chi(T) = 1$, we are able to write (provided that $\tau \leqslant \varepsilon/3$)

$$\sum_{T_1-\varepsilon/3\leqslant k\tau\leqslant T} \tau U_{m\beta}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta} \leqslant Cm^2\left(mU_m(\bar\rho^{N,\lambda}_T) + \left[m^2 + \frac{1}{\varepsilon^2}\right]\sum_{T_1-\varepsilon\leqslant k\tau\leqslant T}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\right) + Ct_0^m.$$

To transform the boundary term $U_m(\bar\rho^{N,\lambda}_T)$ into an integral term, we use the following lemma, whose proof is also postponed in Appendix A.

Lemma 4.6. *Let $(u^\tau_k)_{k\in\mathbb{Z}}$ be a family of real sequences indexed by a parameter τ. We assume that there exists $\omega \geqslant 0$ such that for any $k \in \mathbb{Z}$ and any τ, one has $u^\tau_k > 0$ and* (4.2)*. We also assume that there exists $b \geqslant 0$ such that for some $N \in \mathbb{Z}$,*

$$\frac{u^\tau_N - u^\tau_{N-1}}{\tau} \leqslant b u_N \tau.$$

Then, there exists C_1 and C_2 universal constants and τ_0 (which depends on ω and b), such that for any $\eta \leqslant \min\{\pi/(32\omega), \pi/(32b)\}$ and any $\tau \leqslant \tau_0$, then

$$u^\tau_N \leqslant \frac{C_1}{\eta}\sum_{kN-\eta\leqslant k\tau\leqslant kN}\tau u^\tau_k, \tag{4.4}$$

and for any $T_1 < N\tau$,

$$\left(\sum_{T_1 \leqslant k\tau \leqslant N\tau} \tau u_k^\tau\right)^{1/\beta} \leqslant C_2 \frac{(\omega+1)(T-T_1+1)^{1+1/\beta}}{\eta} \sum_{T_1-\eta \leqslant k\tau \leqslant N\tau} \tau (u_k^\tau)^{1/\beta}. \tag{4.5}$$

We are in the case where this lemma can be applied because of Corollary 3.4 or Corollary 3.6 and Proposition 3.8 with $u_k^\tau = U_m(\bar\rho_{k\tau}^{N,\lambda})$, $\omega = Cm$ and $b = Cm$. Thus, if $\varepsilon < C/m$, we can guarantee that we can use equation (4.4) of Lemma 4.6 (with $\varepsilon = \eta$), thus

$$\sum_{T_1-\varepsilon/3 \leqslant k\tau \leqslant T} \tau U_{m\beta}(\bar\rho_{k\tau}^{N,\lambda})^{1/\beta} \leqslant C t_0^m + Cm^2\left[\frac{m}{\varepsilon} + m^2 + \frac{1}{\varepsilon^2}\right] \sum_{T_1-\varepsilon \leqslant k\tau \leqslant T} \tau U_m(\bar\rho_{k\tau}^{N,\lambda}).$$

Then we use equation (4.5) of Lemma 4.6 (but this time with $u_k^\tau = U_{\beta m}(\bar\rho_{k\tau}^{N,\lambda})$) to exchange the sum and the power $1/\beta$ on the l.h.s. to conclude that

$$\left(\sum_{T_1 \leqslant k\tau \leqslant T} \tau U_{m\beta}(\bar\rho_{k\tau}^{N,\lambda})\right)^{1/\beta} \leqslant C\frac{m^3}{\varepsilon}\left[\frac{m}{\varepsilon} + m^2 + \frac{1}{\varepsilon^2}\right]\left(t_0^m + \sum_{T_1-\varepsilon \leqslant k\tau \leqslant T} \tau U_m(\bar\rho_{k\tau}^{N,\lambda})\right).$$

Again, we have put $m^3\varepsilon^{-1}(m\varepsilon^{-1} + m^2 + \varepsilon^{-2})$ in factor of t_0^m, which is legit because this factor is larger than 1 for ε small enough. Taking the power $1/m$ on each side, using the identity $(a+b)^{1/m} \leqslant C\max(a^{1/m}, b^{1/m})$, and letting $N \to +\infty$ and $\lambda \to 0$, we get the result. □

It is then very easy to iterate this result, which looks exactly like Proposition 4.2. Thus, the proof of the following proposition, which is exactly the same as Proposition 4.4, is left to the reader.

Proposition 4.7. *Suppose that either Assumption 1 holds or Assumptions 2 and 4 both hold, and that Assumption 5 also holds. Then, for any $0 < T_1 < T$, there exists C (that depends on T_1, T, f, V and Ω) such that*

$$\limsup_{m\to+\infty} L_{T_1,T}^m \leqslant C\max\left(L_{0,T}^{\alpha+2}, t_0\right).$$

4.3 The weak congestion case

The scheme is very similar in the weak congestion case, even though the iteration is not as direct as in the strong congestion case. Moreover, we will directly prove an L^∞ bound up to $t = T$, because, as we will see, Assumption 5 will be needed anyway to initialize the iterative process. The proofs will be less detailed in this case: the reading on the two previous subsections is advised to understand this one.

Proposition 4.8. *Suppose Assumptions 3, 4 and 5 hold. Then there exist constants C_1 and C_2 (depending on f, V, T and Ω) such that, for any $0 < \varepsilon \leqslant C_1/m$ and any $0 < T_1 < T$ such that $0 < T_1 - \varepsilon$, then for any $m \geqslant m_0$,*

$$L_{T_1,T}^{\beta(m+1+\alpha)} \leqslant \left[C_2\frac{m^{5/2}}{\varepsilon}\left(\frac{m}{\varepsilon} + m + \frac{1}{\varepsilon^2}\right)\right]^{1/(m+1+\alpha)} \max\left[\left(L_{T_1-\varepsilon,T}^m\right)^{m/(m+1+\alpha)}, t_0\right].$$

Proof. The proof starts the same way: starting from Corollary 3.5, we write

$$U_{\beta(m+1+\alpha)}(\bar\rho_{k\tau}^{N,\lambda})^{1/\beta} \leqslant Cm^2\left[\frac{U_m(\bar\rho_{(k-1)\tau}^{N,\lambda}) + U_m(\bar\rho_{(k+1)\tau}^{N,\lambda}) - 2U_m(\bar\rho_{k\tau}^{N,\lambda})}{\tau^2} + CmU_m(\bar\rho_{k\tau}^{N,\lambda})\right] + Ct_0^{m+1+\alpha}. \tag{4.6}$$

Because of Assumption 5, we can also write, tanks to Proposition 3.8, that

$$\frac{U_m(\bar\rho_{T-\tau}^{N,\lambda}) - U_m(\bar\rho_T^{N,\lambda})}{\tau} \geqslant -(m-1)\|\Delta W\|_\infty U_m(\bar\rho_T^{N,\lambda}).$$

We use the same cutoff function χ that in the proof of Proposition 4.5. We multiply (4.6) by $\tau\chi(k\tau)$, perform a discrete integration by parts and end up with

$$\sum_{T_1-\varepsilon/3\leqslant k\tau\leqslant T}\tau U_{\beta(m+1+\alpha)}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta}\leqslant Cm^2\left(\frac{U_m(\bar\rho^{N,\lambda}_T)-U_m(\bar\rho^{N,\lambda}_{T-\tau})}{\tau}+\left[m+\frac{1}{\varepsilon^2}\right]\sum_{T_1-\varepsilon\leqslant k\tau\leqslant T}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\right)+Ct_0^{m+1+\alpha}$$
$$\leqslant Cm^2\left(mU_m(\bar\rho^{N,\lambda}_T)+\left[m+\frac{1}{\varepsilon^2}\right]\sum_{T_1-\varepsilon\leqslant k\tau\leqslant T}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\right)+Ct_0^{m+1+\alpha}.$$

We also use Lemma 4.6 but this time with $\omega^2=Cm$ (this is Corollary 3.5) and $b=Cm$. The frequency ω^2 is smaller than in the strong congestion case (where it was of the order m^2) because we have made stronger assumptions on the potential V, though this is not important as we only use the fact that ω grows not faster than a polynomial of m. With this lemma we can both transform the boundary term into an integral term and exchange the sum and the power $1/\beta$: there exists C_1 such that if $0<\varepsilon\leqslant C_1/m$ and if τ is small enough,

$$\left(\sum_{T_1\leqslant k\tau\leqslant T}\tau U_{\beta(m+1+\alpha)}(\bar\rho^{N,\lambda}_{k\tau})\right)^{1/\beta}\leqslant C\frac{(\sqrt{m}+1)T^{1+1/\beta}}{\varepsilon}\sum_{T_1-\varepsilon/3\leqslant k\tau\leqslant T}\tau U_{\beta(m+1+\alpha)}(\bar\rho^{N,\lambda}_{k\tau})^{1/\beta}$$
$$\leqslant C\frac{m^{5/2}}{\varepsilon}\left(\frac{m}{\varepsilon}+m+\frac{1}{\varepsilon^2}\right)\left(t_0^{m+1+\alpha}+\sum_{T_1-\varepsilon\leqslant k\tau\leqslant T}\tau U_m(\bar\rho^{N,\lambda}_{k\tau})\right).$$

We take the power $1/(m+1+\alpha)$ on both sides, use the fact that $(a+b)^{1/(m+1+\alpha)}\leqslant C\max(a^{1/(m+1+\alpha)},b^{1/(m+1+\alpha)})$, and let $N\to+\infty,\lambda\to 0$ to get the result. □

We proceed the same way by iterating the inequality, even though this expressions are slightly more complicated. Let us underline that the condition (2.5) on β is precisely the one that ensures that $\beta(m+1+\alpha)>m$ as soon as $m\geqslant m_0$: it is only thanks to this condition that the iteration of Proposition 4.8 will give useful information.

Proposition 4.9. *Suppose Assumptions 3, 4 and 5 hold. Then, there exists $\gamma<+\infty$ such that, for any $0<T_1<T$, there exists C (that depends on T_1,T,f,V and Ω) such that*

$$\limsup_{m\to+\infty}L^m_{T_1,T}\leqslant C\left(\max\left[L^{m_0}_{0,T},t_0\right]\right)^\gamma.$$

Proof. As we know, thanks to our normalization choices, that $L^{m_0}_{0,T}\geqslant 1$, it is not restrictive that assume that $t_0\geqslant 1$ (indeed, if this is not the case, Assumption 3 is still valid with $t_0=1$ and the content of Proposition 4.9 does not change).

Once we have chosen $\varepsilon_0<\beta T_1/(\beta-1)$, we define T_1^n by the same formula as in the proof of Proposition 4.4. We also define m_n by recurrence: for any $n\in\mathbb{N}$, we take $m_{n+1}=\beta(m_n+1+\alpha)$. Thus, we have the explicit expression

$$m_n=\left(m_0+(\alpha+1)\frac{\beta}{\beta-1}\right)\beta^n-(\alpha+1)\frac{\beta}{\beta-1}.$$

In particular, $(m_n)_{n\in\mathbb{N}}$ diverges exponentially fast to $+\infty$ as $n\to+\infty$. Using Proposition 4.8 and as $t_0\geqslant 1$, we get

$$L^{m_{n+1}}_{T_1^{n+1},T}\leqslant\left[C_2\frac{m_n^{5/2}}{\varepsilon_0\beta^{-n}}\left(\frac{m_n}{\varepsilon_0\beta^{-n}}+m_n+\frac{1}{(\varepsilon_0\beta^{-n})^2}\right)\right]^{1/(m_n+1+\alpha)}\max\left[\left(L^{m_n}_{T_1^n,T}\right)^{m_n/(m_n+1+\alpha)},t_0\right]$$
$$\leqslant\left[C\beta^{11n/2}\right]^{1/(m_n+\alpha+1)}\max\left(\left[L^{m_n}_{T_1^n,T},t_0\right]\right)^{m_n/(m_n+\alpha+1)}.$$

Denoting by $l_n:=\ln\left(\max\left[L^{m_n}_{T_1^n,T},t_0\right]\right)$, we see that

$$l_{n+1}\leqslant C_3\frac{11n}{2(m_n+\alpha+1)}+\frac{C_4}{m_n+\alpha+1}+\frac{m_n}{m_n+\alpha+1}l_n.$$

Given the exponential asymptotic growth of $(m_n)_{n\in\mathbb{N}}$, we leave it to the reader to check that is enough to conclude that $\limsup_{n\to+\infty} l_n \leqslant \gamma l_0 + C_5$ for some $\gamma < +\infty$ and $C_4 < +\infty$. Taking the exponential gives

$$\limsup_{n\in\mathbb{N}} L^{m_n}_{T_1^n,T} \leqslant C\left(\max\left[L^{m_0}_{0,T}, t_0\right]\right)^{\gamma}.$$

To conclude, we use again (4.3), which is valid independently of Assumption 1 or Assumption 3. □

However, in the weak congestion case, the fact that $L^{m_0}_{0,T} < +\infty$ will require a little bit more of work and relies on the particular form of the boundary conditions.

Proposition 4.10. *Suppose Assumptions 3, 4 and 5 hold. Then there exists T_{max} (which depends on f, V, Ψ and Ω) such that, if $T \leqslant T_{max}$,*

$$L^{m_0}_{0,T} < +\infty.$$

Proof. Again we will use the almost convexity of $U_{m_0}(\bar\rho^{N,\lambda})$. Indeed, we rely on the following lemma, which has the same flavor as the "reverse Jensen inequality" and whose proof is postponed in Appendix A.

Lemma 4.11. *Let $a > 0$, $b \geqslant 0$ and $\omega \geqslant 0$ and set $T_{\max} = \min\{\pi/(32\omega), \pi/(32b)\}$. Then there exist some constants $C < +\infty$ and $\tau_0 > 0$ (all depending on a, b and ω) such that for any $T \leqslant T_{max}$, any $N > 1/\tau_0$ ($\tau := T/N$) and for any sequence $(u^\tau_k)_k \in \mathbb{Z}$ of strictly positive numbers satisfying* (4.2) *for $k \in \{1, 2, \dots, N-1\}$, and such that $u^\tau_0 = a$ and*

$$\frac{u^\tau_{N-1} - u^\tau_N}{\tau} \geqslant -b u^\tau_N,$$

one has $u^\tau_k \leqslant C$ for any $k \in \{1, 2, \dots, N\}$.

We use this lemma with $u^\tau_k = U_{m_0}(\bar\rho^{N,\lambda}_{k\tau})$. Equation (4.2) is satisfied with $\omega^2 = Cm_0$ (Corollary 3.5); one can take

$$a = U_{m_0}(\bar\rho^{N,\lambda}_0) = U_{m_0}(\overline{\rho_0}) = \frac{1}{m_0(m_0-1)}\int_\Omega \overline{\rho_0}^{m_0};$$

and we take $b = (m_0 - 1)\|\Delta W\|_\infty$ (cf. Proposition 3.8). Thus, one can conclude that if $T \leqslant T_{\max}$, then $U_{m_0}(\bar\rho^{N,\lambda}_{k\tau})$ is bounded independently on N. This is enough to conclude that $L^{m_0}_{0,T}$ is finite. □

5 Limit of the discrete problems

In this section, we will see that the solutions $\bar\rho^{N,\lambda}$ of the discrete problems (DiscrPb) converge to the solution $\bar\rho$ of the continuous one (ContPb) when $N \to +\infty$ and $\lambda \to 0$. Then, using the results of the previous sections, we will be able to show the L^∞ bound on $\bar\rho$.

5.1 Building discrete curves from continuous one

In our construction we will need to work with curves with finite entropy. This is easy under Assumption 1 of 2, but requires an approximation in the case of Assumption 3. Hence, we will show that in this case we can approximate curves in Γ by curves in Γ with finite entropy. In order to do this, we will use the heat flow, whose definition and some useful properties are recalled below. For any $s \geqslant 0$ and any $\mu \in \mathcal{P}(\Omega)$, let us define $\Phi_s\mu := u(s)$, where u is the solution of the Cauchy problem

$$\begin{cases} \partial_t u = \Delta u & \text{in } (0\,,+\infty)\times\mathring{\Omega} \\ \nabla u\cdot\mathbf{n} = 0 & \text{on } (0\,,+\infty)\times\partial\Omega \\ \lim_{t\to 0} u(t) = \mu & \text{in } \mathcal{P}(\Omega) \end{cases}.$$

In the equation above, $\mathbf{n}$ stands for the normal vector to the boundary $\partial\Omega$. Provided that the boundary $\partial\Omega$ of Ω is smooth, it is well known (see for instance [6, Section 7] and [28]) that this Cauchy problem is well-posed and admits a unique solution. Because of the no-flux boundary conditions, $\Phi_s\mu \in \mathcal{P}(\Omega)$ for any $s \geqslant 0$. Let us summarize the properties of the heat flow that will be useful to us.

Proposition 5.1.

1. *There exists C (only depending on Ω) such that for all $\mu \in \mathcal{P}(\Omega)$ and any $s > 0$,*
$$\|\Phi_s \mu\|_{L^\infty} \leqslant C s^{-d/2} + 1.$$
2. *If $h : \mathbb{R} \to \mathbb{R}$ is any convex function bounded from below, $\rho \in \mathcal{P}(\Omega) \cap L^1(\Omega)$, and $s \geqslant 0$,*
$$\int_\Omega h\left[(\Phi_s \rho)(x)\right] \, \mathrm{d}x \leqslant \int_\Omega h[\rho(x)] \, \mathrm{d}x;$$
If h is not superlinear, the same stays true for any $\rho \in \mathcal{P}(\Omega)$ by replaing the integral $\int_\Omega h[\rho(x)] \, \mathrm{d}x$ with the expression in (2.3).
3. *If μ and $\nu \in \mathcal{P}(\Omega)$, and $s \geqslant 0$,*
$$W_2(\Phi_s \mu, \Phi_s \nu) \leqslant W_2(\mu, \nu). \tag{5.1}$$
4. *Let $\mu \in \mathcal{P}(\Omega)$ with $U_1(\mu) < +\infty$. Then the curve $s \mapsto \Phi_s \mu$ is 2-absolutely continuous and for any $s \geqslant 0$,*
$$\int_0^s |\dot{\Phi_r \mu}|^2 \, \mathrm{d}r = U_1(\mu) - U_1(\Phi_s(\mu)). \tag{5.2}$$

Proof. The first point is a classic $L^\infty - L^1$ estimate for the heat equation, see for instance [6, Section 7].

To prove the second point in the case of $\rho \in L^1$, let us denote by $K_t(x, y)$ the heat kernel (see [6, Section 7]). Using in particular Jensen's inequality and the fact that $\int_\Omega K_t(x, y) \, \mathrm{d}x = 1$ for any y and t (because $\mathcal{L}$ is invariant under the heat flow),

$$\begin{aligned}
\int_\Omega h\left[(\Phi_s \rho)(x)\right] \, \mathrm{d}x &= \int_\Omega h\left(\int_\Omega K_s(x, y) \rho(y) \, \mathrm{d}y\right) \, \mathrm{d}x \\
&\leqslant \int_{\Omega \times \Omega} K_s(x, y) h(\rho(y)) \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_\Omega h[\rho(y)] \, \mathrm{d}y.
\end{aligned}$$

The proof in the case where h is not superlinear and ρ is not absolutely continuous is obtained by writing $\rho =: \rho^{ac} \mathcal{L} + \rho^{sing}$. Observing that $h'(\infty)$ is the Lipschitz constant of h, we have

$$\int_\Omega h\left[(\Phi_s \rho)(x)\right] \, \mathrm{d}x - \int_\Omega h\left[(\Phi_s \rho^{ac})(x)\right] \, \mathrm{d}x \leqslant h'(\infty) \int_\Omega |\Phi_s \rho^{sing}|(x) \, \mathrm{d}x = h'(\infty) \rho^{sing}(\Omega).$$

The proofs of the third and last points rely on the fact that the heat flow is the gradient flow of the entropy U_1 in the Wasserstein space and can be found in [3, Theorem 11.2.1]. □

Proposition 5.2. *Suppose Assumption 5 holds and that $\overline{\rho_0}$ is such that $U_1(\overline{\rho_0}), F(\overline{\rho_0}) < +\infty$, and let $\rho \in \Gamma$ with $\rho_0 = \overline{\rho_0}$. Then, for any $\varepsilon > 0$, there exists $\tilde{\rho} \in \Gamma$ with $\tilde{\rho}_0 = \overline{\rho_0}$ and $C < +\infty$ such that $\mathcal{A}(\tilde{\rho}) \leqslant \mathcal{A}(\rho) + \varepsilon$ and $U_1(\tilde{\rho}_t) \leqslant C$ for any $t \in [0, 1]$.*

Proof. Without loss of generality, we assume $\mathcal{A}(\rho) < +\infty$. The idea is to use the heat flow to regularize solutions. But we cannot apply the heat flow uniformly, as we would loose the boundary condition $\rho_0 = \overline{\rho_0}$. Consequently, for any $0 < s < T$, we define $\tilde{\rho}^s \in \mathcal{P}(\Gamma)$ by

$$\tilde{\rho}^s(t) := \begin{cases} \Phi_t(\rho_0) & \text{if } 0 \leqslant t \leqslant s \\ \Phi_s\left(\rho\left[\dfrac{t - s}{T - s} T\right]\right) & \text{if } s \leqslant t \leqslant T \end{cases}.$$

In other words, we take the curve $\Phi_s \rho$, squeeze it into $[s, T]$, and use the heat flow to join ρ_0 to ρ_s on $[0, s]$. In particular, $\tilde{\rho}^s_0 = \rho_0 = \overline{\rho_0}$ and $\tilde{\rho}^s_T = \Phi_s \rho_T$. From $U_1(\rho_0) < +\infty$ and the fact that U_1 is decreasing along the heat flow (see Proposition 5.1), $U_1(\tilde{\rho}_t)$ is bounded by $U_1(\rho_0)$ if $t \in [0, s]$ and by a constant C_s (depending only on s and Ω) if $t \in [s, T]$. Hence, for any $s > 0$, there exists $C < +\infty$ such that $U_1(\tilde{\rho}^s_t) \leqslant C$ for any $t \in [0, 1]$.

It remains to show that $\mathcal{A}$ does not increase too much because of our regularization process. Using the second point of Proposition 5.1, one can see that

$$\int_0^T F(\tilde{\rho}_t^s)\,\mathrm{d}t + G(\tilde{\rho}_T^s) \leqslant sF(\overline{\rho_0}) + \frac{(T-s)}{T}\int_0^T F(\rho_t)\,\mathrm{d}t + G(\rho_T).$$

To handle the action of $\tilde{\rho}^s$, we remark thanks to the third point of Proposition 5.1 and the representation formula (2.4) that applying uniformly the heat flow decreases the action. Hence, performing a affine change of variables on $[s\,,T]$,

$$\begin{aligned}\int_0^T |\dot{\tilde{\rho}}_t^s|^2\,\mathrm{d}t &= \int_0^s |\dot{\Phi_t\rho_0}|^2\,\mathrm{d}t + \frac{T}{T-s}\int_0^T |\dot{\Phi_s\rho_t}|^2\,\mathrm{d}t\\ &\leqslant U_1(\rho_0) - U_1(\Phi_s\rho_0) + \frac{T}{T-s}\int_0^T |\dot{\rho}_t|^2\,\mathrm{d}t.\end{aligned}$$

By lower semi-continuity of U_1 and as $U_1(\rho_0) = U_1(\overline{\rho_0})$ is finite, one concludes that

$$\limsup_{s\to 0}\int_0^T |\dot{\tilde{\rho}}_t^s|^2\,\mathrm{d}t \leqslant \int_0^T |\dot{\rho}_t|^2\,\mathrm{d}t.$$

Finally to handle the term involving the potentials, one uses, by continuity of the heat flow, that $\tilde{\rho}_t^s$ converges to ρ_t for any $t\in[0\,,1]$ as s goes to 0. As $\int_0^T |\dot{\tilde{\rho}}_t^s|^2\,\mathrm{d}t$ is uniformly bounded, the family $(\tilde{\rho}^s)_{0\leqslant s<T}$ is uniformly equicontinuous, hence $\tilde{\rho}^s$ converges uniformly to ρ as $s\to 0$. This allows us to write

$$\lim_{s\to 0}\left[\int_0^T\int_\Omega V\,\mathrm{d}\tilde{\rho}_t^s\,\mathrm{d}t + \int_\Omega W\,\mathrm{d}\tilde{\rho}_T^s\right] = \int_0^T\int_\Omega V\,\mathrm{d}\rho_t\,\mathrm{d}t + \int_\Omega W\,\mathrm{d}\rho_T.$$

Gluing all the inequalities that we have collected, we see that $\limsup_{s\to 0}\mathcal{A}(\tilde{\rho}^s) \leqslant \mathcal{A}(\rho)$. Hence, it is enough to take $\tilde{\rho} := \tilde{\rho}^s$ for s small enough. □

Now, let us show how one can sample a continuous curve to get a discrete one that approximates it.

Proposition 5.3. *Let $\rho\in\Gamma$ with $\rho_0 = \overline{\rho_0}$ be such that $\int_0^T U_1(\rho_t)\,\mathrm{d}t < +\infty$ and $\lambda > 0$ be fixed. For any $N\geqslant 1$ we can build a curve $\rho^N\in\Gamma_N$ with $\rho_0^N = \overline{\rho_0}$ in such a way that*

$$\limsup_{N\to+\infty}\mathcal{A}^{N,\lambda}(\rho^N) \leqslant \mathcal{A}(\rho) + \lambda\int_0^T U_1(\rho_t)\,\mathrm{d}t + \lambda_N U_1(\rho_T).$$

We recall that $\lambda_N = 0$ by default except if Assumption 5 holds.

Proof. We can assume $\mathcal{A}(\rho) < +\infty$. The idea is to sample ρ on a grid translated w.r.t. T^N. We start with the following observation.

$$\begin{aligned}\int_0^\tau \sum_{k=1}^{N-1}\left(F(\rho_{k\tau+s}) + \lambda U_1(\rho_{k\tau+s})\right)\,\mathrm{d}s &= \int_0^{T-\tau}\left(F(\rho_t) + \lambda U_1(\rho_t)\right)\,\mathrm{d}t\\ &\leqslant \int_0^T\left(F(\rho_t) + \lambda U_1(\rho_t)\right)\,\mathrm{d}t + C\tau,\end{aligned}$$

where C depends only on the lower bounds of F and U_1. Therefore, there exists $s_N\in(0\,,\tau)$ such that

$$\tau\sum_{k=1}^{N-1}\left(F(\rho_{k\tau+s_N}) + \lambda U_1(\rho_{k\tau+s_N})\right) \leqslant \int_0^T\left(F(\rho_t) + \lambda U_1(\rho_t)\right) + C\tau.$$

Let us define $\rho^N\in\Gamma_N$ by sampling ρ on the grid translated by s_N: for any $k\in\{0,1,\dots,N\}$,

$$\rho^N := \begin{cases}\rho_0 & \text{if } k = 0\\ \rho_T & \text{if } k = N\\ \rho_{k\tau+s_N} & \text{if } 1\leqslant k\leqslant N-1\end{cases}.$$

As the boundary values are left unchanged and given the choice of s_N, it is clear that

$$\left(\mathcal{A}(\rho)+\lambda\int_0^T U_1(\rho_t)\,\mathrm{d}t+\lambda_N U_1(\rho_T)\right)-\mathcal{A}^{N,\lambda}(\rho^N)\geqslant\int_0^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t-\sum_{k=1}^N\frac{W_2^2(\rho^N_{(k-1)\tau},\rho^N_{k\tau})}{2\tau}-C\tau.$$

The r.h.s. of the above equation is delicate to evaluate because of the non uniformity of the grid near the boundaries. Recall that if $t\leqslant s$ then $W_2^2(\rho_t,\rho_s)\leqslant(s-t)\int_t^s|\dot\rho_r|^2\,\mathrm{d}r$, hence

$$\begin{aligned}\sum_{k=1}^N\frac{W_2^2(\rho^N_{(k-1)\tau},\rho^N_{k\tau})}{2\tau}&=\frac{W_2^2(\rho_0,\rho_{\tau+s_N})}{2\tau}+\sum_{k=2}^{N-1}\frac{W_2^2(\rho_{(k-1)\tau+s_N},\rho_{k\tau+s_N})}{2\tau}+\frac{W_2^2(\rho_{(k-1)\tau+s_N},\rho_T)}{2\tau}\\&\leqslant\frac{\tau+s_N}{2\tau}\int_0^{\tau+s_N}\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t+\sum_{k=2}^{N-1}\int_{(k-1)\tau+s_n}^{k\tau+s_N}\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t+\frac{\tau-s_N}{2\tau}\int_{T-\tau+s_N}^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t\\&\leqslant\int_0^{\tau+s_N}|\dot\rho_t|^2\,\mathrm{d}t+\int_{\tau+s_N}^{T-\tau+s_N}\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t+\int_{T-\tau+s_N}^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t\\&\leqslant\int_0^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t+\int_0^{2\tau}\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t.\end{aligned}$$

In particular, we have used $\tau+s_N\leqslant 2\tau$ and $\tau-s_N\leqslant\tau$. Letting $N\to+\infty$ (hence $\tau\to 0$), we end up with

$$\limsup_{N\to+\infty}\sum_{k=1}^N\frac{W_2^2(\rho^N_{(k-1)\tau},\rho^N_{k\tau})}{2\tau}\leqslant\int_0^T\frac{1}{2}|\dot\rho_t|^2\,\mathrm{d}t,$$

and this is enough to conclude. □

Corollary 5.4. *Under the assumptions of Theorems 2.10, 2.11 or 2.12, there exists $C<+\infty$ such that, uniformly in $N\geqslant 1$ and $\lambda\in(0\,,1]$, one has*

$$\mathcal{A}^{N,\lambda}(\bar\rho^{N,\lambda})\leqslant C.$$

Proof. If we are under the assumptions of Theorems 2.10 ot 2.11, we take $\rho\in\Gamma$ such that $\mathcal{A}(\rho)<+\infty$. As $U_1\leqslant C_f F+C$, we see that $\int_0^T U_1(\rho_t)\,\mathrm{d}t<+\infty$. If we are under the assumptions of 2.12, we take $\rho\in\Gamma$ such that $\mathcal{A}(\rho)<+\infty$ and regularize it thanks to Proposition 5.2. For this regularized curve, one has $\int_0^T U_1(\rho_t)\,\mathrm{d}t+\lambda U_1(\rho_T)<+\infty$.

In any of these two cases, we construct ρ^N as in Proposition 5.3 and define $C:=\sup_{N\geqslant 1}\mathcal{A}^{N,\lambda}(\rho^N)$, then we use the fact that $\mathcal{A}^{N,\lambda}(\bar\rho^{N,\lambda})\leqslant\mathcal{A}^{N,\lambda}(\rho^N)\leqslant C$. □

5.2 Solution of the continuous problem as limit of discrete curves

We will build a suitable interpolation of the discrete curves $\bar\rho^{N,\lambda}$ that will converge to some continuous curve $\bar\rho$ as $N\to+\infty$ and $\lambda\to 0$, and we will show that $\bar\rho$ is a solution of (ContPb).

As the order in which the limits $N\to+\infty$ and $\lambda\to 0$ are taken does not matter, we will do them in the same time. We take two sequences $(N_n)_{n\in\mathbb{N}}$ and $(\lambda_n)_{n\in\mathbb{N}}$ that go respectively to $+\infty$, and 0 (the second one being strictly positive). We will not relabel the sequences when extracting subsequences. Moreover, to avoid heavy notations, we will drop the index n, and $\lim_{n\to+\infty}$ will be denoted by $\lim_{N\to+\infty,\lambda\to 0}$. We will need to define two kind of interpolations: one filling the gaps with constant-speed geodesics, and the other one by using piecewise constant curves.

Definition 5.5. *If $N\geqslant 1$ and $\lambda>0$, we define $\hat\rho^{N,\lambda}\in\Gamma$ as the curve such that $\hat\rho^{N,\lambda}$ coincides with $\bar\rho^{N,\lambda}$ on T^N, and such that for any $k\in\{0,1,\dots,N-1\}$, the restriction of $\hat\rho^{N,\lambda}$ to $[k\tau\,,(k+1\tau)]$ is the constant-speed geodesic joining $\bar\rho^{N,\lambda}_{k\tau}$ to $\bar\rho^{N,\lambda}_{(k+1)\tau}$.*

As $\bar\rho^{N,\lambda}_{k\tau}$ is absolutely continuous w.r.t. $\mathcal{L}$ for any $k\in\{1,2,\dots,N-1\}$, the constant-speed geodesic between $\bar\rho^{N,\lambda}_{k\tau}$ and $\bar\rho^{N,\lambda}_{(k\pm1)\tau}$ is always unique. From the characterization of constant-speed geodesics, one has, for any $k\in\{0,1,\dots,N-1\}$,

$$\int_{k\tau}^{(k+1)\tau}\frac{1}{2}\left|\dot{\hat\rho}^{N,\lambda}_t\right|^2\,\mathrm{d}t=\frac{W_2^2(\bar\rho^{N,\lambda}_{k\tau},\bar\rho^{N,\lambda}_{(k+1)\tau})}{2\tau}.$$

Summing these identities over k,

$$\int_0^T \frac{1}{2}\left|\dot{\hat{\rho}}_t^{N,\lambda}\right|^2 \,\mathrm{d}t = \sum_{k=1}^{N} \frac{W_2^2(\bar{\rho}_{(k-1)\tau}^{N,\lambda}, \bar{\rho}_{k\tau}^{N,\lambda})}{2\tau}. \tag{5.3}$$

In other words, the continuous action of the interpolated curve $\hat{\rho}^{N,\lambda}$ is equal to the discrete action of the discrete curve $\bar{\rho}^{N,\lambda}$.

Definition 5.6. *If $N \geqslant 1$ and $\lambda > 0$, we define $\tilde{\rho}^{N,\lambda} : [0\,,T] \to \mathcal{P}(\Omega)$ as the function such that $\tilde{\rho}^{N,\lambda}$ coincides with $\bar{\rho}^{N,\lambda}$ on T^N, and such that for any $k \in \{0,1,\dots,N-1\}$, the restriction of $\tilde{\rho}^{N,\lambda}$ to $[k\tau\,,(k+1\tau))$ is equal to $\bar{\rho}_{k\tau}^{N,\lambda}$.*

The curve $\tilde{\rho}^{N,\lambda}$ is not continuous as it might admit discontinuities at every point in T^N. Let us underline that the following identity trivially holds:

$$\sum_{k=1}^{N-1} \tau\left(F(\bar{\rho}_{k\tau}^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\bar{\rho}_{k\tau}^{N,\lambda}\right) = \int_0^{T-\tau}\left(F(\tilde{\rho}_t^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\tilde{\rho}_t^{N,\lambda}\right)\mathrm{d}t \tag{5.4}$$

Proposition 5.7. *Under the assumptions of Theorems 2.10, 2.11 or 2.12, there exists $\bar{\rho} \in \Gamma$ such that $\hat{\rho}^{N,\lambda}$ and $\tilde{\rho}^{N,\lambda}$ converge uniformly to $\bar{\rho}$ as $N \to +\infty$ and $\lambda \to 0$.*

Proof. Let us denote by C the constant given in Corollary 5.4. As all the terms in $\mathcal{A}^{N,\lambda}$ are bounded from below and given identity (5.3), one can see that there exists C_1 such that

$$\int_0^T \frac{1}{2}\left|\dot{\hat{\rho}}_t^{N,\lambda}\right|^2 \,\mathrm{d}t \leqslant C_1$$

uniformly in $N \geqslant 1$ and $\lambda \in (0\,,1]$. Thus, by compactness of the sublevel sets of the action (Proposition 2.7), one concludes of the existence of $\bar{\rho} \in \Gamma$ such that $\hat{\rho}^{N,\lambda}$ converges uniformly (up to extraction) to $\bar{\rho}$ as $N \to +\infty$ and $\lambda \to 0$. Moreover, one can see that for any $t \in [0\,,T]$ and any $N \geqslant 1$, by setting k to be the largest integer such that $k\tau \leqslant t$, one has

$$W_2\left(\hat{\rho}_t^{N,\lambda}, \tilde{\rho}_t^{N,\lambda}\right) = W_2\left(\hat{\rho}_t^{N,\lambda}, \hat{\rho}_{k\tau}^{N,\lambda}\right) \leqslant \sqrt{\tau}\sqrt{\int_{k\tau}^t \left|\dot{\hat{\rho}}_s^{N,\lambda}\right|^2 \,\mathrm{d}s} \leqslant \sqrt{2C_1\tau}.$$

This allows to conclude that $\tilde{\rho}^{N,\lambda}$ also converges uniformly to $\bar{\rho}$ as $N \to +\infty$ and $\lambda \to 0$. □

Proposition 5.8. *Under the assumptions of Theorems 2.10, 2.11 or 2.12, the curve $\bar{\rho}$ is the solution to the continuous problem* (ContPb)*.*

Proof. Taking the limit $N \to +\infty$ and $\lambda \to 0$ in (5.3), as the action is l.s.c., we end up with

$$\int_0^T \frac{1}{2}|\dot{\bar{\rho}}_t|^2 \,\mathrm{d}t \leqslant \liminf_{N\to+\infty,\lambda\to 0} \sum_{k=1}^{N} \frac{W_2^2(\bar{\rho}_{(k-1)\tau}^{N,\lambda}, \bar{\rho}_{k\tau}^{N,\lambda})}{2\tau}.$$

Then, to handle the terms with the potential and the congestion, one can notice that for any $t \in [0\,,T]$, by lower semi-continuity of F and the convergence of $\tilde{\rho}_t^{N,\lambda}$ to $\bar{\rho}_t$,

$$F(\bar{\rho}_t) + \int_\Omega V \,\mathrm{d}\bar{\rho}_t \leqslant \liminf_{N\to+\infty,\lambda\to 0} F(\tilde{\rho}_t^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\tilde{\rho}_t^{N,\lambda}.$$

Thus, using Fatou's lemma, as F, V and U_1 are bounded from below, one has for any $\tau_0 > 0$,

$$\begin{aligned}
\int_0^{T-\tau_0}\left(F(\bar{\rho}_t) + \int_\Omega V \,\mathrm{d}\bar{\rho}_t\right) &\leqslant \liminf_{N\to+\infty,\lambda\to 0} \int_0^{T-\tau}\left(F(\tilde{\rho}_t^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\tilde{\rho}_t^{N,\lambda}\right)\mathrm{d}t \\
&= \liminf_{N\to+\infty,\lambda\to 0} \sum_{k=1}^{N-1} \tau\left(F(\bar{\rho}_{k\tau}^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\bar{\rho}_{k\tau}^{N,\lambda}\right)\mathrm{d}t \\
&\leqslant \liminf_{N\to+\infty,\lambda\to 0} \sum_{k=1}^{N-1} \tau\left(F(\bar{\rho}_{k\tau}^{N,\lambda}) + \int_\Omega V \,\mathrm{d}\bar{\rho}_{k\tau}^{N,\lambda} + \lambda U_1(\bar{\rho}_{k\tau}^{N,\lambda})\right).
\end{aligned}$$

In the equation above, τ_0 is arbitrary thus it is still valid for $\tau_0 = 0$. As moreover the boundary penalization Ψ is l.s.c. and the entropic penalization $\lambda_N U_1(\rho_T)$ is positive, one is allowed to write that

$$\mathcal{A}(\bar{\rho}) \leqslant \liminf_{N\to+\infty,\lambda\to 0} \mathcal{A}^{N,\lambda}(\bar{\rho}^{N,\lambda}).$$

Let us assume by contradiction that there exists $\rho \in \Gamma$ such that $\mathcal{A}(\rho) < \mathcal{A}(\bar{\rho})$. Using, if needed, Proposition 5.2, we can assume without loss of generality that $\mathcal{A}(\rho) < \mathcal{A}(\bar{\rho})$ and $\int_0^T U_1(\rho_t)\,\mathrm{d}t + \lambda_N U_1(\rho_T) < +\infty$. Using Proposition 5.3, for any $N \geqslant 1$, we can build $\rho^N \in \Gamma_N$ in such a way that

$$\limsup_{N\to+\infty} \mathcal{A}^{N,\lambda}(\rho^N) \leqslant \mathcal{A}(\rho) + \lambda \int_0^T U_1(\rho_t)\,\mathrm{d}t + \lambda_N U_1(\rho_T).$$

Taking the limit $\lambda \to 0$, one can see that

$$\limsup_{N\to+\infty,\lambda\to 0} \mathcal{A}^{N,\lambda}(\rho^N) \leqslant \mathcal{A}(\rho) < \mathcal{A}(\bar{\rho}) \leqslant \liminf_{N\to+\infty,\lambda\to 0} \mathcal{A}^{N,\lambda}(\bar{\rho}^{N,\lambda}).$$

Taking N large enough and λ small enough, we conclude that $\mathcal{A}^{N,\lambda}(\rho^N) < \mathcal{A}^{N,\lambda}(\bar{\rho}^{N,\lambda})$, which is a contradiction with the optimality of $\bar{\rho}^{N,\lambda}$. □

5.3 Uniform bounds on $\bar{\rho}$

To conclude and prove the Theorems 2.10, 2.11 and 2.12, it is enough to show the L^∞ bounds on $\bar{\rho}$, which of course we will prove using the discrete solutions $\bar{\rho}^{N,\lambda}$. The key is the following proposition.

Proposition 5.9. *Let $0 < T_1 < T_2 \leqslant T$. Then for any $0 < T_1' < T_1$ and any $T_2 < T_2' < T$ (or $T_2' = T_2 = T$ in the case $T_2 = T$),*

$$\operatorname*{ess\,sup}_{T_1\leqslant t\leqslant T_2, x\in\Omega} |\bar{\rho}_t(x)| \leqslant \limsup_{m\to+\infty} L^m_{T_1',T_2'}.$$

Proof. We rely on the well-known identity

$$\operatorname*{ess\,sup}_{T_1\leqslant t\leqslant T_2,\ x\in\Omega} |\bar{\rho}_t(x)| = \limsup_{m\to+\infty} \left(\int_{T_1}^{T_2}\int_\Omega \bar{\rho}_t^m\,\mathrm{d}t\right)^{1/m} = \limsup_{m\to+\infty} \left(\int_{T_1}^{T_2} U_m(\bar{\rho}_t)\,\mathrm{d}t\right)^{1/m}.$$

For a fixed $m > 1$ and for $\tau > 0$ small enough, one has

$$\int_{T_1}^{T_2} U_m(\tilde{\rho}_t^{N,\lambda})\,\mathrm{d}t \leqslant \sum_{T_1'\leqslant k\tau\leqslant T_2'} \tau U_m(\bar{\rho}_{k\tau}^{N,\lambda}).$$

When sending $N \to \infty$ and $\lambda \to 0$, by lower semi-continuity of U_m and by convergence of $\tilde{\rho}^{N,\lambda}$ to $\bar{\rho}$, we know that

$$\begin{aligned}\int_{T_1}^{T_2} U_m(\bar{\rho}_t)\,\mathrm{d}t &\leqslant \liminf_{N\to+\infty,\lambda\to 0} \int_{T_1}^{T_2} U_m(\tilde{\rho}_t^{N,\lambda})\,\mathrm{d}t \\ &\leqslant \liminf_{N\to+\infty,\lambda\to 0} \sum_{T_1'\leqslant k\tau\leqslant T_2'} \tau U_m(\bar{\rho}_{k\tau}^{N,\lambda}).\end{aligned}$$

Taking the power $1/m$ on each side and by definition of $L^m_{T_1',T_2'}$, one gets

$$\left(\int_{T_1}^{T_2} U_m(\bar{\rho}_t)\,\mathrm{d}t\right)^{1/m} \leqslant L^m_{T_1',T_2'}.$$

It is enough to take the limit $m \to +\infty$ to get the announced inequality. □

We can now conclude the desired bounds:

Proof of Theorem 2.10. Combining Proposition 5.9 and Proposition 4.4, it is enough to show that $L^{\alpha+2}_{0,T} < +\infty$. Because of Assumption 1 or 2, we know that $U_{\alpha+2} \leqslant C_1 F + C_2$ with $C_1 > 0$. Hence, in order to conclude that $L^{\alpha+2}_{0,T} < +\infty$, it is enough to use Corollary 5.4, which provides a constant $C < +\infty$ such that for any $N \geqslant 1$ and any $\lambda \in (0\,,1]$ we have

$$\sum_{k=1}^{N-1} \tau F(\bar{\rho}^{N,\lambda}_{k\tau}) \leqslant C. \qquad \square$$

Proof of Theorem 2.11. We combine Proposition 5.9 and Proposition 4.7, as we saw that $L^{\alpha+2}_{0,T} < +\infty$ (in the proof of Theorem 2.10). $\square$

Proof of Theorem 2.12. It is enough to combine Proposition 5.9 with Propositions 4.9 and 4.10. $\square$

A Reverse Jensen inequality

In this section, we prove Lemma 4.3 (the "reverse Jensen inequality") as well as Lemmas 4.6 and 4.11, whose proofs were postponed in order not to overload the key arguments of the paper. In all the sequel, we consider a family of sequences $(u^\tau_k)_{k\in\mathbb{Z}}$ indexed by a parameter $\tau > 0$. We assume that there exists $\omega \geqslant 0$ such that for any $k \in \mathbb{Z}$, one has $u^\tau_k > 0$ and

$$\frac{u^\tau_{k+1} + u^\tau_{k-1} - 2u^\tau_k}{\tau^2} + \omega^2 u^\tau_k \geqslant 0. \tag{A.1}$$

This inequation is a discrete counterpart of the differential inequality $u'' + \omega^2 u \geqslant 0$. Let us remark, by the positivity of u^τ_k, that we can assume without loss of generality that $\omega > 0$, even though the proofs are considerably simpler if $\omega = 0$: the constants would be better, and the strategy of the proof would be slightly different. The key point to handle u^τ_k is to compare it with explicit sequences realizing the opposite inequality in (A.1).

Definition A.1. *For any $\tau > 0$, let $\mathcal{T}^\tau$ be the set of sequences $(v_k)_{k\in\mathbb{Z}}$ of the form $v_k = A\cos(2\omega k\tau + \delta)$.*

Lemma A.2. *There exists $\tau_0 > 0$ such that for any $\tau \leqslant \tau_0$, if $(v_k)_{k\in\mathbb{Z}} \in \mathcal{T}^\tau$ and k is such that $v_k > 0$ then*

$$\frac{v_{k+1} + v_{k-1} - 2v_k}{\tau^2} + \omega^2 v_k < 0$$

Proof. This is a consequence of the trigonometric identity

$$\frac{v_{k+1} + v_{k-1} - 2v_k}{\tau^2} + \omega^2 v_k = \left(2\frac{\cos(2\omega\tau) - 1}{\tau^2} + \omega^2\right) v_k$$

and the fact that $2\dfrac{\cos(2\omega\tau) - 1}{\tau^2} + \omega^2 \sim -3\omega^2$ as τ goes to 0. $\square$

We also note the following properties on the sequences in $\mathcal{T}^\tau$, that we do not prove and leave to the reader as an exercise.

- if $k_1 < k_2$ are fixed with $|k_2 - k_1|\tau\omega < \pi/8$ and τ is small enough, then for every fixed positive values $a_1, a_2 > 0$ there exists a unique sequence in $\mathcal{T}^\tau$ with $v_{k_1} = a_1$ and $v_{k_2} = a_2$. Moreover, such a sequence $(v_k)_{k\in\mathbb{Z}}$ is such that there exists an open interval I of the form either $(k_0\tau, k_1\tau)$ or $(k_2\tau, k_3\tau)$, with length at least $\pi/(8\omega)$, with $v_k > 0$ for all the indices k such that $k\tau \in I$.
- if $k_1 < N$ and $b \geqslant 0$ are fixed and $|N - k_1|\tau < \min\{\pi/(8\omega), \pi/(8b)\}$ and τ is small enough, then for every $a > 0$ there exists a unique sequence in $\mathcal{T}^\tau$ with $v_{k_1} = a$ and $(v_N - v_{N-1})/\tau = bv_N$. Moreover, such a sequence $(v_k)_{k\in\mathbb{Z}}$ is such that there exists an open interval I of the form $(k_0\tau, k_1\tau)$ with length at least $\min\{\pi/(32\omega), \pi/(32b)\}$, with $v_k > 0$ for all the indices k such that $k\tau \in I$.

Note that, for the purpose of Lemma A.2 and of the subsequent observations other choices of v_k were possible, such as $v_k = A\cos((1+\varepsilon)\omega k\tau + \delta)$ for some $\varepsilon > 0$, but we chose $\varepsilon = 1$ for simplicity in the next computations (more generally in this appendix we have not been looking for the sharpest constants). Indeed, all these results are not surprising: at the continuous level v solves $v'' + 4\omega^2 v = 0$ and most of the discrete results are just an adaptation of this property. The important point is the following comparision principle between $(u^\tau_k)_{k\in\mathbb{Z}}$ and $(v_k)_{k\in\mathbb{Z}}$.

Lemma A.3. *Let $k_1 < k_2$ such that $|k_2 - k_1|\tau\omega < \pi/8$ and assume $\tau \leqslant \tau_0$. Let $v \in \mathcal{T}^\tau$ the unique element of $\mathcal{T}^\tau$ such that $v_{k_1} = u^\tau_{k_1}$ and $v_{k_2} = u^\tau_{k_2}$. Let k_0 (resp. k_3) be the largest (resp. the smallest) index smaller that k_1 (resp. larger than k_2) such that $v_{k_0-1} < 0$ (resp. $v_{k_3+1} < 0$).*

Then $u^\tau_k \leqslant v_k$ for any $k_1 \leqslant k \leqslant k_2$ and $u^\tau_k \geqslant v_k$ for any $k_0 \leqslant k \leqslant k_1$ and any $k_2 \leqslant k \leqslant k_3$.

In other words, u^τ is below v between k_1 and k_2 and above outside k_1 and k_2 (as long as $v \geqslant 0$).

Proof. The fact that there exists only one $v \in \mathcal{T}^\tau$ such that $v_{k_1} = u^\tau_{k_1}$ and $v_{k_2} = u^\tau_{k_2}$ has been already observed above. Let us define $w_k = u^\tau_k - v_k$. By (A.1) and Lemma A.2,

$$\frac{w_{k+1} + w_{k-1} - 2w_k}{\tau^2} + \omega^2 w_k > 0 \tag{A.2}$$

for any $k_0 \leqslant k \leqslant k_3$ and $w_{k_1} = w_{k_2} = 0$. We want to prove $w_k \leqslant 0$ for every $k_1 \leqslant k \leqslant k_2$. We consider the piecewise affine interpolation $\bar{w}$ of the values w_k: a function which is affine on each interval $[k\tau, (k+1)\tau]$ and is equal to w_k at the point $k\tau$. The condition (A.2) translates on $\bar{w}$ as differential inequality in the sense of distributions:

$$\bar{w}'' + \omega^2 \sum_k \tau w_k \delta_{k\tau} \geqslant 0. \tag{A.3}$$

Let us assume by contradiction that there is an open interval $I \subset (k_1\tau, k_2\tau)$ on which $\bar{w} > 0$, with $\bar{w} = 0$ on ∂I. We denote by $|I|$ the length of such an interval, and we have $|I| \leqslant |k_2 - k_1|\tau$. By multiplying the above inequality by $\bar{w}$ and integrating by parts we get

$$\int_I |\bar{w}'|^2 = -\int_I \bar{w}''\bar{w} \leqslant \omega^2 \sum_{k\,:\,k\tau\in I} \tau |w_k|^2.$$

Then, we observe that we have, for each k s.t. $k\tau \in I$,

$$|w_k| \leqslant \frac{1}{2}\int_I |\bar{w}'| \leqslant \frac{1}{2}\sqrt{|I|\int_I |\bar{w}'|^2}.$$

The reason for the factor $1/2$ in the above inequality is the possibility to choose to integrate $\bar{w}'$ on an interval at the right or at the left of $k\tau$, and to choose the one where the integral of $|\bar{w}'|$ is smaller. This implies

$$\int_I |\bar{w}'|^2 \leqslant \omega^2\tau\, \#\{k : k\tau \in I\}\frac{1}{4}|k_2 - k_1|\tau \int_I |\bar{w}'|^2.$$

Since $\{k : k\tau \in I\} \subset \{k : k_1 < k < k_2\}$, we have $\#\{k : k\tau \in I\} < |k_2 - k_1|$ and the contradiction comes from the assumption $\omega\tau|k_2 - k_1| < \pi/8 < 2$.

In order to prove $w_k \geqslant 0$ for $k_0 \leqslant k \leqslant k_1$, we first observe that (A.2) for $k = k_1$, now that we know $w_{k_1+1} \leqslant 0$, implies $w_{k_1-1} > 0$. If for some k with $k_0 \leqslant k \leqslant k_1$ we had $w_k < 0$, then we could find an open interval $J \subset (k_0\tau, k_1\tau)$ where $\bar{w} > 0$ with $\bar{w} = 0$ on ∂J. We then apply the same approach as above, thus obtaining

$$\int_J |\bar{w}'|^2 \leqslant \omega^2\tau\, \#\{k : k\tau \in J\}\frac{1}{4}|J| \int_J |\bar{w}'|^2.$$

It is important to not that J is contained in an interval of positivity of a function of the form $A\cos(2\omega t + \delta)$, whose length is $\pi/(2\omega)$; the number of points of the form $k\tau$ contained in an interval of such a length is at most $\pi/(2\omega\tau) + 1$ but for $k = k_1, k_2$ the point $k\tau$ does not belong to the open interval J. Hence $\#\{k : k\tau \in J\} \leqslant \pi/(2\omega\tau)$, and we have a contradiction since $\pi^2 < 16$. □

We provide now a variant in the case where on the interval $(k_1\tau, k_2\tau)$ we impose a different boundary condition on the right end side.

Lemma A.4. *Let $k_1 < N$ and $b \geqslant 0$ such that $|N - k_1|\tau < \min\{\pi/(8\omega), \pi\,(8b)\}$ and assume $\tau \leqslant \tau_0$. Suppose $(u_N - u_{N-1})/\tau \leqslant bu_N$. Let $v \in \mathcal{T}^\tau$ the unique element of $\mathcal{T}^\tau$ such that $v_{k_1} = u^\tau_{k_1}$ and $(v_N - v_{N-1})/\tau = bv_N$. Let k_0 be the largest (resp. the smallest) index smaller that $k_1$such that $v_{k_0-1} < 0$.*

Then $u^\tau_k \leqslant v^\tau_k$ for any $k_1 \leqslant k \leqslant N$ and $u^\tau_k \geqslant v^\tau_k$ for any $k_0 \leqslant k \leqslant k_1$.

Proof. The argument is very similar to the one in Lemma A.3. We first define $w_k = u_k - v_k$, as well as the piecewise affine interpolation $\bar{w}$ of the values w_k, which satisfies again (A.3), but also $w'(T) \leqslant bw(T)$, where $T = N\tau$.

Then, we assume by contradiction that there is an open interval $I \subset (k_1\tau, N\tau)$ on which $\bar{w} > 0$. If $\bar{w} = 0$ on ∂I (i.e., on both points on the boundary), the argument is really the same. Otherwise, we can assume $I = (t, T)$, with $\bar{w}(t) = 0$. By multiplying by $\bar{w}$ and integrating by parts we get

$$\int_I |\bar{w}'|^2 = \bar{w}(T)\bar{w}'(T) - \int \bar{w}''\bar{w} \leqslant b|\bar{w}(T)|^2 + \omega^2 \sum_{k\,:\,k\tau\in I} \tau|w_k|^2.$$

Then, we use that on I we have

$$|\bar{w}| \leqslant \int_I |\bar{w}'| \leqslant \sqrt{|I| \int_I |\bar{w}'|^2}.$$

We do not have anymore the factor $1/2$ because $\bar{w}$ only vanishes at one end, now. This implies

$$\int_I |\bar{w}'|^2 \leqslant |I| \left(\omega^2\tau\, \#\{k : k\tau \in I\} + b\right) \int_I |\bar{w}'|^2.$$

Since $\#\{k : k\tau \in I\} < |N - k_1|$ and $|I| \leqslant |N - k_1|\tau$, using the assumptions on $|N - k_1|$ we have

$$\int_I |\bar{w}'|^2 \leqslant \left(\frac{\pi}{8} + \left(\frac{\pi}{8}\right)^2\right) \int_I |\bar{w}'|^2.$$

This is a contradiction, since

$$\frac{\pi}{8} + \left(\frac{\pi}{8}\right)^2 < \frac{1}{2} + \frac{1}{4} < 1. \qquad \square$$

With the two lemma above, we are able to deduce some Harnack-type inequality, which means that we can control the values of a u satisfying (A.1) in the interior of an interval with the values of u outside the interval.

Lemma A.5. *Let $k_1 < k_2$ such that $|k_2 - k_1|\tau\omega < \pi/8$ and assume $\tau \leqslant \tau_0$. Let k_0 (resp. k_3) be the smallest (resp. largest) integer smaller than k_1 (resp. larger than k_2) such that $(k_1 - k_0)\tau\omega < \pi/8$ (resp. $(k_3 - k_2)\tau\omega < \pi/8$). Then one has*

$$\sup_{k_1\leqslant k\leqslant k_2} u_k^\tau \leqslant C \max\left(\inf_{k_0\leqslant k\leqslant k_1} u_k^\tau\,,\ \inf_{k_2\leqslant k\leqslant k_3} u_k^\tau\right),$$

where the constant C is universal.

Proof. Given the symmetry of the property we want to prove w.r.t. to time reversal, we can assume that $u_{k_1}^\tau \leqslant u_{k_2}^\tau$. Let $v \in \mathcal{T}^\tau$ be the unique element of $\mathcal{T}^\tau$ such that $v_{k_1} = u_{k_1}^\tau$ and $v_{k_2} = u_{k_2}^\tau$. We know that it can be written in the form $v_k = A\cos(k\tau\omega + \delta)$ with $A \geqslant 0$. In particular, $A \geqslant |v_k|$ for any $k \in \mathbb{Z}$. Up to a time translation, we can assume that $\delta = 0$ and $k_1 \leqslant 0 \leqslant k_2$. By the hypothesis $u_{k_1}^\tau \leqslant u_{k_2}^\tau$, and $|k_2 - k_1|\tau\omega < \pi/8$, we can even say that $|k_2| \leqslant |k_1|$; thus, one has $k_2\tau \leqslant \pi/(16\omega)$. In particular, for any $k_2 \leqslant k \leqslant k_3$, we can say more than $v_k > 0$:

$$\begin{aligned} v_k &\geqslant A\cos(2\omega k_3\tau) \\ &\geqslant A\cos(2k_2\omega\tau + 2(k_3 - k_2)\omega\tau) \\ &\geqslant \cos\left(\frac{\pi}{8} + \frac{\pi}{4}\right) \sup_{k'\in\mathbb{Z}} |v_{k'}| \\ &\geqslant \frac{1}{C} \sup_{k'\in\mathbb{Z}} |v_{k'}|, \end{aligned}$$

with $C = \cos(3\pi/8)^{-1} < +\infty$. Thus, by using the comparison between u^τ and v (Lemma A.3), one can say that, for any $k_2 \leqslant k \leqslant k_3$,

$$u_k^\tau \geqslant \frac{1}{C} \sup_{k_1\leqslant k'\tau\leqslant k_2} u_{k'}^\tau,$$

which easily implies the claim. $\square$

We also provide the same type of lemma but where a different condition is imposed on the right end side, namely a Neumann-type boundary condition.

Lemma A.6. *Let $k_1 < N$ and $b \geqslant 0$ such that $|N-k_1|\tau \leqslant \min\{\pi/(32\omega), \pi/(32b)\}$ and assume $\tau \leqslant \tau_0$. Suppose $(u_N - u_{N-1})/\tau \leqslant bu_N$. Let k_0 be the smallest integer smaller than k_1 such that $(k_1-k_0)\tau \leqslant \min\{\pi/(32\omega), \pi/(32b)\}$. Then one has*

$$\sup_{k_1 \leqslant k \leqslant N} u_k^\tau \leqslant C \inf_{k_0 \leqslant k \leqslant k_1} u_k^\tau,$$

where the constant C is universal.

Proof. The strategy of the proof is the same than for Lemma A.5. We take v to be the unique element of $\mathcal{T}^\tau$ such that $v_{k_1} = u_{k_1}^\tau$ and $(v_N - v_{N-1})/\tau = bv_N$. We know that v is of the form $v_k = A\cos(2k\tau\omega + \delta)$. Up to a time translation, we can assume that $N\tau = 0$ and take $\delta \in (-\pi/2\,, \pi/2)$. Starting from $(v_N - v_{N-1})/\tau = bv_N$ and using well known factorization formulas, one ends up with

$$b = -2\omega\tan(\delta) + O(\omega\tau).$$

Thus, if $\tau \leqslant \tau_0$, one can say that $\arctan(-b/\omega) \leqslant \delta \leqslant \arctan(-b/(4\omega))$. Hence, using the fact that $\arctan(t) + \arctan(1/t) = -\pi/2$ (if $t < 0$) and that $\min\{\pi t/4, \pi/4\} \leqslant \arctan(t) \leqslant t$ (if $t \geqslant 0$), one concludes that

$$\min\left\{-\frac{\pi}{2} + \frac{\pi\omega}{4b}, -\frac{\pi}{4}\right\} \leqslant \delta \leqslant \min\left\{-\frac{\pi}{2} + \frac{4\omega}{b}, 0\right\}.$$

In other words, δ cannot be too close to $-\pi/2$, the point where the cosine vanishes. Given the information that we have on k_1 and k_0, one can check that

$$\begin{aligned}\delta - 2\omega\tau k_0 &= \delta - 2\omega\tau k_1 - 2\omega\tau(k_0 - k_1)\\ &\geqslant \min\left\{-\frac{\pi}{2} + \frac{\pi\omega}{4b}, -\frac{\pi}{4}\right\} - 2\min\left\{\frac{\pi}{16}, \frac{\pi\omega}{16b}\right\}\\ &\geqslant \min\left\{-\frac{\pi}{2} + \frac{\pi\omega}{8b}, -\frac{3\pi}{8}\right\}.\end{aligned}$$

As, for every $k_0 \leqslant k \leqslant N$, one has

$$A\cos(\delta - 2\omega\tau k_0) \leqslant v_k \leqslant A\cos(\delta),$$

it is easy to conclude that

$$\frac{\sup_{k_0 \leqslant k \leqslant N} v_k}{\inf_{k_0 \leqslant k \leqslant N} v_k} \leqslant \frac{\cos(\min\{-\frac{\pi}{2} + \frac{4\omega}{b}, 0\})}{\cos(\min\{-\frac{\pi}{2} + \frac{\pi\omega}{8b}, \frac{3\pi}{8}\})} \leqslant C,$$

where the value of C can be estimated by noting that if $\omega/b \ll 1$ both the numerator and the denominator are of the order of ω/b and if ω/b is not small the denominator is far from 0 and the numerator is bounded by 1. This proves that C is a universal constant. It remains to use Lemma A.4 to transfer the above inequality into an information on u^τ. □

To conclude, we can prove the Lemmas 4.3, 4.6 and 4.11 that we used throughout the paper, by using the above results. To prove Lemma 4.3, we cut the interval $[T_1\,, T_2]$ into several pieces of length of order $1/\omega$, on each piece we use the Harnack inequality to exchange the sum and the power $1/\beta$, and we use rough comparisons to put the pieces together.

Proof of Lemma 4.3. Let M be the smallest integer larger than $8\omega(T_2 - T_1)/\pi + 1$. We cut the interval $[T_1\,, T_2]$ into M closed intervals $I_1, I_2, \ldots, I_M$ of equal length (all equal to $(T_2 - T_1)/M < \pi/(8\omega)$). Let us choose an interval I_i, we can use Lemma A.5 to write

$$\begin{aligned}\left(\sum_{k\,:\,k\tau \in I_i} \tau u_k^\tau\right)^{1/\beta} &\leqslant (|I_i| + \tau)^{1/\beta} \sup_{k\tau \in I_i} (u_\tau^k)^{1/\beta}\\ &\leqslant C(|I_i| + \tau)^{1/\beta}\left(\inf_{T_1^i - \eta \leqslant k\tau \leqslant T_1^i} (u_k^\tau)^{1/\beta} + \inf_{T_2^i \leqslant k\tau \leqslant T_2^i + \eta} (u_k^\tau)^{1/\beta}\right)\\ &\leqslant C\frac{(|I_i| + \tau)^{1/\beta}}{\eta} \sum_{k\,:\,k\tau \in I_i \pm \eta} (u_k^\tau)^{1/\beta},\end{aligned}$$

where $I_i \pm \eta$ denotes the set of real numbers which are at a distance at most η of I_i. Then we put together the estimate for each I_i:

$$\begin{aligned}
\left(\sum_{T_1\leqslant k\tau\leqslant T_2} \tau u_k^\tau\right)^{1/\beta} &\leqslant \left(\sum_{i=1}^{M}\sum_{k\,:\,k\tau\in I_i} \tau u_k^\tau\right)^{1/\beta}\\
&\leqslant M^{1/\beta}\sum_{i=1}^{M}\left(\sum_{k\,:\,k\tau\in I_i} \tau u_k^\tau\right)^{1/\beta}\\
&\leqslant CM^{1/\beta}\left(\frac{T_2-T_1}{M}+\tau\right)^{1/\beta}\frac{1}{\eta}\sum_{i=1}^{M}\sum_{k\,:\,k\tau\in I_i\pm\eta}\tau\,(u_k^\tau)^{1/\beta}\\
&\leqslant C\frac{M(T_2-T_1+M\tau)^{1/\beta}}{\eta}\sum_{T_1-\eta\leqslant k\tau\leqslant T_2+\eta}\tau\,(u_k^\tau)^{1/\beta}\\
&\leqslant C\frac{(\omega+1)(T_2-T_1+1)^{1+1/\beta}}{\eta}\sum_{T_1-\eta\leqslant k\tau\leqslant T_2+\eta}\tau\,(u_k^\tau)^{1/\beta},
\end{aligned}$$

where we have used the fact that $M\tau \leqslant 1$ if $\tau \leqslant \tau_0$ (where τ_0 depends on ω) and also that M can be estimated by a constant times $\omega+1$. □

Proof of Lemma 4.6. For the first part, we apply Lemma A.6 with $k_1 = N$. With the choice of η, one has $(k_1 - k_0)\tau \leqslant \min\{\pi/(32\omega), \pi/(32b)\}$. Thus, one can write that

$$u_N^\tau \leqslant C \inf_{T-\eta\leqslant k\tau\leqslant T} u_k^\tau,$$

which is enough to to conclude as the r.h.s. is bounded by the mean of u_k^τ, for $T-\eta \leqslant k\tau \leqslant T$.

For the second part (which is a variant of Lemma 4.3, but with Neumann boundary conditions on one side), we can say with the help of Lemma A.6 that with k_1 the smallest integer smaller than N such that $|N-k_1|\tau\max\{\omega, b\} < \pi/32$,

$$\begin{aligned}
\left(\sum_{k_1\tau\leqslant k\tau\leqslant T} \tau u_k^\tau\right)^{1/\beta} &\leqslant |T-k_1\tau+\tau|^{1/\beta}\sup_{k_1\tau\leqslant k\tau\leqslant T}(u_k^\tau)^{1/\beta}\\
&\leqslant C|T-k_1\tau+\tau|^{1/\beta}\inf_{k_1\tau-\eta\leqslant k\tau\leqslant k_1\tau}(u_k^\tau)^{1/\beta}\\
&\leqslant \frac{C|T-k_1\tau+\tau|^{1/\beta}}{\eta}\sum_{k_1\tau-\eta\leqslant k\tau\leqslant k_1\tau}\tau(u_k^\tau)^{1/\beta}.
\end{aligned}$$

Then, we combine this estimate with the interior estimate Lemma 4.3 (with $T_2 = T - k_1\tau$) to end up with the announced result. □

Proof of Lemma 4.11. We apply Lemma A.6 with $k_1 = 0$. Thus if $T = kN \leqslant \min\{\pi/(32\omega), \pi/(32b)\}$, one has

$$\sup_{0\leqslant k\tau\leqslant T} u_k^\tau \leqslant Cu_0^\tau = Ca.$$

Thus, the l.h.s. is bounded by a constant which does not depend on N. □

Acknowledgments

Both authors acknowledge the support of the French ANR via the contracts ISOTACE (ANR-12-MONU-0013) and 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] Luigi Ambrosio and Alessio Figalli. On the regularity of the pressure field of Brenier's weak solutions to incompressible Euler equations. *Calculus of Variations and Partial Differential Equations*, 31(4):497–509, 2008.

[2] Luigi Ambrosio and Alessio Figalli. Geodesics in the space of measure-preserving maps and plans. *Archive for rational mechanics and analysis*, 194(2):421–462, 2009.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Birkhäuser Verlag, 2008.

[4] Luigi Ambrosio and Filippo Santambrogio. Necessary optimality conditions for geodesics in weighted Wasserstein spaces. *Rend. Lincei Mat. Appl.*, 18:23–37, 2007.

[5] Luigi Ambrosio and Paolo Tilli. *Topics on Analysis in Metric Spaces*. Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2003.

[6] Wolfgang Arendt. Semigroups and evolution equations: functional calculus, regularity and kernel estimates. *Handbook of Differential Equations: Evolutionary Equations*, 1:1–85, 2002.

[7] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[8] Jean-David Benamou, Guillaume Carlier, and Filippo Santambrogio. *Variational Mean Field Games*, pages 141–171. Springer International Publishing, Cham, 2017.

[9] Guy Bouchitté and Giuseppe Buttazzo. New lower semicontinuity results for nonconvex functionals defined on measures. *Nonlinear Analysis*, 15:679–692, 1990.

[10] Alessio Brancolini, Giuseppe Buttazzo, and Filippo Santambrogio. Path functionals over Wasserstein spaces. *Journal of the European Mathematical Society*, 8(3):415–434, 2006.

[11] Yann Brenier. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *Journal of the American Mathematical Society*, 2(2):225–255, 1989.

[12] Yann Brenier. Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Communications on pure and applied mathematics*, 52(4):411–452, 1999.

[13] Giuseppe Buttazzo, Chloé Jimenez, and Edouard Oudet. An optimization problem for mass transportation with congested dynamics. *SIAM Journal on Control and Optimization*, 48(3):1961–1976, 2009.

[14] Giuseppe Buttazzo and Filippo Santambrogio. A model for the optimal planning of an urban area. *SIAM journal on mathematical analysis*, 37(2):514–530, 2005.

[15] Pierre Cardaliaguet. Notes on mean field games. Technical report, CEREMADE, Univ. Paris-Dauphine, 2010.

[16] Pierre Cardaliaguet and P Jameson Graber. Mean field games systems of first order. *ESAIM: Control, Optimisation and Calculus of Variations*, 21(3):690–722, 2015.

[17] Pierre Cardaliaguet, Alpár R Mészáros, and Filippo Santambrogio. First order mean field games with density constraints: pressure equals price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016.

[18] Minyi Huang, Roland P Malhamé, and Peter E Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.

[19] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[20] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I–le cas stationnaire. *Comptes Rendus Mathématique*, 343(9):619–625, 2006.

[21] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II–Horizon fini et contrôle optimal. *Comptes Rendus Mathématique*, 343(10):679–684, 2006.

[22] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[23] Hugo Lavenant. Time-convexity of the entropy in the multiphasic formulation of the incompressible Euler equation. *arXiv preprint arXiv:1701.06391*, 2017.

[24] Pierre-Louis Lions. Jeux à champ moyen et applications. videos available on the website of Coll. de France, 2006-12.

[25] Daniel Matthes, Robert J McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34(11):1352–1397, 2009.

[26] Robert J McCann. A convexity principle for interacting gases. *Advances in mathematics*, 128(1):153–179, 1997.

[27] Jürgen 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):457–468, 1960.

[28] Michel Pierre. Uniqueness of the solutions of $u_t - \Delta\phi(u) = 0$ with initial datum a measure. *Nonlinear Analysis: Theory, Methods & Applications*, 6(2):175–187, 1982.

[29] Adam Prosinski and Filippo Santambrogio. Global-in-time regularity via duality for congestion-penalized Mean Field Games . *Stochastics*, in press.

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

[31] Filippo Santambrogio. Regularity via duality in calculus of variations and degenerate elliptic PDEs . *Journal of Mathematical Analysis and Applications*, in press.

[32] Cédric Villani. *Topics in optimal transportation*. Number 58 in Graduate Studies in Mathematics. American Mathematical Soc., 2003.